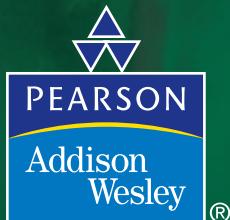
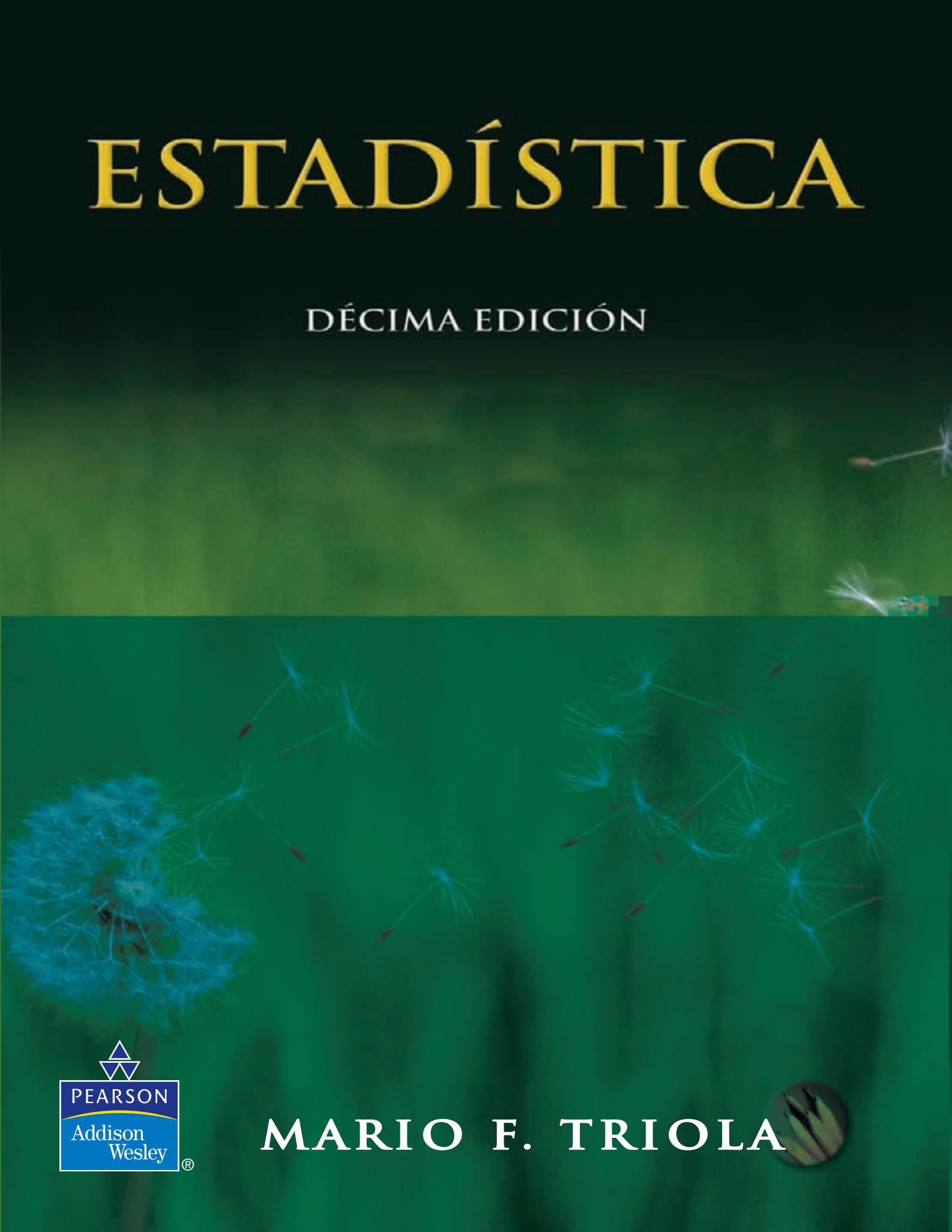


ESTADÍSTICA

DÉCIMA EDICIÓN



MARIO F. TRIOLA



ESTADÍSTICA

Décima edición

ESTADÍSTICA

Décima edición

Mario F. Triola

TRADUCCIÓN

Leticia Esther Pineda Ayala

Traductora profesional

REVISIÓN TÉCNICA

Roberto Hernández Ramírez

Departamento de Matemáticas

Universidad de Monterrey



México • Argentina • Brasil • Colombia • Costa Rica • Chile • Ecuador
España • Guatemala • Panamá • Perú • Puerto Rico • Uruguay • Venezuela

Datos de catalogación bibliográfica

TRIOLA, MARIO F.

Estadística. Décima edición

PEARSON EDUCACIÓN, México, 2009

ISBN: 978-970-26-1287-2

Área: Matemáticas

Formato: 21 × 27 cm

Páginas: 904

Authorized translation from the English Language edition, entitled *Elementary Statistics with Multimedia Study Guide, 10th Edition* by *Mario F. Triola*, published by Pearson Education Inc., publishing as Addison-Wesley, Copyright ©2008. All rights reserved.

ISBN 9780321460929

Versión en español de la obra titulada, *Elementary Statistics with Multimedia Study Guide, 10^a Edición* por *Mario F. Triola*, publicada originalmente en inglés por Pearson Education Inc., publicada como Addison-Wesley, Copyright ©2008. Todos los derechos reservados.

Esta edición en español es la única autorizada.

Edición en español

Editor: Rubén Fuerte Rivera

e-mail: ruben.fuerte@pearsoned.com

Editor de desarrollo: Felipe Hernández Carrasco

Supervisor de producción: Juan José García Guzmán

Edición en inglés

Publisher: Greg Tobin

Executive Editor: Deirdre Lynch

Executive Project Manager: Christine O'Brien

Assistant Editor: Sara Oliver

Managing Editor: Ron Hampton

Senior Production Supervisor: Peggy McMahon

Senior Designer: Barbara T. Atkinson

Photo Researcher: Beth Anderson

Digital Assets Manager: Marianne Groth

Production Coordinator, Supplements: Emily Portwood

Media Producer: Cecilia Fleming

Software Development: Ted Hartman and Janet Wann

Marketing Manager: Phyllis Hubbard

Marketing Assistant: Celena Carr

Senior Author Support/Technology Specialist: Joe Vetere

Senior Prepress Supervisor: Caroline Fell

Rights and Permissions Advisor: Dana Weightman

Senior Manufacturing Buyer: Evelyn Beaton

Text and Cover Design: Leslie Haimes

Production Services, Composition and Illustration: Nesbitt Graphics

Cover Photo: Getty Images/Jean Louis Batt

DÉCIMA EDICIÓN, 2009

D.R.© 2009 por Pearson Educación de México, S.A. de C.V.

Atlaculco Núm. 500, 5º Piso

Col. Industrial Atoto

53519, Naucalpan de Juárez, Edo. de México

Cámara Nacional de la Industria Editorial Mexicana. Reg. Núm. 1031

Addison-Wesley es una marca registrada de Pearson Educación de México, S.A. de C.V.

Reservados todos los derechos. Ni la totalidad ni parte de esta publicación pueden reproducirse, registrarse o transmitirse, por un sistema de recuperación de información, en ninguna forma ni por ningún medio, sea electrónico, mecánico, fotoquímico, magnético o electroóptico, por fotocopia, grabación o cualquier otro, sin permiso previo por escrito del editor.

El préstamo, alquiler o cualquier otra forma de cesión de uso de este ejemplar requerirá también la autorización del editor o de sus representantes.

ISBN 10: 970-26-1287-X

ISBN 13: 978-970-26-1287-2

Impreso en México. *Printed in Mexico.*

1 2 3 4 5 6 7 8 9 0 - 11 10 09 08

Para Marc y Scott

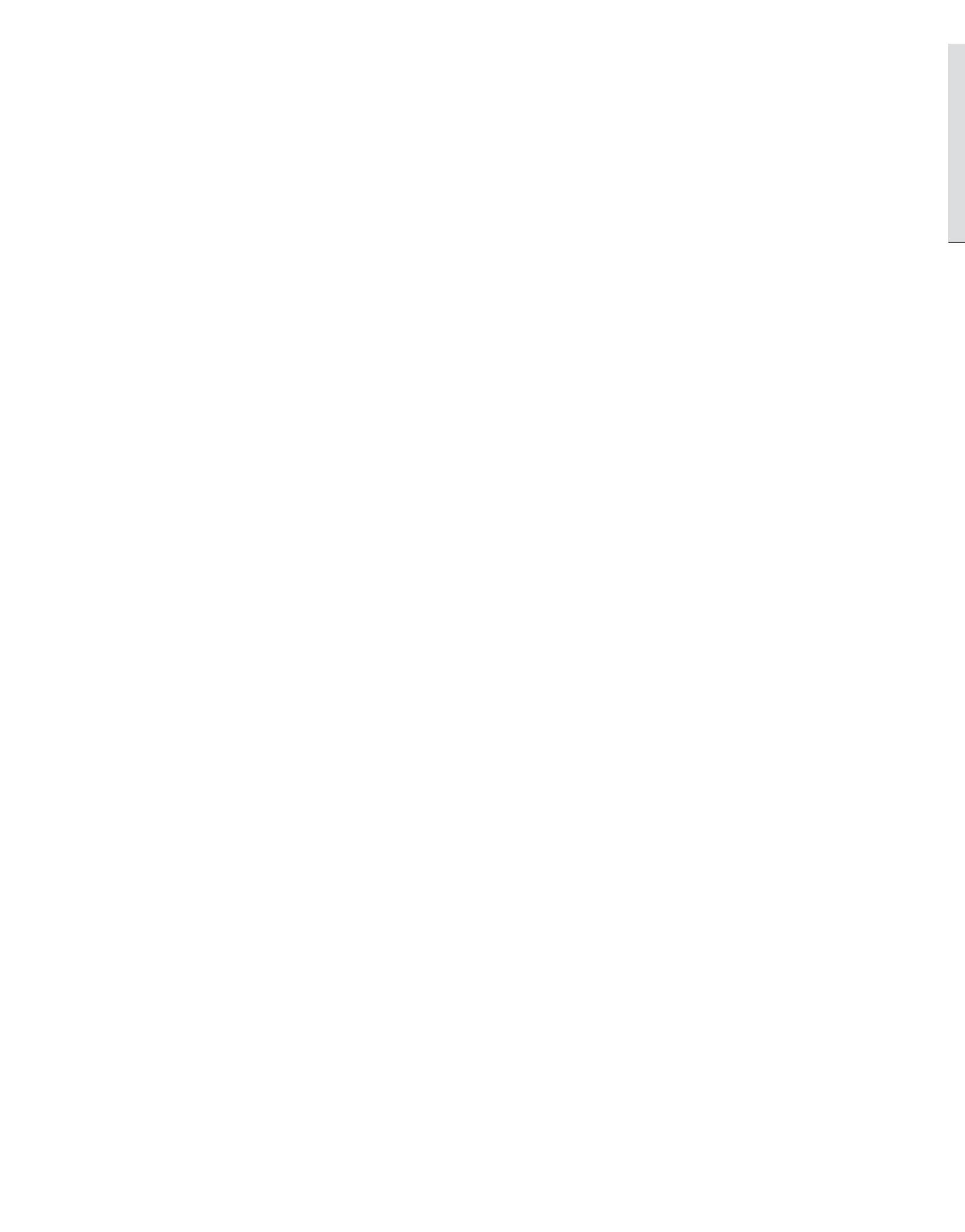


Acerca del autor

Mario F. Triola es profesor emérito de matemáticas en el Dutchess Community College, donde ha enseñado estadística durante más de 30 años. Marty es autor de las obras *Essentials of Statistics*, *Elementary Statistics Using Excel* y *Elementary Statistics Using the Graphic Calculator*; también es coautor de los libros *Biostatistics for the Biological and Health Sciences*, *Statistical Reasoning for Everyday Life* y *Business Statistics*. Ha escrito diversos manuales y libros de trabajo para educación en estadística con apoyos tecnológicos. Fuera del salón de clases, Marty ha sido orador en muchas conferencias y universidades. Su trabajo de consultoría incluye el diseño de máquinas tragamonedas para casinos y de cañas de pescar; ha trabajado con abogados en la determinación de probabilidades en casos de demandas de paternidad, en la identificación de desigualdades salariales entre géneros, en el análisis de resultados de elecciones en disputa, en el análisis de datos médicos y en el análisis de encuestas de escuelas de medicina. Por otro lado, fungió como testigo experto en la Suprema Corte del estado de Nueva York. La Text and Academic Authors Association otorgó a Mario F. Triola el premio *Texty* de Excelencia por su trabajo en el libro *Estadística*.

Tabla de contenido abreviada

1	Introducción a la estadística	2
2	Resumen y gráficas de datos	40
3	Estadísticos para describir, explorar y comparar datos	74
4	Probabilidad	136
5	Distribuciones de probabilidad	198
6	Distribuciones de probabilidad normal	244
7	Estimaciones y tamaños de muestra	318
8	Prueba de hipótesis	384
9	Inferencias a partir de dos muestras	454
10	Correlación y regresión	514
11	Experimentos multinomiales y tablas de contingencia	588
12	Análisis de varianza	634
13	Estadística no paramétrica	674
14	Control estadístico de procesos	732
15	Proyectos, procedimientos y perspectivas	760
	Apéndices	767
	Apéndice A: Tablas	768
	Apéndice B: Conjuntos de datos	785
	Apéndice C: Glosario	808
	Apéndice D: Bibliografía	816
	Apéndice E: Soluciones de los ejercicios impares (y de todos los ejercicios de repaso y de los ejercicios de repaso acumulativo)	817
	Créditos	855
	Índice	857





Contenido

1 Introducción a la estadística 2

- 1-1 Panorama general 4
- 1-2 Tipos de datos 5
- 1-3 Pensamiento crítico 11
- 1-4 Diseño de experimentos 21

2 Resumen y gráficas de datos 40

- 2-1 Panorama general 42
- 2-2 Distribuciones de frecuencias 42
- 2-3 Histogramas 51
- 2-4 Gráficas estadísticas 56

3 Estadísticos para describir, explorar y comparar datos 74

- 3-1 Panorama general 76
- 3-2 Medidas de tendencia central 76
- 3-3 Medidas de variación 92
- 3-4 Medidas de posición relativa 110
- 3-5 Análisis exploratorio de datos (AED) 119

4 Probabilidad 136

- 4-1 Panorama general 138
- 4-2 Fundamentos 138
- 4-3 Regla de la suma 151
- 4-4 Regla de la multiplicación: Fundamentos 159
- 4-5 Regla de la multiplicación: Complementos y probabilidad condicional 168
- 4-6 Probabilidades por medio de simulaciones 174
- 4-7 Conteo 179
- 4-8 Teorema de Bayes (en CD-ROM) 190

5 Distribuciones de probabilidad discreta 198

- 5-1 Panorama general 200
- 5-2 Variables aleatorias 201
- 5-3 Distribuciones de probabilidad binomial 213
- 5-4 Media, varianza y desviación estándar para la distribución binomial 225
- 5-5 Distribuciones de probabilidad de Poisson 230

6 Distribuciones de probabilidad normal 244

- 6-1 Panorama general 246
- 6-2 La distribución normal estándar 247
- 6-3 Aplicaciones de las distribuciones normales 259
- 6-4 Distribuciones muestrales y estimadores 269
- 6-5 El teorema del límite central 280
- 6-6 La distribución normal como aproximación de la distribución binomial 291
- 6-7 Determinación de la normalidad 302

7 Estimados y tamaños de muestra 318

- 7-1 Panorama general 320
- 7-2 Estimación de la proporción de una población 320
- 7-3 Estimación de una media poblacional: σ conocida 338
- 7-4 Estimación de una media poblacional: σ desconocida 349
- 7-5 Estimación de la varianza de una población 363

8 Prueba de hipótesis 384

- 8-1 Panorama general 386
- 8-2 Fundamentos de la prueba de hipótesis 387
- 8-3 Prueba de una aseveración respecto de una proporción 407
- 8-4 Prueba de una aseveración respecto de una media: σ conocida 418
- 8-5 Prueba de una aseveración respecto de una media: σ desconocida 426
- 8-6 Prueba de una aseveración respecto de una desviación estándar o de una varianza 436

9 Inferencias a partir de dos muestras 454

- 9-1 Panorama general 456
- 9-2 Inferencias acerca de dos proporciones 456
- 9-3 Inferencias acerca de dos medias: Muestras independientes 469
- 9-4 Inferencias a partir de datos apareados 484
- 9-5 Comparación de la variación en dos muestras 495

10 Correlación y regresión 514

- 10-1 Panorama general 517
- 10-2 Correlación 517
- 10-3 Regresión 541
- 10-4 Variación e intervalos de predicción 557
- 10-5 Regresión múltiple 566
- 10-6 Elaboración de modelos 576

11 Experimentos multinomiales y tablas de contingencia 588

- 11-1 Panorama general 590
- 11-2 Experimentos multinomiales: Bondad de ajuste 591
- 11-3 Tablas de contingencia: Independencia y homogeneidad 606
- 11-4 Prueba de McNemar para datos apareados 621

12 Análisis de varianza 634

- 12-1 Panorama general 636
- 12-2 ANOVA de un factor 637
- 12-3 ANOVA de dos factores 655

13 Estadística no paramétrica 674

- 13-1 Panorama general 676
- 13-2 Prueba del signo 678
- 13-3 Prueba de rangos con signo de Wilcoxon para datos apareados 689
- 13-4 Prueba de la suma de rangos de Wilcoxon para dos muestras independientes 695
- 13-5 Prueba de Kruskal-Wallis 702
- 13-6 Correlación de rangos 708
- 13-7 Prueba de rachas para detectar aleatoriedad 717

14 Control estadístico de procesos 732

- 14-1 Panorama general 734
- 14-2 Gráficas de control para la variación y la media 734
- 14-3 Gráficas de control para atributos 748

15 Proyectos, procedimientos y perspectivas 760

- 15-1 Proyectos 760
- 15-2 Procedimientos 765
- 15-3 Perspectivas 767

Apéndices 767

Apéndice A:	Tablas 768
Apéndice B:	Conjuntos de datos 785
Apéndice C:	Glosario 808
Apéndice D:	Bibliografía 816
Apéndice E:	Soluciones de los ejercicios impares (y de todos los ejercicios de repaso y de los ejercicios de repaso acumulativo) 817

Créditos 855**Índice 857**

Prefacio



Filosofía

Estadística, décima edición, es el resultado de más de 30 años de enseñanza, investigación e innovación en la instrucción de la estadística. La meta de este libro es que se convierta en una introducción interesante y detallada a la estadística para los estudiantes. Aunque a lo largo del texto se encuentran fórmulas y procedimientos formales, destaca el desarrollo de conocimientos estadísticos y de un pensamiento crítico. Este libro fomenta el pensamiento más allá del uso irreflexivo de procedimientos mecánicos.

Estadística ha sido el principal libro de texto de introducción a la estadística en Estados Unidos durante muchos años. Al llegar a millones de estudiantes, se ha convertido en el libro de estadística más vendido de todos los tiempos. Las siguientes son algunas características importantes que han contribuido a su éxito continuo:

- Énfasis en los conocimientos estadísticos y en un pensamiento crítico
- Énfasis en la comprensión de los conceptos y no en la realización de cálculos de forma mecánica
- Uso abundante de datos *reales*
- Un estilo de escritura claro, fácil de entender y en ocasiones con sentido del humor
- Componentes pedagógicos abundantes y diversos
- Una gama de complementos útiles para los estudiantes y los profesores
- Profesionales de ventas, técnicos, de apoyo y editoriales de Addison-Wesley con un compromiso y experiencia excepcionales

Además de enseñar estadística, otro objetivo importante de *Estadística*, décima edición, es brindar un marco de referencia que fomente el crecimiento personal a través del uso de la tecnología, el trabajo con los compañeros, el pensamiento crítico y el desarrollo de habilidades de comunicación. *Estadística* permite que los estudiantes apliquen las habilidades adquiridas fuera del salón de clases, en un contexto del mundo real.

Este texto obedece las recomendaciones y lineamientos de la American Statistical Association, la Mathematical Association of America, la American Mathematical Association of Two-Year Colleges y el National Council of Teachers of Mathematics.

Público/Prerrequisitos

El libro *Estadística* se escribió para estudiantes de cualquier carrera. Aun cuando el uso del álgebra es mínimo, los estudiantes deben haber cursado al menos una materia de álgebra elemental en la preparatoria o la universidad. En muchos casos se incluyen teorías subyacentes, pero este libro no enfatiza el rigor matemático que es más adecuado para carreras en matemáticas. Puesto que la gran cantidad de

ejemplos y ejercicios cubren una amplia variedad de aplicaciones estadísticas distintas e interesantes, *Estadística* es apropiado para estudiantes de una gran diversidad de disciplinas, que van desde las ciencias sociales, la psicología y la sociología, hasta áreas tales como la educación, los campos de la salud, negocios, economía, ingeniería, humanidades, ciencias físicas, periodismo, comunicación y artes.

Herramientas tecnológicas

Estadística, décima edición, puede utilizarse fácilmente sin referencia a alguna tecnología específica. Muchos profesores continúan usando las distintas ediciones de este libro con estudiantes, con tan sólo una variedad de calculadoras científicas. Sin embargo, para aquellos que deciden complementar el curso con herramientas tecnológicas, se incluye material específico dentro del texto, aunque también existen materiales complementarios disponibles.

Cambios en esta edición

- La sección de Visualización de datos se ha dividido en dos secciones, dando así un mayor énfasis en las gráficas estadísticas:

Sección 2-3: Histogramas

Sección 2-4: Gráficas estadísticas

- El capítulo en ediciones anteriores referente a la Descripción, exploración y comparación de datos se dividió en dos capítulos:

Capítulo 2: Resumen y gráficas de datos

Capítulo 3: Estadísticos para describir, explorar y comparar datos

- Nueva sección: **Prueba de McNemar para datos apareados** (sección 11-4)
- En algunas secciones, el libro se ha dividido en Parte 1 (aspectos básicos) y Parte 2 (más allá de lo básico) para facilitar su enfoque en conceptos centrales.
- El análisis de ciertos temas se ha ampliado: potencia (sección 8-2); gráficas residuales (sección 10-3); regresión logística (sección 10-5); y gráficas de interacción (sección 12-3).
- **Verificación de requisitos:** Cuando se considera pertinente, las soluciones inician con una verificación formal de los requisitos que deben cubrirse antes de utilizar un método en particular.
- **Conocimientos estadísticos y pensamiento crítico:** Cada sección de ejercicios inicia con cuatro ejercicios que implican específicamente conocimientos estadísticos y un pensamiento crítico. Asimismo, al final a cada capítulo se incluyen cuatro ejercicios más de este tipo.
- **Respuestas de herramientas tecnológicas:** Las respuestas en el apéndice E se basan en el uso de tablas, pero también se incluyen respuestas de las herramientas tecnológicas cuando existen discrepancias. Por ejemplo, una respuesta aparece como “valor P : 0.2743 (herramienta tecnológica: 0.2739)”, donde “herramienta tecnológica” indica la respuesta que se obtendría utilizando un programa como STATDISK, Minitab, Excel o una calculadora TI-83/84 Plus. Además, siempre que es posible, se utilizan los valores P en la mayoría de las respuestas.
- **Conjuntos pequeños de datos:** En esta edición se incluye un número mucho mayor de ejercicios con conjuntos pequeños de datos.

- **Nuevos ejercicios de ejemplos:** El 68% de los ejercicios son nuevos, y el 53% de los ejercicios incluyen datos reales. El 66% de los ejemplos son nuevos.
- **Los 20 temas más importantes (Top 20):** En esta edición identificamos los **20 temas más importantes** en cualquier curso de introducción a la estadística, los cuales aparecen marcados en el texto como .

Contenido flexible al plan de estudios

La organización de este libro refleja las preferencias de la mayoría de los profesores de estadística, pero existen dos variaciones comunes que pueden ser fácilmente utilizadas en esta décima edición:

- **Pronta cobertura de correlación/regresión:** Algunos profesores prefieren cubrir los aspectos básicos de la correlación y la regresión al inicio del curso, inmediatamente después de los temas del capítulo 3. *Las secciones 10-2 (correlación) y 10-3 (regresión) pueden cubrirse en las primeras etapas del curso.* Simplemente omita la cobertura de la Parte 1 (conceptos básicos) en cada una de las dos secciones.
- **Poco contenido del tema de probabilidad:** Algunos profesores consideran que el tema de probabilidad debe cubrirse en forma extensa, mientras que otros prefieren cubrirlo en forma mínima. Estos últimos pueden incluir la sección 4-2 y omitir las secciones restantes del capítulo 4, ya que no son esenciales para los capítulos posteriores. Muchos profesores prefieren cubrir sólo los fundamentos de la probabilidad, junto con los aspectos básicos de las reglas de la suma y la multiplicación; estos temas se pueden cubrir con las secciones 4-1 a 4-4. La sección 4-5 incluye la probabilidad condicional, y las secciones posteriores se ocupan de los métodos de simulación y conteo (incluyendo las permutaciones y las combinaciones).

Ejercicios

Se incluyen más de 1750 ejercicios *¡y el 68% de éstos son nuevos!* Un número mayor de ejercicios utilizan conjuntos más pequeños de datos y muchos de ellos requieren la *interpretación* de los resultados. En virtud de que los ejercicios son de gran importancia en cualquier libro de estadística, se ha tenido gran cuidado para asegurar su utilidad, relevancia y exactitud. Tres especialistas en estadística leyeron el libro en sus etapas finales para verificar la precisión del material del texto y de las respuestas a los ejercicios. Los ejercicios se acomodaron en orden de dificultad creciente dividiéndolos en dos grupos: **1.** Destrezas y conceptos básicos, y **2.** Más allá de lo básico, los cuales incluyen conceptos más difíciles o requieren de un acervo matemático más sólido. En pocos casos estos ejercicios también presentan un concepto nuevo.

Datos reales: *El 53% de los ejercicios utilizan datos reales.* (Puesto que esta edición tiene muchos más ejercicios en la sección de Conocimientos estadísticos y pensamiento crítico, el porcentaje de ejercicios que utilizan datos reales es menor que en la novena edición, pero el número de ejercicios que utilizan datos reales es aproximadamente el mismo). Como el uso de datos reales es tan importante para los estudiantes, se dedicaron cientos de horas para encontrar información real, sig-

nificativa e interesante. Además de los datos reales incluidos a lo largo del libro, muchos ejercicios se refieren a los 18 conjuntos grandes de datos listados en el apéndice B.

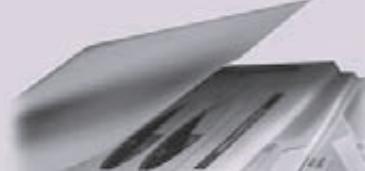
Características distintivas

Se ha tenido mucho cuidado para asegurar que cada capítulo de *Estadística* ayude a los estudiantes a comprender los conceptos presentados. Las siguientes características se diseñaron para lograr este objetivo:



- **Características del inicio de cada capítulo:** Se incluye una lista de secciones que presentan el capítulo al estudiante; un problema que inicia el capítulo, basado en datos reales, motiva el estudio del material presentado, y la primera sección es un panorama general que establece los objetivos del capítulo.
- **Características del final de cada capítulo:** Un Repaso del capítulo resume los conceptos y temas principales; los ejercicios sobre Conocimientos estadísticos y pensamiento crítico enfatizan los conceptos del capítulo; los Ejercicios de repaso permiten revisar los conceptos y procedimientos del capítulo.
- **Los Ejercicios de repaso acumulativo** refuerzan el material que se estudió con anterioridad.
- **De los datos a la decisión: Pensamiento crítico** es un problema final que requiere de pensamiento crítico y de habilidades de redacción;

De los datos a la decisión

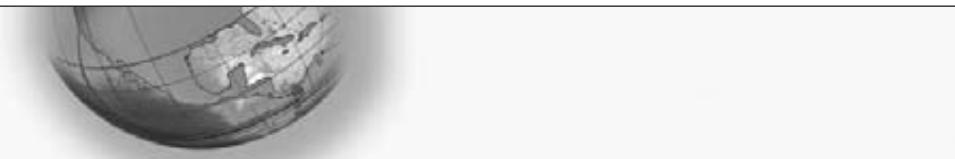


- Las **Actividades de cooperación en equipo** fomentan el aprendizaje activo en grupos;
- Los **Proyectos tecnológicos** requieren del uso de STATDISK, Minitab, Excel o de una calculadora TI-83/84 Plus.

Uso de la tecnología



STATDISK
MINITAB
EXCEL
T1-83/84 PLUS



Proyecto de Internet

- **Ensayos al margen:** El texto incluye 122 ensayos al margen, que ilustran los usos y abusos de la estadística en aplicaciones reales, prácticas e interesantes. Incluyen temas como *¿Prevalece un género en las familias?*, *¿Los zurdos mueren antes?* y *Elección de números de lotería*.
- **Diagramas de flujo:** Éstos aparecen a lo largo del texto para simplificar y aclarar conceptos y procedimientos más complejos. Como novedad en esta edición, los diagramas de flujo están animados y pueden revisarse en la página de Internet de MyStatLab de este libro (www.mystatlab.com).
- **Programas estadísticos de cómputo:** A lo largo del libro se encuentran instrucciones y resultados de STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus.
- **Conjuntos de datos reales:** Se usan extensamente en todo el libro. En el apéndice B se listan 18 conjuntos de datos, 4 de los cuales son nuevos y 3 que incluyen datos nuevos. Estos conjuntos aparecen de forma impresa en el apéndice B, y en forma electrónica en el sitio de Internet. Los conjuntos de datos se refieren a temas tan variados como el consumo de alcohol y tabaco en películas infantiles de dibujos animados, las erupciones del géiser Old Faithful y mediciones relacionadas con el tabaquismo pasivo.
- **Entrevistas:** Cada capítulo incluye entrevistas realizadas por el autor a hombres y mujeres profesionales de diversos campos que utilizan la estadística en su trabajo diario.
- **Tablas de referencia rápida:** Las tablas A-2 y A-3 (referentes a la distribución normal y distribución *t*) están reproducidas en la guarda al final del libro y en la tercera de forros. Al principio del libro se incluye una tabla de símbolos, para poder consultar con rapidez los símbolos clave.
- **Inserto de fórmulas y tablas desprendible:** Este material está organizado por capítulos y ofrece a los estudiantes una referencia rápida para estudio o, si lo permiten los profesores, para contestar exámenes.
- **CD-ROM complementario:** El CD-ROM fue elaborado por Mario F. Triola y viene incluido con cada nuevo ejemplar del texto; incluye los conjuntos de datos del apéndice B, que vienen almacenados como archivos de texto, hojas de cálculo de Minitab, archivos de SPSS, archivos de SAS, hojas de cálculo de Excel y aplicaciones de la calculadora TI-83/84 Plus. El disco compacto también incluye una sección sobre el teorema de Bayes, programas para la calculadora graficadora TI-83/84 Plus®, el programa estadístico STATDISK (versión 10.1) y el recurso “Add-Inn” de Excel, diseñado para incrementar las capacidades de los programas estadísticos de Excel.

Complementos

Los paquetes complementarios para el profesor tienen el objetivo de conformar el sistema de aprendizaje más completo y útil disponible para un curso de introducción a la estadística. Los profesores deben ponerse en contacto con su representante local de ventas de Pearson Educación o enviar un correo electrónico a la compañía, a la dirección editorialmx@pearsoned.com, para recibir copias de los exámenes.

Para el profesor

- **Manual de soluciones para el profesor**, escrito por Milton Loyer (Penn State University), contiene soluciones a todos los ejercicios y ejemplos del curso.
- **Guía de enseñanza para la serie de Estadística de Triola**, escrita por Mario F. Triola, contiene ejemplos de planes de estudio de estadística y consejos para incorporar proyectos, así como también panoramas generales de las lecciones, ejemplos adicionales, objetivos breves y tareas recomendadas para cada capítulo.
- **MyStatLab** (que forma parte de la familia de productos MyMathLab y MathXL) es un curso en línea específico para el libro y fácil de adaptar, que integra una instrucción multimedia interactiva con el contenido del libro de texto. MyStatLab está fortalecido por CourseCompass™—el entorno de enseñanza en línea de Pearson Educación—y por MathXL®—nuestro sistema de tareas, tutorial y evaluación en línea. MyStatLab le ofrece las herramientas necesarias para impartir todo su curso o una parte de él en línea, ya sea que los estudiantes se encuentren en un ambiente de laboratorio o en su hogar. MyStatLab ofrece un conjunto rico y flexible de materiales para el curso, incluyendo ejercicios de respuesta libre para una práctica y dominio ilimitados. Los profesores pueden utilizar los administradores de tareas y exámenes de MyStatLab para seleccionar y asignar ejercicios en línea relacionados directamente con el libro; también pueden crear y asignar sus propios ejercicios en línea, así como importar exámenes TestGen para añadir flexibilidad. El libro de calificaciones en línea de MyStatLab—diseñado específicamente para matemáticas y estadística—registra automáticamente los resultados de las tareas y los exámenes de los estudiantes, y permite que el profesor determine la forma de calcular las calificaciones finales. Los profesores también pueden añadir calificaciones no obtenidas en línea (sino con lápiz y papel) al libro de calificaciones. MyStatLab está disponible para practicantes autorizados.

Para mayor información póngase en contacto con su representante de ventas de Pearson Educación.

- **Sistema de evaluación:** Se tuvo gran cuidado en asegurar el sistema de evaluación más sólido para la nueva edición de *Estadística*. Además de un banco de exámenes impreso, también existe un generador de exámenes computarizado, el **TestGen**, que permite al profesor ver y editar preguntas del banco de exámenes, transferirlas a otros exámenes y realizar impresiones en diversos formatos. El programa también ofrece muchas opciones para organizar y presentar los bancos de exámenes y los exámenes mismos. Gracias a su capacidad de elaboración aleatoria y a su generador de exámenes, el **TestGen** resulta ideal para crear múltiples versiones de exámenes, ya que ofrece mayor posibilidad de reactivos de exámenes que las preguntas del banco de reactivos impresas. Los usuarios pueden exportar los exámenes para que sean compatibles con diversos sistemas de administración de cursos o incluso para que aparezca en un navegador de Internet. Además, las

pruebas creadas con TestGen pueden utilizarse con el **QuizMaster**, el cual permite al estudiante resolver exámenes a través de una computadora.

Para el estudiante

- **MathXL® para Estadística** es un poderoso sistema que complementa los libros de texto de estadística y matemáticas de Pearson Educación, el cual ofrece tareas, evaluaciones y tutoriales en línea. Con la herramienta MathXL para Estadística, los profesores pueden crear, editar y asignar tareas en línea, creadas específicamente para el libro de texto de Triola, así como pruebas que utilizan ejercicios generados de manera algorítmica, correlacionados con el nivel de los objetivos de este libro. Todo el trabajo de los estudiantes se registra en el libro de calificaciones en línea de MathXL. Los estudiantes pueden resolver exámenes de capítulos en esta herramienta y recibir planes de estudio personalizados a partir de sus resultados. El plan de estudio diagnostica debilidades y vincula a los estudiantes directamente con ejercicios tutoriales para los objetivos que necesitan estudiar y reevaluar. Los estudiantes también pueden revisar animaciones y clips de video de Triola directamente a partir de ejercicios seleccionados. MathXL para Estadística está disponible para los practicantes autorizados. Para mayor información, póngase en contacto con su representante de ventas de Pearson Educación.
- **Página de Internet de Estadística de Triola:** Se puede acceder a este sitio en <http://www.pearsoneducacion.net/triola>. Este sitio ofrece proyectos de Internet relacionados con cada uno de los capítulos del texto, así como los conjuntos de datos.

RECONOCIMIENTOS

Esta décima edición de *Estadística* es particularmente especial. Estoy muy agradecido con los miles de profesores de estadística que han contribuido al éxito de este libro. Agradezco en particular a mis alumnos, quienes desempeñaron un papel fundamental en la creación de un método de enseñanza efectivo que pudiera traducirse en un libro de texto, y a los numerosos estudiantes que han aprendido con este libro y que gentilmente han expresado muchos comentarios útiles.

El éxito de *Estadística* se puede atribuir al compromiso y dedicación de todo el equipo de Pearson Educación, y expreso mi más sincero agradecimiento a Deirdre Lynch, Christine O'Brien, Greg Tobin, Peggy McMahon, Barbara Atkinson, Phyllis Hubbard, Ceci Fleming, Celena Carr, Sara Oliver, Joe Vetere, Beth Anderson y Dana Weightman. También agradezco a Janet Nuciforo de Nesbitt Graphics por su excelente trabajo de producción.

Este libro no habría sido posible sin el apoyo de mi familia. Agradezco a mi esposa Ginny por su constante apoyo y guía, a mi hijo Scott por animarme continuamente, y a mi hijo Marc Triola, doctor en medicina, por reprogramar y fortalecer el STATDISK, que ahora es un programa poderoso y de calidad.

De entre los muchos trabajadores de Pearson Educación, me gustaría agradecer y reconocer personalmente las contribuciones de los representantes de ventas y de los gerentes de ventas que han sido muy útiles al atender a los profesores que utilizan este libro. Ha sido un placer absoluto trabajar con los siguientes profesionales durante 10 años o más:

Paul Altier
Jay Beckenstein
Eileen Burke
John Cross
Andrew Crowley
Julie Davis
Karin DeJamaer
Margaret Dzierzanowski

Peter Harris
Nancy Hart
Jim Lawler
Bill Leonard
Steve May
Tom Shaffer
Otis Taylor
Julie Ward

También me gustaría agradecer especialmente a los siguientes representantes de ventas veteranos que han vendido diversas ediciones de *Estadística*:

Nola Akala
Allison Andrews
Naomi Bahary
Michael Bailey
Corinn Berman
Carol Britz
Kathy Campbell
Dave Chwalik
Jamie Commissaris
Michelle Cook

Susan Coughlin
Tami Dreyfus
Jane Fleming
Matthew Genaway
Rhonda B. Goedeker
Lori Hales
Leigh Jacka
Jay Johnson
Laura C. Johnson
Jennifer Koehler

Ann Kuick	Leah Newman
Dara Lanier	Teri Orr
Mary Kaye Leonard	Amanda Perdaris
Donna Loughman	Scott Perrine
Martha McDonald	Marisa Raffaele
Richard McMenamy	Nick Rumpff
Lee Monroe	Karen Scholz
Lorri Morgan	Eugene Smith
Tracy Morse	Pam Snow
Linda Nelson	Frank Steed

Me gustaría agradecer a las siguientes personas por su ayuda a la décima edición:

Revisores de estilo

Emily Keaton	Tim Mogill
David R. Lund, University of Wisconsin at Eau Claire	Kimberly Polly, Parkland College
	Tom Wegleitner

Revisores de la décima edición

Raid W. Amin, University of West Florida	Carla Monticelli, Camden County Community College
Keith Carroll, Benedictine University	Julia Norton, California State University Hayward
Monte Cheney, Central Oregon Community College	Michael Oriolo, Herkimer Community College
Christopher Donnelly, Macomb Community College	Jeanne Osborne, Middlesex Community College
Theresa DuRapau, Our Lady of Holy Cross	Ali Saadat, University of California—Riverside
Billy Edwards, University of Tennessee—Chattanooga	Radha Sankaran, Passaic County Community College
Marcos Enriquez, Moorpark College	Pradipta Seal, Boston University
Angela Everett, Chattanooga State Technical Community College	Sharon Testone, Onondaga Community College
Joe Franko, Mount San Antonio College	Dave Wallach, University of Findlay
Sanford Geraci, Broward Community College	Cheng Wang, Nova Southeastern University
Laura Heath, Palm Beach Community College	Gail Wiltse, St. John River Community College
Laura Hillerbrand, Broward Community College	Claire Wladis, Borough of Manhattan Community College
Gary King, Ozarks Technical Community College	Yong Zeng, University of Missouri at Kansas City
Mickey Levendusky, Pima County Community College	Jim Zimmer, Chattanooga State Technical Community College
Tristan Londre, Blue River Community College	Cathleen Zucco-Teveloff, Trinity College
Alma Lopez, South Plains College	Mark Z. Zuiker, Minnesota State University, Mankato

Por su ayuda y sugerencias en áreas especiales, agradezco a las siguientes personas:

Vincent DiMaso	David Straayer, Sierra College
Rod Elsdon, Chaffey College	Glen Weber, Christopher Newport University

Por su ayuda al probar y mejorar el programa STATDISK, agradezco a los siguientes colaboradores:

Justine Baker
Henry Feldman, M.D.
Robert Jackson
Caren McClure
Sr. Eileen Murphy

John Reeder
Carolyn Renier
Cheryl Slayden
Victor Strano
Gary Turner

Por sus sugerencias, quiero expresar mi más sincero agradecimiento a los siguientes revisores y usuarios de ediciones anteriores de este libro:

Dan Abbey, Broward Community College
Mary Abkemeier, Fontbonne College
William A. Ahroon, Plattsburgh State
Scott Albert, College of Du Page
Jules Albertini, Ulster County Community College
Tim Allen, Delta College
Stu Anderson, College of Du Page
Jeff Andrews, TSG Associates, Inc.
Mary Anne Anthony, Rancho Santiago Community College
William Applebaugh, University of Wisconsin–Eau Claire
James Baker, Jefferson Community College
Justine Baker, Peirce College, Philadelphia, PA
Anna Bampton, Christopher Newport University
Donald Barrs, Pellissippi State Technical Community College
James Beatty, Burlington County College
Philip M. Beckman, Black Hawk College
Marian Bedee, BGSU, Firelands College
Marla Bell, Kennesaw State University
Don Benbow, Marshalltown Community College
Michelle Benedict, Augusta College
Kathryn Benjamin, Suffolk County Community College
Ronald Bensema, Joliet Junior College
David Bernklau, Long Island University
Maria Betkowski, Middlesex Community College
Shirley Blatchley, Brookdale Community College
David Balueuer, University of Findlay
Randy Boan, Aims Community College
John Bray, Broward Community College-Central
Denise Brown, Collin County Community College

Patricia Buchanan, Pennsylvania State University
John Buchl, John Wood Community College
Michael Butler, Mt. San Antonio College
Jerome J. Cardell, Brevard Community College
Don Chambliss, Auburn University
Rodney Chase, Oakland Community College
Bob Chow, Grossmont College
Philip S. Clarke, Los Angeles Valley College
Darrell Clevidence, Carl Sandburg College
Paul Cox, Ricks College
Susan Cribelli, Aims Community College
Imad Dakka, Oakland Community College
Arthur Daniel, Macomb Community College
Gregory Davis, University of Wisconsin, Green Bay
Tom E. Davis, III, Daytona Beach Community College
Charles Deeter, Texas Christian University
Joseph DeMaio, Kennesaw State University
Joe Dennin, Fairfield University
Nirmal Devi, Embry Riddle Aeronautical University
Richard Dilling, Grace College
Rose Dios, New Jersey Institute of Technology
Dennis Doverspike, University of Akron
Paul Duchow, Pasadena City College
Bill Dunn, Las Positas College
Marie Dupuis, Milwaukee Area Technical College
Evelyn Dwyer, Walters State Community College
Jane Early, Manatee Community College
Wayne Ehler, Anne Arundel Community College
Sharon Emerson-Stonnell, Longwood College
P. Teresa Farnum, Franklin Pierce College
Ruth Feigenbaum, Bergen Community College
Vince Ferlini, Keene State College
Maggie Flint, Northeast State Technical Community College

Reconocimientos

Bob France, Edmonds Community College
 Christine Franklin, University of Georgia
 Richard Fritz, Moraine Valley Community College
 Maureen Gallagher, Hartwick College
 Joe Gallegos, Salt Lake Community College
 Mahmood Ghamsary, Long Beach City College
 Tena Golding, Southeastern Louisiana University
 Elizabeth Gray, Southeastern Louisiana University
 Jim Graziose, Palm Beach Community College
 David Gurney, Southeastern Louisiana University
 Francis Hannick, Mankato State University
 Sr. Joan Harnett, Molloy College
 Kristin Hartford, Long Beach City College
 Leonard Heath, Pikes Peak Community College
 Peter Herron, Suffolk County Community College
 Mary Hill, College of Du Page
 Larry Howe, Rowan College of New Jersey
 Lloyd Jaisingh, Morehead State University
 Lauren Johnson, Inver Hills Community College
 Martin Johnson, Gavilan College
 Roger Johnson, Carleton College
 Herb Jolliff, Oregon Institute of Technology
 Francis Jones, Huntington College
 Toni Kasper, Borough of Manhattan Community College
 Alvin Kaumeyer, Pueblo Community College
 William Keane, Boston College
 Robert Keever, SUNY, Plattsburgh
 Alice J. Kelly, Santa Clara University
 Dave Kender, Wright State University
 Michael Kern, Bismarck State College
 John Klages, County College of Morris
 Marlene Kovaly, Florida Community College at Jacksonville
 John Kozarski, Community College of Baltimore County–Catonsville
 Tomas Kozubowski, University of Tennessee
 Shantra Krishnamachari, Borough of Manhattan Community College
 Richard Kulp, David Lipscomb University
 Linda Kurz, SUNY College of Technology
 Christopher Jay Lacke, Rowan University
 Tommy Leavelle, Mississippi College
 Tzong-Yow Lee, University of Maryland
 R. E. Lentz, Mankato State University
 Timothy Lesnick, Grand Valley State University
 Dawn Lindquist, College of St. Francis
 George Litman, National-Louis University
 Benny Lo, Ohlone College
 Sergio Loch, Grand View College
 Debra Loeffler, Community College of Baltimore County–Catonsville
 Vincent Long, Gaston College
 Barbara Loughead, National-Louis University
 David Lund, University of Wisconsin-Eau Claire
 Rhonda Magel, North Dakota State University–Fargo
 Gene Majors, Fullerton College
 Hossein Mansouri, Texas State Technical College
 Virgil Marco, Eastern New Mexico University
 Joseph Mazzonec, Delta College
 Caren McClure, Santa Ana College
 Phillip McGill, Illinois Central College
 Marjorie McLean, University of Tennessee
 Austen Meek, Canada College
 Robert Mignone, College of Charleston
 Glen Miller, Borough of Manhattan Community College
 Kermit Miller, Florida Community College at Jacksonville
 Kathleen Mittag, University of Texas–San Antonio
 Mitra Moassessi, Santa Monica College
 Charlene Moeckel, Polk Community College
 Theodore Moore, Mohawk Valley Community College
 Rick Moscatello, Southeastern Louisiana University
 Gerald Mueller, Columbus State Community College
 Sandra Murrell, Shelby State Community College
 Faye Muse, Asheville-Buncombe Technical Community College
 Gale Nash, Western State College
 Felix D. Nieves, Antillean Adventist University
 Lyn Noble, Florida Community College at Jacksonville–South
 DeWayne Nyman, University of Tennessee
 Patricia Oakley, Seattle Pacific University
 Keith Oberlander, Pasadena City College
 Patricia Odell, Bryant College
 James O'Donnell, Bergen Community College

Alan Olinksy, Bryant College
Nasser Ordoukhani, Barry University
Ron Pacheco, Harding University
Lindsay Packer, College of Charleston
Kwadwo Paku, Los Medanos College
Deborah Paschal, Sacramento City College
S. A. Patil, Tennessee Technological University
Robin Pepper, Tri-County Technical College
David C. Perkins, Texas A&M University—
Corpus Christi
Anthony Piccolino, Montclair State University
Kim Polly, Parkland College
Richard J. Pulskamp, Xavier University
Diann Reischman, Grand Valley State University
Vance Revennaugh, Northwestern College
C. Richard, Southeastern Michigan College
Don Robinson, Illinois State University
Sylvester Roebuck, Jr., Olive Harvey College
Ira Rosenthal, Palm Beach Community
College—Eissey Campus
Kenneth Ross, Broward Community College
Charles M. Roy, Camden County College
Kara Ryan, College of Notre Dame
Fabio Santos, LaGuardia Community College
Richard Schoenecker, University of Wisconsin,
Stevens Point
Nancy Schoeps, University of North Carolina,
Charlotte
Jean Schrader, Jamestown Community College
A. L. Schroeder, Long Beach City College
Phyllis Schumacher, Bryant College
Sankar Sethuraman, Augusta College
Rosa Seyfried, Harrisburg Area Community
College
Calvin Shad, Barstow College
Carole Shapero, Oakton Community College
Adele Shapiro, Palm Beach Community
College
Lewis Shoemaker, Millersville University
Joan Sholars, Mt. San Antonio College
Galen Shorack, University of Washington
Teresa Siak, Davidson County Community
College
Cheryl Slayden, Pellissippi State Technical
Community College
Arthur Smith, Rhode Island College
Marty Smith, East Texas Baptist University
Laura Snook, Blackhawk Community College
Aileen Solomon, Trident Technical College
Sandra Spain, Thomas Nelson Community
College
Maria Spinacia, Pasco-Hernandez Community
College
Paulette St. Ours, University of New England
W. A. Stanback, Norfolk State University
Carol Stanton, Contra Costa College
Richard Stephens, Western Carolina College
W. E. Stephens, McNeese State University
Terry Stephenson, Spartanburg Methodist
College
Consuelo Stewart, Howard Community College
David Stewart, Community College of
Baltimore County—Dundalk
Ellen Stutes, Louisiana State University at
Eunice
Sr. Loretta Sullivan, University of Detroit Mercy
Tom Sutton, Mohawk College
Andrew Thomas, Triton College
Evan Thweatt, American River College
Judith A. Tully, Bunker Hill Community
College
Gary Van Velsir, Anne Arundel Community
College
Paul Velleman, Cornell University
Randy Villa, Napa Valley College
Hugh Walker, Chattanooga State Technical
Community College
Charles Wall, Trident Technical College
Glen Weber, Christopher Newport College
David Weiner, Beaver College
Sue Welsch, Sierra Nevada College
Roger Willig, Montgomery County
Community College
Odell Witherspoon, Western Piedmont
Community College
Jean Woody, Tulsa Junior College
Carol Yin, LeGrange College
Thomas Zachariah, Loyola Marymount
University
Elyse Zois, Kean College of New Jersey

M.F.T.
LaGrange, Nueva York,
julio de 2005

Índice de aplicaciones

Los significados de las letras entre paréntesis son los siguientes: **PC** = problema del capítulo, **EJT** = ejemplo en el texto, **M** = ejemplo al margen, **E** = ejercicio, **MB** = más allá de lo básico, **R** = ejercicio de repaso, **RA** = ejercicio de repaso acumulativo, **DD** = de los datos a la decisión, **ACE** = actividad de cooperación en equipo, **PT** = proyecto tecnológico, **ET** = la estadística en el trabajo.

Agricultura

- Dientes de león (E), 234
- Experimento de crecimiento de árboles (E), 187
- Fenotipos de chícharos (E), 87, 105
- Fertilizante (RA), 132; (EJT), 488-489
- Gallinas que ponen huevos (EJT), 6, 202
- Leche de vaca (EJT), 6, 202
- Longevidad de árboles tratados con fertilizante (MB), 91
- Mediciones de árboles (R), 131; (RA), 131
- Nuevo fertilizante y crecimiento de árboles (EJT), 24-25
- Pesos de álamos (E), 88-89, 107-108, 650, 661, 662; (PC), 635; (EJT), 639-640, 655-660, 703-704; (MB), 653; (PT), 670
- Prueba de semillas de maíz (E), 128, 493, 687, 691; (EJT), 681-682
- Semillas de paja (R), 508

Alimentos/bebidas

- Azúcar en cereales (E), 434
- Azúcar en naranjas (M), 355
- Azúcar y calorías en cereales (E), 553
- Barras energéticas de proteínas (R), 36
- Carbohidratos en la comida (RA), 312-313
- Cereales (E), 87, 105
- Coca-Cola frente a Pepsi (ACE), 380, 449; (EJT), 498-500
- Coca-Cola regular y Coca-Cola dietética (E), 50, 55, 89-90, 107, 288
- Comida saludable de chocolate (E), 18
- Comparación entre Pepsi regular y Pepsi dietética (E), 362
- Dulces M&M (E), 148, 228, 300, 336, 417, 604, 651, 707; (MB), 363, 418, 423, 442, 653; (EJT), 400, 419, 427-428; (RA), 668-669
- Escala para calificar alimentos (MB), 11
- Huevos rotos (E), 188
- Llenado de latas de bebidas (E), 745-746
- Pastel de frutas (R), 192
- Pesos de Coca-Cola regular y Coca-Cola dietética (E), 483, 502, 503
- Pesos de paquetes de azúcar (R), 446

Ambiente

- Cantidades de precipitación pluvial (E), 443, 467
- Choque de meteoritos (EJT), 142
- Contaminación de automóviles (E), 362; (R), 666-667

Contaminación del aire (EJT), 16

- Datos sobre el clima (E), 50, 55
- Decaimiento radiactivo (E), 234
- Errores de pronóstico (E), 348, 425, 435; (M), 542
- Exactitud de pronóstico y temperaturas (E), 336-337, 417
- Géiser Old Faithful (E), 66-67, 87, 105, 494, 535, 554, 575; (PC), 515-516 (EJT), 523, 524, 527-529, 544, 546, 559, 562, 567-568; (R), 582-583
- Incendios y acres quemados (E), 537, 555
- Incidencia del radón (E), 466-467
- Lluvia (E), 50, 55, 504-505, 723
- Peso de basura desechada por los hogares (R), 376

Precipitación pluvial (E), 49, 90, 108, 309

- Precipitación pluvial en Boston (E), 336, 359, 417, 467, 747; (EJT), 720; (MB), 752

Precisión de los pronósticos del clima (E), 88, 89, 106, 107

- Radón en hogares (R), 36
- Temperatura mínima diaria (E), 49, 90, 108
- Temperaturas (E), 11, 309, 537, 555
- Temperaturas de los Everglades (E), 423
- Temperaturas máximas (E), 66; (RA), 194
- Temperaturas reales y pronosticadas (E), 360, 489, 491, 493, 538-539, 556, 694, 716; (EJT), 485-488
- Terremotos (E), 235
- Verificación del plomo en el aire (EJT), 78, 79, 81; (E), 310, 434, 443

Biología

- Anchura de cráneos (E), 362, 652-653, 706
- Bacteria *E. Coli* (E), 173
- Chirridos de grillos y temperatura (EJT), 60-61; (E), 537, 555, 565, 715
- Cigarras (E), 10
- Clonación de seres humanos (EJT), 142-143
- Datos de osos (E), 128, 489, 533, 536, 554-555, 574, 576
- Ecología, comportamiento animal y ecotoxicología (ET), 383
- Experimento de hibridación (E), 157-158, 177, 178, 300, 414; (MB), 605
- Experimento del color de los ojos (E), 603
- Experimento genético (E), 210, 404, 604; (MB), 224
- Género de osos (EJT), 719-720; (E), 722, 723
- Genes de ojos azules (RA), 669
- Genética: color de ojos, edad y género (E), 149, 150, 665

Genética mendeliana (E), 147, 228, 334; (EJT), 411-413

- Genotipos (EJT), 142
- Investigación arqueológica (ET), 513
- Largo de huevos de ave (E), 109
- Mamíferos más pequeños del mundo (E), 347-348, 361-362, 374, 425, 434, 442
- Método de captura y recaptura (ACE), 194
- Moscas de la fruta (E), 88, 106, 279
- Muerte de manatíes (E), 580; (R), 582
- Nucleótidos del ADN (E), 187
- Plantas cultivadas en casas (RA), 132
- Plantas vasculares y no vasculares (EJT), 163
- Sociabilidad y población de cachalotes (ET), 39
- Tamaño de poblaciones de vida silvestre (M), 339
- Variación en volúmenes cerebrales (R), 508

Deportes

- Anotación de un tiro libre (EJT), 141
- Clasificación de gimnastas (E), 714
- Competidores del triatlón olímpico (E), 280
- Congelar al pateador (M), 572
- Derby de Kentucky (E), 150
- Distancias de *home runs* (PT) 133; (E), 653
- Edades de corredores de maratón (E), 534
- Educación y deportes (E), 32
- Estatura de corredores (E), 34
- Estatura de jugadores de los Lakers de LA (E), 310
- Ganadores de medallas olímpicas de oro (E), 722
- Ganadores olímpicos (E), 435
- Género de atletas profesionales (EJT), 6
- Hits* de jugadores de béisbol (MB), 301-302
- Intercepción de lanzamientos del mariscal de campo (EJT), 271
- Juegos de la serie mundial (E), 603
- Lanzamiento en paracaídas (M), 293
- Manchas solares y puntos en el Súper Bowl (E), 715
- Maratones (E), 537, 555, 663-664
- Número de la camiseta de jugadores de básquetbol (E), 10
- Pelotas de béisbol (E), 433
- Pesos de timoneles y remeros en una carrera de canotaje (E), 54-55
- Porcentaje de *strikes* declarados por árbitros (R), 447
- Promedio de bateo (M), 96
- Rachas de suerte en los deportes (M), 719
- Salarios y desempeño de la NBA (M), 570

Serie mundial de béisbol (E), 211, 723, 724
 Súper Bowl (M), 476, 558; (E), 534, 553;
 (RA), 583-584
 Tiros libres de básquetbol (ACE), 754
 Torneo de básquetbol de la NCAA (E),
 188-189
 Ventaja del equipo local (E), 464, 620;
 (M), 612

Derecho

Cámara vigilante (DD), 381
 Condenado por probabilidad (M), 163
 Crimen en campus (E), 20
 Crimen y extraños (R), 629
 Crímenes por drogas (E), 414
 Demandas por negligencia médica (E), 405
 Detección de fraude (E), 301, 335; (PC),
 589; (EJT), 597-598
 Detectores de mentiras (M), 388
 Estado de Arizona vs. Wayne James Nelson
 (DD), 632
 Fraude de tarjeta de crédito (E), 172, 301
 Identificación de ladrones (M), 473
 Identificación de voz de un criminal (E), 166
 Multas por exceso de velocidad (E), 90, 108
 Pena de muerte (E), 405; (ACE), 449, 510;
 (M), 462
 Precisión de pruebas con polígrafo (E), 618
 Puesto de revisión de sobriedad (E), 32
 Robo de identidad (EJT), 180
 Selección de miembros de un jurado
 (PC), 199; (EJT), 201, 203, 207, 208,
 215-217, 219, 225-227; (E), 211, 220,
 301, 335, 417
 Sentencia independiente de la declaración
 de inocencia (E), 618
 Soborno en el Jai Alai (M), 743
 Testimonio en la Suprema Corte (M), 459
 Trampa en impuestos (E), 32
 Velocidades de conductores multados en
 carretera interestatal (E), 347

Economía y negocios

Acciones (E), 534, 553, 580-581; (M), 558
 Accionistas de The Coca Cola Company
 (R), 35
 Alto costo de la baja calidad (M), 750
 Análisis de ventas (ET), 757
 Analista del IRS (E), 32
 Artículos defectuosos (E), 166, 171, 173,
 227, 298, 752; (MB), 174, 213, 230; (R),
 753-754; (RA), 754; (PT), 755
 Auditorías del IRS (E), 223
 Aumento de la calidad (E), 224
 Canciones descargadas (E), 336
 Centro telefónico (E), 278-279
 Códigos de barras (M), 185; (R), 193
 Comerciales (M), 419
 Compañía farmacéutica (ET), 197
 Compras en Internet (E), 335
 Consumidores engañados (R), 446

Consumo diario de petróleo (EJT), 14-15
 Consumo eléctrico (R), 753
 Control de calidad (E), 31, 33, 172, 279-
 280, 403, 431, 438-439; (EJT), 427-428;
 (M), 748
 Diferencia en los valores del hogar (E), 482
 Discos compactos fabricados por Sony
 (M), 568
 Empresa editorial (ET), 73
 Estadística y administración de la calidad
 (ET), 452-453
 Etiquetas de empaques de dulces M&M
 (E), 290
 Exactitud del IRS (E), 148
 Fabricación de altímetros para aeronaves
 (E), 442
 Fabricación de latas de aluminio (DD), 756
 Fabricación de teléfonos celulares (E), 178;
 (PT), 195
 Factura fiscal media (E), 359
 Filas de espera en los bancos (E), 443
 Ley de Moore (MB), 581
 Máquina para recubrimiento de papel
 (M), 740
 Máquinas expendedoras (E), 290; (R) 377
 Medios de comunicación masiva y publicidad
 (E), 20
 Mercado de acciones (E), 723
 Mercado de acciones y ventas de automóviles
 (E), 714
 Muestreo de aceptación (E), 223, 300-301
 Navegación en Internet (E), 415
 Necesidad de agua caliente en un hotel
 (E), 289-290
 Números del seguro social (E), 10
 Precio de lista y precio de venta (E),
 538, 556
 Precios de casas (E), 494, 575
 Precisión de escáner (E), 416, 619
 Predicción de costos de electricidad (R), 583
 Predicción de precios de condominios
 (E), 545
 Producto de consumo (E), 10
 Promedio industrial Dow Jones (RA),
 583-584; (ACE), 754-755
 Pronóstico y análisis de Walt Disney World
 (ET), 587
 Propinas (M), 121
 Publicidad (ACE), 380
 Publicidad televisiva (E), 299-300
 Quejas ante compañía telefónica
 (EJT), 59, 60
 Reconocimiento de marca (E), 148, 211
 Reemplazo de televisores (MB), 301
 Seis Sigma en la industria (M), 749
 Tasa de defectos (E), 751
 Telemarketing (E), 279
 Tiempos de espera de consumidores (E), 89;
 (EJT), 92-93, 95, 96, 97, 102, 107;
 (ACE), 510
 Toxicólogo (ET), 673
 Ventas de casas (E), 213

Educación

Asistencia a clases y calificaciones
 (M), 680
 Ausencias (ACE), 755
 Bloqueo en exámenes (E), 482, 504
 Calificación perfecta en la prueba SAT
 (M), 164
 Calificación y lugar para sentarse (E), 602
 Calificaciones de examen (EJT), 84, 351;
 (E), 118, 650; (RA), 630
 Calificaciones de mujeres en la prueba de
 matemáticas del SAT (RA), 448
 Calificaciones en la prueba SAT (E),
 662-663; (RA), 668-669
 Calificaciones en un curso (EJT), 7-8
 Carrera y género (ACE), 670, 728
 CI de estudiantes de estadística (E), 348, 360
 Clasificación de escuelas de medicina
 (R), 726
 Clasificación de escuelas de negocios y
 leyes (R), 725
 Clasificación de universidades (PC), 675;
 (EJT), 710-711, 712
 Conjeturas en un examen (E), 165, 220,
 221, 227-228; (R), 237
 Cursos de preparación para la prueba SAT
 (E), 492; (MB), 494
 Curva de aprendizaje (MB), 716-717
 Edad de profesores (E), 108
 Educación y deportes (E), 32
 El crecimiento de la estadística (M), 113
 Estudiantes de estadística presentes en una
 clase (EJT), 202
 Estudiantes suspendidos (EJT), 13
 Exámenes de opción múltiple (RA), 630
 Gastos para el regreso a clases (E), 359
 Las evaluaciones de maestros se correlacio-
 nan con las calificaciones (M), 522
 Las personas que se gradúan de la universidad
 viven más tiempo (E), 18
 Longitud de un salón de clases (ACE), 380,
 449, 669-670; (R), 666
 Mejores resultados con tamaños de clases
 más pequeños (M), 477
 Muestra de estudiantes (E), 33
 Normalización de calificaciones de un
 examen (MB), 269
 Nueva política de asistencia (EJT), 23
 Número de clases (E), 51
 Orden de los asientos en clase (ACE), 727
 Paradoja del tamaño de la clase (M), 79
 Precios de libros de texto universitarios
 (EJT), 9
 Predicción de éxito (M), 567
 Preparación para la prueba SAT (E), 288-289
 Programa de preparación para la prueba
 SAT (RA), 727
 Promedio de calificación (E), 423
 Pruebas SAT y ACT (MB), 269
 Puntuaciones de CI (E), 34, 117, 266, 298,
 404, 489, 553, 706; (EJT), 100, 101,
 261-262, 391; (M), 94, 736; (MB), 268,

- 605-606; (R), 310, 311, 446, 447; (PT), 450, 511
- Puntuaciones de CI de gemelos idénticos (PT) 585
- Puntuaciones de CI de profesores (E), 344-345, 432; (R), 446
- Selección de estudiantes (E), 165
- Sesgo por género en una pregunta de examen (R), 312
- Tamaño de la clase (EJT), 247-249; (E), 257
- Tiempo de estudio y calificaciones (E), 18
- Tiempo para obtener un título universitario (E), 347
- Encuestas y sondeos de opinión**
- Adultos que se oponen a los impuestos estatales (E), 444
- Aplicación de encuestas (E), 19
- Creencia de que existe vida en otros lugares de la galaxia (E), 156
- Cuestionarios para grupos de mujeres (M), 392
- Detección de datos falsos (M), 16
- El medio de encuesta puede afectar los resultados (M), 613
- Encuesta a clientes de Merrill Lynch (E), 19
- Encuesta a consumidores (RA), 70
- Encuesta de American Online (ACE), 37
- Encuesta de estudiantes (MB), 33; (E), 332; (R), 628
- Encuesta de la revista *Glamour* (R), 446
- Encuesta de MTV (E), 32
- Encuesta de parejas casadas en centros comerciales (MB), 694
- Encuesta de salud (E), 467
- Encuesta de trabajadores (E), 414
- Encuesta de votantes (E), 278, 299
- Encuesta Gallup (EJT), 4; (E), 20, 404, 416, 464
- Encuesta por correo (E), 11, 19
- Encuesta por correo electrónico (EJT), 329-330
- Encuesta por Internet (E), 130, 414
- Encuesta preelectoral (M), 458
- Encuesta pública (ET), 243
- Encuesta sobre bebida (E), 10, 416
- Encuesta sobre clonación (E), 301, 335
- Encuesta sobre consumo de leche (R), 377
- Encuesta sobre tabaquismo y educación universitaria (R), 378; (E), 417
- Encuesta sobre teléfonos celulares (R), 34
- Encuestas de salida (E), 32; (RA), 132
- Encuestas de televidentes (E), 222-223
- Encuestas sobre temas delicados (ACE), 194-195; (M), 262
- Encuestas y sondeos telefónicos (E), 32, 212-213, 465, 752
- Estimados para mejorar el censo (M), 356
- Exactitud de encuesta (MB), 337
- Falsificación de datos (MB), 21; (M), 326
- Influencia del género (EJT), 612-613
- La ética en los reportes (M), 643
- Minería de datos (M), 119
- Nivel de confianza de encuesta (E), 166
- Porcentaje de usuarios del teléfono (E), 414
- Pregunta incorrecta (E), 19
- Prueba de la influencia del género (E), 619
- ¿Qué está mal en este asunto? (MB), 21
- Rechazo de encuestas y grupo de edad (E), 620
- Reporte de resultados de encuesta en periódico (R), 376
- Resistencia a la encuesta (E), 158; (M), 642
- Respuestas de encuesta (EJT), 7
- Respuestas de encuesta confusas (E), 334
- Visitas repetidas (M), 215
- Entretenimiento**
- Asistencia a parque temático (RA), 379
- Audiencia televisiva (R), 237
- Audiencias y ventas de canción (E), 535, 554, 715-716
- Autocinemas (EJT), 61-62; (MB), 118
- Boston's Women's Club (MB), 291
- Calificación de facilidad de lectura (R), 666
- Calificaciones de películas (E), 11; (R), 34
- Clasificación de críticos de cine (R), 35
- Clasificación de Nielsen (E), 11, 31
- Comparación de facilidad de lectura (E), 507-509
- Compra de una audiencia televisiva (E), 536-537, 555
- Concierto de rock (R), 34
- Consumo de alcohol y tabaco en películas (E), 336, 417, 435, 467, 494, 505, 694; (MB), 717
- Edades de actores y actrices ganadores del Óscar (PC), 41, 75; (EJT), 43-46, 48, 52, 57-59, 81, 83, 99, 112-115, 119-122; (E), 55-56, 603; (MB), 68; (R), 69-70
- Montaña rusa (MB), 174
- Número posible de melodías (E), 188
- Presupuestos e ingresos brutos de películas (E), 535-536, 554, 564, 565, 716
- Rutas a los juegos en Disney World (EJT), 182
- Salas de cine (E), 68, 723
- Sitio Web de Napster (EJT), 13
- Televidentes (E), 10, 348
- Viajes en un parque de diversiones (E), 346
- Finanzas**
- Bonos especulativos (MB), 213
- Cajeros automáticos (E), 189
- Calificación de crédito (E), 361, 374, 433, 442
- Cambio de un dólar (MB), 190
- Datos de ingresos (E), 32
- Desempeño de inversiones (ET), 633
- Deuda de crédito (E), 372; (R), 506-507
- Dinero gastado en automóviles nuevos en Estados Unidos (E), 348-349
- Elección de códigos personales de seguridad (M), 182
- Ingreso personal (EJT), 13-14; (E), 88, 106
- Ingresos de tiempo completo de estudiantes universitarios (R), 446
- Ingresos promedio anuales (E), 19
- Más acciones, menos riesgos (M), 99
- Montos de cheques del autor (E), 604
- NSS e ingreso (E), 532
- Presupuesto tardío del estado de Nueva York (R), 312; (E), 538, 556
- Tarjetas de crédito (E), 32, 130; (RA), 239; (R), 446
- Individuos y psicología**
- Children's Defense Fund (M), 322
- Contacto visual (E), 267
- Datos del censo (E), 19
- Detectores humanos de mentiras (M), 126
- Discriminación racial (E), 18, 615-616; (R), 507
- Dominancia de la mano izquierda (E), 178; (RA), 313
- Dominancia de la mano izquierda y género (E), 617
- Estatura de supermodelos (E), 434, 442-443, 536, 555
- Estatura requerida para las mujeres soldados (E), 266
- Estatura requerida por el Club Beanstalk (E), 266; (R), 312
- Estaturas de estudiantes de estadística (E), 50, 55; (R), 507
- Estaturas de hombres (E), 49, 109, 117, 492, 687; (EJT), 103-104, 110-111, 303-306; (R), 131; (ACE), 510
- Estaturas de mujeres (EJT), 46-47; (E), 117, 287-288, 307, 309, 372; (MB), 268; (R), 447; (ACE), 510
- Estaturas de padre e hijo (E), 537, 555
- Estaturas de presidentes (RA), 448; (EJT), 484; (E), 491, 687
- Estudio de mortalidad (E), 211
- Estudio nacional prospectivo de los niños (M), 26
- Florence Nightingale (M), 57
- Género de hijos (EJT), 143, 168-169, 270-271; (E), 147, 149, 166, 333-334; (MB), 167; (ACE), 313, 380
- Género en una familia (M), 270
- Identificación de trastornos psiquiátricos (E), 481
- Intensidad del dolor (DD), 586
- La gente más adinerada (E), 279
- Lectura de la palma de la mano (M), 519
- Longevidad (R), 667-668; (RA), 668
- Medición de la desobediencia (M), 7
- Medición de la inteligencia en niños (E), 492-493; (R), 725
- Novias de junio (E), 603
- Número de hijos (ACE), 240
- Número de niñas (E), 210, 212
- Percepción del tiempo (E), 87, 105, 347, 424
- Percepción extrasensorial (ACE), 240, 449

Periodos de vida (ACE), 669
 Pesos de hombres (EJT), 103-104
 Pesos de supermodelos (EJT), 6; (RA), 378-379; (E), 442, 536, 555
 Política de seguro de vida (R), 193; (E), 212
 Posposición de la muerte (E), 334, 416, 543, 688
 Predicción del color de ojos (E), 573
 Pruebas de estrés (E), 118
 Psicología del trauma (E), 31
 Riqueza y CI (EJT), 17
 Selección del género (E), 147, 188, 228, 298, 333-334, 415, 465, 686, 688; (EJT), 164, 174, 179, 184, 190, 386-387, 388-389, 682-683; (R), 192-193; (MB), 230; (DD), 241; (ACE), 380, 448
 Tamaño de la familia (E), 10, 19
 Terapia de contacto (E), 31, 229; (PC), 319; (EJT), 321, 327-328
 Tiempo de reacción (ACE), 510, 584
 Trastornos psiquiátricos relacionados con factores biológicos (E), 700
 Volumen cerebral y enfermedades psiquiátricas (R), 508

Ingeniería

Carga axial de una lata de aluminio (E), 424-425
 Consumo de energía eléctrica (E), 108
 Corriente eléctrica (R), 193
 Diseño de asientos (MB), 291
 Diseño de asientos de aeronave (DD), 315
 Diseño de ataúdes (E), 267
 Diseño de cascos (E), 268, 289
 Diseño de entradas (E), 267
 Diseño de luces estroboscópicas (E), 289
 Diseño de tableros para automóvil (EJT), 264-265
 Energía solar (E), 651, 706
 Ensamble de partes de teléfonos (R), 446
 Ingeniero de control de calidad (E), 130
 Mars Climate Orbiter (M), 739
 Rediseño de asientos expulsores (E), 290
 Voltaje para un detector de humo (EJT), 203
 Voltajes y corrientes (MB), 91

Interés general

Acuñación de monedas de 25 centavos (E), 373, 441-442, 746-747
 Anchuras del codo de mujeres (E), 309
 Anclaje de números (ACE), 132; (E), 700
 Antigüedad de libros (ACE), 380, 449
 Años (EJT), 8
 Candados de combinación (E), 188
 Cargas axiales de latas de aluminio (E), 51, 109, 746; (MB), 56
 Circunferencia de la cabeza y longitud del antebrazo (ACE), 584
 Códigos de área (E), 188
 Códigos postales (E), 116-117
 Coincidencias (M), 170
 Confiableidad de sistemas (M), 159

Consumo de energía y temperatura (E), 67
 Correo electrónico y privacidad (R), 629
 Costo del índice de la risa (M), 110
 Cumpleaños (M), 145; (E), 148, 177; (MB), 167, 173, 179; (EJT), 175
 Desciframiento de mensajes (E), 228
 Día de Acción de Gracias (EJT), 144
 Día Nacional de la Estadística (R), 192
 Distancia de asiento (E), 268
 Edad del presidente de Estados Unidos (ACE), 380, 449
 Edades de polizones (E), 88, 106; (EJT), 357
 Edades de solicitantes (EJT), 352-353; (E), 361; (MB), 362-363
 Efecto del peso al nacer sobre el CI (E), 479-480, 504
 El estado de la estadística (M), 5
 Errores de medición de peso (R), 311
 Escasez de números telefónicos (M), 180
 Estatura y envergadura de brazos (ACE), 584, 727
 Estatura y estatura del ombligo (ACE), 584, 728
 Estaturas de marcianos (MB), 363
 Ética en los experimentos (M), 427
 Experimento de correo (E), 10
 Experimento de física (E), 580
 Fuerza de agarre (E), 311
 Gane \$1,000,000 si tiene poderes extrasensoriales (M), 393
 Gemelos en Twinsburg (M), 499
 Identificación de autores (M), 44
 La secretaria aleatoria (M), 184
 La vida en Alfa Romeo (E), 191
 Lanzamiento y giro de centavos (EJT), 613-614; (E), 619
 Longitud de pajillas (ACE), 37
 Longitud del muslo (E), 108
 ¿Los zurdos mueren más pronto? (M), 428
 Lunas de Júpiter (E), 280
 Medidas de maniquíes y de mujeres (M), 77
 Monedas ocultas (MB), 174
 Monos mecanógrafos (M), 177
 Moscas sobre una naranja (MB), 151
 Muertes por coches de caballos (E), 235
 Números telefónicos en Port Jefferson (MB), 417-418
 Periodista (ET), 317
 Período de vida de un conductor (E), 434
 Peso al nacer y graduación (E), 617
 Peso y uso del control remoto (EJT), 61
 Pesos de anillos de compromiso (EJT), 9
 Pesos de centavos (E), 10, 50, 55, 89, 90, 107, 108, 309, 504, 653, 701, 707; (RA), 36; (EJT), 47-48, 369-370
 Pesos de monedas de 25 centavos (E), 128, 348, 374-375, 425, 435, 443, 482, 504, 688
 Pesos de plástico desechado (E), 49
 Plástico desechado y tamaño de la familia (E), 538, 556

Probabilidades que desafían la intuición (M), 139
 Prueba de guantes de laboratorio (E), 464
 Pruebas de inflamabilidad de tela (E), 651, 707
 Puntos en un palo (MB), 151
 Recaudación de fondos (E), 32
 Redundancia de despertadores (E), 166, 173
 Saludos y mesas redondas (MB), 189
 Ser alcanzado por un relámpago (E), 147
 Tamaño del sombrero y CI (E), 547
 Tarjetas de crédito y llaves (ACE), 584-585
 Tiempo de llamadas telefónicas (MB), 310
 Uso de ropa naranja de cazador (E), 165-166
 Vida extraterrestre (PT), 728
 Viernes 13 (E), 491, 687
 Vocabulario de Shakespeare (M), 152

Juegos

Apuestas en hipódromos (M), 142
 ¿Cuántas veces hay que barajar? (M), 183
 Dado cargado (E), 50, 55, 601
 Dados de casino (E), 212
 Esquemas para vencer a la lotería (M), 292
 Ganadores múltiples de la lotería (M), 260
 Ganar centavos de la lotería (M), 181
 Ganar la lotería (E), 186, 189
 ¿La lotería es aleatoria? (R), 726
 Lanzamiento de monedas (RA), 509
 Lanzamiento de un dado (E), 187; (EJT), 200
 Los estados controlan las selecciones de lotería (M), 303
 Lotería Fantasy 5 de California (EJT), 179-180; (E) 186
 Lotería Pick 4 de Kentucky (EJT), 209, 233
 Lotto 54 (MB), 224
 Máquina tragamonedas (E), 223, 601
 Principio fundamental del juego (M), 175
 Problema de Monty Hall (MB), 179; (ACE), 194
 Puedes apostarlo (M), 140
 Puntuaciones en el *pinball* (EJT), 712-713
 Recomendación para la lotería (M), 161
 Reparto de naipes (MB), 167
 Rifa organizada por una revista (E), 212
 Ruleta (RA), 70; (EJT), 146; (E), 150, 212, 601; (MB), 301
 Seis grados de Kevin Bacon (PC), 3
 Selección de números de la lotería (M), 202, 271
 Solitario (E), 150
 Sopa de letras (E), 188

Salud

Alquitrán y cigarrillos (E), 482
 Alquitrán y monóxido de carbono en cigarrillos (E), 67, 556
 Alquitrán y nicotina en cigarrillos (E), 538, 556, 564, 575

- Aspirina y prevención de ataques cardíacos (M), 460
- Baterías usadas en marcapasos cardíacos (E), 130
- Brecha de género en las pruebas de fármacos (M), 690
- Captopril para disminuir la presión sanguínea sistólica (E), 490
- Ceguera al color (E), 156; (MB), 418
- Comparación de dietas (E), 480
- Comparación de TEP/TC con IRM (E), 626
- Comparación de tratamientos (R), 629-630
- Concentración de alcohol en la sangre (R), 726
- Costos hospitalarios de choques (R), 378
- Cotinina en fumadores (EJT), 180-181; (E), 347, 424, 537
- Cura para el resfriado común (R), 445
- Datos de salud (EJT), 571-572; (E), 574-575
- Dieta Atkin (E), 348, 424
- Diseño de experimentos (E), 187
- Duración del embarazo (E), 117, 267
- Efecto adverso del Viagra (E), 149; (MB), 150-151
- Efecto placebo (M), 285
- Efectos adversos del Clarinex (E), 417, 466
- Efectos cardiovasculares (MB), 33
- Efectos de la cocaína en niños (E), 479
- Efectos del alcohol (E), 481, 503; (MB), 483-484
- Efectos del consumo de marihuana en estudiantes universitarios (E), 480, 503
- Eficacia de la equinácea (E), 479, 502
- Eficacia de la hipnosis para reducir el dolor (E), 492
- Eficacia de la pasta dental Crest para disminuir las caries (M), 486
- Eficacia de la vacuna de Salk (EJT), 23-24, 26; (M), 461
- Eficacia de parches de nicotina (E), 416
- Eficacia de un programa de tratamiento para el VIH (RA), 238
- Eficacia de una dieta (E), 433-434, 479
- Eficacia de vacuna (E), 617
- Eficacia del Prilosec (E), 479
- Eficacia del Sleepeze (R), 35
- Ejercicio y estrés (E), 652, 706-707
- Enfermedades infecciosas (R), 753
- Ensayos clínicos (E), 32, 604; (EJT), 183, 185; (M), 261
- Estatura y ejercicio (E), 18
- Estatura y pulso (E), 538, 536
- Experimento de poliomielitis (M), 461
- Fármaco para disminuir la presión sanguínea (EJT), 25
- Fármaco que reduce el colesterol (E), 191, 229, 300, 403, 479; (R), 665
- Frecuencias cardíacas al trabajar con la pala (E), 360-361, 373
- Género al nacer (EJT), 139, 144, 155; (M), 151; (E), 172, 177, 210, 227, 298, 299
- Gráficas de crecimiento actualizadas (M), 47
- Grupos y tipos sanguíneos (E), 157, 300
- Hawthorne y efectos en el experimentador (M), 24
- IMC y género; (E), 89, 90, 107, 108; (EJT), 358, 484
- Índice de masa corporal (E), 50, 55, 128, 309, 374; (ACE), 71; (EJT), 697-698; (MB), 701
- Infecciones de VIH (E), 173
- Internista especializado en enfermedades infecciosas (ET), 135
- Interpretación de la eficacia de un tratamiento (E), 149
- Latidos cardíacos (ACE), 449
- Lectura diástólica (E), 565
- Lipitor (E), 20, 149, 222; (M), 59; (R), 192, 446; (EJT), 469
- Lipoproteína de baja densidad (EJT), 484
- ¿Los mosquiteros reducen la malaria? (E), 465
- ¿Los pacientes quirúrgicos que están tibios se recuperan mejor? (R), 508
- Métodos para dejar de fumar (E), 616-617, 625-626; (MB), 627
- Nacimientos (E), 602
- Negligencia médica (E), 334
- Nicotina en cigarrillos (E), 50, 55; (R), 447
- Niveles de colesterol (E), 118, 309; (R), 312
- Pérdida de peso (R), 444
- Pérdida de peso con diferentes dietas (E), 360, 650
- Perros para identificar el cáncer (DD), 451; (E), 618
- Peso (RA), 238-239; (EJT), 591-592, 594-596; (E), 600-601
- Pesos al nacer (E), 267, 360, 373, 433, 441; (EJT), 356-357; (RA), 669
- Píldora de dieta costosa (E), 474
- Píldoras de vitaminas (E), 33
- Píldoras defectuosas (E), 187; (R), 193
- Presión sanguínea (E), 87, 105, 289, 309, 347, 425, 491, 536, 555, 564
- Presión sanguínea sistólica (E), 268, 359
- Proceso de aprobación de fármaco (M), 408
- Prueba de la eficacia de una vacuna (E), 466, 479
- Prueba de sangre (E), 172
- Prueba de sífilis (M), 171
- Prueba de un tratamiento (E), 626
- Pruebas de audición (E), 663
- Pulsos (EJT), 47, 99, 124-126, 339, 341-342; (PT), 71; (E), 349, 362, 403, 435, 604, 663, 701; (ACE), 510, 584, 727, 754
- Reacción adversa a un fármaco (E), 532
- Relación entre el tabaquismo y el cáncer (M), 711
- Remedio para la gripe (E), 18
- Resultados de prueba de embarazo (DD), 196
- Síndrome de túnel carpiano: entubillado o cirugía (PC), 455; (EJT), 458-460, 461-462; (E), 618
- SMSI (E), 20
- Sustancias dañinas en cigarrillos (E), 716
- Sustituto de la nicotina (E), 178-179
- Tabaquismo y nicotina (E), 555
- Tabaquismo y resistencia física (E), 615
- Tabaquismo, temperatura corporal y género (R), 666
- Tabletas de Bufferin (EJT), 400
- Tasa de clamidia (R), 193
- Tasa de nacimientos (E), 751
- Teléfonos celulares y cáncer (E), 147, 229, 300, 334, 415-416
- Temperaturas corporales (EJT), 8, 285-286, 370, 684; (E), 87, 90, 105-106, 108, 117, 128, 267, 278, 360, 373, 423, 433, 687-688; (R), 447; (MB), 695
- Terapia de parches de nicotina de dosis alta (EJT), 330-331
- Terapia hormonal (M), 23
- Trastorno vinculado al cromosoma X (E), 172, 210; (RA), 379
- Tratamiento de enfermedad del movimiento (E), 490
- Tratamiento de la sífilis (E), 31
- Tratamiento del pie de atleta (EJT), 621-624; (E), 626; (MB), 627
- Tratamiento del síndrome de fatiga crónica (RA), 193-194; (E), 433
- Tratamiento magnético del dolor (E), 480-481, 503
- Tratamiento para la depresión bipolar (E), 480, 503
- Xynamine para disminuir la frecuencia de pulso (DD), 671

Tecnología

- Componente de computadora defectuoso (MB), 224
- Configuraciones de teclado (DD), 133; (E), 493
- Contraseña para computadora (E), 165
- Diseño de computadora (E), 187
- Inteligencia de las computadoras (MB), 190
- Navegación en Internet (E), 688
- Nombres de variables de cómputo (MB), 189
- Periodo de vida de teléfonos celulares (E), 432
- Porcentaje de usuarios del correo electrónico (E), 415
- Reparación de computadora (R), 753
- Termómetros científicos (EJT), 251-253, 255-256; (E), 257-258
- Uso de Internet (EJT), 296-297
- Vida de una computadora de escritorio (E), 360, 432

Temas sociales

- Aceptación de una cita (E), 172
 Actitudes hacia el matrimonio (E), 467
 Asesinatos y tamaño de la población (E), 536, 555
 Campaña de Napoleón para invadir Moscú en 1812 (EJT), 63-64; (E), 68
 Ciudades clasificadas según su "habitabilidad" (EJT), 8
 Conducción bajo los efectos del alcohol (E), 31, 167
 Control poblacional (MB), 179
 Correo electrónico y privacidad (E), 465
 Dinero gastado en la asistencia social (EJT), 16
 Discriminación por edad (E), 187, 478; (EJT), 471-474
 Discriminación por género (E), 223, 300; (R), 312, 726
 Eficacia de las prohibiciones del tabaquismo (E), 465-466
 Ergonomía (E), 32
 Estudiantes que beben (E), 32
 Filas de espera (E), 374
 Habilidades sociales (E), 166
 Hacer fila (M), 231
 Hogares con teléfono (E), 335
 Hogares en Estados Unidos (EJT), 17
 Llamadas telefónicas (E), 234
 Muertes (R), 238
 Muertes de peatones (E), 156-157, 167
 Muertes en hospitales militares británicos (EJT), 65; (ACE), 71
 Muertes en vehículos automotores y asesinatos (E), 715
 Muertes por homicidio (E), 235
 Peatones intoxicados (E), 167
 Pistolas y tasa de asesinatos (E), 533-534
 Población de Estados Unidos (EJT), 578-579
 Población en 2050 (MB), 581
 Poblaciones cambiantes (M), 78
 Política de servicio de alcohol (R), 377
 Programa de acción afirmativa (E), 223
 Propiedad de armas (E), 10, 11
 Prueba de drogas (PC), 137; (E), 149, 172; (EJT), 152-153, 154, 160-161, 170-171
 Reconstrucción de las torres del World Trade Center (R), 34
 Tamaño de la multitud (M), 357
 Tamaño de la población (E), 579
 Tasa de divorcios (E), 751
 Tasa de matrimonios (ACE), 755
 Uso de la basura para predecir el tamaño de la población (E), 575

Trabajo

- Comparación de ingresos (R), 507
 Conseguir trabajo por medio de contactos (PC), 385; (EJT), 391, 392, 398, 408

Contratación de solicitantes de empleo (MB), 689

- Desempleo (EJT), 29
 Empleo (E), 32
 Empleos en el campo de la estadística (M), 577
 Error de tipografía en solicitud de empleo (EJT), 5
 Errores de entrevista (E), 446
 Fuentes de empleo (E), 67
 Ganancias de hombres y mujeres (DD), 38
 Lesiones laborales fatales (E), 67
 Prueba de drogas a solicitantes de empleo (E), 415
 Razones de despido (R), 238
 Riesgos laborales (E), 619
 Salario de maestros (EJT), 82
 Salario y demanda física (E), 715
 Salario y estrés (E), 715
 Salarios de mujeres ejecutivas (E), 10
 Satisfacción laboral (R), 131; (E), 404

Transporte

- Analista de tráfico (ET), 731
 Anchura de cadera y asientos de aeronaves (E), 268
 Antigüedad de automóviles conducidos por estudiantes (R), 131
 Antigüedad de automóviles de profesores y estudiantes (E), 504
 Autobuses de casino (E), 288
 Beber y conducir (R), 665
 Botones para paso de peatones (EJT), 5; (E), 10, 148, 464
 Brazaletes magnéticos para pasajeros de crucero (E), 31
 Cargas seguras en aviones y barcos (PC), 245; (EJT), 291-292, 293-294
 Cascos de motocicleta (E), 19
 Cascos y lesiones faciales en accidentes de bicicleta (EJT), 614-615
 Choques de automóviles (DD), 72; (E), 414, 415, 602, 688
 Colores de automóviles (EJT), 7
 Conducir al trabajo (E), 466
 Datos de confiabilidad de automóviles (EJT), 63
 Desaceleración del pecho en un choque automovilístico (E), 652
 Diferencia de antigüedad de automóviles y taxis (E), 482, 700
 Diferencia de género en el uso del cinturón de seguridad (E), 466
 Edades de motociclistas muertos en choques (E), 347
 El uso del cinturón de seguridad es independiente del tabaquismo (E), 619-620
 Equipaje perdido (EJT), 14-15; (E), 466
 Equipo de navegación usado en aviones (M), 351
- Errores de altímetro de aeronave (PC), 733; (EJT), 735-736, 740-741, 743-744, 749-750
 Estándares de seguridad de aeronaves (E), 290
 Fallas de aterrizaje (E), 67
 Ford y Mazda producen transmisiones similares (M), 497
 Hombre dueño de motocicleta (E), 172
 ¿Las bolsas de aire salvan vidas? (M), 487
 Lesiones y color de casco de motocicleta (EJT), 606-610; (E), 620
 Longevidad de acumuladores para automóviles (MB), 128-129
 Los asientos más seguros en un avión (M), 594
 Miedo a volar (DD), 512
 Motores de aviones (M), 160; (EJT), 205
 Muertes de peatones (E), 617
 Muertes en motocicleta (ACE), 71; (E), 602
 Muertes y estado de ebriedad los fines de semana (R), 629
 Neumático desinflado y clase perdida (E), 602
 Nitrógeno en neumáticos (E), 336
 Pasajeros de aerolíneas con equipaje de mano (EJT), 391
 Peso de automóviles y consumo de combustible (E), 536, 554, 564
 Peso de automóviles y lesiones (E), 706; (R), 726
 Pesos de pasajeros de taxi acuático (EJT), 260-261, 263-264, 283-285
 Probabilidad del choque de un automóvil (E), 148
 Rutas aéreas (E), 187
 Seguridad de motocicletas (E), 177, 178
 Seguridad de teleférico (E), 288
 Sistema en línea para el registro de conductores (E), 33
 Sobreventa de boletos en vuelos (E), 223, 300; (PT), 241; (MB), 302
 Sobrevivientes del Titanic (E), 10
 Tasa de consumo de combustible (E), 565
 Techo corredizo y bolsas de aire laterales (E), 336
 Teléfonos celulares y choques (RA), 509
 Tiempo de propiedad de automóviles (R), 377-378
 Traumatismo craneal en un choque de automóvil (EJT), 652
 Valor de un automóvil (E), 34
 Velocidad promedio (MB), 91
 Velocidades en carretera (RA), 509
 Vida operativa de un avión (M), 141

ESTADÍSTICA

Décima edición



Introducción a la estadística

1



- 1-1 Panorama general**
- 1-2 Tipos de datos**
- 1-3 Pensamiento crítico**
- 1-4 Diseño de experimentos**



Seis grados de Kevin Bacon: ¿el estudio original utilizó buenos datos?

“Seis grados de Kevin Bacon” es un juego popular reciente, que consiste en identificar a un actor o a una actriz de cine, y luego vincularlo con el actor Kevin Bacon. (En el momento en que se escribió esto, el juego podía jugarse en el sitio Web www.cs.virginia.edu/oracle). Consideremos a Richard Gere como ejemplo. Gere actuó en la película *Cotton Club* con Laurence Fishburne, que trabajó en la película *Mystic River* con Kevin Bacon. El vínculo Gere-Fishburne-Bacon tiene *dos* grados de separación porque no se cuenta la persona meta. Este juego, creado por tres estudiantes (Craig Fass, Brian Turtle y Mike Ginelly) de Albright College, es una versión más especializada de “Small World Problem”, que plantea la siguiente pregunta: ¿Cuántos intermediarios (amigos, parientes y otros conocidos) se necesitan para conectar a cualesquiera dos personas elegidas al azar en la Tierra? Es decir, para cualesquiera dos personas en nuestro planeta, ¿cuál es el número de grados de separación? Este problema de conexión tiene aplicaciones prácticas en muchos campos, como las redes eléctricas, el uso de Internet, las neuronas del cerebro y la propagación de enfermedades.

El concepto de “seis grados de separación” surgió de un estudio realizado en 1967 por el psicólogo Stanley Milgram, quien originalmente describió que en Estados Unidos dos residentes al azar están conectados por un promedio de seis intermediarios. En su primer experimento, Milgram envió 60 cartas a personas de Wichita, Kansas, a quienes les pidió que reenviaran esas cartas a una mujer específica en Cambridge, Massachusetts. A esas personas se les dio la instrucción de entregar en

mano las cartas a conocidos que, según ellos, podrían contactar a la persona indicada, ya fuera directamente o a través de otros conocidos. Participaron 50 de las 60 personas, y tres cartas llegaron a su destino. Dos experimentos posteriores tuvieron tasas de terminación más bajas; pero finalmente Milgram alcanzó una tasa del 35 por ciento, y describió que cada cadena completa tenía un promedio de alrededor de seis intermediarios. Como consecuencia, los datos originales de Milgram produjeron el concepto “seis grados de separación”.

Veamos dos preguntas clave: ¿Eran adecuados los datos originales de Milgram? ¿Los datos originales de Milgram justifican el concepto de “seis grados de separación?” Un principio *extremadamente* importante en este capítulo, en este libro, y en la estadística en general, es que el método que se utiliza para reunir datos de muestras puede construir o destruir la validez de las conclusiones basadas en los datos.

En la actualidad, a todos nosotros se nos bombardea con encuestas y resultados de encuestas. Algunas reúnen datos de muestras que son útiles porque describen de manera exacta características importantes de poblaciones. Otras encuestas usan datos muestrales recolectados de tal forma que condenan los resultados a la creciente pila de basura de la mala información.

En este capítulo, examinamos la pregunta sobre la calidad de los datos del experimento de Stanley Milgram, y analizamos y destacamos la importancia de reunir datos usando métodos sólidos que puedan llevar a conclusiones que sean válidas.

1-1 Panorama general

El problema del capítulo en la página anterior se refiere a un estudio que produjo datos muestrales. Una meta común de este tipo de estudios consiste en reunir datos de una pequeña parte de un grupo más grande, para aprender algo acerca de este último. Una meta común e importante de la materia de la estadística es la siguiente: aprender acerca de un grupo grande examinando los datos de algunos de sus miembros. En dicho contexto, los términos *muestra* y *población* adquieren importancia. Las definiciones formales de estos y otros términos básicos se presentan a continuación.

Definiciones

Datos son las observaciones recolectadas (como mediciones, géneros, respuestas de encuestas).

Estadística es un conjunto de métodos para planear estudios y experimentos, obtener datos y luego organizar, resumir, presentar, analizar, interpretar y llegar a conclusiones basadas en los datos.

Población es el conjunto completo de todos los elementos (puntuaciones, personas, medidas, etcétera) que se va a estudiar. El conjunto es completo porque incluye a todos los sujetos que se estudiarán.

Censo es el conjunto de datos de *cada uno* de los miembros de la población.

Muestra es un *subconjunto* de miembros seleccionados de una población.

Por ejemplo, una encuesta Gallup preguntó a 1087 adultos: “¿Consumió usted bebidas alcohólicas como licor, vino o cerveza, o es totalmente abstemio?” Los 1087 sujetos de la encuesta constituyen una *muestra*: en tanto que la *población* consiste en el conjunto completo de los 202,682,345 adultos estadounidenses. Cada 10 años, el gobierno de Estados Unidos intenta obtener un *censo* de cada ciudadano, pero no lo logra debido a que es imposible localizar a cada uno de ellos. Una polémica actual gira en torno al intento de emplear métodos estadísticos sólidos para aumentar la exactitud del censo, aunque los aspectos políticos constituyen un factor clave para que los miembros del Congreso se resistan a tal mejoría. Quizás algún día algunos lectores de este libro sean miembros del Congreso y tengan la sabiduría de trasladar el censo al siglo XXI.

Una actividad importante de este libro es demostrar cómo usar los datos muestrales para obtener conclusiones sobre poblaciones. Veremos que es *sumamente* importante obtener datos muestrales que sean representativos de la población de la que se obtienen. Por ejemplo, si usted encuesta a los alumnos que se graduaron de su universidad, y les pide que anoten su ingreso anual y que se lo envíen por correo, es probable que las respuestas no sean representativas de la población de todos los alumnos. Quienes tengan ingresos más bajos serán menos proclives a responder, y los que respondan tal vez se sientan inclinados a exagerar. Al avanzar en este capítulo debemos enfocarnos en los siguientes conceptos clave:

- Los datos muestrales deben reunirse de una forma adecuada, como a través de un proceso de selección *aleatoria*.
- Si los datos muestrales no se reúnen de forma apropiada, resultarán tan inútiles que ninguna cantidad de tortura estadística podrá salvarlos.

Ante todo, le pedimos que inicie su estudio de la estadística con una mente abierta. No dé por hecho que el estudio de la estadística es comparable con un procedimiento inflexible. Según la experiencia del autor, a menudo los estudiantes se sorprenden por la interesante naturaleza de la estadística, y también por el hecho de que en realidad pueden dominar los principios básicos sin gran dificultad, incluso si no sobresalen en otros cursos de matemáticas. Estamos convencidos de que cuando termine este curso introductorio, tendrá la firme creencia de que la estadística es una materia rica e interesante, con aplicaciones extensas, reales y significativas. También estamos convencidos de que con la asistencia a clases y la dedicación constantes, usted tendrá éxito al dominar los conceptos básicos de la estadística presentados en este curso.



El estado de la estadística

El término *estadística* se deriva de la palabra latina *status* (que significa “estado”). Los primeros usos de la estadística implicaron la recopilación de datos y la elaboración de gráficas, para describir diversos aspectos de un estado o de un país. En 1662 John Graunt publicó información estadística acerca de los nacimientos y los decesos. Al trabajo de Graunt siguieron estudios de tasas de mortalidad y de enfermedad, tamaño de poblaciones, ingresos y tasas de desempleo. Los hogares, gobiernos y empresas se apoyan mucho en datos estadísticos para dirigir sus acciones. Por ejemplo, se reúnen datos de manera cuidadosa y con regularidad para establecer las tasas de desempleo, las tasas de inflación, los índices del consumidor y las tasas de nacimientos y muertes; en tanto que los líderes empresariales utilizan los datos resultantes para tomar decisiones que afectan futuras contrataciones, los niveles de producción y la expansión hacia nuevos mercados.

1-2 Tipos de datos

Concepto clave En la materia de estadística se trata principalmente de utilizar datos muestrales para hacer inferencias (o generalizaciones) sobre una población completa. Debemos saber y entender las definiciones de *población*, *muestra*, *parámetro* y *estadístico*, ya que son básicas y fundamentales. También necesitamos reconocer la diferencia entre *datos cuantitativos* y *datos cualitativos*. Tenemos que entender que algunos números, como los códigos postales, no son cantidades en el sentido de que realmente midan o cuenten algo. Los códigos postales son, en realidad, ubicaciones geográficas, por lo que no tiene sentido hacer cálculos con ellos, como calcular su promedio. En esta sección se describen distintos aspectos de la naturaleza de los datos muestrales, los cuales pueden afectar de manera importante los métodos estadísticos que se utilicen con ellos.

En la sección 1-1 definimos los términos *población* y *muestra*. Los siguientes dos términos se utilizan para distinguir entre los casos en que tenemos datos de una población completa y los casos donde sólo tenemos datos de una muestra.

Definiciones

Parámetro es una medición numérica que describe algunas características de una *población*.

Estadístico es una medición numérica que describe algunas características de una *muestra*.

EJEMPLOS

- Parámetro:** En la ciudad de Nueva York hay 3250 botones para caminar, que los peatones emplean en las intersecciones de tránsito. Se descubrió que el 77% de dichos botones no funciona (según datos del artículo “For Exercise in New York Futility, Push Button”, de Michael Luo, *New York Times*). La cifra del 77% es un *parámetro* porque está basada en la población de todos los 3250 botones para peatones.
- Estadístico:** Con base en una muestra de 877 ejecutivos encuestados, se encontró que el 45% de ellos no contrataría a alguien con un error ortográfico en su solicitud de empleo. Esta cifra del 45% es un *estadístico*, ya que está basada en una muestra y no en la población completa de todos los ejecutivos.

Algunos conjuntos de datos consisten en números (como alturas de 66 y 72 pulgadas); mientras que otros son no numéricos (como los colores de ojos verde y café). Los términos *datos cuantitativos* y *datos cualitativos* suelen utilizarse para distinguir entre ambos tipos.

Definiciones

Los **datos cuantitativos** consisten en números que representan conteos o mediciones.

Los **datos cualitativos** (o **categóricos** o **de atributo**) se dividen en diferentes categorías que se distinguen por algunas características no numéricas.

EJEMPLOS

1. **Datos cuantitativos:** Los pesos de las supermodelos.

2. **Datos cualitativos:** El género (hombre/mujer) de atletas profesionales.

Cuando se trabaja con datos cuantitativos, es importante utilizar las unidades de medida apropiadas, como dólares, horas, pies, metros, etcétera. Debemos ser especialmente cuidadosos para observar aquellas referencias como “todas las cantidades están en *miles de dólares*” o “todos los tiempos están en *centésimas de segundo*” o “las unidades están expresadas en *kilogramos*”. Ignorar unidades de medida como éstas nos llevaría a conclusiones incorrectas. La NASA perdió su Mars Climate Orbiter de \$125 millones* cuando la sonda se estrelló debido a que el programa de control tenía los datos de aceleración en unidades *inglesas*, pero ellos incorrectamente consideraron que estaban en unidades *métricas*.

Los datos cuantitativos se describen con mayor detalle distinguiendo entre los tipos *discreto* y *continuo*.

Definiciones

Los **datos discretos** resultan cuando el número de valores posibles es un número finito o un número que “puede contarse” (es decir, el número de valores posibles es 0, 1, 2, etcétera).

Los **datos continuos (numéricos)** resultan de un infinito de posibles valores que corresponden a alguna escala continua que cubre un rango de valores sin huecos, interrupciones o saltos.

EJEMPLOS

1. **Datos discretos:** El número de huevos que ponen las gallinas son datos *discretos* porque representan conteos.

2. **Datos continuos:** Las cantidades de leche que producen las vacas son datos *continuos* porque son mediciones que pueden tomar cualquier valor dentro de un continuo. Durante un intervalo de tiempo dado, una vaca produce una cantidad de leche que puede ser cualquier valor entre 0 y 5 galones. Es posible obtener 2.343115 galones, porque la vaca no está restringida a cantidades discretas de 0, 1, 2, 3, 4 o 5 galones.

*En esta obra, el signo \$ hará referencia a dólares estadounidenses, a menos que se especifique otra unidad monetaria.

Un ejemplo más: el número de latas de bebidas de cola son datos discretos; en tanto que el volumen real de la bebida de cola es un dato continuo.

Otra forma común de clasificar los datos consiste en usar cuatro niveles de medición: nominal, ordinal, de intervalo y de razón. Cuando se aplica la estadística a problemas reales, el nivel de medición de los datos es un factor importante para determinar el procedimiento a utilizar. (Véase la figura 15-1 en la página 764). En este libro encontraremos algunas referencias a estos niveles de medición; sin embargo, lo importante aquí se basa en el sentido común: no hay que hacer cálculos ni utilizar métodos estadísticos que no sean apropiados para los datos. Por ejemplo, no tendría sentido calcular el promedio de los números del seguro social, ya que estos números son datos que se utilizan como identificación, y no representan mediciones o conteos de algo. Por la misma razón, no tendría sentido calcular un promedio de los números que aparecen en las camisetas de los jugadores de básquetbol.

Definición

El **nivel de medición nominal** se caracteriza por datos que consisten exclusivamente en nombres, etiquetas o categorías. Los datos no se pueden acomodar en un esquema de orden (como del más bajo al más alto).

EJEMPLOS Veamos algunos ejemplos de datos muestrales a nivel de medición nominal.

1. **Sí/no/indeciso:** Respuestas de *sí*, *no* e *indeciso* en una encuesta
2. **Colores:** Los colores de los automóviles conducidos por estudiantes universitarios (rojo, negro, azul, blanco, magenta, púrpura, etcétera)

Puesto que los datos nominales carecen de orden y no tienen un significado numérico, no se deben utilizar para hacer cálculos. En ocasiones se asignan números a las distintas categorías (especialmente cuando los datos se codifican para utilizarse en computadoras), pero estos números no tienen un significado computacional real y cualquier promedio que se calcule carece de sentido.

Definición

Los datos están en el **nivel de medición ordinal** cuando pueden acomodarse en algún orden, aunque no es posible determinar diferencias entre los valores de los datos o tales diferencias carecen de significado.

EJEMPLOS Veamos algunos ejemplos de datos muestrales a nivel de medición ordinal.

1. **Las calificaciones de un curso:** Un profesor universitario asigna calificaciones de A, B, C, D, E o F. Tales calificaciones se pueden ordenar, aunque *continúa*



Medición de la desobediencia

¿De qué manera se recolectan datos sobre algo que parece que no es medible, como el nivel de desobediencia de la gente? El psicólogo Stanley Milgram diseñó el siguiente experimento: Un investigador enseñó a un sujeto voluntario a operar un tablero de control que administraba “choques eléctricos” cada vez más dolorosos a una tercera persona. En realidad no se daban tales choques, y la tercera persona era un actor. El voluntario iniciaba con 15 voltios y fue instruido para incrementar los choques en aumentos de 15 voltos. El nivel de desobediencia era el punto donde el sujeto se negaba a incrementar el voltaje. Fue sorprendente que dos terceras partes de los sujetos obedecieron las órdenes, aun cuando el actor gritaba y fingía sufrir un ataque cardiaco.

no es posible determinar diferencias entre tales calificaciones. Por ejemplo, sabemos que A es mayor que B (por lo que hay un orden); pero no podemos restar B de A (por lo que no se puede calcular la diferencia).

2. **Rangos:** Con base en varios criterios, una revista ordena las ciudades de acuerdo con su “habitabilidad”. Dichos rangos (primero, segundo, tercero, etcétera) determinan un orden. Sin embargo, las diferencias entre los rangos no tienen ningún significado. Por ejemplo, una diferencia “del segundo menos el primero” sugeriría $2 - 1 = 1$, pero esta diferencia de 1 no tiene significado porque no es una cantidad exacta que sea comparable con otras diferencias de este tipo. La diferencia entre la primera y la segunda ciudades no es la misma que la diferencia entre la segunda y la tercera ciudades. Utilizando los rangos de la revista, la *diferencia* entre la ciudad de Nueva York y Boston no se puede comparar de forma cuantitativa con la *diferencia* entre San Luis y Filadelfia.

Los datos ordinales proporcionan información sobre comparaciones relativas, pero no las magnitudes de las diferencias. Por lo general, los datos ordinales no deben utilizarse para hacer cálculos como promedios, aunque en ocasiones esta norma se infringe (como sucede cuando utilizamos calificaciones con letras para calcular una calificación promedio).



Definición

El **nivel de medición de intervalo** se parece al nivel ordinal, pero con la propiedad adicional de que la diferencia entre dos valores de datos cualesquiera tiene un significado. Sin embargo, los datos en este nivel no tienen punto de partida cero *natural* inherente (donde *nada* de la cantidad está presente).

EJEMPLOS Los siguientes ejemplos ilustran el nivel de medición de intervalo.

1. **Temperaturas:** Las temperaturas corporales de 98.2°F y 98.6°F son ejemplos de datos a nivel de medición de intervalo. Dichos valores están ordenados, y podemos determinar su diferencia de 0.4°F . Sin embargo, no existe un punto de inicio natural. Pareciera que el valor de 0°F es un punto de inicio; sin embargo, éste es arbitrario y no representa la ausencia total de calor. Puesto que 0°F no es un punto de partida cero natural, sería incorrecto decir que 50°F es *dos veces* más caliente que 25°F .
2. **Años:** Los años 1000, 2008, 1776 y 1492. (El tiempo no inició en el año 0, por lo que el año 0 es arbitrario y no constituye un punto de partida cero natural que represente “la ausencia de tiempo”).



Definición

El **nivel de medición de razón** es similar a nivel de intervalo, pero con la propiedad adicional de que sí tiene un punto de partida cero natural (donde el cero indica que *nada* de la cantidad está presente). Para valores a este nivel, tanto las diferencias como las proporciones tienen significado.

EJEMPLOS Los siguientes son ejemplos de datos al nivel de medición de razón. Observe la presencia de un valor cero natural, así como el uso de proporciones que significan “dos veces” y “tres veces”.

1. **Pesos:** Los pesos (en quilates) de anillos de compromiso de diamante (el 0 realmente representa la ausencia de peso y 4 quilates es dos veces el peso de 2 quilates).
2. **Precios:** Los precios de libros de texto universitarios (\$0 realmente representa ningún costo y un libro de \$90 es tres veces más caro que un libro de \$30).

Este nivel de medición se denomina de razón porque el punto de partida cero hace que las razones o cocientes tengan significado. Entre los cuatro niveles de medición, la principal dificultad surge al distinguir entre los niveles de intervalo y de razón. *Sugerencia:* Para simplificar esta diferencia, utilice una sencilla “prueba de razón”: Consideré dos cantidades en las cuales un número es dos veces el otro y pregúntese si “dos veces” sirve para describir correctamente las cantidades. Puesto que un peso de 200 libras es *dos veces* más pesado que un peso de 100 libras, pero 50°F *no es dos veces* más caliente que 25°F, los pesos están en el nivel de razón, mientras que las temperaturas Fahrenheit están en el nivel de intervalo. Para una comparación y un repaso concisos, estudie la tabla 1-1 que señala las diferencias entre los cuatro niveles de medición.

Tabla 1-1 Niveles de medición de datos			
Nivel	Resumen	Ejemplo	
Nominal	Sólo categorías Los datos no pueden acomodarse en un esquema de orden.	Origen de estudiantes: 5 californianos 20 texanos 40 neoyorquinos	{ Sólo categorías o nombres.
Ordinal	Las categorías están ordenadas, pero no hay diferencias o carecen de significado.	Automóviles de estudiantes: 5 compactos 20 medianos 40 grandes	Orden determinado por “compacto, mediano, grande”.
De intervalo	Las diferencias tienen un significado, pero no hay punto de partida cero natural, y los cocientes no tienen significado.	Temperaturas del campus: 5°F 20°F 40°F	0°F no significa “sin calor”. 40°F no es dos veces más caliente que 20°F.
De razón	Hay un punto de partida cero natural y los cocientes tienen significado.	Distancias de traslado de los estudiantes: 5 km 20 km 40 km	{ 40 km es <i>dos veces</i> más lejos que 20 km.

1-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Parámetro y estadístico.** ¿Cuál es la diferencia entre un parámetro y un estadístico?
2. **Datos cualitativos y cuantitativos.** ¿Cuál es la diferencia entre los datos cualitativos y los datos cuantitativos?
3. **Datos discretos y continuos.** ¿Cuál es la diferencia entre los datos discretos y los datos continuos?
4. **Datos continuos y cuantitativos.** Si un experimento produce datos que son de naturaleza continua, ¿los datos también deben ser cuantitativos o pueden ser cualitativos?

En los ejercicios 5 a 8, determine si el valor dado es un estadístico o un parámetro.

5. **Tamaño de la familia.** Se selecciona una muestra de hogares y el número promedio (media) de personas por familia es de 2.58 (según datos de la Oficina censal estadounidense).
6. **Política.** En la actualidad, el 42% de los gobernadores de las 50 entidades de Estados Unidos son demócratas.
7. **Titanic.** En un estudio de los 2223 pasajeros del *Titanic*, se encontró que 706 sobrevivieron cuando se hundió.
8. **Audiencia televisiva.** Se selecciona una muestra de estadounidenses y se descubre que la cantidad de tiempo promedio (media) que ven la televisión es de 4.6 horas al día.

En los ejercicios 9 a 12, determine si los valores dados provienen de un conjunto de datos discreto o continuo.

9. **Experimento de correo.** En el problema del capítulo se señaló que cuando se enviaron 50 cartas como parte de un experimento, tres de ellas llegaron a la dirección indicada.
10. **Botones para peatones.** En la ciudad de Nueva York hay 3250 botones para cruzar, que los peatones presionan en las intersecciones de tránsito, y 2500 de ellos no funcionan (según datos del artículo “For Exercise in New York Futility, Push Button”, de Michael Luo, *New York Times*).
11. **Peso de peniques.** El peso promedio de los peniques o centavos que actualmente se acuñan en Estados Unidos es de 2.5 gramos.
12. **Propiedad de armas.** En una encuesta realizada con 1059 adultos, se encontró que el 39% de ellos tienen armas en sus hogares (basado en una encuesta Gallup).

En los ejercicios 13 a 20, determine cuál de los cuatro niveles de medición (nominal, ordinal, de intervalo, de razón) es el más apropiado.

13. **Maratón.** Los números en las camisetas de los corredores de maratones.
14. **Producto de consumo.** Las calificaciones que da la revista *Consumer Reports* de “la mejor compra, recomendado, no recomendado”.
15. **NSS.** Los números de seguridad social.
16. **Encuesta de bebidas.** El número de respuestas “sí” recibidas cuando se les preguntó a 500 estudiantes si alguna vez se habían embriagado en la universidad.
17. **Cigarras.** Los años de aparición de cigarras: 1936, 1953, 1970, 1987 y 2004.
18. **Mujeres ejecutivas.** Los salarios de mujeres que son directoras generales de corporaciones.

19. Calificaciones. Calificaciones de las películas de una estrella, dos estrellas, tres estrellas y cuatro estrellas.

20. Temperaturas. Las temperaturas actuales en las capitales de las 50 entidades de Estados Unidos.

En los ejercicios 21 a 24, identifique a) la muestra y b) la población. Además, determine si la muestra parece ser representativa de la población.

21. Proyecto de investigación. Un científico político selecciona al azar a 25 de los 100 senadores que actualmente conforman el Congreso, y luego calcula la cantidad de tiempo que han prestado servicio.

22. Nivel de audiencia de Nielsen. Durante el juego del Super Tazón, una encuesta de 5108 hogares elegidos al azar revela que el 44% de ellos tienen sus televisores sintonizados en el juego (según datos de Nielsen Media Research).

23. Propiedad de armas. En una encuesta Gallup de 1059 adultos seleccionados al azar, el 39% respondió que “sí” cuando se les preguntó “¿tiene un arma en su casa?”

24. Encuesta por correo. Una estudiante de posgrado de la Universidad de Newport realiza un proyecto de investigación sobre la comunicación. Ella envía por correo una encuesta a los 500 adultos que conoce, y les pide que respondan y regresen por correo la siguiente pregunta: “¿Prefiere utilizar el correo electrónico o el correo ordinario (el servicio postal)?” Ella recibe 65 respuestas, y 42 de ellas indican una preferencia por el correo ordinario.

1-2 MÁS ALLÁ DE LO BÁSICO

25. Interpretación de los incrementos de temperatura. En la tira cómica “Born Loser” de Art Samson, Brutus se alegra por un incremento en la temperatura de 1° a 2° . Cuando él le pregunta qué tiene de bueno estar a 2° , él responde que “hace dos veces más calor que en la mañana”. Explique por qué Brutus se equivoca nuevamente.

26. Interpretación de encuesta política. Un encuestador le pregunta a 200 personas por el partido político de su preferencia, y codifica las respuestas como 0 (para los demócratas), 1 (para los republicanos), 2 (para los independientes) o 3 (para cualquier otra respuesta). Luego calcula el promedio (media) de los números y obtiene 0.95. ¿Cómo se interpreta este valor?

27. Escala para calificar comida. Un grupo de estudiantes elabora una escala para calificar la calidad de los alimentos en la cafetería, donde 0 representa “neutral: ni buena ni mala”. A las comidas malas se asignan números negativos y a las comidas buenas números positivos, y la magnitud del número corresponde al grado de lo bueno o lo malo. Las primeras tres comidas se califican con 2, 4 y -5. ¿Cuál es el nivel de medición de este tipo de calificaciones? Explique su respuesta.

1-3 Pensamiento crítico

Concepto clave El éxito en el curso de introducción a la estadística generalmente requiere más *sentido común* que conocimientos matemáticos (a pesar de la advertencia de Voltaire de que “el sentido común no es tan común”). Puesto que ahora tenemos acceso a calculadoras y computadoras, las aplicaciones modernas de la estadística ya no nos exigen el dominio de algoritmos complejos de operaciones matemáticas. En cambio, nos enfocamos en la *interpretación* de datos y resultados.

Esta sección está diseñada para ilustrar la forma en que se utiliza el sentido común cuando pensamos críticamente acerca de datos y estadísticos. En esta sección, en vez de memorizar métodos o procedimientos específicos, hay que enfocarse en el pensamiento y el uso del sentido común al analizar datos. Es importante saber que cuando los datos muestrales se reúnen de manera inapropiada, como cuando se utiliza una *muestra de respuestas voluntarias* (que se define más adelante en esta sección), ningún método estadístico es capaz de producir resultados válidos.

Hace cerca de un siglo, el célebre estadista Benjamin Disraeli dijo: “Existen tres clases de mentiras: mentiras, viles mentiras y estadísticas”. Se ha dicho que “las cifras no mienten, pero los mentirosos también usan cifras”. El historiador Andrew Lang dijo que algunas personas usan la estadística “como un borracho utiliza los postes de alumbrado—como apoyo más que como iluminación”. El caricaturista político Don Wright nos anima diciendo “retome el misterio de la vida: mienta a un encuestador”. El autor Franklin P. Jones escribió que “la estadística se puede utilizar para apoyar cualquier cosa—especialmente a los estadísticos”. En el *Esar's Comic Dictionary* encontramos que la definición de un estadístico es “un especialista que reúne cifras y luego hace que se pierdan”. Estas aseveraciones se refieren a casos donde los métodos estadísticos se utilizaron de forma errónea, de manera que resultaron engañosos. Hay dos fuentes principales de este engaño: 1. el intento malintencionado por parte de personas deshonestas y 2. errores no intencionales por parte de personas que no saben mucho. Sin importar la fuente, como ciudadanos responsables y como empleados profesionales más valiosos, deberíamos tener la capacidad básica de distinguir entre conclusiones estadísticas que pueden ser válidas y aquellas que están llenas de errores.

Para mantener esta sección en la perspectiva correcta, es necesario que sepa que este libro no trata sobre los malos usos de la estadística. El resto de este texto incluye usos muy importantes de métodos estadísticos válidos. Aprenderemos métodos generales para usar datos muestrales y así hacer inferencias relevantes acerca de poblaciones. Aprenderemos acerca de encuestas y tamaños de muestras; conoceremos medidas importantes de características fundamentales de los datos. Junto con las discusiones de estos conceptos generales, conoceremos muchas aplicaciones reales específicas, como los efectos del tabaquismo pasivo, la prevalencia del alcohol y el tabaco en las películas de dibujos animados para niños, y la calidad de productos de consumo como los dulces M&M, cereales, Coca-Cola y Pepsi. Pero incluso en esas aplicaciones reales e importantes, debemos tener cuidado de interpretar correctamente los resultados de métodos estadísticos válidos.

Comenzamos nuestro desarrollo del pensamiento crítico considerando muestras erróneas, las cuales son erróneas en el sentido de que el método de muestreo arruina la muestra, de modo que es posible que esté *sesgada* (que no sea representativa de la población de la que se obtuvo). En la siguiente sección analizaremos con mayor detalle los métodos de muestreo, y se describirá la importancia de la *aleatoriedad*. El primer ejemplo describe un procedimiento de muestreo que carece gravemente de la aleatoriedad, la cual es muy importante. La siguiente definición se refiere a uno de los usos incorrectos de la estadística más comunes y graves.



Definición

Una **muestra de respuesta voluntaria** (o **muestra autoseleccionada**) es aquella en que los propios sujetos deciden ser incluidos.

Por ejemplo, la revista *Newsweek* hizo una encuesta sobre el controvertido sitio Web Napster, que estuvo permitiendo el libre acceso a la copia de CD musicales. A los lectores se les planteó la siguiente pregunta: “¿Continuaría utilizando Napster si tuviera que pagar una cuota?” Los lectores podían registrar sus respuestas en el sitio Web newsweek.msnbc.com. De las 1873 respuestas recibidas, el 19% dijo que sí, porque continúa siendo más barato que comprar los CD. Otro 5% dijo que sí, que se sentirían más cómodos utilizándolo por una cuota. Cuando *Newsweek* o alguien más hace una encuesta por Internet, los propios individuos deciden participar, por lo que constituyen una muestra de respuesta voluntaria. Sin embargo, las personas con opiniones extremas son más proclives a participar, por lo que sus respuestas no son representativas de toda la población. A continuación se presentan algunos ejemplos de muestras de respuesta voluntaria que, por su naturaleza, tienen graves errores y no deberíamos obtener conclusiones sobre una población con base en muestras sesgadas como éstas:

- Encuestas por Internet, donde los sujetos deciden si responden o no.
- Encuestas por correo, donde los sujetos deciden si responden o no.
- Encuestas telefónicas, donde anuncios televisivos, de radio o de periódicos le piden que llame voluntariamente a un número especial para registrar su opinión.

Con este tipo de muestras de respuesta voluntaria sólo se logran conclusiones válidas sobre el grupo de gente específico que decidió participar; aunque una práctica común consiste en afirmar o sacar conclusiones incorrectas sobre una población más grande. Desde un punto de vista estadístico, una muestra de este tipo es defectuosa y no debe usarse para hacer afirmaciones generales sobre una población más grande.

Muestras pequeñas Las conclusiones no se deben basar en muestras demasiado pequeñas. Como ejemplo, el Children's Defense Fund publicó Children Out of School in America, donde se reportó que, de los estudiantes de secundaria suspendidos en una región, el 67% fueron suspendidos al menos tres veces. ¡Pero esta cifra está basada en una muestra de tan sólo tres estudiantes! Los informes en los medios de comunicación no mencionaron que el tamaño de la muestra era muy pequeño. (En los capítulos 7 y 8 veremos que en ocasiones podemos hacer algunas inferencias valiosas a partir de muestras pequeñas, pero debemos tener el cuidado de verificar si se satisfacen los requisitos necesarios).

En ocasiones una muestra parecería relativamente grande (como en una encuesta de “2000 adultos estadounidenses elegidos al azar”); no obstante, si se sacan conclusiones sobre subgrupos, como los hombres republicanos de 21 años de edad de Pocatello, este tipo de conclusiones podrían estar basadas en muestras demasiado pequeñas. Aunque es importante tener una muestra que sea lo suficientemente grande, también es importante contar con datos muestrales que se obtengan de manera apropiada, como en la selección aleatoria. Incluso muestras grandes llegan a ser muestras inadecuadas.

Gráficas Las gráficas, como las de barras y las de pastel (circulares), se pueden utilizar para exagerar o subestimar la verdadera naturaleza de los datos. (En el capítulo 2 analizaremos diferentes tipos de gráficas). Las dos gráficas de la figura 1-1 describen los *mismos datos* obtenidos del U.S. Bureau of Economic Analysis, aunque el inciso b) está diseñado para exagerar la diferencia entre el ingreso personal

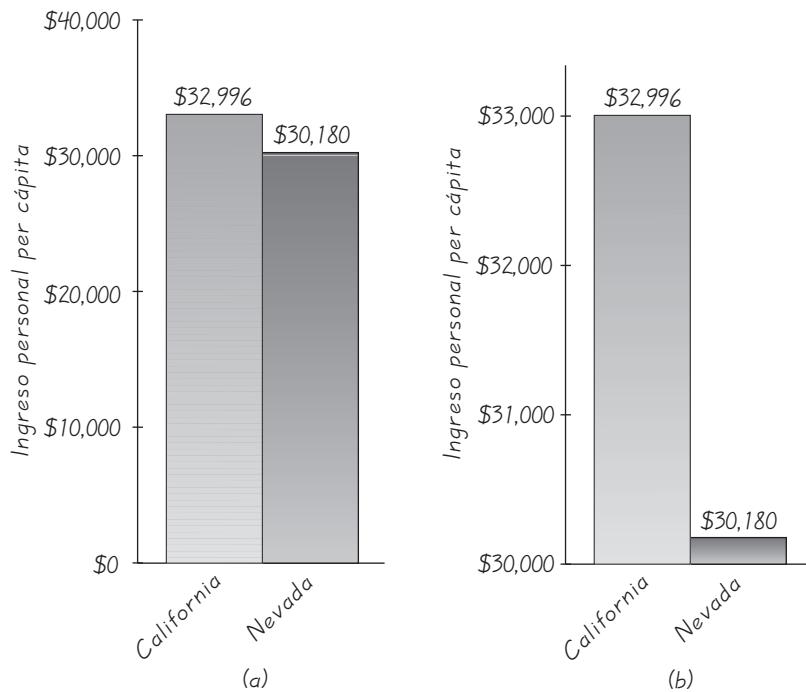


La encuesta del Literary Digest

En la contienda presidencial de 1936, la revista *Literary Digest* realizó una encuesta y predijo la victoria de Alf Landon; sin embargo, Franklin D. Roosevelt obtuvo una victoria abrumadora. Maurice Bryson señala: “Se enviaron por correo 10 millones de papeletas de muestra para votar a electores potenciales; pero sólo se devolvieron 2.3 millones. Como todo el mundo debía saber, este tipo de muestras prácticamente siempre están sesgadas”. Bryson también afirma que “la respuesta voluntaria a cuestionarios enviados por correo es quizás el método más común que los estadísticos han encontrado para recolectar datos en las ciencias sociales, y quizá también sea el peor”. (Véase el artículo de Bryson “The Literary Digest Poll: Making of a Statistical Mith”, *The American Statistician*, vol. 30, núm. 4).

Figura 1-1

Comparación entre California y Nevada: Ingreso personal per cápita



per cápita en California y en Nevada, su estado vecino. Como el eje vertical no inicia en cero, la gráfica del inciso b) tiende a producir una impresión subjetiva engañosa, provocando que los lectores crean de manera incorrecta que la diferencia es mucho mayor de lo que realmente es. La figura 1-1 enseña una lección importante: Para interpretar de manera correcta una gráfica, tenemos que analizar la información *numérica* que presenta, para no engañarnos por su forma general.

Pictogramas Los dibujos de objetos, llamados *pictogramas*, también suelen ser confusos. Algunos objetos que se utilizan comúnmente para describir datos son los objetos tridimensionales, como las bolsas de dinero, sacos de monedas y tanques del ejército (para gastos militares), barriles (para la producción de petróleo) y casas (para la construcción de viviendas). Al dibujar este tipo de objetos, los artistas podrían crear falsas impresiones que distorsionen las diferencias. Si usted duplica cada lado de un cuadrado, el área no tan sólo se duplica, sino que aumenta en un factor de cuatro. Si usted duplica cada lado de un cubo, el volumen no solamente se duplica, sino que aumenta en un factor de ocho. Vea la figura 1-2, donde el inciso a) está dibujado para describir correctamente la relación entre el consumo diario de petróleo en Estados Unidos y Japón. En la figura 1-2a) parece que Estados Unidos consume aproximadamente cuatro veces más petróleo que Japón. Sin embargo, el inciso b) de la figura 1-2 se dibujó con barriles, en los que *cada dimensión* está dibujada en proporción a las cantidades reales. Vea como la figura 1-2b) exagera mucho la diferencia al crear la falsa impresión de que el consumo de petróleo de Estados Unidos es aproximadamente 50 veces mayor que el de Japón.

Porcentajes En ocasiones se utilizan porcentajes confusos o poco claros. Si usted toma el 100% de alguna cantidad, la está tomando *toda*. (No se requiere del 110% de esfuerzo para entender esta afirmación). Al referirse a la pérdida de equipaje, Continental Airlines publicó anuncios que afirmaban que se trataba de una área que “mejoraron un 100% durante los últimos seis meses”. En una editorial

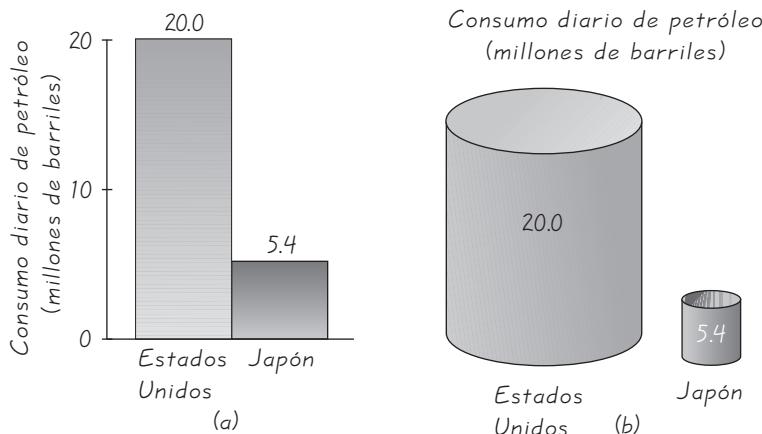


Figura 1-2 Comparación entre Estados Unidos y Japón: Consumo diario de petróleo (millones de barriles)

El inciso b) está diseñado para exagerar la diferencia al incrementar cada dimensión en proporción a las cantidades reales del consumo de petróleo.

que criticaba dicha estadística, el *New York Times* interpretó correctamente que la cifra de mejora en un 100% significa que ya no se pierde el equipaje: un logro que todavía no disfruta Continental Airlines.

Los siguientes son algunos principios clave que se aplican cuando tratamos con porcentajes. Todos estos principios usan la noción básica de que % o “por ciento” realmente significa “dividido entre 100”. Este primer principio se empleará con frecuencia en este libro.

- **Porcentaje de:** Para encontrar el *porcentaje de* una cantidad, excluya el símbolo % y divida el valor del porcentaje entre 100, y después multiplique por la cantidad. Este ejemplo muestra que el 6% de 1200 es 72:

$$6\% \text{ de } 1200 \text{ respuestas} = \frac{6}{100} \times 1200 = 72$$

- **Fracción → porcentaje:** Para convertir de una fracción a un porcentaje, divida el denominador entre el numerador para obtener un número decimal equivalente, después multiplíquelo por 100 y agregue el símbolo %. Este ejemplo muestra que la fracción $\frac{3}{4}$ es equivalente al 75%:

$$\frac{3}{4} = 0.75 \rightarrow 0.75 \times 100\% = 75\%$$

- **Decimal → porcentaje:** Para convertir de un número decimal a un porcentaje multiplíquelo por 100%. Este ejemplo muestra que 0.250 es equivalente a 25.0%:

$$0.250 \rightarrow 0.250 \times 100\% = 25.0\%$$

- **Porcentaje → decimal:** Para convertir de un porcentaje a un número decimal, elimine el símbolo % y divida entre 100. Este ejemplo muestra que el 85% es equivalente a 0.85:

$$85\% = \frac{85}{100} = 0.85$$



Datos estadísticos engañosos en el periodismo

El reportero del *New York Times* Daniel Okrant escribió que, a pesar de que cada oración de su periódico se revisa para lograr claridad y una buena redacción, “los números, tan extraños para muchos, no reciben este trato. El periódico no exige una capacitación específica para mejorar las nociones aritméticas elementales, ni cuenta con especialistas que se dediquen a mejorarlas”. El periodista cita un ejemplo del *New York Times*, donde se informó que se estima que los neoyorquinos gastan más de \$23,000 millones de dólares al año en bienes falsificados. Okrant escribe que “un cálculo aritmético rápido habría demostrado que \$23,000 millones de dólares resultarían en aproximadamente \$8,000 dólares por familia, una cifra evidentemente absurda”.



Detección de datos falsos

Un maestro asigna la tarea de registrar los resultados de lanzar una moneda al aire 500 veces. Un estudiante deshonesto decide ahorrar tiempo inventando los resultados, en vez de realmente lanzar una moneda. Como las personas generalmente no pueden inventar resultados que en realidad sean aleatorios, con frecuencia identificamos datos falsos como éstos. En 500 lanzamientos de una moneda real, es bastante probable que usted obtenga una serie de seis caras o de seis cruces, aunque la gente casi nunca incluye una racha como ésa cuando inventa resultados.

Otra forma de detectar datos fabricados consiste en establecer que los resultados violan la ley de Benford: Para muchos conjuntos de datos, los primeros dígitos no se distribuyen de manera uniforme, sino que los primeros dígitos de 1, 2,..., 9 ocurren con tasas del 30%, 18%, 12%, 10%, 8%, 7%, 6%, 5% y 5%, respectivamente. (Véase “The Difficulty of Faking Data” por Theodore Hill, *Chance*, vol. 12, núm. 3).

Preguntas predispuestas Hay muchos aspectos que afectan las preguntas de encuesta. Éstas pueden estar “cargadas” o redactadas intencionalmente para obtener la respuesta deseada. Observe las tasas reales de la respuesta “sí” para las diferentes formas de redacción de una pregunta:

97% sí: “¿Debería el presidente utilizar su poder de voto para eliminar los desperdicios?”

57% sí: “¿Debería el presidente utilizar su poder de voto o no?”

En *The Superpollster*, David W. Moore describe un experimento en el que se preguntó a diferentes sujetos si estaban de acuerdo con las siguientes afirmaciones:

- Se gasta muy poco dinero en subsidios del Estado.
- Se gasta muy poco dinero en asistencia a los pobres.

Aun cuando los pobres son quienes reciben el subsidio del estado, sólo el 19% estuvo de acuerdo cuando se usaron las palabras “subsidio del Estado”, aunque el 63% estuvo de acuerdo con “asistencia a los pobres”.

Orden de las preguntas En ocasiones las preguntas de una encuesta se sesgan de manera no intencional debido a factores como el orden de los reactivos que se someten a consideración. Observe estas preguntas de una encuesta aplicada en Alemania:

- ¿Cree usted que el tránsito vehicular contribuye a la contaminación del aire más o menos que la industria?
- ¿Cree usted que la industria contribuye a la contaminación del aire más o menos que el tránsito vehicular?

Cuando se presentó primero el tránsito, el 45% culpó a este factor, y el 27% a la industria; cuando la industria se presentó primero, el 24% culpó al tránsito y el 57% culpó a la industria.

Falta de respuesta Existe una *falta de respuesta* cuando alguien se rehúsa a responder una pregunta de encuesta o cuando la persona no está disponible. Cuando se plantean preguntas de encuesta a los individuos, algunos se rehúsan firmemente a responder. La tasa de negativas ha crecido en los últimos años, en parte debido a que muchos televendedores persistentes tratan de vender bienes o servicios, iniciando con un argumento de venta similar a una encuesta de opinión. (A esta estrategia de “vender bajo el disfraz” de una encuesta suele conocerse como *sugging*). En *Lies, Damn Lies, and Statistics*, el autor Michael Wheeler señala de manera correcta que “las personas que se rehúsan a hablar con encuestadores tienden a ser diferentes de las personas que sí hablan con ellos. Algunos sienten temor por los extraños y otros protegen su privacidad; sin embargo, su negativa a hablar demuestra que la perspectiva que tienen del mundo que les rodea es notablemente diferente de la que tienen las personas que reciben a los encuestadores en sus casas”.

Datos faltantes En ocasiones los resultados se ven muy afectados por datos faltantes. A veces faltan datos muestrales por el azar, lo cual implica que la posibilidad de que falte un dato no tiene ninguna relación con sus valores u otros valores. Sin embargo, algunos datos faltan debido a factores especiales, como los individuos con bajos ingresos que son menos proclives a reportar cuánto dinero ganan. Es bien sabido que en el censo de Estados Unidos hay personas faltantes, quienes a menudo pertenecen a los grupos de bajos ingresos o sin hogar. Hace algunos años, las encuestas realizadas por teléfono solían ser inexactas, ya que no incluían a las personas que no tenían dinero suficiente para contar con un teléfono.

Correlación y causalidad En el capítulo 10 de este libro analizaremos la asociación estadística entre dos variables, como la riqueza y el CI. Usaremos el término *correlación* para indicar que las dos variables están relacionadas. Sin embargo, en el capítulo 10 hacemos esta importante aclaración: *correlación no implica causalidad*. Esto significa que cuando encontramos una asociación estadística entre dos variables, no podemos concluir que una de las variables sea causa de (o lo que afecta directamente a) la otra variable. Si encontramos una correlación entre la riqueza y el CI, no podemos concluir que el coeficiente intelectual de una persona afecta directamente su riqueza, ni que la riqueza de una persona afecta directamente su puntuación de CI. En los medios de comunicación masiva son muy comunes los reportes de una correlación recién encontrada con una redacción que indica o implica directamente que una de las variables es causa de la otra.

Estudios para el propio beneficio En ocasiones los estudios reciben el patrocinio de grupos con intereses específicos que buscan promover. Por ejemplo, Kiwi Brands, un fabricante de lustrador para calzado, encargó un estudio que dio como resultado la siguiente aseveración impresa en algunos periódicos: “De acuerdo con una encuesta nacional de 250 empleadores profesionales, la razón más común por la que un solicitante de empleo no logró dar una buena impresión fue llevar los zapatos desaseados”. Debemos ser muy cautos con encuestas como éstas, donde el patrocinador puede obtener ganancias monetarias con base en los resultados. En los últimos años ha aumentado la preocupación por la práctica de las empresas farmacéuticas por financiar a médicos que realizan experimentos clínicos y reportan sus resultados en revistas de prestigio, como *Journal of the American Medical Association*.

Números precisos “En la actualidad hay 103,215,027 hogares en Estados Unidos”. Puesto que esta cifra es muy precisa, mucha gente considera erróneamente que también es *exacta*. En este caso, el número es un estimado y sería mejor decir que el número de hogares es aproximadamente de 103 millones.

Imágenes parciales “El 90% de todos nuestros automóviles, vendidos en este país en los últimos 10 años, continúa circulando”. Millones de consumidores escucharon ese mensaje comercial y no se dieron cuenta de que el 90% de los automóviles que el anunciante vendió en este país se vendieron durante los últimos tres años, por lo que la mayoría de esos automóviles que circulaban estaban casi nuevos. La afirmación era técnicamente correcta, aunque muy engañosa pues no presentaba los resultados completos.

Distorsiones deliberadas En el libro *Tainted Truth*, Cynthia Crossen cita un ejemplo de la revista *Corporate Travel* que publicó resultados que mostraban que, entre las compañías de renta de automóviles, Avis fue la ganadora en una encuesta realizada a personas que utilizan dicho servicio. Cuando Hertz solicitó información detallada sobre la encuesta, las respuestas originales desaparecieron y el coordinador de encuesta de la revista renunció. Hertz demandó a Avis (por publicidad falsa basada en la encuesta) y a la revista; al final se llegó a un acuerdo.

Además de los casos citados anteriormente, existen muchos otros usos inadecuados de la estadística. Algunos de ellos se encuentran en libros como el clásico de Darrel Huff *How to Lie with Statistics*; el de Robert Richard, *The Figure Finaglers*; y el de Cynthia Crossen, *Tainted Truth*. Comprender tales prácticas resultará sumamente útil en la evaluación de los datos estadísticos que se encuentran en situaciones cotidianas.



Sesgo de publicación

Hay un “sesgo de publicación” en las revistas científicas, que es la tendencia a publicar resultados positivos (como demostrar que algún tratamiento es efectivo) con mucha mayor frecuencia que resultados negativos (como demostrar que cierto tratamiento no tiene efecto alguno). En el artículo “Registering Clinical Trials” (*Journal of the American Medical Association*, vol. 290, núm. 4), los autores Kay Dickersin y Drummond Rennie afirman que “el resultado de no saber quién tuvo cuál desempeño (ensayo clínico) es la pérdida y distorsión de la evidencia, el desperdicio y la duplicación de ensayos, la incapacidad de planear de las agencias patrocinadoras, y un sistema caótico del que sólo ciertos patrocinadores se pueden beneficiar, lo cual invariablemente va en contra de los intereses de quienes se ofrecieron a participar en los ensayos y de los pacientes en general”. Esos autores apoyan un proceso donde *todos* los ensayos clínicos queden registrados en un sistema central.

1-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Muestra de respuestas voluntarias.** ¿Qué es una muestra de respuestas voluntarias, y por qué generalmente no es adecuada para los métodos estadísticos?
- 2. Correlación y causalidad.** Si realizamos un análisis estadístico y encontramos que hay una correlación (o una asociación) entre la cantidad de tiempo dedicado a estudiar y las calificaciones en distintos cursos, ¿podemos concluir que un mayor tiempo de estudio genera calificaciones más altas? ¿Por qué?
- 3. Falta de respuestas.** Un investigador determina que necesita los resultados de por lo menos 300 sujetos para realizar un estudio. Para compensar la baja tasa de respuestas devueltas por correo, envía una encuesta a 10,000 sujetos por este medio y recibe 320 respuestas. ¿El grupo de 320 respuestas es una buena muestra?
- 4. Pregunta predisposta.** El teléfono suena y una voz automatizada pregunta si usted está dispuesto a votar por un candidato “con una larga historia de aumento de impuestos y de malversación del dinero de los contribuyentes”. Suponiendo que las llamadas se hacen a individuos elegidos al azar, ¿es probable que los resultados reflejen la preferencia de los votantes por dicho candidato? ¿Por qué?

En los ejercicios 5 a 8, utilice el pensamiento crítico para elaborar una conclusión alternativa. Por ejemplo, considere un informe de los medios de comunicación masiva de que los automóviles BMW causan que las personas estén saludables, ya que se ha descubierto que los individuos que conducen automóviles BMW están más saludables que quienes no lo hacen. Es probable que sea errónea la conclusión de que los automóviles BMW provocan un mejor estado de salud. Veamos una conclusión más adecuada: Los conductores de automóviles BMW tienden a ser más adinerados que otros adultos, y una mayor riqueza está relacionada con un mejor estado de la salud.

- 5. Estatura y ejercicio.** Con base en un estudio de las estaturas de hombres y mujeres que juegan básquetbol, un investigador concluye que la ejercitación que se logra al jugar básquetbol causa que la gente crezca más.
- 6. Las personas que se gradúan de la universidad viven más tiempo.** Con base en un estudio que revela que las personas que se gradúan de la universidad viven más tiempo que quienes no lo hacen, un investigador concluye que el estudio provoca que la gente viva más tiempo.
- 7. ¿Perfil racial?** Un estudio reveló que en el condado de Orange se expiden más multas por exceso de velocidad a los individuos de grupos minoritarios que a los caucásicos. Conclusión: En el condado de Orange los individuos de grupos minoritarios exceden la velocidad límite más que los blancos.
- 8. Remedio contra el resfriado.** En un estudio sobre los síntomas del resfriado, se encontró que todos los sujetos que estaban resfriados mejoraron dos semanas después de tomar píldoras de jengibre. Conclusión: Las píldoras de jengibre curan el resfriado.

En los ejercicios 9 a 20, utilice el pensamiento crítico para indicar lo que se le pide.

- 9. El chocolate es un alimento saludable.** El *New York Times* publicó un artículo que incluía la siguiente afirmación: “Por fin, el chocolate ocupa el lugar que merece en la pirámide alimenticia, junto a sus vecinos de clase alta: el vino tinto, las frutas, los vegetales y el té verde. Varios estudios, reportados en el *Journal of Nutrition* revelaron que, después de comer chocolates, los sujetos a prueba incrementaron los niveles de antioxidantes en su sangre. El chocolate contiene flavonoides, antioxidantes asociados con la disminución del riesgo de enfermedades cardíacas y embolias. Mars, Inc., la empresa de dulces, y la Chocolate Manufacturers Association financiaron gran parte de la investigación”. ¿Qué es incorrecto en este estudio?

- 10. Datos del censo.** Después de la realización del último censo nacional, Poughkeepsie Journal imprimió el siguiente titular de primera página: “281,421,906 en Estados Unidos”. ¿Qué es incorrecto en ese titular?
- 11. Encuesta por correo.** Cuando la autora Shere Hite escribió *Woman and Love: A Cultural Revolution in Progress*, basó sus conclusiones en las 4500 respuestas que recibió después de enviar por correo 100,000 cuestionarios a diversos grupos de mujeres. ¿Es probable que sus conclusiones sean válidas, en el sentido de que puedan aplicarse a la población general de todas las mujeres? ¿Por qué?
- 12. Números “900”.** En una encuesta de “Nightline” de la ABC, 186,000 televidentes pagaron 50 centavos cada uno para llamar a un número telefónico “900” y dar su opinión acerca de mantener la sede de las Naciones Unidas en Estados Unidos. Los resultados mostraron que el 67% de quienes llamaron estaban a favor de que las Naciones Unidas salieran de Estados Unidos. Interprete los resultados identificando lo que concluiríamos acerca del sentir de la población general, con respecto a mantener la sede de las Naciones Unidas en Estados Unidos.
- 13. Realización de encuestas.** Usted planea realizar una encuesta para conocer el porcentaje de personas que vive en su estado que puedan identificar al asistente del gobernador, el cual planea postularse para el Senado de Estados Unidos. Usted obtiene direcciones y teléfonos de directorios y envía por correo una encuesta a 850 personas elegidas al azar. ¿Por qué no es correcto utilizar el directorio telefónico como fuente de los sujetos para la encuesta?
- 14. Pictogramas.** Durante los últimos 10 años, los impuestos en Newport se han duplicado, y un candidato a alcalde desea elaborar una gráfica que destaque ese aspecto. El candidato representa los impuestos de hace 10 años utilizando una caja con anchura, longitud y altura de 1 pulgada. Luego, duplica cada dimensión para mostrar una caja más grande que represente los impuestos actuales. ¿Cuál es el volumen de la caja pequeña? ¿Cuál es el volumen de la caja grande? ¿Este pictograma describe correctamente la relación entre los impuestos de hace 10 años y los impuestos actuales?
- 15. Cascos para motociclista.** El Senado del estado de Hawái entró en audiencia para considerar una ley que obligaba a los motociclistas a usar cascos. Algunos motociclistas testificaron que habían participado en choques donde los cascos resultaron inútiles. ¿Qué grupo importante no fue capaz de testificar? (Véase “A Selection of Selection Anomalies” de Wainer, Palmer y Bradlow en *Chance*, vol. 11, núm. 2).
- 16. Encuesta a un cliente de Merrill Lynch.** El autor recibió una encuesta de la empresa de inversiones Merrill Lynch. La encuesta fue diseñada para medir su satisfacción como cliente, y contenía preguntas específicas para calificar al consultor financiero personal del autor. La portada de la carta incluía la siguiente afirmación: “Sus respuestas son extremadamente valiosas para su consultor financiero, Russell R. Smith y para Merrill Lynch... Compartiremos su nombre y las respuestas con su consultor financiero” ¿Qué es incorrecto en esta encuesta?
- 17. Promedio de promedios.** Un economista selecciona al azar a 10 individuos asalariados de cada una de las 50 entidades de Estados Unidos. Para cada estado calcula el promedio de los ingresos anuales y luego suma esos 50 valores y los divide entre 50. ¿Es probable que el resultado sea un buen estimado del promedio (media) de todos los individuos asalariados de Estados Unidos? ¿Por qué?
- 18. Pregunta incorrecta.** Una encuesta incluye este reactivo: “Añote su estatura en pulgadas”. A partir de ese dato se espera obtener las estaturas reales de los encuestados y analizarlas; no obstante, hay dos problemas importantes en este reactivo. Identifíquelo.
- 19. Número de miembros de la familia.** Usted necesita hacer un estudio para determinar el tamaño promedio de una familia en el estado donde vive. Para esto reúne datos que consisten en el número de hermanos y hermanas de los estudiantes de su universidad. ¿Qué grupo de familias se pierde con este enfoque? ¿Los resultados serán representativos de todas las familias del estado?

20. SMSI. En una carta al editor del *New York Times*, la ciudadana de Moorestown, New Jersey, Jean Mercer criticó la declaración de que “colocar a los bebés en posición supina ha disminuido las muertes por SMSI”. SMSI son las siglas del síndrome de muerte súbita infantil, y la posición *supina* implica estar acostado sobre la espalda con la cara hacia arriba. Ella sugirió que la siguiente afirmación es mejor: “Los pediatras aconsejaron la posición supina durante un periodo en el que disminuyeron las tasas de SMSI”. ¿Qué es incorrecto al decir que la posición supina *disminuyó* las muertes por SMSI?

Porcentajes. En los ejercicios 21 a 26, responda las preguntas que se hacen con respecto a los porcentajes.

21. Porcentajes

- a. Convierta la fracción $\frac{3}{20}$ a un porcentaje equivalente.
- b. Convierta 56.7% a su equivalente decimal.
- c. ¿Cuál es el 34% de 500?
- d. Convierta 0.789 a un porcentaje equivalente.

22. Porcentajes

- a. ¿Cuál es el 15% de 620?
- b. Convierta 5% a su equivalente decimal.
- c. Convierta 0.01 a un porcentaje equivalente.
- d. Convierta la fracción $\frac{987}{1068}$ a un porcentaje equivalente. Exprese la respuesta a la décima más cercana del porcentaje.

23. Porcentajes en una encuesta Gallup

- a. En una encuesta Gallup, el 52% de 1038 adultos encuestados manifestó que el tabaquismo pasivo es “muy dañino”. ¿Cuál es el número real de adultos que dijeron que el tabaquismo pasivo es “muy dañino”?
- b. De los 1038 adultos encuestados, 52 dijeron que el tabaquismo pasivo “no es dañino en absoluto”. ¿Cuál es el porcentaje de personas que escogió “no es dañino en absoluto”?

24. Porcentajes en un estudio de Lipitor

- a. En un estudio del fármaco Lipitor contra el colesterol, a 270 pacientes se les dio un placebo, y 19 de esos 270 pacientes reportaron dolor de cabeza. ¿Qué porcentaje de este grupo placebo reportó dolor de cabeza?
- b. De los 270 pacientes del grupo placebo, el 3.0% reportó dolor de espalda. ¿Cuál es el número real de pacientes que reportó dolor de espalda?

25. Porcentajes delictivos en el campus. En un estudio sobre los delitos cometidos por estudiantes bajo la influencia de alcohol o drogas en los planteles universitarios, se aplicó una encuesta por correo a 1875 estudiantes. Un artículo de *USA Today* señaló que “el 8% de los estudiantes que respondieron de forma anónima afirmó haber cometido un delito en el campus. En tanto que el 62% de ese grupo dijo que lo hizo bajo la influencia de alcohol o drogas”. Considerando que el número de estudiantes que respondió de manera anónima es 1875, ¿cuántos cometieron realmente un delito en el campus mientras estaban bajo la influencia de alcohol o drogas?

26. Porcentajes en los medios de comunicación y en la publicidad

- a. Un editorial del *New York Times* criticó un gráfico que describía un enjuague bucal que “reduce la placa dental en más del 300% ”. ¿Qué es incorrecto en esta afirmación?
- b. En el *New York Times Magazine*, un informe acerca de la disminución de la inversión occidental en Kenia afirmó que “después de años de vuelos diarios, Lufthansa y Air France han interrumpido el servicio de pasajeros. La inversión extranjera cayó el 500% durante la década de 1990”. ¿Qué es incorrecto en esta afirmación?
- c. En un anuncio de Club, un dispositivo que se utiliza para disminuir los robos de automóviles, se afirmó que “Club reduce las probabilidades de robo de su automóvil en un 400 por ciento”. ¿Qué es incorrecto en esta afirmación?

1-3 MÁS ALLÁ DE LO BÁSICO

27. **Falsificación de datos.** En una ocasión, un investigador del Sloan-Kettering Cancer Research Center fue criticado por falsificar datos. Entre sus datos había cifras obtenidas de seis grupos de ratones, con 20 ratones en cada grupo. Los siguientes valores se dieron para el porcentaje de éxito en cada grupo: 53%, 58%, 63%, 46%, 48%, 67%. ¿Cuál es la principal falla?
28. **¿Qué está mal en este asunto?** Intente identificar cada una de las cuatro fallas principales en lo siguiente. Un periódico realizó una encuesta pidiendo a los lectores que llamaran y respondieran esta pregunta: “¿Apoya usted el desarrollo de armas atómicas que podrían matar a millones de personas inocentes?” Se reportó que 20 electores respondieron, y que el 87% dijo que “no”, mientras que 13% respondió que “sí”.



1-4 Diseño de experimentos

Concepto clave Esta sección contiene mucha información y muchas definiciones. De todas las definiciones, el concepto de una *muestra aleatoria simple* es especialmente importante, por el papel que tiene en el resto de este libro y en la estadística en general, de modo que asegúrese de entender muy bien su definición. Además, comprenda con claridad que el método que se utilice para reunir datos es crítico y absolutamente importante. Reconozca que, para diseñar experimentos, se requiere de una gran reflexión y cuidado para asegurarse de obtener resultados válidos.

Si los datos muestrales no se reúnen de manera adecuada, éstos podrían resultar inútiles por completo, de tal forma que ninguna cantidad de tortura estadística los salvaría.

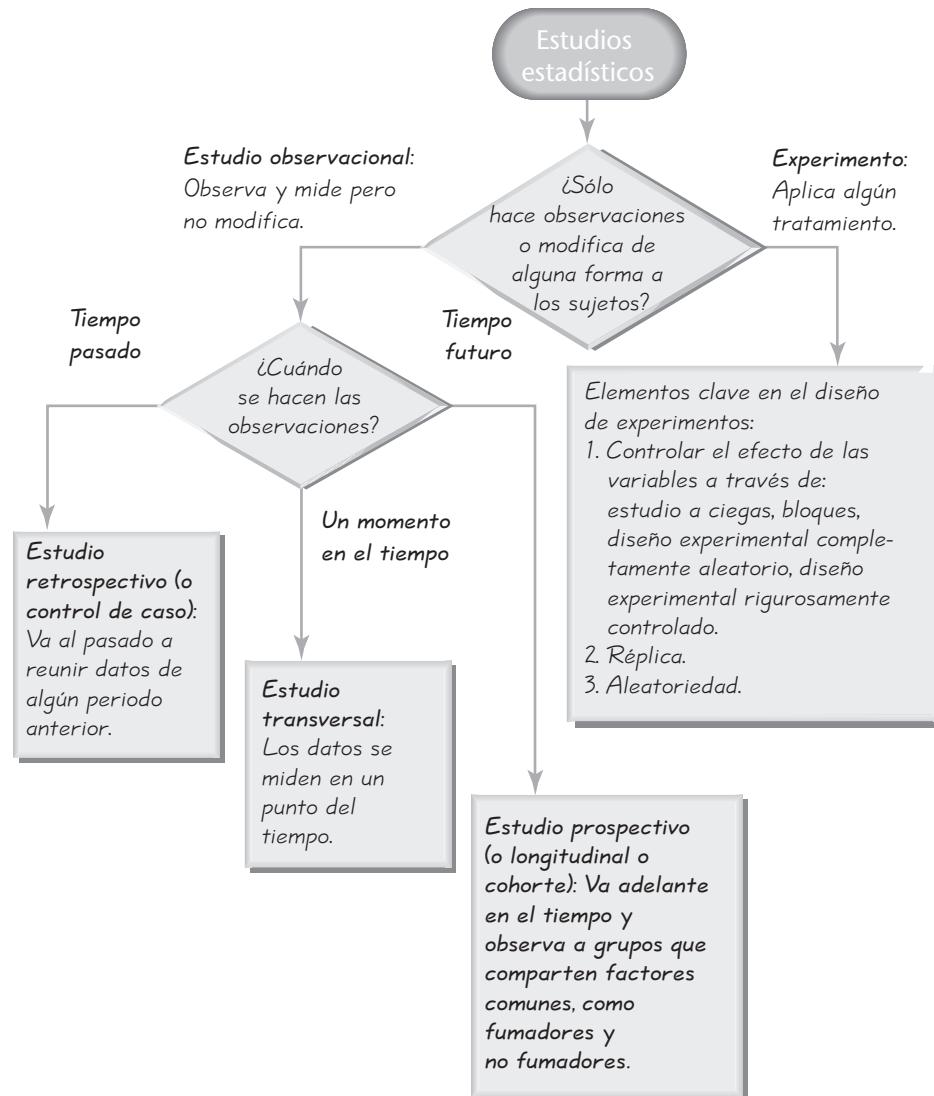
Los métodos estadísticos se rigen por los datos. Por lo regular obtenemos datos de dos fuentes distintas: los *estudios observacionales* y los *experimentos*.

Definiciones

En un **estudio observacional**, vemos y medimos características específicas, pero no intentamos *modificar* a los sujetos que estamos estudiando.

En un **experimento** aplicamos algunos *tratamientos* y luego procedemos a observar sus efectos sobre los sujetos (en los experimentos, a los sujetos se les denomina **unidades experimentales**).

Tipos de estudios observacionales Una encuesta Gallup es un buen ejemplo de un estudio observacional, en tanto que el ensayo clínico del fármaco Lipitor es un buen ejemplo de un experimento. La encuesta Gallup es observacional en el sentido de que simplemente se observan personas (a menudo haciendo entrevistas) sin modificarlas de ninguna forma. Sin embargo, el ensayo clínico de Lipitor implica el tratamiento de algunas personas con el fármaco, de manera que se modifica a los individuos tratados. Existen distintos tipos de estudios observacionales, como se ilustra en la figura 1-3. Estos términos, que suelen utilizarse en muchas revistas científicas diferentes, se definen a continuación.

Figura 1-3**Estudios estadísticos****Definiciones**

En un **estudio transversal**, los datos se observan, miden y reúnen en un solo momento.

En un **estudio retrospectivo (o de control de caso)**, los datos se toman del pasado (mediante el examen de registros, entrevistas y otros).

En un **estudio prospectivo (o longitudinal o de cohorte)**, los datos se reunirán en el futuro y se toman de grupos (llamados *cohortes*) que comparten factores comunes.

Hay una diferencia importante entre el muestreo realizado en estudios retrospectivos y estudios prospectivos. En los estudios retrospectivos regresamos en el tiempo a reunir datos acerca de características resultantes de interés, como un grupo de conductores que murieron en accidentes automovilísticos y otro grupo de conductores que no murieron en este tipo de accidentes. En los estudios prospectivos recurrimos al futuro siguiendo grupos con un factor potencialmente causal y grupos que no lo tienen, como un grupo de conductores que utilizan teléfonos celulares y un grupo de conductores que no los utilizan.

Las tres definiciones anteriores se aplican a los estudios observacionales. No obstante, por ahora nos enfocaremos en los experimentos. Los resultados de los experimentos en ocasiones se estropean debido a la *confusión*.

Definición

La **confusión** ocurre en un experimento cuando uno no es capaz de distinguir entre los efectos de diferentes factores.

Trate de planear el experimento para que no se presente confusión.

Por ejemplo, suponga que un profesor de Vermont experimenta con una nueva política de asistencia (“su calificación promedio en el curso bajará un punto por cada clase que falte”); sin embargo, llega un invierno excepcionalmente benigno, sin nieve y sin temperaturas muy frías, que son factores que habían limitado la asistencia en años anteriores. Si la asistencia mejora, no podemos determinar si ello se debe a la nueva política de asistencia o al invierno benigno. Los efectos de la política de asistencia y del clima se han confundido. Generalmente es muy importante controlar los efectos de las variables. Además de la confusión, los experimentos también se pueden arruinar por otros factores, como el hecho de no lograr reunir una muestra que sea representativa de la población. En general, la organización de los experimentos requiere de un gran cuidado y una extensa planeación. En la figura 1-3 se observa que los siguientes tres aspectos son muy importantes para el diseño de experimentos:

1. Control de los efectos de las variables.
2. Uso de la réplica.
3. Empleo de la aleatoriedad.

Ahora nos enfocaremos en el papel que juegan estos tres factores en el diseño de experimentos.

Control de los efectos de las variables

En la figura 1-3 se muestra que uno de los elementos clave en el diseño de experimentos es el control de los efectos de las variables. Se adquiere control al utilizar dispositivos como un estudio a ciegas, un diseño de bloques aleatorio, un diseño experimental completamente aleatorio o un diseño experimental rigurosamente controlado.

Estudio a ciegas En 1954 se diseñó un experimento masivo para probar la efectividad de la vacuna de Salk para prevenir la poliomielitis que había matado o paralizado a miles de niños. En ese experimento, a un grupo de tratamiento se le administró la vacuna real de Salk; mientras que a un segundo grupo se le dio un placebo que no contenía ningún fármaco. En experimentos que implican el uso de



LA ESTADÍSTICA EN LAS NOTICIAS

Pruebas clínicas versus estudios observacionales

En un artículo del *New York Times* acerca de la terapia hormonal para las mujeres, la reportera Denise Grady escribió acerca de un reporte de tratamientos probados en ensayos clínicos controlados aleatorios. Ella declaró que “pruebas como ésta, donde los pacientes se designan al azar para un tratamiento o un placebo, se consideran el estándar por excelencia en la investigación médica. En cambio, los estudios observacionales, donde los pacientes deciden por sí mismos si toman un fármaco, se consideran menos confiables [...] Los investigadores dicen que los estudios observacionales tal vez han dado una falsa imagen color de rosa del reemplazo hormonal, ya que las mujeres que optan por recibir los tratamientos son más saludables y tienen mejores hábitos, que las mujeres que no lo hacen”.



Los efectos Hawthorne y del experimentador

El conocido efecto placebo ocurre cuando un sujeto no tratado cree incorrectamente que está recibiendo un tratamiento real, y reporta una mejoría en sus síntomas. El efecto Hawthorne ocurre cuando, por alguna razón, los sujetos tratados responden de manera diferente por el simple hecho de formar parte del experimento. (Este fenómeno se denominó “efecto Hawthorne” porque se observó por primera vez en un estudio realizado con obreros en la planta Hawthorne, de Western Electric). Ocurre un efecto del experimentador (a veces llamado efecto Rosenthal) cuando el investigador o experimentador influye involuntariamente en los sujetos mediante factores como la expresión facial, el tono de voz o la actitud.

placebos, a menudo existe un **efecto placebo** que ocurre cuando un sujeto no tratado reporta una mejoría en los síntomas. (La mejoría reportada en el grupo placebo puede ser real o imaginaria). El efecto placebo se puede disminuir o evitar usando un **estudio a ciegas**, una técnica en que el sujeto no sabe si está recibiendo un tratamiento o un placebo. El estudio a ciegas nos permite determinar si el efecto del tratamiento es significativamente diferente del efecto placebo. El experimento de la poliomielitis fue un **estudio a ciegas doble**, lo que quiere decir que el estudio se realizó en dos niveles: 1. los niños inyectados no sabían si estaban recibiendo la vacuna de Salk o un placebo, y 2. los médicos que aplicaron las inyecciones y evaluaron los resultados tampoco lo sabían.

Bloques Cuando se diseña un experimento, con frecuencia sabemos de antemano que existen algunos factores que pueden tener un efecto importante sobre las variables que se están considerando. Por ejemplo, al diseñar un experimento para probar la eficacia de un nuevo fertilizante sobre el crecimiento de los árboles, sabemos que el contenido de humedad del suelo (seco o húmedo) puede afectar el crecimiento de los árboles. En este caso, debemos utilizar bloques. Un **bloque** es un conjunto de sujetos que son similares, pero los bloques son diferentes en aspectos que pueden afectar el resultado del experimento. La prueba de los efectos de un fertilizante podría incluir un bloque de árboles en tierra seca y un bloque de árboles en tierra húmeda. Al diseñar un experimento para probar la eficacia de un nuevo fármaco para enfermedades cardíacas, podríamos crear un bloque de hombres y un bloque de mujeres, ya que se sabe que el corazón de los hombres y el de las mujeres se comporta de manera diferente. Después de identificar los bloques, se procede a asignar *aleatoriamente* los tratamientos a los sujetos de cada bloque.

Diseño de bloque aleatorio: Utilice este diseño experimental si está realizando un experimento para probar uno o más tratamientos diferentes, si existen distintos grupos de sujetos similares y si los grupos difieren en aspectos que podrían afectar las respuestas a los tratamientos:

1. Forme bloques (o grupos) de sujetos con características similares.
2. Asigne los tratamientos de manera aleatoria a los sujetos dentro de cada bloque.

Por ejemplo, suponga que desea probar la eficacia de un fertilizante sobre el peso de 12 álamos, y que tiene dos terrenos: uno de los terrenos es seco y el otro es húmedo. Un diseño experimental inadecuado utilizaría el fertilizante en 6 árboles plantados en suelo húmedo; mientras que los 6 árboles sin tratamiento se plantarían en terreno seco. De esta forma no sabríamos si las diferencias se deben al tratamiento con el fertilizante o al tipo de tierra o a ambos. Véase la figura 1-4a). Sería mucho mejor plantar 6 árboles en el terreno seco y 6 árboles en el terreno húmedo. Luego, se usaría la aleatoriedad para identificar 3 árboles que se tratarían con el fertilizante y 3 que no recibirían el tratamiento, como se observa en la figura 1-4b). Los resultados nos permitirían ver la eficacia del fertilizante. (La elección de una muestra con un tamaño de 12 árboles es totalmente arbitraria. Un mayor número de árboles produciría mejores resultados. En capítulos posteriores abordamos el tema de la determinación del tamaño de la muestra. Asimismo, el experimento debe diseñarse de forma cuidadosa, para que el fertilizante que se aplique a uno de los árboles no se esparza a los árboles adyacentes. Otros factores importantes como el riego, la temperatura y la luz del sol se deben controlar cuidadosamente. El diseño experimental

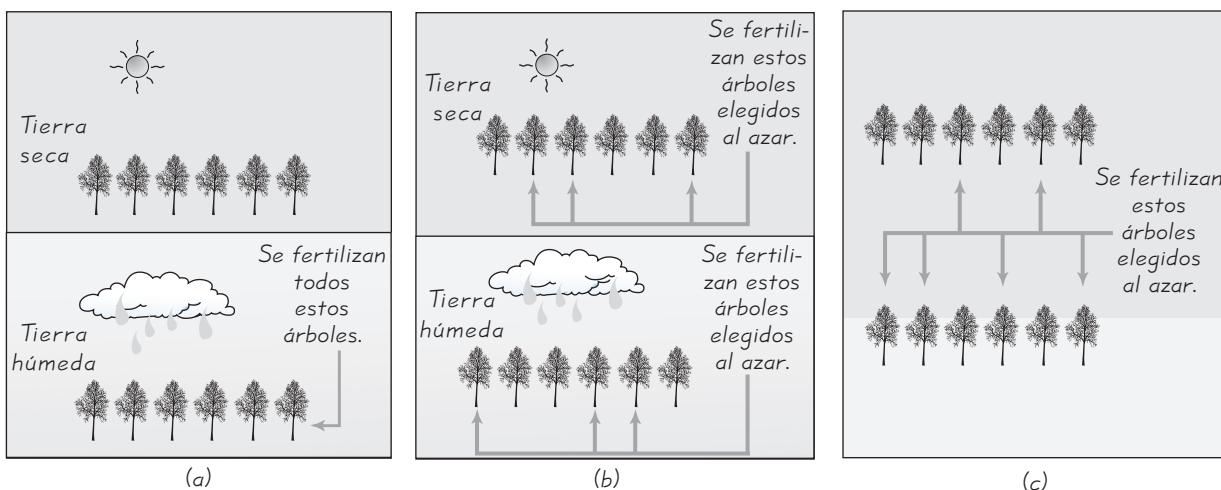


Figura 1-4 Diseño de un experimento con diferentes tratamientos

- (a) **Diseño experimental inadecuado:** Utilizar el fertilizante para tratar a todos los árboles de la tierra húmeda y no utilizar el fertilizante para los árboles de la tierra seca. No habrá forma de saber si los resultados se deben al tratamiento o al tipo de tierra.
- (b) **Diseño de bloques aleatorio:** Formar un bloque de árboles de tierra húmeda y otro bloque de árboles de tierra seca. Dentro de cada bloque se utiliza la aleatoriedad para determinar cuáles se tratarán con el fertilizante y cuáles no.
- (c) **Diseño experimental completamente aleatorio:** De manera aleatoria se determina cuáles árboles se tratarán con el fertilizante y cuáles no.

general requiere de mucha mayor reflexión y cuidado de lo que aquí se ha descrito. Tomar un curso sobre el diseño de experimentos sería una forma muy adecuada de aprender mucho más de lo que podemos describir aquí).

Diseño experimental completamente aleatorio En un **diseño experimental completamente aleatorio**, los sujetos se asignan a distintos grupos de tratamiento mediante un proceso de *selección aleatoria*. Observe la figura 1-4c), donde se muestra que, de los 12 árboles, se seleccionan 6 al azar para el tratamiento con el fertilizante. Otro ejemplo de un diseño experimental completamente aleatorio es el experimento de la poliomielitis: Los niños se asignaron al grupo de tratamiento o al grupo placebo, mediante un proceso de selección aleatoria (equivalente a lanzar una moneda).

Diseño rigurosamente controlado Otro método para asignar sujetos a los tratamientos consiste en utilizar el **diseño rigurosamente controlado**, donde los sujetos se *eligen cuidadosamente*, de manera que quienes reciban cada tratamiento sean similares en los aspectos que son importantes para el experimento. En un experimento para probar la efectividad de un fármaco diseñado para disminuir la presión sanguínea, si el grupo placebo incluye a un hombre de 30 años de edad con sobrepeso, fumador, que consume grandes cantidades de bebidas alcohólicas, sal y grasas, el grupo de tratamiento también debería incluir a una persona con características similares (lo cual, en este caso, sería fácil de conseguir).

Réplica y tamaño de la muestra

Además de controlar los efectos de las variables, otro elemento fundamental del diseño experimental es el tamaño de las muestras. Las muestras deben ser lo



Estudio nacional prospectivo de los niños

Un buen ejemplo de un estudio prospectivo es el National Children's Study que comenzó en 2005, en el cual se realiza un seguimiento de 100,000 niños, desde el nacimiento hasta los 21 años de edad. Los niños son de 96 regiones geográficas diferentes. Su objetivo consiste en mejorar la salud de los niños al identificar los efectos de factores ambientales como la dieta, la exposición a sustancias químicas, la vacunación, las películas y la televisión. Se planea que alrededor de 2008 haya algunos resultados preliminares. El estudio abordara preguntas como éstas: ¿De qué manera interactúan los genes y el ambiente para fomentar o prevenir la conducta violenta en los adolescentes? ¿La falta de ejercicio y una dieta inadecuada son las únicas razones del sobrepeso de muchos niños? ¿Las infecciones afectan el progreso del desarrollo, el asma, la obesidad y las enfermedades cardíacas? ¿De qué manera la planeación y construcción de las ciudades y de los vecindarios fomentan o disminuyen las lesiones?

suficientemente grandes para que el comportamiento errático, que es característico de muestras muy pequeñas, no disfraze los efectos verdaderos de los distintos tratamientos. La repetición de un experimento sobre un grupo de sujetos suficientemente grande se conoce como **réplica**, la cual se utiliza con eficacia cuando tenemos los sujetos suficientes como para reconocer las diferencias que resultan de los diferentes tratamientos. (En otro contexto, la *réplica* se refiere a la repetición o duplicación de un experimento para confirmar o verificar los resultados). Con la réplica, se incrementa la posibilidad de reconocer diferentes efectos del tratamiento en las muestras de tamaño grande. Sin embargo, una muestra grande no necesariamente es buena. Aunque es importante tener una muestra que sea lo suficientemente grande, es más importante tener una muestra en la que los datos se hayan obtenido de una forma apropiada, como la selección aleatoria (que se describe más adelante).

Utilice una muestra de un tamaño que sea lo bastante grande para distinguir la verdadera naturaleza de cualquiera de los efectos, y obtenga la muestra utilizando un método apropiado, como uno basado en la aleatoriedad.

En el experimento diseñado para probar la vacuna de Salk, a 200,000 niños se les administró la vacuna real de Salk y a otros 200,000 niños se les dio un placebo. Fue posible observar la eficacia de la vacuna porque en realidad en el experimento real se utilizaron muestras de un tamaño lo suficientemente grande. No obstante, aun cuando los grupos placebo y de tratamiento fueron muy grandes, el experimento habría sido un fracaso si los sujetos no se hubieran asignado a los dos grupos para que fueran similares en los aspectos importantes para el experimento.



Aleatoriedad y otras estrategias de muestreo

En la estadística, como en la vida, uno de los peores errores consiste en reunir datos de una forma inapropiada. Insistiremos en este punto tan importante:

Si los datos muestrales no se reúnen de forma adecuada, resultarán tan inútiles que ninguna cantidad de tortura estadística podrá salválos.

En la sección 1-3 vimos que una muestra de respuestas voluntarias es aquella donde los sujetos deciden por sí mismo si responden o no. Este tipo de muestras son muy comunes, aunque sus resultados suelen ser inútiles para hacer inferencias válidas acerca de poblaciones más grandes.

Ahora definiremos algunos de los procedimientos de muestreo más comunes.

Definiciones

En una **muestra aleatoria** los miembros de la población se seleccionan de forma que cada *miembro individual* tenga la misma posibilidad de ser elegido.

Una **muestra aleatoria simple** de n sujetos se selecciona de manera que cada posible *muestra del mismo tamaño n* tenga la misma posibilidad de ser elegida.

Una **muestra probabilística** implica seleccionar miembros de una población de forma que cada miembro tenga una posibilidad conocida (aunque no necesariamente la misma) de ser elegido.

EJEMPLO Muestra aleatoria, muestra aleatoria simple, muestra probabilística Imagine un salón de clases con 60 estudiantes acomodados en seis filas de 10 estudiantes cada una. Suponga que el profesor selecciona una muestra de 10 estudiantes lanzando un dado y seleccionando la fila correspondiente al resultado. ¿El resultado es una muestra aleatoria? ¿Una muestra aleatoria simple? ¿Una muestra probabilística?

SOLUCIÓN La muestra es aleatoria porque cada estudiante individual tiene la misma posibilidad (una posibilidad en seis) de resultar seleccionado. La muestra *no* es aleatoria simple porque no todas las muestras de tamaño 10 tienen la misma posibilidad de ser escogidas. Por ejemplo, este diseño muestral de usar un dado para seleccionar una fila hace imposible la selección de 10 estudiantes que estén en filas diferentes (aunque hay una posibilidad en seis de seleccionar la muestra que consiste en los 10 estudiantes de la primera fila). Se trata de una muestra probabilística porque cada estudiante tiene una posibilidad conocida (una posibilidad en seis) de ser elegido.

Importante: A lo largo de este libro, utilizaremos una variedad de procedimientos estadísticos diferentes, y muchas veces tendremos como requisito reunir una *muestra aleatoria simple*, como se definió anteriormente. Con el muestreo aleatorio esperamos que todos los componentes de la población estén representados (aproximadamente) de manera proporcional. Las muestras aleatorias se seleccionan usando muchos procedimientos diferentes, incluyendo el uso de computadoras para generar números aleatorios. (Antes de la llegada de las computadoras, generalmente se usaban las tablas de números aleatorios. Si quiere leer algo verdaderamente interesante, consulte el libro *A Million Random Digits*, publicado por Free Press, que contiene un millón de dígitos generados aleatoriamente). A diferencia del muestreo realizado con descuido o por casualidad, el muestreo aleatorio exige una planeación y ejecución muy cuidadosas.

Además del muestreo aleatorio, existen otras técnicas de muestreo en uso, y aquí se describen las más comunes. Observe la figura 1-5, que describe los diferentes métodos de muestreo. No olvide que en el resto del libro sólo se utilizarán el muestreo aleatorio y el muestreo aleatorio simple.

Definiciones

En el **muestreo sistemático**, elegimos algún punto de partida y luego seleccionamos cada k -ésimo (por ejemplo, cada quincuagésimo) elemento en la población.

En el **muestreo de conveniencia**, simplemente se utilizan resultados que sean muy fáciles de obtener.

En el **muestreo estratificado** subdividimos a la población en al menos dos subgrupos (o estratos) diferentes, de manera que los sujetos que pertenecen al mismo subgrupo comparten las mismas características (como el género o la categoría de edad), y luego obtenemos una muestra de cada subgrupo (o estrato).

En el **muestreo por conglomerados** primero dividimos el área de la población en secciones (o conglomerados), y luego elegimos al azar algunos de estos conglomerados, y después elegimos *a todos* los miembros de los conglomerados seleccionados.

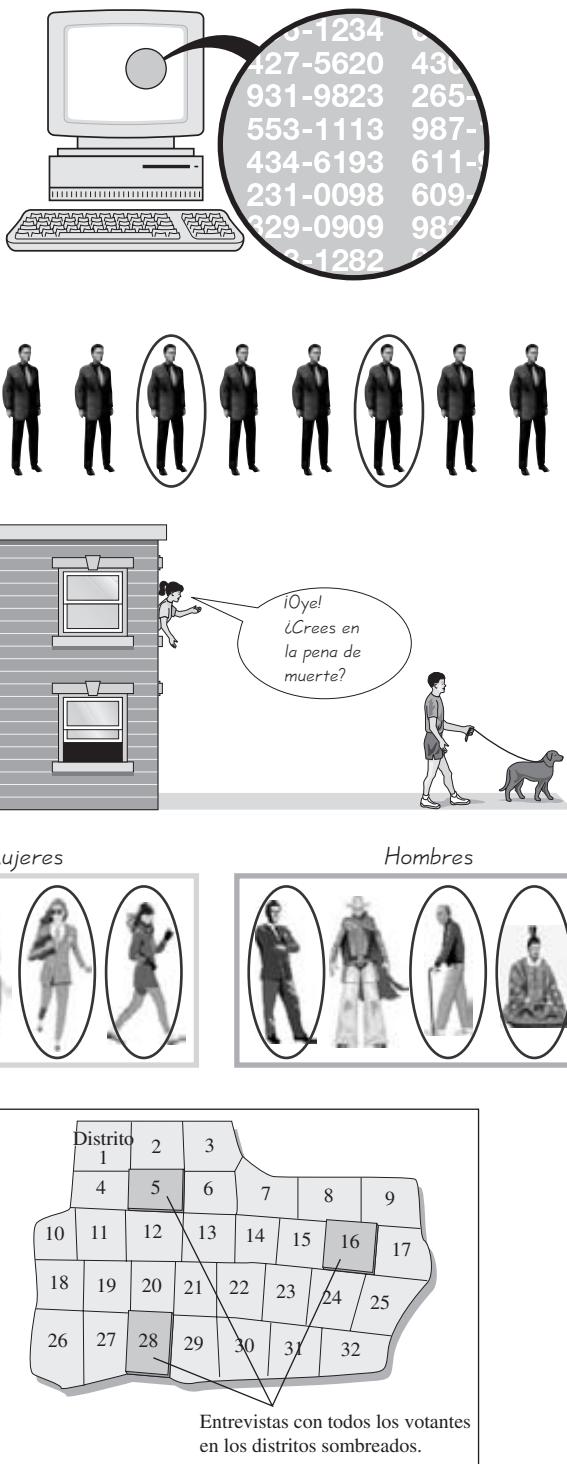


Figura 1-4 Procedimientos de muestreo comunes

Muestreo aleatorio:

Cada miembro de la población tiene la misma probabilidad de resultar seleccionado. A menudo se usan computadoras para generar números telefónicos aleatorios.

Muestreo aleatorio simple:

Se selecciona una muestra de n sujetos, de manera que cada posible muestra del mismo tamaño n tenga la misma posibilidad de ser elegida.

Muestreo sistemático:

Se selecciona un punto de partida, después se elige cada k -ésimo (por ejemplo, cada quincuagésimo) elemento de la población.

Muestreo de conveniencia:

Se utilizan resultados que son fáciles de obtener.

Muestreo estratificado:

Se subdivide a la población en al menos dos subgrupos (o estratos diferentes), de manera que los sujetos del mismo subgrupo comparten las mismas características (como el género o la categoría de edad), y después se obtiene una muestra de cada subgrupo.

Muestreo por conglomerados:

Se divide el área de la población en secciones (o conglomerados), luego se eligen al azar algunos de estos conglomerados, y después se elige a todos los miembros de los conglomerados seleccionados.

Es fácil confundir el muestreo estratificado y el muestreo por conglomerados, ya que ambos implican la formación de subgrupos. Sin embargo, el muestreo por conglomerados usa a *todos* los miembros de una *muestra* de conglomerados; en tanto que el muestreo estratificado emplea una *muestra* de los miembros de *todos* los estratos. Un ejemplo de muestreo por conglomerados es una encuesta previa a las elecciones, donde se seleccionan aleatoriamente 30 distritos electorales de un número mayor de distritos, y luego se encuesta a todas las personas de cada uno de esos distritos elegidos. Este método es mucho más rápido y mucho menos costoso que elegir a una persona de cada uno de los muchos distritos del área de la población. Los resultados del muestreo estratificado o por conglomerados se ajustan o ponderan para corregir cualquier representación desproporcionada de los grupos.

Para una muestra de tamaño fijo, si usted selecciona al azar sujetos de diferentes estratos, es probable que obtenga resultados más consistentes (y menos variables), que si simplemente seleccionó una muestra al azar de la población general. Por tal razón, el muestreo estratificado se utiliza con frecuencia para reducir la variación en los resultados. Muchos de los procedimientos que se analizarán después en este libro tienen como requisito que los datos muestrales formen una *muestra aleatoria simple*, y ni el muestreo estratificado ni el muestreo por conglomerados satisfacen tal requisito.

Muestreo de etapas múltiples La figura 1-5 ilustra cinco procedimientos de muestreo comunes; sin embargo, los encuestadores profesionales y los investigadores gubernamentales a menudo recolectan datos utilizando cierta combinación de los cinco métodos. Un **diseño de muestreo de etapas múltiples** implica la selección de una muestra en diferentes pasos, los cuales suelen incluir distintos procedimientos de muestreo. Consideremos las estadísticas gubernamentales sobre el desempleo que se obtienen de hogares encuestados. No es práctico hacer visitas personales a cada miembro de una muestra aleatoria simple, ya que los hogares individuales están distribuidos por todo el país. Sería demasiado costoso en tiempo y en dinero encuestar a cada miembro de una muestra aleatoria simple. En cambio, el U.S. Census Bureau y el Bureau of Labor Statistics se combinan para realizar una encuesta llamada Current Population Survey, que se utiliza para obtener datos que describen factores tales como la tasa de desempleo, la matrícula de universidades y los salarios semanales. La encuesta incorpora un diseño de etapas múltiples, que se describe de manera general a continuación:

1. Todo el país (Estados Unidos) se divide en 2007 regiones diferentes llamadas *unidades muestrales primarias* (UMP). Las unidades muestrales primarias son áreas metropolitanas, grandes condados o grupos de condados más pequeños.
2. En cada uno de los 50 estados, se selecciona una muestra de unidades muestrales primarias. Para la Current Population Survey se utilizan 792 de las unidades muestrales primarias. (Se usan las 432 unidades muestrales primarias de las poblaciones más grandes, y se seleccionan al azar 360 unidades muestrales primarias de las 1575 restantes).
3. Cada una de las 792 unidades muestrales primarias elegidas se divide en bloques, y se utiliza un muestreo estratificado para seleccionar una muestra de bloques.
4. En cada bloque seleccionado, se identifican conglomerados de hogares cercanos entre sí. Los conglomerados se eligen al azar, y se entrevista a todos los hogares de los conglomerados elegidos.

Observe que este diseño de muestreo de etapas múltiples incluye los muestreos aleatorio, estratificado y por conglomerados en diferentes etapas. El resultado final



¿Debe usted creer en un estudio estadístico?

En el libro *Statistical Reasoning for Everyday Life*, segunda edición, los autores Jeff Bennett, William Briggs y Mario Triola proporcionan los siguientes ocho lineamientos para evaluar de forma crítica un estudio estadístico. **1.** Identificar el objetivo del estudio, la población considerada y el tipo de estudio. **2.** Considerar la fuente, especialmente con respecto a la posibilidad de un sesgo. **3.** Analizar el procedimiento de muestreo. **4.** Buscar problemas en la definición o medida de las variables de interés. **5.** Tener cuidado con variables confusas que podrían invalidar las conclusiones, **6.** Considerar el contexto y la redacción de cualquier encuesta. **7.** Verificar que las gráficas representen los datos de forma adecuada y que las conclusiones estén justificadas. **8.** Considerar si las conclusiones logran los objetivos del estudio, si tienen sentido y si poseen un significado práctico.

es un diseño de muestreo complejo; no obstante, es mucho más práctico y menos costoso que utilizar un diseño más sencillo, como el muestreo aleatorio simple.

Errores de muestreo

No importa lo bien que usted planee y ejecute el proceso de recolección de muestras, es probable que ocurra algún error en los resultados. Por ejemplo, seleccione 1000 adultos al azar, pregúntele si se graduaron de bachillerato y registre el porcentaje de respuestas afirmativas en la muestra. Si usted elige otra muestra de 1000 adultos al azar, es probable que obtenga un porcentaje *diferente* en esa muestra.

Definiciones

Un **error de muestreo** es la diferencia entre el resultado de una muestra y el verdadero resultado de la población; este error es consecuencia de las fluctuaciones por el azar.

Un **error que no es de muestreo** sucede cuando los datos muestrales se obtienen, registran o analizan de forma incorrecta (como cuando se selecciona una muestra sesgada, cuando se usa un instrumento de medición defectuoso o cuando se copian los datos de forma incorrecta).

Si recolectamos con cuidado una muestra, de manera que sea representativa de la población, entonces podemos utilizar los procedimientos que se describen en este libro para analizar el error muestral; sin embargo, debemos ser muy cuidadosos de no minimizar el error que no es de muestreo.

En el “Problema del capítulo”, nos preguntamos si los datos experimentales originales que derivaron el concepto de los “seis grados de separación” eran datos *buenos*. Señalamos que Stanley Milgram envió 60 cartas a individuos de Wichita, Kansas, y que les pidió que reenviaran a las cartas a una mujer específica en Cambridge, Massachusetts. Participaron 50 de los 60 sujetos, pero sólo tres cartas llegaron al objetivo. Dos experimentos posteriores también mostraron bajas tasas de terminación, aunque finalmente Milgram logró una tasa de terminación del 35%, y calculó que el número promedio de intermediarios era de alrededor de 6. Los experimentos de Milgram tuvieron grandes fallas en la recolección de datos muestrales, ya que éstos no eran representativos de la población de residentes de Estados Unidos. El experimento original incluyó a pobladores de Wichita y sus tasas de respuesta y terminación fueron muy bajas. Incluso con una tasa de terminación del 35%, bien sería posible que sólo hayan participado los individuos interesados en el experimento, los cuales tal vez conocían a más personas, dando como resultado un número menor de intermediarios. Sus anuncios fueron diseñados para reclutar personas que estuvieran “orgullosas de sus habilidades sociales y que tuvieran confianza en su capacidad de encontrar a alguien atravesando las barreras de clase”. Los sujetos reclutados eran de una región geográfica específica, eran más sociables y más adinerados que la gente común, y la tasa de respuesta fue muy baja. En consecuencia, el resultado de 6 como el número de grados de separación no estaba justificado por los datos (esto no necesariamente significa que el valor de 6 sea incorrecto). El diseño experimental tenía graves errores y cualquier conclusión basada en los resultados sería muy sospechosa. Por desgracia, no se ha realizado un experimento bien diseñado para verificar si el número 6 es correcto o incorrecto. (En 2001, Duncan Watts, de Columbia University, inició un experimento por correo electrónico a gran escala; sin embargo, las personas que utilizan el correo electrónico no necesariamente son representativas de la población mundial. No obstante, se aprende mucho de sus resultados).

Después de leer esta sección, es normal sentirse un poco abrumado por la variedad de las diferentes definiciones. Sin embargo, recuerde un punto clave: El procedimiento usado para reunir datos es sumamente importante, y debemos recordar que la *aleatoriedad* es especialmente relevante. Si los datos muestrales no se reúnen de manera adecuada, resultarán inútiles por completo, de forma que ninguna cantidad de tortura estadística podrá salvarlos. Además, el diseño de un experimento debe realizarse con gran reflexión y cuidado.

1-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. ¿Cuál es la diferencia entre una muestra aleatoria y una muestra aleatoria simple?
2. ¿Cuál es la diferencia entre un estudio observacional y un experimento?
3. Cuando se realiza un experimento para probar la eficacia de una nueva vacuna, ¿qué es un estudio a ciegas y porque es importante?
4. Al realizar un experimento para probar la eficacia de una nueva vacuna, un investigador decidió utilizar los bloques, con un bloque de hombres y un bloque de mujeres. ¿Cómo ayudaría el uso de los bloques al experimento?

En los ejercicios 5 a 8, determine si la descripción dada corresponde a un estudio observacional o a un experimento.

5. **Terapia de contacto.** Emily Rosa, de 9 años de edad, se convirtió en la autora de un artículo en el *Journal of the American Medical Association*, después de poner a prueba a terapeutas de contacto profesionales. Usando una mampara de cartón, ella colocabía la mano encima de la mano del terapeuta, quien debía de identificar la mano que Emily había elegido.
6. **Tratamiento contra la sífilis.** Ha surgido una gran controversia en torno del estudio de pacientes con sífilis que *no* recibieron un tratamiento que los habría curado. Su salud fue vigilada por años después de que se descubrió que padecían esa enfermedad.
7. **Control de calidad.** La Food and Drug Administration de Estados Unidos elige al azar una muestra de grageas de aspirina Bayer, y mide la exactitud de la cantidad de aspirina en cada gragea.
8. **Brazaletes magnéticos.** A los pasajeros de un barco de crucero se les dan brazaletes magnéticos, que aceptan usar en un intento por disminuir o eliminar los efectos del mareo.

En los ejercicios 9 a 12, identifique el tipo de estudio observacional (transversal, retrospectivo o prospectivo).

9. **Psicología del trauma.** Un investigador del hospital Monte Sinaí de la ciudad de Nueva York, planea obtener datos al hacer un seguimiento (hasta el año 2015) a los hermanos de las víctimas que perecieron en el ataque terrorista al Word Trade Center el 11 de septiembre de 2001.
10. **Investigación de los conductores en estado de ebriedad.** Un investigador de la Universidad Johns Hopkins obtiene datos sobre los efectos del alcohol al conducir, examinando informes de accidentes automovilísticos de los últimos cinco años.
11. **Audiencias televisivas.** Nielsen Media Research Company encuesta a 5000 hogares para determinar la proporción de éstos que sintonizan el programa *Saturday Night Live*.
12. **Estadísticas del éxito.** Un economista reúne datos de ingresos al seleccionar y entrevistar actualmente a un grupo de sujetos; después se remonta al pasado para ver si tuvieron la sabiduría de tomar un curso de estadística entre 1980 y 2005.

En los ejercicios 13 a 24, identifique el tipo de muestreo que se utilizó: aleatorio, sistemático, de conveniencia, estratificado o por conglomerados.

13. **Puesto de revisión de sobriedad.** El autor fue un observador en un puesto de revisión de sobriedad de la policía, donde se detenía y entrevistaba a cada quinto conductor. (El autor fue testigo del arresto de un ex alumno).
14. **Encuestas de salida.** En épocas de elecciones presidenciales, los medios noticiosos organizan una encuesta de salida, en la que se eligen estaciones de sondeo al azar y se encuesta a todos los votantes conforme abandonan el lugar.
15. **Educación y deportes.** Un investigador de la empresa de equipo deportivo Spaulding estudia la relación entre el nivel académico y la participación en cualquier deporte. El investigador hace una encuesta a 40 golfistas, 40 tenistas y 40 nadadores, todos elegidos al azar.
16. **Ergonomía.** Un estudiante de ingeniería mide la fuerza de los dedos necesaria para presionar botones al probar a miembros de familias.
17. **Hacer trampa.** Un investigador del Internal Revenue Service estudia las trampas en las declaraciones de impuestos, al encuestar a todos los meseros y las meseras de 20 restaurantes seleccionados al azar.
18. **Encuesta de MTV.** Un experto en marketing de MTV está planeando una encuesta en la que se elegirá a 500 personas al azar de cada uno de los siguientes grupos de edad: 10-19, 20-29 y así sucesivamente.
19. **Datos de tarjetas de crédito.** El autor encuestó a todos sus estudiantes para obtener datos muestrales que consistían en el número de tarjetas de crédito que posee cada uno.
20. **Recaudación de fondos.** Los recaudadores de fondos de la Universidad de Newport prueban una nueva campaña de telemarketing, obteniendo una lista de todos los alumnos y eligiendo cada centésimo nombre de dicha lista.
21. **Encuestas telefónicas.** En una encuesta de Gallup de 1059 adultos, los sujetos encuestados fueron seleccionados, usando de una computadora para generar aleatoriamente los números telefónicos a los que después se llamó.
22. **Investigación de mercados.** Un investigador de mercados separó a todos los residentes de California en las categorías de desempleado, empleado de tiempo completo y empleado de tiempo parcial. El investigador encuesta a 50 personas de cada categoría.
23. **Estudiantes que beben.** La Universidad de Newport, motivada por un estudiante que murió en estado de ebriedad, realizó una investigación de estudiantes que beben, seleccionando al azar 10 diferentes salones de clase y entrevistando a todos los estudiantes en cada uno de esos grupos.
24. **Ensayo clínico de un tratamiento sanguíneo.** En la fase II de la prueba de un nuevo fármaco diseñado para incrementar el conteo de glóbulos rojos, una investigadora encuentra sobres con los nombres y las direcciones de todos los sujetos tratados. Ella desea incrementar la dosis en una submuestra de 12 sujetos, por lo que revuelve exhaustivamente todos los sobres en una caja, y luego saca 12 de ellos para identificar a los sujetos que recibirán el incremento en la dosis.

Muestras aleatorias y muestras aleatorias simples. Los ejercicios 25 a 30 se refieren a muestras aleatorias y a muestras aleatorias simples.

25. **Muestra por conglomerados.** Un analista de la IRS procesa una devolución de impuestos cada 10 minutos, de manera que en su primera semana de trabajo procesa un total de 240 devoluciones. El gerente verifica su trabajo al seleccionar al azar un día de la semana y revisar todas las devoluciones que se procesaron ese día. ¿Tal plan de muestreo da como resultado una muestra aleatoria? ¿Una muestra aleatoria simple? Explique.

- 26. Muestra de conveniencia.** Un profesor de estadística obtiene una muestra de estudiantes, al seleccionar a los primeros 10 que entran a su salón de clases. ¿Este plan de muestreo da como resultado una muestra aleatoria? ¿Una muestra aleatoria simple? Explique.
- 27. Muestra sistemática.** Un ingeniero de control de calidad selecciona cada diezmilésimo dulce M&M que se produce. ¿Este plan de muestreo da como resultado una muestra aleatoria? ¿Una muestra aleatoria simple? Explique.
- 28. Muestra estratificada.** Un investigador del Departamento de Vehículos Motorizados del condado de Orange intenta probar un nuevo sistema en línea para el registro de conductores, utilizando una muestra de 20 hombres y 20 mujeres seleccionados al azar. (El condado de Orange tiene el mismo número de conductores hombres y de conductores mujeres). ¿Este plan de muestreo da como resultado una muestra aleatoria? ¿Una muestra aleatoria simple? Explique.
- 29. Muestreo de estudiantes.** Un salón de clases tiene 36 estudiantes sentados en seis filas diferentes, con seis estudiantes en cada fila. El profesor tira un dado para determinar una fila, y luego lo tira nuevamente para elegir a un estudiante específico de la fila. Este proceso se repite hasta completar una muestra de 6 estudiantes. ¿Este plan de muestreo da como resultado una muestra aleatoria? ¿Una muestra aleatoria simple? Explique.
- 30. Muestreo de píldoras de vitaminas.** Un inspector de la Food and Drug Administration de Estados Unidos obtiene píldoras de vitaminas producidas en una hora en la empresa Health Supply Company. Luego las mezcla exhaustivamente y extrae una muestra de 10 píldoras para probar la cantidad exacta del contenido vitamínico. ¿Este plan de muestreo da como resultado una muestra aleatoria? ¿Una muestra aleatoria simple? Explique.

1-4 MÁS ALLÁ DE LO BÁSICO

- 31. Diseño de muestreo.** La compañía de publicaciones Addison-Wesley le ha comisionado a usted para encuestar a 100 estudiantes usuarios de este libro. Describa procedimientos para obtener una muestra de cada tipo: aleatoria, sistemática, de conveniencia, estratificada, por conglomerados.
- 32. Confusión.** Mencione un ejemplo (diferente del que está en el texto) que ilustre la forma en que ocurre la confusión.
- 33. Diseño muestral.** En el artículo “Cardiovascular Effects of Intravenous Triiodothyronine in Patients Undergoing Coronary Artery Bypass Graft Surgery” (*Journal of the American Medical Association*, vol. 275, núm. 9), los autores explican que los pacientes fueron asignados a uno de tres grupos: **1.** un grupo tratado con triyodotironina, **2.** un grupo tratado con una píldora de sal normal y dopamina, y **3.** un grupo placebo al que se le dio una píldora de sal normal. Los autores resumen el diseño muestral como un “experimento prospectivo, aleatorio, a ciegas doble, placebo y controlado”. Describa el significado de cada uno de estos términos en el contexto de este estudio.

Repaso

Este capítulo incluyó algunas definiciones y conceptos que son muy importantes para la materia de estadística. Se presentaron definiciones fundamentales como *muestra*, *población*, *estadístico* y *parámetro*. En la sección 1-2 se analizaron los diferentes tipos de datos, así como la diferencia entre datos cualitativos y datos cuantitativos, que debe quedar completamente clara. La sección 1-3 trató el uso del pensamiento crítico para analizar y evaluar los resultados estadísticos. En específico, deberíamos saber que, por razones estadísticas, algunos ejemplos (como las muestras de respuestas voluntarias) son muy inadecuados. En la sección 1-4 se presentaron elementos importantes del diseño de experimentos. Debemos comprender la definición de una muestra aleatoria simple. También debemos reconocer la importancia de

un diseño experimental planeado de forma cuidadosa. Se analizaron brevemente algunos principios básicos del diseño de experimentos; no obstante, hay libros y cursos completos dedicados a este tema. Al terminar el estudio de este capítulo, usted debe ser capaz de:

- Distinguir entre una población y una muestra, así como entre un parámetro y un estadístico.
- Entender la importancia de un buen diseño experimental, incluyendo el control de los efectos de las variables, la réplica y la aleatoriedad.
- Reconocer la importancia de seguir buenos procedimientos de muestreo en general, y reconocer la importancia de una *muestra aleatoria simple* en particular. Entender que si los datos muestrales no son reunidos de una forma apropiada, los datos pueden ser tan completamente inútiles que ninguna cantidad de tortura estadística los salvará.

Conocimientos estadísticos y pensamiento crítico

1. **Tamaño de la muestra.** ¿Una muestra grande es necesariamente una buena muestra? ¿Por qué?
2. **Naturaleza de los datos.** Un fisiólogo elige al azar a 16 corredores que terminan el maratón de Nueva York, y luego mide la estatura de cada persona elegida.
 - a. ¿Los datos son cualitativos o cuantitativos?
 - b. ¿Los datos son discretos o continuos?
 - c. Si el investigador utiliza los datos muestrales para inferir algo acerca de la población, ¿cuál es la población?
3. **Muestreo.** Un estudiante de posgrado está realizando una investigación en psicología y necesita obtener las puntuaciones de CI de 50 individuos. Él coloca un anuncio en el periódico local pidiendo voluntarios, cada uno de los cuales recibirá \$50 por responder una prueba de CI. ¿Se trata de una muestra adecuada? ¿Por qué?
4. **Muestreo.** Una investigadora de mercados quiere determinar el valor promedio de un automóvil que posee un residente en Estados Unidos. Ella elige aleatoriamente a 10 propietarios de automóviles de cada estado y los encuesta; obtiene una lista de 500 valores muestrales. Luego suma los 500 valores y los divide entre 500 para sacar un promedio. ¿El resultado es una buena estimación del valor promedio de un automóvil poseído por un residente en Estados Unidos? ¿Por qué?

Ejercicios de repaso

1. **Muestreo.** Poco después de que las torres del World Trade Center fueran destruidas por terroristas, America Online aplicó una encuesta a sus suscriptores de Internet y preguntó lo siguiente: “¿Deberían reconstruirse las torres del World Trade Center?” Entre 1,304,240 personas que respondieron, 768,731 dijeron que “sí”, 286,756 contestaron que “no”, y 248,753 dijeron que era “demasiado pronto para decidir”. Puesto que esta muestra es extremadamente grande, ¿se puede considerar que las respuestas son representativas de la población de Estados Unidos? Explique.
2. **Diseño de muestreo.** Usted ha sido contratado por Verizon para realizar una encuesta acerca del uso del teléfono celular, entre los estudiantes de tiempo completo que asisten a su universidad. Describa un procedimiento para obtener una muestra de cada tipo: aleatoria, sistemática, de conveniencia, estratificada y por conglomerados.
3. Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) que se aplica a cada uno de los siguientes ejemplos:
 - a. Los pesos de las personas que son lanzadas al aire en un apasionado concierto de rock.
 - b. Una clasificación de crítica de cine de “debe verse; recomendada; no recomendada; ni se le ocurra verla”.

- c. Una clasificación de crítica de cine de “drama, comedia y aventura”.
- d. Bob, que es distinto en muchas formas, mide el tiempo en días, donde 0 corresponde a su fecha de nacimiento. El día anterior a su nacimiento es -1 , el día después de su nacimiento es $+1$, y así sucesivamente. Bob ha convertido fechas de eventos históricos importantes a su sistema de numeración. ¿Cuál es el nivel de medición de estos números?
- 4. Coca-Cola.** Coca-Cola Company tiene 366,000 accionistas y efectúa una encuesta mediante la selección aleatoria de 30 accionistas de cada una de las 50 entidades de Estados Unidos. Se registra el número de acciones de cada accionista de la muestra.
- ¿Los valores obtenidos son discretos o continuos?
 - Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.
 - ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por conglomerados) se usa?
 - Si se calcula el número promedio (la media) de acciones, ¿el resultado es un estadístico o un parámetro?
 - Si usted fuera el ejecutivo en jefe de Coca-Cola Company, ¿qué característica del conjunto de datos consideraría usted que es extremadamente importante?
 - ¿Qué es lo que está incorrecto al evaluar la opinión de los accionistas enviando un cuestionario por correo, que éstos podrían llenar y regresar por el mismo medio?
- 5. Más Coca-Cola.** Identifique el tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por conglomerados) que se utiliza cuando una muestra de 366,000 accionistas de Coca-Cola se obtiene como ya se describió. Después determine si el esquema de muestreo produciría una muestra que sea representativa de la población de todos los 366,000 accionistas.
- Se reúne una lista completa de todos los accionistas y se selecciona cada 500-ésimo nombre.
 - En la junta anual de accionistas, se realiza una encuesta a todos los asistentes.
 - Se seleccionan al azar 50 diferentes corredores de bolsa, y se hace una encuesta acerca de todos sus clientes que tienen acciones de Coca-Cola.
 - Se reúne un archivo de computadora de todos los accionistas, de modo que a todos se les numera de forma consecutiva, y luego se usan números aleatorios generados por computadora para seleccionar la muestra de accionistas.
 - Se reúnen todos los códigos postales de los accionistas, y se seleccionan al azar a cinco accionistas de cada código postal.
- 6. Diseño de experimento.** Usted planea realizar un experimento para probar la eficacia del Sleepeze, un nuevo fármaco que se supone que reducirá el efecto del insomnio. Usted usará una muestra de sujetos que han sido tratados con el fármaco y otra muestra de sujetos a los que se les dio un placebo.
- ¿Qué es el “estudio a ciegas” y como se utilizaría en este experimento?
 - ¿Por qué es importante el uso del estudio a ciegas en este experimento?
 - ¿Qué es un diseño completamente aleatorio?
 - ¿Qué es un diseño rigurosamente controlado?
 - ¿Qué es la réplica y por qué es importante?
- 7. El honesto Abe**
- Cuando Abraham Lincoln fue elegido por primera vez para la presidencia, recibió el 39.82% de los 1,865,908 votos emitidos. El conjunto de todos esos votos es la población a considerar. ¿El 39.82% es un parámetro o un estadístico?
 - El inciso a) indica el total de votos emitidos en la elección presidencial de 1860. Considere el número total de votos emitidos en todas las elecciones presidenciales. ¿Estos valores son discretos o continuos?
 - ¿Cuántos votos obtuvo Lincoln cuando fue elegido presidente por primera vez?

8. Porcentajes

- Las etiquetas de las barras energéticas de proteína U-Turn incluyen la afirmación de que contienen “125% menos grasa que las principales marcas de dulces de chocolate” (según datos de la revista *Consumer Reports*). ¿Cuál es el error en esta afirmación?
- En un estudio sobre la presencia de radón en las casas, se midieron 237 viviendas en LaGrange, Nueva York, y se encontró que el 19% de ellas tenían niveles de radón mayores que 0.4 pCi/L (picocurios por litro), que es el nivel máximo de seguridad aceptado por la Environmental Protection Agency. ¿Cuál es el número real de casas evaluadas que tienen niveles de radón inseguros?
- Cuando se probaron los niveles de radón en 186 casas de Hyde Park, Nueva York, se descubrió que 30 de ellas superaban el nivel máximo de seguridad descrito en el inciso b). ¿Qué porcentaje de las casas de Hyde Park rebasan el nivel máximo de seguridad del radón?

Ejercicios de repaso acumulativo

Los ejercicios de repaso acumulativos de este libro están diseñados para incluir temas de capítulos anteriores. Para los capítulos 2 a 15, estos ejercicios incluyen temas de capítulos anteriores. Para este capítulo, presentamos ejercicios de calentamiento para calculadora, con expresiones similares a las que se encuentran a lo largo de este texto. Utilice su calculadora para obtener los valores indicados.

- Remítase al conjunto de datos 14 del apéndice B y considere sólo los pesos de las monedas con la cabeza india. ¿Qué valor se obtiene cuando se suman estos pesos y el total se divide entre la cantidad de esas monedas? (Este resultado, llamado *media*, se analiza en el capítulo 3).

$$2. \frac{98.20 - 98.60}{0.62}$$

$$3. \frac{98.20 - 98.60}{0.62} \\ \sqrt{106}$$

$$4. \left[\frac{1.96 \cdot 0.25}{0.03} \right]^2$$

$$5. \frac{(50 - 45)^2}{45}$$

$$6. \frac{(2 - 4)^2 + (3 - 4)^2 + (7 - 4)^2}{3 - 1}$$

$$7. \sqrt{\frac{(2 - 4)^2 + (3 - 4)^2 + (7 - 4)^2}{3 - 1}}$$

$$8. \frac{8(151,879) - (516.5)(2176)}{\sqrt{8(34,525.75) - 516.5^2} \sqrt{8(728,520) - 2176^2}}$$

En los ejercicios 9 a 12, las expresiones dadas están diseñadas para dar resultados expresados en notación científica. Por ejemplo, el resultado de la pantalla de la calculadora de 1.23E5 puede expresarse como 123,000, y el resultado de 4.56E-4 (o 4.65^{-4} en algunas calculadoras) puede expresarse como 0.000456. Realice la operación que se indica y exprese el resultado como un número ordinario que no esté en notación científica.

$$9. 0.5^{10} \quad 10. 2^{40} \quad 11. 7^{12} \quad 12. 0.8^{50}$$

Actividades de cooperación en equipo

- 1. Actividad fuera de la clase** Busque en periódicos y revistas un ejemplo de una gráfica engañosa. (Véase, por ejemplo, las figuras 1-1 y 1-2). Describa por qué la gráfica es engañosa y reconstrúyala para que presente la información de forma correcta.
- 2. Actividad en clase** Obtenga 18 pajillas (popotes) de la cafetería. Corte 6 de ellos por la mitad, corte 6 de ellos en cuartos, y los otros 6 déjelos como están. Ahora debe haber 42 popotes de diferentes longitudes. Póngalos en una bolsa, revuélvalos, luego seleccione una pajilla, mida su longitud, y póngalo de nuevo en la bolsa. Repita esto hasta seleccionar 20 pajillas. *Importante:* Seleccione las pajillas sin mirar al interior de la bolsa y saque la primera que toque. Calcule el promedio (media) de la muestra de 20 pajillas. Ahora saque todas las pajillas y encuentre la media de la población. ¿Dio la muestra un promedio cercano al promedio de la población real? ¿Por qué?
- 3. Actividad en clase** A mediados de diciembre de un año reciente, el proveedor de servicios de Internet America Online (AOL) aplicó una encuesta a sus usuarios. Se les preguntó lo siguiente acerca de los árboles de Navidad: “¿Cuál prefiere usted?” Las respuestas eran “un árbol natural” o “un árbol artificial”. De las 7073 respuestas recibidas de los usuarios de Internet, 4650 indicaron un árbol natural, y 2423 indicaron un árbol artificial. Ya señalamos que como la muestra es de respuesta voluntaria, no es posible obtener conclusiones acerca de una población mayor que las 7073 personas que respondieron. Identifique otros problemas en esta pregunta de encuesta.
- 4. Actividad en clase** Identifique los problemas de lo siguiente.
 - Un reporte televisado recientemente por *CNN Headline News* incluyó el comentario de que la criminalidad en Estados Unidos disminuyó en la década de 1980 debido al incremento de abortos en la década de 1970, que resultó en un menor número de niños no deseados.
 - La revista *Consumer Reports* envió por correo un cuestionario anual acerca de automóviles y otros productos de consumo. También se incluyó la petición de una contribución económica voluntaria y una votación para el Consejo de Administración de la revista. Las respuestas tenían que enviarse por correo en sobres que requerían timbres postales.

Proyecto tecnológico

El objetivo de este proyecto consiste en presentar los recursos tecnológicos que usted usará en su curso de estadística. Remítase al conjunto de datos 13 en el apéndice B y use únicamente los pesos de los M&M rojos. Utilizando su programa de estadística o la calculadora TI-83/84 Plus, introduzca los 13 valores y luego imprima un listado de ellos.

STATDISK: Haga clic en **Datasets** en la parte superior de la pantalla, después seleccione **M&M** para abrir el conjunto de datos y luego seleccione **Print Data**.

Minitab: Introduzca los datos en la columna C1, después haga clic en **File** y seleccione **Print Worksheet**.

- Excel:** Registre los datos en la columna A, después haga clic en **File** y seleccione **Print**.
- TI-83/84 Plus:** La impresión de la pantalla de la TI-83/84 Plus sólo es posible mediante el uso de la conexión a una computadora, y los procedimientos para distintas conexiones varían. Consulte su manual para seguir el procedimiento correcto.

De los datos a la decisión

Pensamiento crítico

En la actualidad una mujer gana 76 centavos por cada dólar que gana un hombre. Construya una gráfica que describa esta información de manera imparcial y sin sesgos. Elabore una segunda gráfica que exagere la diferencia de las ganancias entre hombres y mujeres.

Análisis de los resultados

Suponiendo que usted debe defender la discrepancia entre lo que ganan los hombres y

las mujeres, identifique al menos un factor que sirva para justificar esta discrepancia. Las investigaciones y los análisis han demostrado que parte de esta discrepancia se puede explicar usando factores legítimos, aunque gran parte de ella no pueda explicarse con tales factores. Suponiendo que existe un factor importante de discriminación con base en el género, ¿sería adecuado tratar de luchar contra esa discriminación publicando una gráfica que exagere las diferencias entre las ganancias de hombres y mujeres?



Proyecto de Internet

El sitio en Internet de Estadística elemental

En esta sección de cada capítulo, se le pedirá que visite la página principal en Internet de este libro de texto. Desde ahí usted puede llegar a las páginas que sirven para todos los proyectos de Internet que vienen en la *décima edición de Estadística*. Vaya a este sitio ahora y familiarícese con todas las características a las que puede tener acceso para este libro.

Cada proyecto de Internet incluye actividades como la exploración de conjuntos de datos, la ejecución de modelos de simulación y la investigación de ejemplos de la vida real, que se encuentran en varios sitios de Internet. Estas actividades le ayudarán a explorar y entender la rica naturaleza de la estadística y su importancia en nuestro mundo. ¡Visite el sitio del libro ahora y disfrute de las exploraciones!

<http://www.pearsoneducacion.net/triola>

La estadística en el trabajo

“Utilizamos la estadística para determinar el grado de aislamiento entre grupos putativos”.



Sarah Mesnick

Ecologista conductual y molecular

Sara Mesnick es miembro posdoctorado del National Research Council. En su trabajo como bióloga en mamíferos marinos lleva a cabo investigación en el mar y en el Laboratorio de Ecología Molecular. Sus estudios se enfocan en la organización social y estructura poblacional de los cachalotes. Obtuvo su doctorado en biología evolutiva en la Universidad de Arizona.

¿A qué se dedica?

Mi investigación se enfoca en la relación que existe entre la sociabilidad y la estructura poblacional de los cachalotes. Nosotros usamos esta información para crear mejores modelos administrativos, para la conservación de esta y otras especies de mamíferos marinos en peligro de extinción.

¿Qué conceptos de la estadística utiliza?

En la actualidad empleo la chi cuadrada y el estadístico *F* para examinar la estructura poblacional, y medidas de regresión para estimar el grado de relación entre los individuos de la manada de ballenas. Utilizamos la chi cuadrada y el estadístico *F* para determinar la cantidad de poblaciones discretas de ballenas en el Pacífico. Tales poblaciones se manejan como grupos independientes. El análisis de regresión de la relación se usa para determinar el parentesco dentro de los grupos.

¿Podría citar un ejemplo específico que ilustre el uso de la estadística?

Actualmente estoy trabajando con muestras de tejido que obtengo de tres encallamientos masivos de cachalotes. Utilizamos marcadores genéticos para determinar el grado de parentesco entre los individuos encallados. Se trata de un comportamiento sorprendente: manadas completas nadaron hacia la playa siguiendo a un ballenato hembra, encallaron y después murieron. Pensamos que para hacer algo tan extremo como esto, los individuos implicados debieron tener una relación muy cercana; sin embargo, estamos descubriendo que no es así. La estadística nos permite determinar la probabilidad de

que dos individuos sean parientes, dado el número de alelos que comparten. Además, los cachalotes y muchas otras especies de mamíferos marinos, aves, y tortugas se lastiman o mueren incidentalmente en operaciones de pesca. Necesitamos conocer el tamaño de la población de la que provienen estos animales; si la población es pequeña, y las muertes incidentales abundantes, la población de mamíferos marinos se verá amenazada. Empleamos la estadística para determinar el grado de aislamiento que hay entre los grupos putativos. Si resultara que los grupos están aislados, usaríamos esa información para preparar planes de manejo diseñados específicamente para conservar a los mamíferos marinos de la región. Tal vez sean necesarias actividades humanas que protejan la salud del ambiente marino y a sus habitantes.

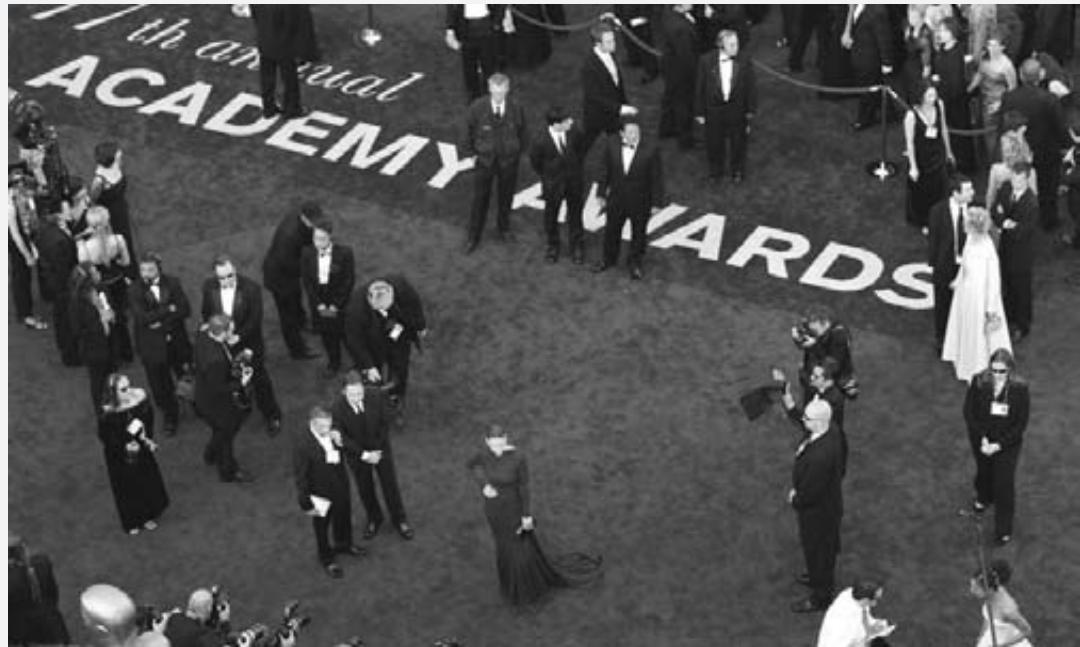
¿De qué manera enfoca su investigación?

Tratamos de evitar ideas preconcebidas acerca de la forma en que los animales están distribuidos en su medio ambiente. Dado que los mamíferos marinos, en particular, son tan difíciles de estudiar, suelen existir ideas aceptadas sobre lo que estos animales hacen, aún cuando esto no se haya investigado de manera crítica. En lo que se refiere al parentesco entre individuos dentro de grupos de cachalotes, alguna vez se pensó que éste era matrilineal y que incluía a un “macho dominante del harem”. Con el advenimiento de la tecnología genética, la dedicación en el trabajo de campo, mentes más abiertas y análisis más críticos (aquí interviene la estadística), somos capaces de examinar de nuevo estas ideas.



Resumen y gráficas de datos

2



- 2-1 Panorama general**
- 2-2 Distribuciones de frecuencias**
- 2-3 Histogramas**
- 2-4 Gráficas estadísticas**

¿Los Premios de la Academia discriminan por la edad?

Cada año se otorgan Óscars a la mejor actriz y al mejor actor. En la tabla 2-1 se presenta una lista con las edades de los galardonados en el momento de la ceremonia de entrega de los premios. Las edades aparecen en orden, empezando con la primera ceremonia de los Premios de la Academia en 1928. [Notas: En 1968 hubo un empate en la categoría de mejor actriz, y se utilizó el promedio (la media) de las dos edades; en 1932 hubo un empate en la categoría de mejor actor, y se utilizó el promedio (la media) de las dos edades. Tales datos se basan en el artículo “Ages of Oscar-winning Best Actors and Actresses”, de Richard Brown y Gretchen Davis, en la revista *Mathematics Teacher*. En ese artículo, el año de nacimiento del ganador del premio se restó del año de la ceremonia; no obstante, las edades de la tabla 2-1 se basan en la fecha de nacimiento del ganador y en la fecha de la ceremonia de premiación].

La pregunta básica que consideraremos es: ¿Hay diferencias importantes entre las edades de las mejores actrices y las edades de los mejores actores? ¿Al parecer los actores y las actrices son juzgados estrictamente por sus habilidades artísticas? O bien, ¿existe discriminación por la edad y las mejores actrices suelen ser más jóvenes que los mejores actores? ¿Hay algunas otras diferencias evidentes? Además de ser interesante, esto es importante porque nos brinda información sobre la forma en que nuestra sociedad percibe a los hombres y a las mujeres en general.

Pensamiento crítico: Una comparación visual entre las edades de la tabla 2-1 sería reveladora para las personas que tienen una habilidad especial para observar un orden en este tipo de listas de números; sin embargo, para nosotros los simples mortales, es probable

que la lista no revele mucha información. Afortunadamente, se dispone de métodos para investigar este tipo de conjuntos de datos, y pronto veremos que tales procedimientos revelan características importantes que nos permiten *entender* los datos. Seremos capaces de hacer comparaciones inteligentes y reveladoras; aprenderemos técnicas para resumir, graficar, describir, explorar y comparar conjuntos de datos como los de la tabla 2-1.

Tabla 2-1 Premios de la Academia: Edades de las mejores actrices y los mejores actores

Las edades (en años) aparecen en orden, empezando con la primera ceremonia de premiación.

Mejores actrices

22	37	28	63	32	26	31	27	27	28
30	26	29	24	38	25	29	41	30	35
35	33	29	38	54	24	25	46	41	28
40	39	29	27	31	38	29	25	35	60
43	35	34	34	27	37	42	41	36	32
41	33	31	74	33	50	38	61	21	41
26	80	42	29	33	35	45	49	39	34
26	25	33	35	35	28				

Mejores actores

44	41	62	52	41	34	34	52	41	37
38	34	32	40	43	56	41	39	49	57
41	38	42	52	51	35	30	39	41	44
49	35	47	31	47	37	57	42	45	42
44	62	43	42	48	49	56	38	60	30
40	42	36	76	39	53	45	36	62	43
51	32	42	54	52	37	38	32	45	60
46	40	36	47	29	43				

2-1 Panorama general

En este capítulo presentamos métodos importantes para organizar, resumir y graficar conjuntos de datos. El objetivo principal no es simplemente obtener alguna tabla o gráfica, sino *entender* los datos. Cuando se describen, exploran y comparan conjuntos de datos, las siguientes características suelen ser sumamente importantes.

Características importantes de los datos

1. **Centro:** Valor promedio o representativo que indica la localización de la mitad del conjunto de los datos.
2. **Variación:** Medida de la cantidad en que los valores de los datos varían entre sí.
3. **Distribución:** La naturaleza o forma de la distribución de los datos (como en forma de campana, uniforme o sesgada).
4. **Valores extremos:** Valores muestrales que están muy alejados de la vasta mayoría de los demás valores de la muestra.
5. **Tiempo:** Características cambiantes de los datos a través del tiempo.

Sugerencia de estudio: La memorización suele ser ineficaz para aprender o recordar información importante. Sin embargo, las cinco características anteriores son tan importantes que pueden recordarse usando una técnica mnémica con las iniciales CVDVT; por ejemplo, “Cuidado con los Virus que Destruyen el Valioso Trabajo”. Este tipo de técnicas de memorización son muy efectivas para recordar palabras clave importantes que evoquen conceptos fundamentales.

Pensamiento crítico e interpretación: más allá de las fórmulas y de los cálculos a mano

Los profesores de estadística generalmente creen que no es importante memorizar fórmulas o realizar cálculos aritméticos complejos a mano. En cambio, suelen enfocarse en obtener resultados utilizando algún tipo de tecnología (calculadoras o programas de cómputo), y después darle un sentido práctico a los resultados a través del pensamiento crítico. Téngalo en cuenta mientras estudia este capítulo, el siguiente y el resto del libro. Aunque este capítulo incluye pasos detallados de procedimientos importantes, no es necesario dominar esos pasos en todos los casos. Sin embargo, le recomendamos realizar algunos cálculos manuales antes de utilizar una calculadora o una computadora. Con ello, logrará una mejor comprensión y apreciará mejor los resultados que se obtienen con la tecnología.

2-2 Distribuciones de frecuencias

Concepto clave Cuando trabajamos con grandes conjuntos de datos, a menudo es útil organizarlos y resumirlos al construir una tabla llamada *distribución de frecuencias*, la cual definimos más adelante. Puesto que los programas de cómputo y las calculadoras pueden generar distribuciones de frecuencias de manera automática, los detalles sobre su elaboración no son tan importantes como entender lo que nos dicen sobre los conjuntos de datos. En particular, una distribución de frecuencias nos ayuda a entender la naturaleza de la *distribución* de un conjunto de datos.

Definición

Una **distribución de frecuencias** (o **tabla de frecuencias**) lista valores de los datos (ya sea de manera individual o por grupos de intervalos), junto con sus frecuencias (o conteos) correspondientes.

La tabla 2-2 es una distribución de frecuencias que resume las edades de las actrices ganadoras del premio Óscar de la tabla 2-1. La **frecuencia** de una clase en particular es el número de valores originales que caen en esa clase. Por ejemplo, la primera clase de la tabla 2-2 tiene una frecuencia de 28, que indica que 28 de las edades originales están entre los 21 y los 30 años inclusive.

Primero presentaremos algunos términos estándar que se utilizan al referirse a la distribución de frecuencias, y luego describiremos la manera en que se construyen e interpretan.

Definiciones

Los **límites de clase inferiores** son las cifras más pequeñas que pueden pertenecer a las diferentes clases. (Los límites de clase inferiores de la tabla 2-2 son 21, 31, 41, 51, 61 y 71).

Los **límites de clase superiores** son las cifras más grandes que pueden pertenecer a las diferentes clases. (Los límites de clase superiores de la tabla 2-2 son 30, 40, 50, 60, 70 y 80).

Las **fronteras de clase** son las cifras que se utilizan para separar las clases, pero sin los espacios creados por los límites de clase. En la figura 2-1 se muestran los espacios creados por los límites de clases de la tabla 2-2. En la figura 2-1 se percibe con facilidad que los valores 30.5, 40.5, . . . , 70.5 están en el centro de esos espacios, y a tales cifras se les conoce como fronteras de clase. Las dos fronteras de clase desconocidas (que en la figura 2-1 se indican con signos de interrogación) se identifican fácilmente al seguir el patrón establecido por las otras fronteras de clase de 30.5, 40.5, . . . , 70.5. La frontera de clase inferior es 20.5, y la frontera de clase superior es 80.5. Por lo tanto, la lista completa de las fronteras de clase es 20.5, 30.5, 40.5, 50.5, 60.5, 70.5 y 80.5. Las fronteras de clase serán muy útiles en la siguiente sección, cuando elaboraremos una gráfica llamada histograma.

Las **marcas de clase** son los puntos medios de las clases. (Las marcas de clase de la tabla 2-2 son 25.5, 35.5, 45.5, 55.5, 65.5 y 75.5). Las marcas de clase se calculan sumando el límite de clase inferior con el límite de clase superior, y dividiendo la suma entre 2.

La **anchura de clase** es la diferencia entre dos límites de clase inferiores consecutivos o dos fronteras de clase inferiores consecutivas. (La anchura de clase de los datos de la tabla 2-2 es 10).

Las definiciones de anchura de clase y fronteras de clase son un poco engañosas. Tenga cuidado para evitar el error común de considerar la anchura de clase como la diferencia entre el límite de clase inferior y el límite de clase superior. Observe que en la tabla 2-2 la anchura de clase es de 10, no de 9. El proceso del

Tabla 2-2

Distribución de frecuencias: Edades de las mejores actrices

Edad de las actrices	Frecuencias
21–30	28
31–40	30
41–50	12
51–60	2
61–70	2
71–80	2



Se identifican autores

Entre 1787 y 1788, Alexander Hamilton, John Jay y James Madison publicaron, de forma anónima, el famoso diario *Federalist Papers*, en un intento por convencer a los neoyorquinos de que deberían ratificar la Constitución. Se conoció la identidad de la mayoría de los autores de los artículos, pero la autoría de 12 de éstos siguió siendo motivo de discusión. Mediante el análisis estadístico de las frecuencias de varias palabras, ahora podemos concluir que *probablemente* James Madison era el autor de esos 12 documentos. En muchos de los artículos disputados, la evidencia en favor de la autoría de Madison es abrumadora, al grado de que estamos casi seguros de que es lo correcto.

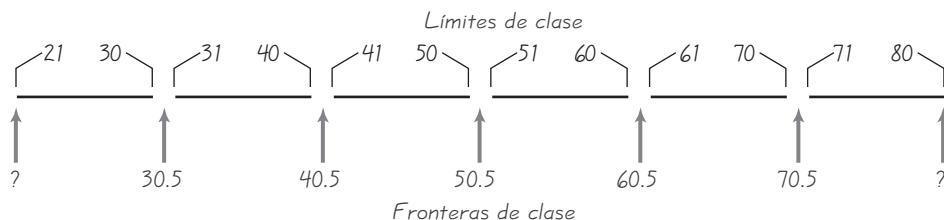


Figura 2-1 Cálculo de las fronteras de clase

cálculo de las fronteras de clase se simplifica al entender que éstos básicamente dividen la diferencia entre el final de una clase y el inicio de la siguiente.

Procedimiento para construir una distribución de frecuencias

Las distribuciones de frecuencias se construyen por las siguientes razones: **1.** Es posible resumir conjuntos grandes de datos, **2.** se logra cierta comprensión sobre la naturaleza de los datos, y **3.** se tiene una base para construir gráficas importantes (como los *histogramas*, que se estudiarán en la siguiente sección). Muchos usos de la tecnología nos permiten obtener distribuciones de frecuencias de manera automática, sin necesidad de tener que hacerlas manualmente; no obstante, a continuación se presenta el procedimiento básico:

1. Decida el número de clases que desea, el cual debe estar entre 5 y 20. El número que elija puede verse afectado por la comodidad de usar cifras enteras.
 2. Calcule
- $$\text{Anchura de clase} \approx \frac{(\text{valor más alto}) - (\text{valor más bajo})}{\text{número de clases}}$$
- Redondee este resultado para obtener un número más adecuado. (Generalmente se redondea *hacia arriba*). Es probable que necesite cambiar el número de clases, pero la prioridad debe ser utilizar valores que sean fáciles de comprender.
3. Punto de partida: Comience por elegir un número para el límite inferior de la primera clase. Elija el valor del dato más bajo o un valor conveniente que sea un poco más pequeño.
 4. Usando el límite inferior de la primera clase y la anchura de clase, proceda a listar los demás límites de clase inferiores. (Sume la anchura de clase al punto de partida para obtener el segundo límite de clase inferior. Despues sume la anchura de clase al segundo límite de clase inferior para obtener el tercero, y así sucesivamente).
 5. Anote los límites inferiores de clase en una columna vertical y luego proceda a anotar los límites superiores de clase, que son fáciles de identificar.
 6. Ponga una marca en la clase adecuada para cada dato. Utilice las marcas para obtener la frecuencia total de cada clase.

Cuando construya una distribución de frecuencias, asegúrese de que las clases no se traslapen, de modo que cada uno de los valores originales pertenezca exactamente a una de las clases. Incluya todas las clases, aun las que tengan una frecuencia de cero. Trate de utilizar la misma anchura para todas las clases, aunque a veces es imposible evitar intervalos con finales abiertos, como “65 años y mayores”.

EJEMPLO Edades de las mejores actrices Use las edades de las mejores actrices de la tabla 2-1 y siga el procedimiento anterior para construir la distribución de frecuencias de la tabla 2-2. Suponga que desea incluir 6 clases.

SOLUCIÓN

Paso 1: Comience seleccionando 6 clases.

Paso 2: Calcule la anchura de clase. En el siguiente cálculo, 9.833 se redondea a 10, ya que es un número más conveniente.

$$\text{Anchura de clase} \approx \frac{(\text{valor más alto}) - (\text{valor más bajo})}{\text{número de clases}}$$

$$= \frac{80 - 21}{6} = 9.833 \approx 10$$

Paso 3: Paso 3: Elegimos un punto de partida de 21, que es el valor más bajo de la lista y un número conveniente, ya que 21-30 se convierte en la primera clase.

Paso 4: Sume la anchura de clase 10 al punto de partida 21 para determinar que el segundo límite inferior de clase es igual a 31. Continúe y sume la anchura de clase 10 para obtener los límites inferiores de clase restantes de 41, 51, 61 y 71.

Paso 5: Liste los límites de clase inferiores de forma vertical, como se muestra al margen. Con esta lista podemos identificar con facilidad los límites de clases superiores correspondientes, que son 30, 40, 50, 60, 70 y 80.

Paso 6: Despues de identificar los límites inferiores y superiores de cada clase, proceda a trabajar con el conjunto de datos asignando una marca a cada valor. Una vez completadas las marcas, súmelas para obtener las frecuencias que se presentan en la tabla 2-2.

21 –
31 –
41 –
51 –
61 –
71 –

Distribución de frecuencias relativas

Una variante importante de la distribución básica de frecuencias utiliza las **frecuencias relativas**, que se obtienen fácilmente dividiendo cada frecuencia de clase entre el total de frecuencias. Una **distribución de frecuencias relativas** incluye los mismos límites de clase que una distribución de frecuencias, pero utiliza las frecuencias relativas en vez de las frecuencias reales. Las frecuencias relativas en ocasiones se expresan como porcentajes.

$$\text{frecuencia relativa} = \frac{\text{frecuencia de clase}}{\text{suma de todas las frecuencias}}$$

En la tabla 2-3 las frecuencias reales de la tabla 2-2 se reemplazan con las frecuencias relativas correspondientes, expresadas en porcentajes. Debido a que 28 de los 76 datos caen en la primera clase, la primera clase tiene una frecuencia relativa de $28/76 = 0.368$ o 36.8%, que a menudo se redondea a 37%. La segunda clase tiene una frecuencia relativa de $30/76 = 0.395$ o 39.5%, y así sucesivamente. Si se construye de manera correcta, la suma de las frecuencias relativas debería totalizar 1 (o 100%), con algunas pequeñas discrepancias debidas al error de redondeo. Los resultados redondeados de la tabla 2-3 hacen que la suma de las frecuencias relativas sea de 101% en vez de 100%.

Tabla 2-3

Distribución de las frecuencias relativas de las edades de las mejores actrices

Edad de la actriz	Frecuencia relativa
21–30	37%
31–40	39%
41–50	16%
51–60	3%
61–70	3%
71–80	3%

Puesto que utilizan porcentajes simples, las distribuciones de frecuencias relativas facilitan la comprensión de la distribución de los datos y permiten comparar diferentes conjuntos de datos.

Distribución de frecuencias acumulativas

Otra variante de la distribución de frecuencias estándar se utiliza cuando se buscan totales acumulativos. La **frecuencia acumulativa** de una clase es la suma de las frecuencias para esa clase y todas las clases anteriores. La tabla 2-4 presenta la distribución de frecuencias acumulativas basada en la distribución de frecuencias de la tabla 2-2. Con el uso de las frecuencias originales de 28, 30, 12, 2, 2 y 2, sumamos 28 + 30 para obtener la segunda frecuencia acumulativa de 58; luego, sumamos 28 + 30 + 12 = 70 para obtener la tercera, y así sucesivamente. En la tabla 2-4 se observa que, además del uso de las frecuencias acumulativas, los límites de clase son reemplazados por expresiones como “menor que”, las cuales describen el nuevo intervalo de valores.

Pensamiento crítico: Interpretación de las distribuciones de frecuencias

La transformación de datos en bruto en una distribución de frecuencias suele ser un medio para un llegar a un gran fin. Un objetivo importante consiste en identificar la naturaleza de la distribución, y las distribuciones “normales” son extremadamente importantes para el estudio de la estadística.

Distribución normal En capítulos posteriores de este libro, haremos referencias frecuentes a los datos con una *distribución normal*. Este uso del término “normal” tiene un significado especial en estadística, que difiere del que se utiliza en el lenguaje cotidiano. El concepto de una distribución normal se describirá más adelante, aunque por ahora usaremos una distribución de frecuencias, para ayudarnos a determinar si los datos tienen una distribución aproximadamente normal. Una característica fundamental de una distribución normal es que, cuando se grafica, el resultado tiene forma de “campana”; y al inicio las frecuencias son bajas, luego se incrementan hasta un punto máximo y luego disminuyen. Por ahora, consideraremos que una distribución de frecuencias es aproximadamente normal al determinar si tiene las siguientes características:

Distribución normal

1. Al inicio las frecuencias son bajas, después se incrementan hasta un punto máximo y luego disminuyen.
2. La distribución debe ser aproximadamente simétrica, y las frecuencias tienen que distribuirse de manera uniforme a ambos lados de la frecuencia máxima. (Las frecuencias de 1, 5, 50, 25, 20, 15, 10, 5, 3, 2, 1 no se distribuyen de forma simétrica alrededor de la puntuación máxima de 50, ni satisfacen los requisitos de simetría).

EJEMPLO Distribución normal Se seleccionaron 1000 mujeres al azar y se midieron sus estaturas. Los resultados se resumen en la distribución de frecuencias de la tabla 2-5. Al inicio las frecuencias son bajas, después se incrementan hasta un punto máximo y luego disminuyen. Además, las frecuencias se distribuyen de manera aproximadamente simétrica alrededor de la frecuencia máxima de 324. Al parecer se trata de una distribución aproximadamente normal.

Tabla 2-5 Estaturas de una muestra de 1000 mujeres

Distribución normal: Al inicio las frecuencias son bajas, luego alcanzan un nivel máximo y disminuyen nuevamente.

Estatura (pulgadas)	Frecuencia	Distribución normal:
56.0–57.9	10	← Al inicio las frecuencias son bajas, . . .
58.0–59.9	64	
60.0–61.9	178	
62.0–63.9	324	← aumentan hasta un punto máximo, . . .
64.0–65.9	251	
66.0–67.9	135	
68.0–69.9	32	
70.0–71.9	6	← disminuyen nuevamente.

La tabla 2-5 presenta datos con una distribución normal. Los siguientes ejemplos ilustran la manera en que las distribuciones de frecuencias se pueden utilizar para describir, explorar y comparar conjuntos de datos. (La siguiente sección indica que la construcción de una distribución de frecuencias suele ser el primer paso para la creación de una gráfica que describa visualmente la naturaleza de la distribución).

EJEMPLO Descripción de datos: ¿Cómo se midió la frecuencia cardíaca? Remítase al conjunto de datos 1 del apéndice B, que se refiere al pulso de 40 adultos varones seleccionados aleatoriamente. La tabla 2-6 presenta los *últimos dígitos* de estos datos. Si la tasa de pulsaciones se mide contando el número de latidos cardiacos por minuto, esperamos que los últimos dígitos tengan frecuencias muy similares. Sin embargo, observe que la distribución de frecuencias muestra que todos los últimos dígitos son números *pares*; no hay números impares, lo cual sugiere que las tasas de pulsaciones no se contaron durante un minuto. Tal vez se contaron durante 30 segundos y después se duplicaron los resultados. (Al examinar más las tasas de pulsaciones *originales*, observamos que cada valor original es un múltiplo de cuatro, lo cual sugiere que el número de latidos por minuto se contó durante 15 segundos y que dicho conteo después se multiplicó por cuatro). Es fascinante aprender algo acerca del método de recolección de datos con la simple descripción de algunas características de los mismos.

EJEMPLO Exploración de datos: ¿Qué nos indica un hueco? La tabla 2-7 es una tabla de frecuencias de los pesos (en gramos) de monedas de un centavo elegidas al azar. Un examen de las frecuencias revela un gran hueco entre las monedas de un centavo más ligeras y las más pesadas. Esto sugiere que tenemos dos poblaciones diferentes. En una investigación posterior, se descubre que las monedas de un centavo hechas antes de 1983 tenían un 97% de cobre y un 3% de zinc; mientras que las monedas de un centavo acuñadas después de 1983 tienen un 3% de cobre y un 97% de zinc, lo cual explicaría el gran hueco entre las monedas de un centavo más ligeras y las más pesadas.



Gráficas de crecimiento actualizadas

Los pediatras acostumbran utilizar gráficas de crecimiento estandarizadas para comparar el peso y la estatura de sus pacientes con una muestra de otros niños. Se considera que los niños están en un intervalo normal, si su peso y estatura caen entre los percentiles 5 y 95. Si están fuera de este intervalo, generalmente se les aplican pruebas para asegurarse de que no tengan dificultades médicas graves. Los pediatras ahora son más conscientes de un problema importante de las gráficas: como éstas se basan en niños que vivieron entre 1929 y 1975, las gráficas de crecimiento estaban resultando imprecisas. Para rectificar este problema, en el año 2000 se actualizaron las gráficas para que reflejaran las medidas actuales de millones de niños. Los pesos y las estaturas de los niños son buenos ejemplos de poblaciones que cambian con el paso del tiempo. Ésta es la razón para incluir las características que cambian en los datos con el paso del tiempo, como un aspecto importante de una población.

Tabla 2-6

Últimos dígitos de la frecuencia cardiaca en hombres

Último dígito	Frecuencia
0	7
1	0
2	6
3	0
4	11
5	0
6	9
7	0
8	7
9	0

Tabla 2-7

Monedas de un centavo elegidas al azar

Pesos de monedas de un centavo (gramos)	Frecuencia
2.40–2.49	18
2.50–2.59	19
2.60–2.69	0
2.70–2.79	0
2.80–2.89	0
2.90–2.99	2
3.00–3.09	25
3.10–3.19	8

Tabla 2-8

Edades de actores y actrices ganadores del Óscar

Edad	Actrices	Actores
21–30	37%	4%
31–40	39%	33%
41–50	16%	39%
51–60	3%	18%
61–70	3%	4%
71–80	3%	1%

Huecos El ejemplo anterior sugiere que la presencia de huecos puede revelar el hecho de que tenemos datos que provienen de dos o más poblaciones diferentes. Sin embargo, lo opuesto no es verdad, ya que los datos de diferentes poblaciones no necesariamente revelan huecos al construir histogramas.

EJEMPLO Comparación de edades de ganadores del Óscar El Problema del capítulo incluye las edades de actrices y actores en el momento en que ganaron el premio Óscar de la Academia. La tabla 2-8 muestra las frecuencias relativas de los dos géneros. Al comparar estas frecuencias relativas, parece que las actrices tienden a ser más jóvenes que los actores. Por ejemplo, observe la primera clase, que indica que el 37% de las actrices se encuentra en la categoría más joven, en comparación con sólo el 4% de los actores.

2-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Distribución de frecuencias.** ¿Qué es una distribución de frecuencias y para qué sirve?
- Recuperación de datos originales.** Al trabajar con una lista conocida de valores muestrales, una investigadora construye una distribución de frecuencias (como la que se presenta la tabla 2-2). Luego, elimina los valores de los datos originales. ¿Puede ella utilizar la distribución de frecuencias para identificar todos los valores muestrales originales?
- Clases que se traslanan.** Al construir una distribución de frecuencias, ¿qué problema crearía el uso de los siguientes intervalos de clase: 0-10, 10-20, 20-30, . . . , 90-100?
- Comparación de distribuciones.** Al comparar dos conjuntos de datos, ¿cuál es la ventaja de usar distribuciones de frecuencias relativas en vez de distribuciones de frecuencias?

En los ejercicios 5 a 8, identifique la anchura de clase, las marcas de clase y las fronteras de clase para las distribuciones de frecuencias dadas.

5. Temperatura mínima diaria (°F)	Frecuencia
35–39	1
40–44	3
45–49	5
50–54	11
55–59	7
60–64	7
65–69	1

6. Precipitación diaria (pulgadas)	Frecuencia
0.00–0.49	31
0.50–0.99	1
1.00–1.49	0
1.50–1.99	2
2.00–2.49	0
2.50–2.99	1

7. Estaturas de hombres (pulgadas)	Frecuencia
60.0–64.9	4
65.0–69.9	25
70.0–74.9	9
75.0–79.9	1
80.0–84.9	0
85.0–89.9	0
90.0–94.9	0
95.0–99.9	0
100.0–104.9	0
105.0–109.9	1

8. Pesos de plásticos de desecho (libras)	Frecuencia
0.00–0.99	8
1.00–1.99	12
2.00–2.99	6
3.00–3.99	0
4.00–4.99	0
5.00–5.99	0
6.00–6.99	0
7.00–7.99	5
8.00–8.99	15
9.00–9.99	20

Pensamiento crítico. En los ejercicios 9 a 12, responda las preguntas relativas a los ejercicios 5 a 8.

9. **Identificación de la distribución.** ¿Parece que la distribución de frecuencias del ejercicio 5 tiene una distribución normal, como lo exigen varios métodos estadísticos que se estudian más adelante en este libro?
10. **Identificación de la distribución.** ¿Parece que la distribución de frecuencias del ejercicio 6 tiene una distribución normal, como lo exigen varios métodos estadísticos que se estudian más adelante en este libro? Si sabemos que las cantidades de precipitación se obtuvieron de días elegidos al azar durante los últimos 200 años, ¿los resultados reflejan el comportamiento actual del clima?
11. **Valor extremo.** Remítase a la distribución de frecuencias del ejercicio 8. ¿Qué se sabe acerca de la estatura del hombre más alto incluido en la tabla? ¿La estatura del hombre más alto puede ser un valor correcto? Si el valor más alto fuera un error, ¿qué se concluiría acerca de la distribución después de eliminar dicho error?
12. **Análisis de la distribución.** Remítase a la distribución de frecuencias del ejercicio 8. Al parecer hay un gran hueco entre los pesos más bajos y los pesos más altos. ¿Qué sugiere este hueco? ¿Cómo se explicaría el hueco?

En los ejercicios 13 y 14, construya la distribución de frecuencias relativas que corresponda a la distribución de frecuencias del ejercicio indicado.

13. Ejercicio 5

14. Ejercicio 6

Tabla del ejercicio 18

Resultado	Frecuencia
1	24
2	28
3	39
4	37
5	25
6	27

En los ejercicios 15 y 16, construya la distribución de frecuencias acumulativas que corresponda la distribución de frecuencias del ejercicio indicado.

15. Ejercicio 5**16. Ejercicio 6**

- 17. Análisis del último dígito.** Se obtuvieron las estaturas de estudiantes de estadística como parte de un experimento para la clase. A continuación se presenta una lista de los últimos dígitos de tales estaturas. Construya una distribución de frecuencias con 10 clases. Con base en la distribución, ¿parece que las estaturas se reportaron o que se midieron realmente? ¿Qué sabe usted acerca de la exactitud de los resultados?

0 0 0 0 0 0 0 0 1 1 2 3 3 4 5 5 5 5 5 5 5 5 5 5 5 5 6 6 8 8 8 9

- 18. Dado cargado.** El autor taladró un hoyo en un dado, lo rellenó con plomo y lo lanzó 180 veces. (Sí, el autor tiene mucho tiempo libre). Los resultados se presentan en la distribución de frecuencias al margen. Construya la distribución de frecuencias relativas para los resultados que esperaría de un dado perfectamente legal. ¿El dado cargado difiere significativamente de un dado legal que no haya sido “cargado”?

- 19. Cantidad de lluvia.** Remítase al conjunto de datos 10 del apéndice B y utilice las 52 cantidades de lluvia del domingo. Construya una distribución de frecuencias iniciando con un límite de clase inferior de 0.00 y utilice una anchura de clase de 0.20. Describa la naturaleza de la distribución. ¿Al parecer se trata de una distribución de frecuencias normal, como se describió en esta sección?

- 20. Nicotina en cigarrillos.** Remítase a conjunto de datos 3 del apéndice B y utilice las 29 cantidades de nicotina medidas. Construya una distribución de frecuencias con 8 clases, iniciando con un límite de clase inferior de 0.0 y utilice una anchura de clase de 0.2. Describa la naturaleza de la distribución. ¿Al parecer se trata de una distribución de frecuencias normal, como se describió en esta sección?

- 21. Valores del IMC.** Remítase a conjunto de datos 1 del apéndice B y utilice los valores del índice de masa corporal (IMC) de las 40 mujeres. Construya una distribución de frecuencias iniciando con un límite de clase inferior de 15.0 y utilice una anchura de clase de 6.0. El IMC se calcula dividiendo el peso en kilogramos entre el cuadrado de la estatura en metros. Describa la naturaleza de la distribución. ¿Al parecer se trata de una distribución de frecuencias normal, como se describió en esta sección?

- 22. Datos del clima.** Remítase al conjunto de datos 8 del apéndice B y utilice las temperaturas mínimas reales para construir una distribución de frecuencias iniciando con un límite de clase inferior de 39 y use una anchura de clase de 6. La distribución de frecuencias del ejercicio 6 representa las cantidades de precipitación del conjunto de datos 8. Compare las dos distribuciones de frecuencias (para las temperaturas mínimas reales y las cantidades de precipitación). ¿En qué difieren fundamentalmente?

- 23. Pesos de monedas de un centavo.** Remítase al conjunto de datos 14 del apéndice B y utilice los pesos de las monedas de un centavo acuñadas antes de 1983. Construya una distribución de frecuencias iniciando con un límite de clase inferior de 2.9500 y una anchura de clase de 0.0500. ¿Al parecer los pesos se distribuyen normalmente?

- 24. Coca-Cola regular y Coca-Cola dietética.** Remítase al conjunto de datos 12 del apéndice B. Construyan la distribución de frecuencias relativas con los pesos de la Coca-Cola regular, iniciando la primera clase en 0.7900 lb, con una anchura de clase de 0.0050 lb. Después, construya otra distribución de frecuencias relativas con los pesos de la Coca-Cola dietética, iniciando la primera clase en 0.7750 lb, con una anchura de clase de 0.0050 lb. Luego, compare los resultados y determine si hay una diferencia significativa. Si es así, dé una posible explicación.

2-2 MÁS ALLÁ DE LO BÁSICO

- 25. Cargas axiales de latas de aluminio.** Remítase al conjunto de datos 15 del apéndice B. Use un programa de cómputo estadístico o una calculadora para construir una distribución de frecuencias relativas para las 175 cargas axiales de las latas de aluminio de 0.0109 pulgadas de grosor; luego, haga lo mismo para las 175 cargas axiales de las latas de aluminio de 0.0111 pulgadas de grosor. Compare las dos distribuciones de frecuencias relativas.
- 26. Interpretación de los efectos de los valores extremos.** Remítase al conjunto de datos 15 del apéndice B y utilice las cargas axiales de las latas de aluminio con un grosor de 0.0111 pulgadas. La carga de 504 lb es un *valor extremo* porque está muy alejado de los otros valores. Construya una distribución de frecuencias que incluya el valor de 504 lb, y luego construya otra distribución de frecuencias sin incluirlo. En ambos casos, inicie la primera clase en 200 lb, con una anchura de clase de 20 lb. Interprete los resultados haciendo una generalización sobre el efecto que tendría un valor extremo en una distribución de frecuencias.
- 27. Número de clases.** Para la construcción de una distribución de frecuencias, los lineamientos de Sturges sugieren que el número ideal de clases puede aproximarse usando $1 + (\log n)/(\log 2)$, donde n es el número de datos. Utilice este lineamiento para completar la tabla determinando el número de clases ideal.



2-3 Histogramas

Concepto clave En la sección 2-2 se presentó la distribución de frecuencias como una herramienta para resumir y conocer la naturaleza de la distribución de un conjunto de datos grande. Esta sección presenta el *histograma* como una gráfica muy importante para describir la naturaleza de la distribución. Puesto que muchos programas de cómputo estadísticos y calculadoras generan histogramas de forma automática, no es tan importante dominar los procedimientos mecánicos para construirlos. En cambio, debemos enfocarnos en *comprender* lo que ganaríamos al examinar histogramas. En particular, debemos desarrollar la habilidad de observar un histograma y entender la naturaleza de la distribución de los datos.

Definición

Un **histograma** es una gráfica de barras donde la escala horizontal representa clases de valores de datos y la escala vertical representa frecuencias. Las alturas de las barras corresponden a los valores de frecuencia; en tanto que las barras se dibujan de manera adyacente (sin huecos entre sí).

El primer paso para la construcción de un histograma es la creación de una tabla de distribución de frecuencias. El histograma es, básicamente, una versión gráfica de dicha tabla. Observe la figura 2-2, que incluye el histograma correspondiente a la distribución de frecuencias de la tabla 2-2 de la sección anterior.

En la escala horizontal se marca cada barra del histograma con su frontera de clase inferior a la izquierda, y su frontera de clase superior a la derecha, como se observa en la figura 2-2. En vez de utilizar las fronteras de clase a lo largo del eje horizontal, a menudo es más práctico el uso de los valores de la marca de clase en

Tabla del ejercicio 27

Número de valores	Número ideal de clases
16–22	5
23–45	6
	7
	8
	9
	10
	11
	12



Datos faltantes

Es común que las muestras carezcan de algunos datos. Los datos faltantes caen en dos categorías generales: **1.** valores faltantes que resultan de causas aleatorias no relacionadas con los valores de los datos, y **2.** valores faltantes que resultan de causas que no son aleatorias. Algunas de las causas aleatorias son factores como la anotación incorrecta de valores muestrales o la pérdida de resultados de encuesta. Este tipo de valores faltantes a menudo puede despreciarse, ya que no ocultan de manera sistemática algunas características que podrían afectar los resultados significativamente. Es difícil enfrentarse a valores faltantes que no se deben al azar. Por ejemplo, los resultados del análisis del ingreso podrían verse seriamente afectados, si la gente con ingresos muy altos se rehusa a proporcionar los valores por temor a las auditorías fiscales. Tales ingresos muy altos faltantes no deben descartarse; más bien, se debe realizar otra investigación para identificarlos.

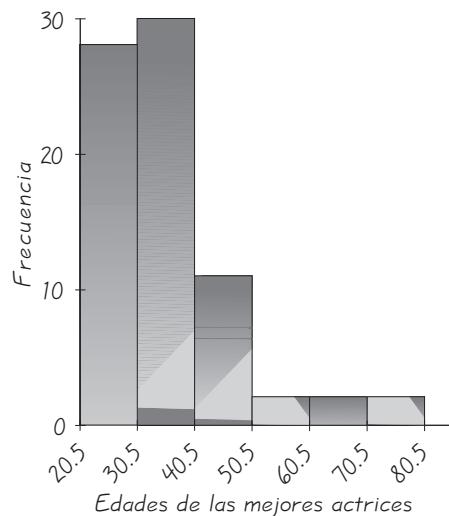


Figura 2-2 Histograma

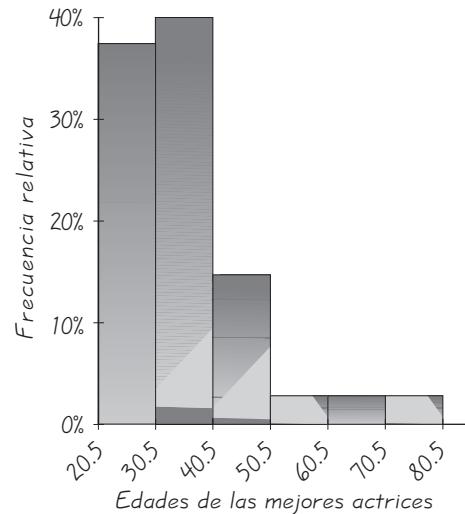


Figura 2-3 Histograma de frecuencias relativas

el centro de las barras correspondientes. El uso de los valores de la marca de clase es muy común en los programas de cómputo que generan histogramas de manera automática.

Escala horizontal: Utilice las fronteras de clase o las marcas de clase.

Escala vertical: Utilice las frecuencias de clase.

Antes de construir un histograma a partir de una distribución de frecuencias terminada, debemos hablar de las escalas que se emplean en los ejes vertical y horizontal. La frecuencia máxima (o el siguiente número más alto conveniente) debería sugerir un valor para la parte superior de la escala vertical; 0 tiene que estar en la parte inferior. En la figura 2-2 diseñamos la escala vertical para que corriera de 0 a 30. La escala horizontal debe subdividirse de tal forma que permita que todas las clases se ajusten bien. De manera ideal, es necesario que intentemos seguir la regla general de que la altura vertical del histograma debe medir alrededor de tres cuartas partes de la anchura total. A ambos ejes se les asignan etiquetas claras.

Histograma de frecuencias relativas

Un **histograma de frecuencias relativas** tiene la misma forma y escala horizontal que un histograma, pero la escala vertical está marcada con las frecuencias relativas en vez de las frecuencias reales, como se observa en la figura 2-3.

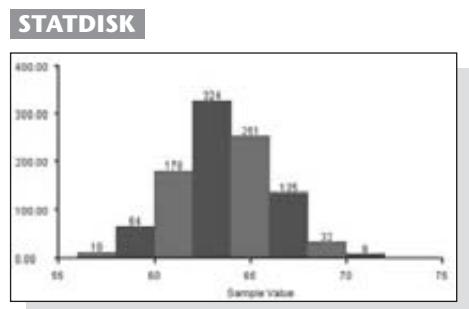
Pensamiento crítico: interpretación de histogramas

Recuerde que el objetivo no es la simple construcción de un histograma, sino *entender* algo acerca de los datos. Analice el histograma para ver qué es posible aprender acerca de “CVDVT”: el centro de los datos, la variación (que se estudiará con detalle en la sección 3-3), la forma de la distribución y la existencia o ausencia de valores extremos (valores que se encuentran lejos de los demás). Al examinar

la figura 2-2, vemos que el histograma se centra alrededor del 35, que los valores varían aproximadamente desde 21 hasta 80, y que la forma de la distribución está más cargada hacia la izquierda, lo cual significa que las actrices que ganan un Óscar tienden a ser desproporcionadamente más jóvenes, y que a pocas actrices mayores se les otorga ese premio.

Distribución normal En la sección 2-2 señalamos que el uso del término “normal” tiene un significado especial en la estadística, el cual difiere del significado que generalmente se le da en el lenguaje cotidiano. Una característica fundamental de una distribución normal es que, cuando se grafica en un histograma, el resultado es una curva en forma de “campana”, como en el histograma generado con STATDISK que aquí se muestra. [Las principales características de la curva en forma de campana son **1.** el aumento de las frecuencias, las cuales alcanzan un punto máximo y luego disminuyen; y **2.** la simetría, donde la mitad izquierda de la gráfica es casi una imagen en espejo de la mitad derecha]. Este histograma corresponde a la distribución de frecuencias de la tabla 2-5, que incluye 1000 estaturas de mujeres elegidas al azar. Muchos procedimientos estadísticos requieren que los datos muestrales provengan de una población que tenga una distribución que no se desvíe mucho de la normalidad, y a menudo podemos usar un histograma para juzgar si se satisface tal requisito de una distribución normal.

Decimos que la distribución es *normal* porque es una curva en forma de campana.



Uso de la tecnología

Los paquetes de cómputo estadísticos actuales son muy eficaces para generar gráficas impresionantes, incluyendo histogramas. En este libro a menudo hacemos referencia a STATDISK, Minitab, Excel, y la calculadora TI-83/84 Plus, que son tecnologías que sirven para generar histogramas. Las instrucciones detalladas varían desde lo extremadamente sencillo a lo muy complejo, por lo que a continuación haremos algunos co-

mentarios pertinentes. Para los procedimientos detallados, consulte los manuales que complementan este libro.

STATDISK genera histogramas con facilidad. Introduzca los datos en la ventana de datos de STATDISK, haga clic en **Data**, luego en **Histogram** y luego en el botón **Plot**. (Si usted prefiere determinar su propia anchura de clase y punto de partida, haga clic en el botón “User defined” antes de seleccionar Plot).

MINITAB genera histogramas con facilidad. Introduzca los datos en una columna, luego haga clic en **Graph** y luego en **Histogram**. Seleccione el histograma “Simple”. Introduzca la columna en la ventana “Graph variables” y seleccione **OK**.

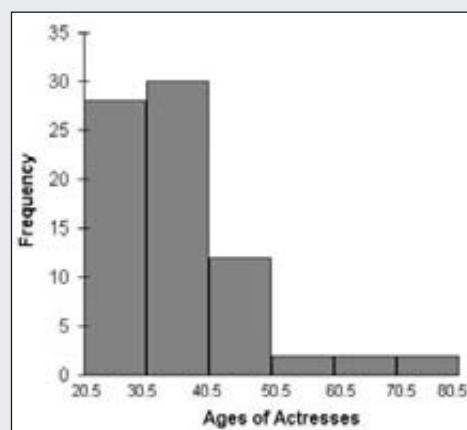
TI-83/84 PLUS Introduzca una lista de datos en L1. Seleccione la función **STAT PLOT** presionando **[2nd] [Y=]**. Presione **[ENTER]** y utilice las teclas de las flechas para poner P1 en la posición de en-

cendido y también elija la gráfica con barras. Presione [ZOOM] [9] para obtener un histograma con los valores predeterminados. (También puede utilizar su propia anchura de clase y fronteras de clase. Consulte el manual de TI-83/84 Plus que complementa este libro).

EXCEL puede generar histogramas como el que se presenta aquí, pero se vuelve algo difícil. Para generarlos con facilidad utilice el complemento DDXL que viene incluido en el CD de este libro. Después de

instalar DDXL dentro de Excel, haga clic en **DDXL**, seleccione **Charts and Plots** y haga clic en “function type” de **Histogram**. Haga clic en el ícono del lápiz e introduzca el intervalo de celdas que contienen los datos, como A1:A500 para 500 valores en los renglones 1 a 500 de la columna A.

Excel



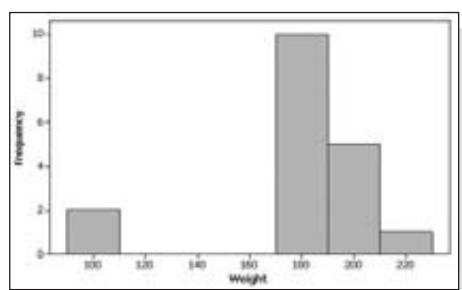
2-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Histograma.** ¿Qué características importantes de los datos se entenderían mejor al examinar un histograma?
- 2. Histograma y distribución de frecuencias.** Puesto que un histograma es básicamente una representación gráfica de los mismos datos de una distribución de frecuencias, ¿qué ventaja importante tiene un histograma sobre una distribución de frecuencias?
- 3. Conjunto de datos pequeño.** Si un conjunto de datos es pequeño, por ejemplo de sólo cinco valores, ¿por qué no debemos tomarnos la molestia de construir un histograma?
- 4. Distribución normal.** Despues de examinar un histograma, ¿qué criterio se puede utilizar para determinar si los datos tienen una distribución aproximadamente normal? ¿Este criterio es totalmente objetivo o implica un juicio subjetivo?

En los ejercicios 5 a 8, responda las preguntas con respecto al histograma que se generó con Minitab. El histograma representa los pesos (en libras) de los timoneles y los remeros de una carrera de embarcaciones entre Oxford y Cambridge. (Con base en datos de A Handbook of Small Data Sets, por D. J. Hand, Chapman y Hall).

Histograma de Minitab

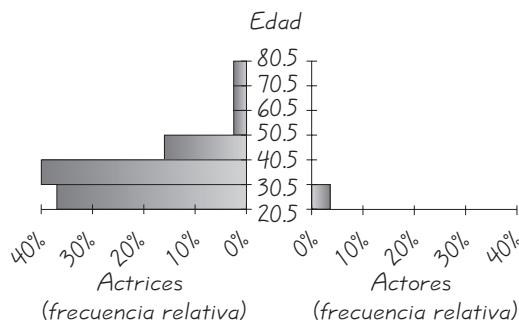


5. **Tamaño de la muestra.** ¿Cuántos miembros de la tripulación están incluidos en el histograma?
6. **Variación.** ¿Cuál es el peso mínimo posible? ¿Cuál es el peso máximo posible?
7. **Hueco.** ¿Cuál sería una explicación razonable para el gran hueco que hay entre la barra del extremo izquierdo y las otras barras?
8. **Anchura de clase.** ¿Cuál es la anchura de clase?
9. **Ánalisis del último dígito.** Remítase al ejercicio 17 de la sección 2-2, sobre los últimos dígitos de las estaturas de los estudiantes de estadística que se obtuvieron para un experimento de la clase. Utilice la distribución de frecuencias del ejercicio para construir un histograma. ¿Qué se puede concluir a partir de la distribución de los dígitos? En específico, ¿parece que las estaturas se reportaron o que realmente se midieron?
10. **Dado cargado.** Remítase al ejercicio 18 de la sección 2-2 sobre los resultados de 180 lanzamientos de un dado que el autor cargó. Utilice la distribución de frecuencias para construir el histograma correspondiente. ¿Cuál debería de ser la apariencia del histograma si el dado fuera perfectamente legal? ¿El histograma de la distribución de frecuencias dada parece diferir significativamente de un histograma obtenido de un dado legal?
11. **Cantidades de lluvia.** Remítase al ejercicio 19 de la sección 2-2 y utilice la distribución de frecuencias para construir un histograma. ¿Al parecer los datos tienen una distribución aproximadamente normal?
12. **Nicotina en cigarrillos.** Remítase al ejercicio 20 de la sección 2-2 y utilice la distribución de frecuencias para construir un histograma. ¿Al parecer los datos tienen una distribución aproximadamente normal?
13. **Valores del IMC.** Remítase el ejercicio 21 de la sección 2-2 y utilice la distribución de frecuencias para construir un histograma. ¿Al parecer los datos tienen una distribución aproximadamente normal?
14. **Datos del clima.** Remítase al ejercicio 22 de la sección 2-2 y utilice la distribución de frecuencias de las temperaturas mínimas reales para construir un histograma. ¿Al parecer los datos tienen una distribución aproximadamente normal?
15. **Pesos de monedas de un centavo.** Remítase al ejercicio 23 de la sección 2-2 y utilice la distribución de frecuencias con los pesos de las monedas de un centavo acuñadas antes de 1983. Construya el histograma correspondiente. ¿Los pesos parecen tener una distribución normal?
16. **Coca-Cola regular y Coca-Cola dietética.** Remítase al ejercicio 24 de la sección 2-2 y utilice las dos distribuciones de frecuencias relativas para construir los dos histogramas de frecuencias relativas correspondientes. Compare los resultados y determine si parece haber una diferencia significativa. Si hay una diferencia, ¿cómo se explicaría?
17. **Comparación de edades de actores y actrices.** Remítase la tabla 2-8 y utilice la distribución de frecuencias relativas de los mejores actores para construir un histograma de frecuencias relativas. Compare los resultados con la figura 2-3, que incluye el histograma de frecuencias relativas para las mejores actrices. ¿Al parecer los dos géneros ganan el premio Óscar a diferentes edades? (Véase también el ejercicio 18 de esta sección).

2-3 MÁS ALLÁ DE LO BÁSICO

18. **Histogramas de frecuencias relativas contiguos.** Cuando se utilizan histogramas para comparar dos conjuntos de datos, en ocasiones es difícil hacer comparaciones al cambiar

la vista de un histograma a otro. Un *histograma de frecuencias relativas contiguo* emplea un formato que facilita las comparaciones. En vez de las frecuencias, deberíamos utilizar las frecuencias relativas para que las comparaciones no se vean distorsionadas por muestras de tamaños diferentes. Complete los histogramas de frecuencias relativas contiguos que se indican a continuación, utilizando los datos de la tabla 2-8 de la sección 2-2. Luego utilice los resultados para comparar los dos conjuntos de datos.



19. **Conjuntos grandes de datos.** Remítase al ejercicio 25 de la sección 2-2 y construya histogramas de frecuencias relativas contiguos para las cargas axiales de las latas de 0.01109 pulgadas de grosor y de las cargas axiales de las latas de 0.0111 pulgadas de grosor. (Los histogramas de frecuencias relativas contiguos se describen en el ejercicio 18). Compare los dos conjuntos de datos. ¿El grosor de las latas de aluminio afecta su resistencia, la cual se midió usando las cargas axiales?
20. **Interpretación de los efectos de los valores extremos.** Remítase al conjunto de datos 15 del apéndice B sobre las cargas axiales de las latas de aluminio de 0.0111 pulgadas de grosor. La carga de 504 lb es un *valor extremo* porque se encuentra muy alejado de los demás valores. Construya un histograma que incluya el valor de 504 lb, y luego construya otro histograma sin este valor. En ambos casos, inicie la primera clase en 200 lb y utilice una anchura de clase de 20 lb. Interprete los resultados planteando una generalización sobre el efecto que tendría un valor extremo en un histograma. (Véase el ejercicio 26 en la sección 2-2).

2-4 Gráficas estadísticas

Concepto clave En la sección 2-3 nos referimos a los histogramas y a los histogramas de frecuencias relativas como gráficas que muestran visualmente las distribuciones de conjuntos de datos. En esta sección presentamos otras gráficas de uso común en los análisis estadísticos, así como también gráficas que describen datos de formas innovadoras. Al igual que en la sección 2-3, el principal objetivo no es la creación de una gráfica, sino *entender* mejor un conjunto de datos usando de una gráfica adecuada, que revele de manera efectiva algunas características importantes. Nuestro mundo requiere de más personas con habilidades para construir gráficas que revelen de manera clara y efectiva las características importantes de los datos. Nuestro mundo también necesita gente con la capacidad de crear gráficas originales e innovadoras que muestren aspectos fundamentales de los datos.

Esta sección comienza con una descripción breve de las gráficas que generalmente se incluyen en los cursos de introducción a la estadística, como polígonos de frecuencias, ojivas, gráficas de puntos, gráficas de tallo y hojas, gráficas de

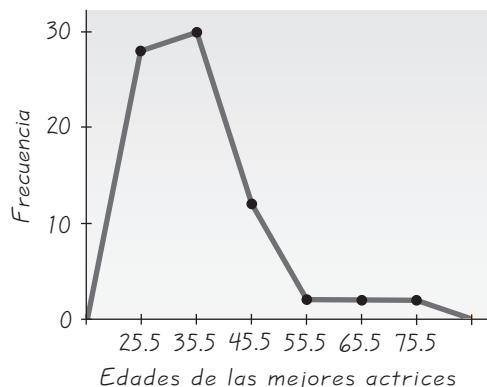


Figura 2-4 Polígono de frecuencias

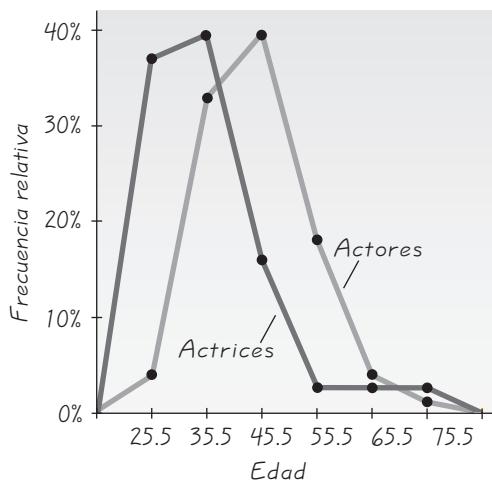


Figura 2-5 Polígonos de frecuencias relativas

Pareto, gráficas circulares, diagramas de dispersión y gráficas de series de tiempo. Luego, estudiaremos algunas gráficas originales y creativas; ahora iniciaremos con los polígonos de frecuencias.

Polígono de frecuencias

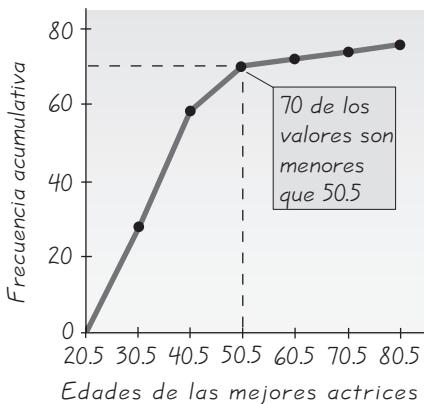
Un **polígono de frecuencias** utiliza segmentos lineales conectados a puntos que se localizan directamente por encima de los valores de las marcas de clase. Observe la figura 2-4, que presenta el polígono de frecuencias correspondiente a la tabla 2-2. Las alturas de los puntos corresponden a las frecuencias de clase; en tanto que los segmentos lineales se extienden hacia la derecha y hacia la izquierda, de manera que la gráfica inicia y termina sobre el eje horizontal.

Una variante del polígono de frecuencias básico es el polígono de frecuencias relativas, que coloca a estas últimas en la escala vertical. Al tratar de comparar dos conjuntos de datos, a menudo es muy útil graficar dos polígonos de frecuencias relativas sobre los mismos ejes. Consulte la figura 2-5, que ilustra polígonos de frecuencias relativas para las edades de las mejores actrices y los mejores actores, listadas en el Problema del capítulo. La figura 2-5 aclara visualmente que las actrices tienden a ser más jóvenes que sus contrapartes masculinos. La figura 2-5 logra algo que es realmente maravilloso: permite una comprensión de los datos que no sería posible mediante el examen visual de las listas de datos de la tabla 2-1. (Es como un buen maestro de poesía que revela el significado real de un poema). Por razones que no se explicarán aquí, parece que hay algún tipo de discriminación por género basada en la edad.

Ojiva

Una **ojiva** es una gráfica lineal que representa frecuencias *acumulativas*, de la misma forma que la distribución de frecuencias acumulativas (véase la tabla 2-4 de la sección anterior) es una lista de éstas. La figura 2-6 es la ojiva correspondiente a la tabla 2-4. Observe que la ojiva utiliza fronteras de clase a lo largo de la escala horizontal, y que la gráfica comienza con la frontera inferior de la primera clase y termina con la frontera superior de la última clase. Las ojivas son útiles

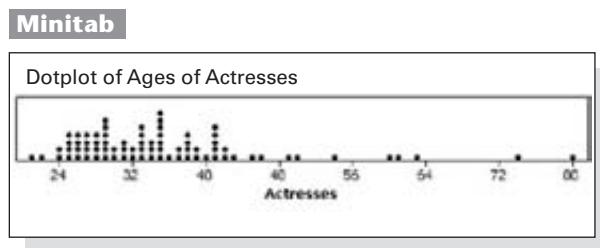
Figura 2-6
Ojiva



para determinar el número de valores que se encuentran por debajo de un valor específico. Por ejemplo, vea la figura 2-6, donde se muestra que 70 de las edades son menores que 50.5.

Gráficas de puntos

Una **gráfica de puntos** es aquella donde se marca cada valor de un dato como un punto a lo largo de una escala de valores. Los puntos que representan valores iguales se apilan. Observe la gráfica de puntos generada con Minitab que presenta las edades de las mejores actrices. (Los datos son de la tabla 2-1 del Problema del capítulo). Los dos puntos que parecen a la izquierda representan las edades de 21 y 22 años. Los siguientes dos puntos se apilan arriba de 24, indicando que dos de las actrices tenían 24 años de edad cuando recibieron el Óscar. En esta gráfica de puntos vemos que las edades por arriba de 48 años son escasas y que están algo separadas.



Gráficas de tallo y hojas

Una **gráfica de tallo y hojas** representa datos que separan cada valor en dos partes: el tallo (el dígito ubicado en el extremo izquierdo) y la hoja (el dígito del extremo derecho). La siguiente ilustración muestra una gráfica de tallo y hojas con las edades de las mejores actrices que aparecen en la tabla 2-1 del problema del capítulo. Dichas edades, ordenadas de forma creciente, son 21, 22, 24, 24, . . . , 80. Es fácil ver como el primer valor de 21 se separó en su tallo de 2 y su hoja de 1. Cada uno de los valores restantes se separa de manera similar. Observe que las hojas se ordenaron de forma creciente y no en el orden en que aparecen en la lista original.

Si colocamos la gráfica de lado, veremos una distribución de esos datos. Una gran ventaja de la gráfica de tallo y hojas radica en que nos permite ver la

Gráfica de tallo y hojas

Tallo (decenas)	Hoja (unidades)
2	12445555666677778888999999
3	001112233333444555555677888899
4	011111223569
5	04 ←Los valores son 50 y 54
6	013
7	4
8	0 ←El valor es 80.

distribución de los datos y, al mismo tiempo, retener toda la información de la lista original. En caso de ser necesario, podríamos reconstruir la lista de valores original. Otra ventaja es que la construcción de una gráfica de tallo y hojas implica una forma fácil y rápida de *ordenar* datos (acomodarlos en orden), y algunos procedimientos estadísticos requieren de un ordenamiento (como el cálculo de la mediana o de los percentiles).

Los renglones de los dígitos en una gráfica de tallo y hojas son similares en naturaleza a las barras del histograma. Uno de los lineamientos para la construcción de histogramas es que se incluyan entre 5 y 20 clases, lo cual se aplica a la gráfica de tallo y hojas por las mismas razones. Por lo general, obtenemos mejores gráficas de tallo y hojas al redondear primero los valores de los datos originales. Además, este tipo de gráficas pueden *expandirse* para incluir más renglones y *condensarse* para disminuir el número de renglones. Véase el ejercicio 26.



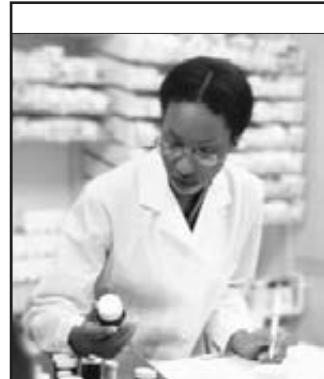
Gráficas de Pareto

La Federal Communications Commission verifica la calidad del servicio telefónico en Estados Unidos. Algunas de las críticas en contra de las compañías telefónicas incluyen los cierres, que implica cambiar de compañía al cliente sin el conocimiento de éste; y los cargos no autorizados. Datos recientes de la FCC mostraron que las quejas en contra de las compañías telefónicas estadounidenses eran las siguientes: 4473 por tarifas y servicios, 1007 por marketing, 766 por llamadas internacionales, 614 por cargos de acceso, 534 por servicios de operadora, 12478 por cierres y 1214 por cargos no autorizados. Si usted fuese reportero de un medio impreso, ¿cómo presentaría dicha información? La simple escritura de oraciones con datos numéricos quizás no lleve a una verdadera comprensión. Una mejor técnica consiste en utilizar una gráfica conveniente, y la gráfica de Pareto sería la adecuada.

Una **gráfica de Pareto** es una gráfica de barras para datos cualitativos, donde las barras se ordenan de acuerdo con las frecuencias. Las escalas verticales de las gráficas de Pareto representan tanto frecuencias como frecuencias relativas. La barra más alta se coloca a la izquierda y las más pequeñas a la derecha. Al ordenar las barras por frecuencias, esta gráfica enfoca la atención en las categorías más importantes. La figura 2-7 es una gráfica de Pareto que muestra con claridad que el cierre es, por mucho, el asunto más grave de las quejas de los clientes respecto a las empresas telefónicas.

Gráficas circulares

Las gráficas circulares también se utilizan para visualizar datos cualitativos. La figura 2-8 es un ejemplo de una **gráfica circular**, que presenta datos cualitativos



El poder de una gráfica

Con ventas anuales cercanas a los \$10,000 millones, y con alrededor de 50 millones de usuarios, el fármaco Lipitor de Pfizer se ha convertido en el medicamento de prescripción más redituable y más utilizado de la historia. Al inicio de su desarrollo, Lipitor se comparó con otros fármacos (Zocor, Mevacor, Lescol y Pravachol), en un proceso que implicó ensayos controlados. El resumen del informe incluyó una gráfica que mostraba una curva del Lipitor con un incremento más pronunciado que las curvas de los otros medicamentos, lo cual demostraba visualmente que Lipitor era más efectivo para reducir el colesterol que los otros fármacos. Pat Kelly, que en ese entonces era un alto ejecutivo de marketing de Pfizer, declaró: "Nunca olvidaré cuando vi esa gráfica [...] En ese momento pensé '¡caray!' Ahora sé de qué se trata. ¡Podemos comunicar esto!" La Food and Drug Administration de Estados Unidos aprobó el Lipitor y permitió a Pfizer incluir la gráfica con cada prescripción. El personal de ventas de la empresa también distribuyó la gráfica entre los médicos.

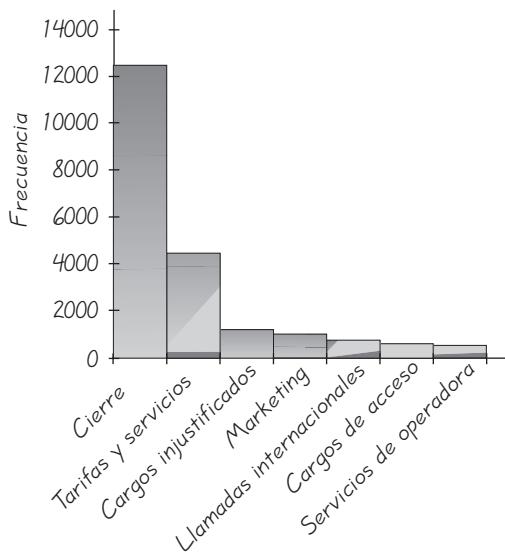


Figura 2-7 Gráfica de Pareto de quejas en contra de las compañías telefónicas

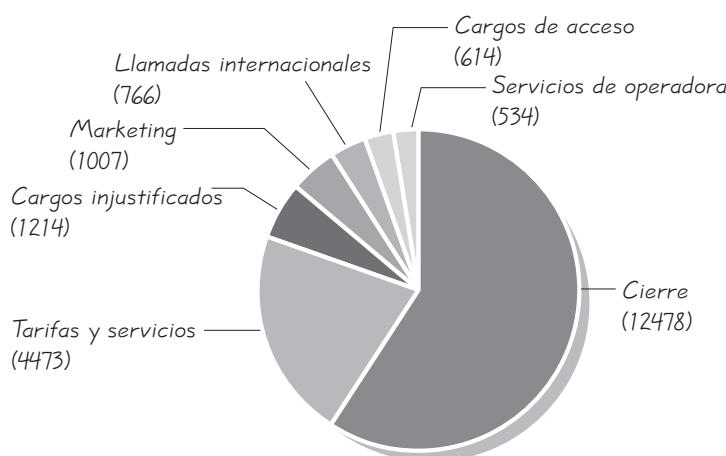


Figura 2-8 Gráfica circular de quejas en contra de las compañías telefónicas

como si fueran rebanadas de un pastel. La figura 2-8 representa los mismos datos de la figura 2-7. Para construir una gráfica circular, se divide el círculo en las proporciones adecuadas. La categoría de quejas por cierre representa un 59% del total, de tal modo que la porción que representa el cierre debería abarcar el 59% del total (con un ángulo central de $0.59 \times 360^\circ = 212^\circ$).

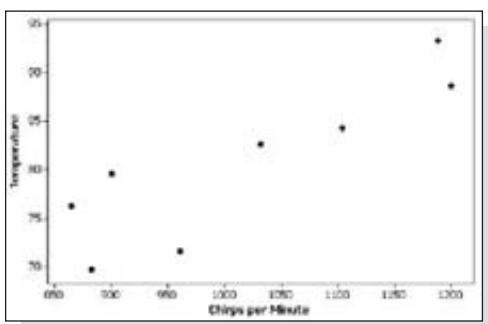
La gráfica de Pareto (figura 2-7) y la gráfica circular (figura 2-8) presentan los mismos datos en formas diferentes, pero una comparación probablemente demuestre que la gráfica de Pareto es mejor para resaltar los tamaños relativos de los distintos componentes. Esto explica por qué muchas compañías, como Boeing Aircraft, utilizan las gráficas de Pareto con gran frecuencia.

Diagramas de dispersión

Un **diagrama de dispersión** es una gráfica de datos apareados (x, y), con un eje x horizontal y un eje y vertical. Los datos se aparean de tal forma que cada valor de un conjunto de datos corresponde a un valor de un segundo conjunto de datos. Para elaborar manualmente un diagrama de dispersión, construya un eje horizontal para los valores de la primera variable, construya un eje vertical para los valores de la segunda variable y después grafique los puntos. El patrón de los puntos graficados suele ser útil para determinar si existe alguna relación entre las dos variables. (Este aspecto se estudia a profundidad en el tema de la correlación, en la sección 10-2).

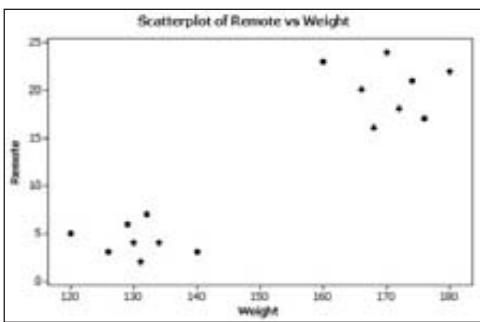
Uno de los usos clásicos que se le dio al diagrama de dispersión es en el cálculo del número de sonidos que emite un grillo por minuto, en relación con la temperatura ($^{\circ}\text{F}$). Utilizando los datos de *The Song of Insects* por George W. Pierce, de Harvard University Press, Minitab produjo el diagrama de dispersión que aquí se presenta. Al parecer existe una relación entre tales sonidos y la temperatura, como lo indica el patrón de los puntos. Por lo tanto, es posible usar a los grillos como termómetros.

Diagrama de dispersión de Minitab



EJEMPLO Grupos Considere el diagrama de dispersión de datos apareados obtenidos de 16 sujetos. Se mide el peso (en libras) de cada sujeto, y también se registra el número de veces que éste utiliza el control remoto del televisor durante una hora. Se utiliza Minitab para generar el diagrama de dispersión de los datos apareados peso/control remoto, que se muestra a continuación. Este diagrama en particular revela dos grupos muy distintos, que pueden explicarse por la inclusión de dos poblaciones diferentes: mujeres (con pesos más bajos y menor uso del control remoto) y hombres (con pesos más altos y mayor uso del control remoto). Si hubiéramos ignorado la presencia de los grupos, pensaríamos de manera incorrecta que hay una relación entre el peso y el uso del control remoto. Sin embargo, observe los dos grupos de manera separada y es mucho más evidente que *no* parece haber una relación entre peso y el uso del control remoto.

Minitab



Gráficas de series de tiempo

Una **gráfica de series de tiempo** incluye *datos de series de tiempo*, los cuales se reúnen en diferentes momentos. Por ejemplo, la gráfica de series de tiempo generada con el programa SPSS, que se presenta abajo, indica el número de pantallas de autocinemas existentes en un periodo reciente de 17 años (basado en datos de la National Association of Theater Owners). Vemos que para este periodo existe una clara tendencia de valores decrecientes. Lo que alguna vez fue parte importante de Estados Unidos, en especial para el autor, está en decadencia. Afortunadamente, la tasa de disminución parece ser menor que a finales de la década de 1980. Con frecuencia es sumamente importante conocer los cambios en los valores de una población a través del tiempo. Muchas compañías cayeron en bancarrota porque no verificaban la calidad de sus bienes o servicios y, de manera incorrecta, creían que trataban con datos estables. No se dieron cuenta de que sus productos se volvían

muy defectuosos conforme cambiaban importantes características de la población. En el capítulo 14 se presentan las *gráficas de control*, que son herramientas eficaces para verificar datos de series de tiempo.



Se solicita diseñador de gráficas estadísticas

Hasta ahora esta sección ha incluido algunas de las gráficas estadísticas estándares importantes, que se utilizan comúnmente en los cursos introductorios a la estadística. Hay muchas otras gráficas, algunas de las cuales aún no se han creado, que sirven para describir datos importantes e interesantes. El mundo necesita desesperadamente más personas con la capacidad de ser creativas y originales para crear gráficas que revelen de manera efectiva la naturaleza de datos. En la actualidad, las gráficas que aparecen en periódicos, revistas y televisión suelen ser creadas por reporteros con un entrenamiento en periodismo o comunicación colectiva, aunque con escasa o ninguna capacitación en el trabajo formal con datos. Esperamos con idealismo, pero de forma realista, que algunos lectores de este libro reconozcan esa necesidad y que, al tener un interés en ese tema, estudien procedimientos para crear gráficas estadísticas. El autor recomienda mucho la lectura cuidadosa de *The Visual Display of Quantitative Information*, 2a. edición, por Edward Tufte (Graphics Press, PO Box 430, Cheshire, CT 06410). Los siguientes son algunos de los principios importantes sugeridos por Tufte:

- Para conjuntos pequeños de datos, de 20 valores o menos, utilice una tabla en vez de una gráfica.
- Una gráfica de datos debería lograr que el observador se enfoque en la verdadera naturaleza de los datos y no en otros elementos, como características de diseño atractivas pero distractoras.
- No distorsione los datos; construya una gráfica para revelar la verdadera naturaleza de los datos.
- Casi toda la tinta de una gráfica debe utilizarse para los datos y no para otros elementos de diseño.
- No utilice imágenes que contengan características como líneas diagonales, puntos o tramas sombreadas, porque crean la incómoda ilusión de movimiento.
- No emplee áreas de volúmenes para datos que en realidad tienen una naturaleza unidimensional. (Por ejemplo, no use dibujos de billetes para representar los presupuestos de diferentes años).

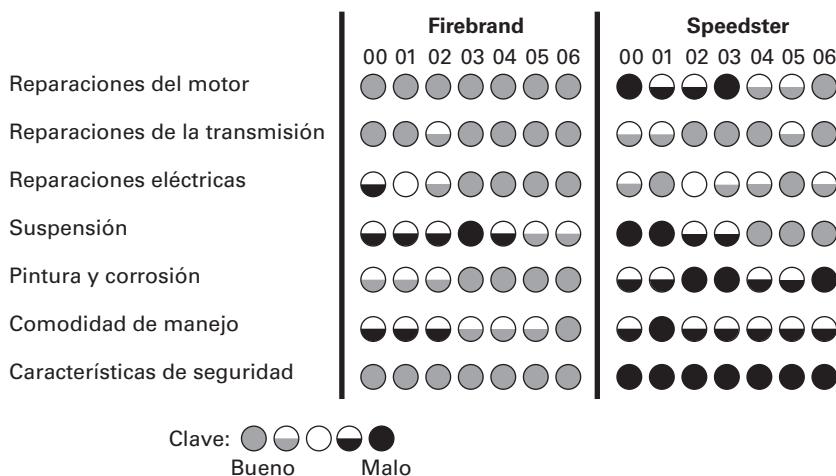


Figura 2-9 Datos sobre confiabilidad de los automóviles

- Nunca publique gráficas circulares porque desperdician tinta en componentes no relacionados con los datos y carecen de una escala apropiada.

La figura 2-9 presenta una comparación de dos automóviles diferentes, y está basada en gráficas de la revista *Consumer's Report*. Esta revista fundamenta sus gráficas en un gran número de encuestas realizadas con propietarios de automóviles. La figura 2-9 ejemplifica un trabajo sobresaliente por su originalidad, creatividad y eficacia al lograr que el lector observe datos complicados en un formato sencillo. La clave en la parte inferior indica que el gris se utilizó para resultados malos y el negro para resultados buenos, de manera que el esquema de color corresponde a una intensidad que puede tener sentido para los conductores. Con facilidad vemos que, durante los últimos años, el automóvil Firebrand ha sido mejor en general que el Speedster. Este tipo de información es valiosa para consumidores que están considerando la compra de un vehículo nuevo o uno seminuevo.

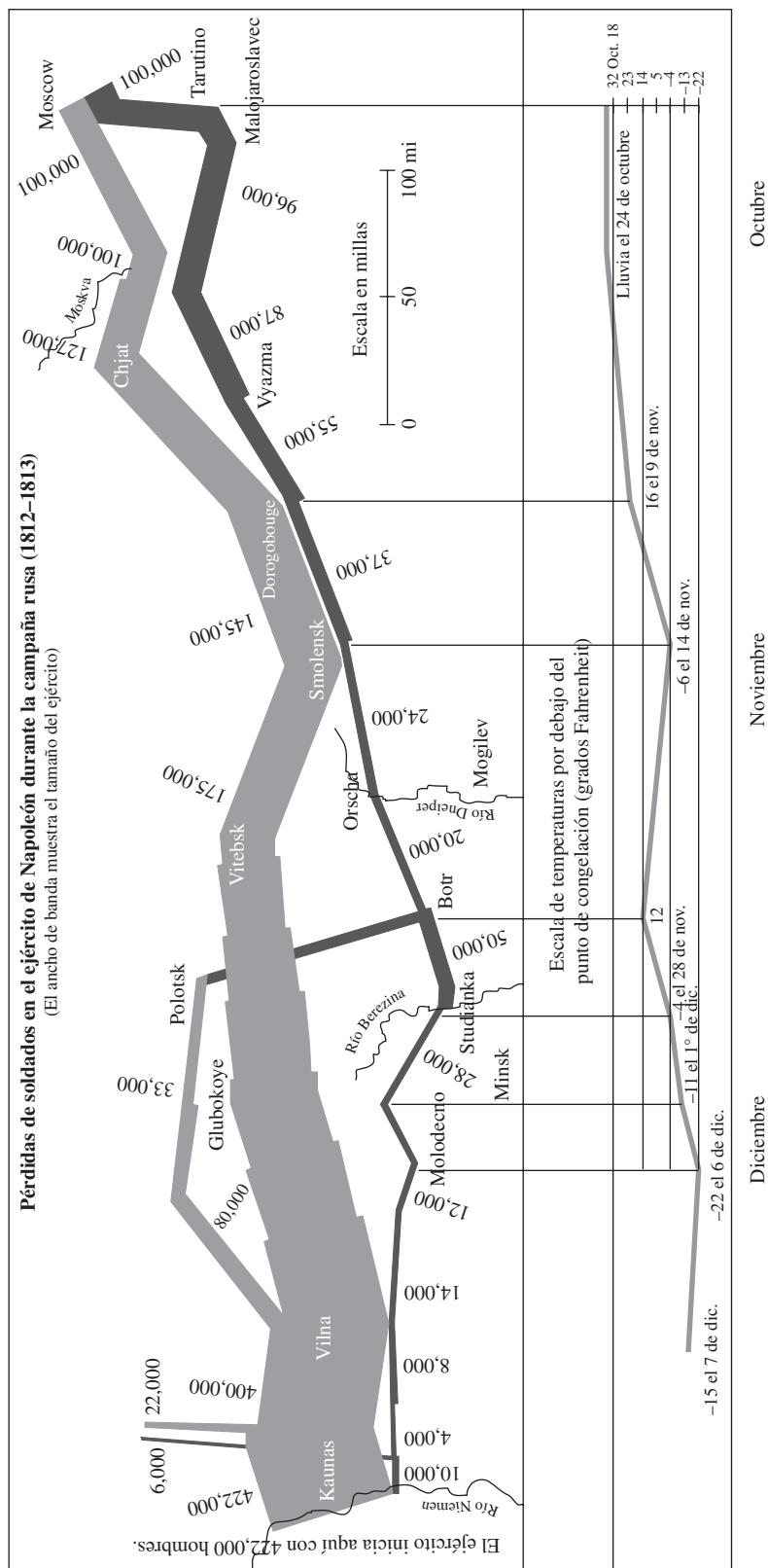
La figura que aparece en la siguiente página se ha descrito como “quizá la mejor gráfica estadística que se haya elaborado jamás”. Esta figura incluye seis variables diferentes respecto de la marcha del ejército de Napoleón hacia Moscú en 1812-1813. La banda gruesa a la izquierda representa el tamaño del ejército cuando, desde Polonia, inició la invasión a Rusia. La banda inferior muestra su tamaño durante la retirada, con las temperaturas y fechas correspondientes. Aunque Charles Joseph Minard la elaboró en 1861, esta gráfica es ingeniosa, aun desde la perspectiva actual.

Otra gráfica notable, de importancia histórica, es la que elaboró la enfermera más famosa del mundo, Florence Nightingale. Esta gráfica, que aparece en la figura 2-10, es particularmente interesante debido a que realmente salvó vidas cuando Nightingale la utilizó para convencer a los oficiales británicos de que los hospitales militares necesitaban mejorar sus condiciones sanitarias, sus tratamientos y su abastecimiento. Su dibujo asemeja a una gráfica circular, excepto que todos los ángulos centrales son iguales y se usan radios diferentes para mostrar los cambios en el número de muertes mensuales. Las regiones externas de la figura 2-10 representan las muertes por enfermedades que pudieron haberse prevenido; las regiones internas, muertes por heridas; y las regiones centrales, muertes por otras causas.

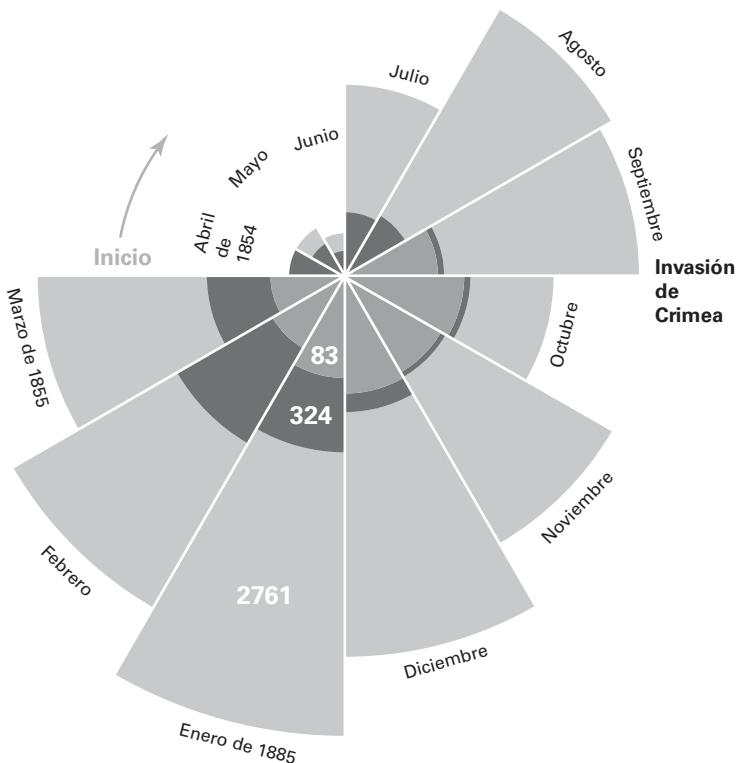


Florence Nightingale

Florence Nightingale (1820-1910) es conocida como fundadora de la profesión de enfermería, aunque también salvó miles de vidas con el uso de la estadística. Cuando encontraba un hospital insalubre y con desabasto, mejoraba tales condiciones y después utilizaba la estadística para convencer a otros de la necesidad de una reforma médica de mayor alcance. Ella diseñó gráficas originales para ilustrar que, durante la Guerra de Crimea, murieron más soldados como resultado de las condiciones insalubres que en combate. Florence Nightingale fue uno de los pioneros en el uso de la estadística social y de las técnicas gráficas.



Creditos: Edward R. Tufte, *The Visual Display of Quantitative Information* (Cheshire, CT: Graphics Press, 1983). Se reproduce con autorización.

**Figura 2-10****Muertes en los hospitales militares británicos durante la guerra de Crimea**

Región externa: Muertes por enfermedades que pudieron prevenirse.

Región central: Muertes que no fueron causadas por heridas ni por enfermedades prevenibles.

Región interna: Muertes por heridas en batalla.

Conclusión

La eficacia de la gráfica de Florence Nightingale ilustra muy bien el siguiente punto importante: Una gráfica no es, en sí misma, un resultado final; es una herramienta para describir, explorar y comparar datos, que consideramos como sigue:

Descripción de datos: En un histograma, por ejemplo, se toman en cuenta el centro, la variación, la distribución y los valores extremos (CVDVE, sin el último elemento del tiempo). ¿Cuál es el valor aproximado del centro de la distribución y cuál es el intervalo aproximado de valores? Considere la forma completa de la distribución. ¿Los valores están distribuidos uniformemente? ¿La distribución está sesgada (ladeada) hacia la izquierda o hacia la derecha? ¿La distribución tiene un pico a la mitad? ¿Hay una brecha grande que sugiere que los datos provendrían de poblaciones diferentes? Identifique cualquier valor extremo y cualquier otra característica notable.

Exploración de datos: Buscamos características de la gráfica que revelen rasgos interesantes y/o útiles del conjunto de datos. Por ejemplo, en la figura 2-10 observamos que morían más soldados por cuidados hospitalarios inadecuados, que por heridas en batalla.

Comparación de datos: Construya gráficas similares que faciliten la comparación de conjuntos de datos. Por ejemplo, si usted grafica un polígono de frecuencias con los pesos de hombres y otro polígono de frecuencias con pesos de mujeres, sobre el mismo conjunto de ejes, el polígono de los hombres debería aparecer a la derecha del polígono de mujeres, lo cual indica que los hombres tienen pesos mayores.

Uso de la tecnología

Ahora existen programas de cómputo poderosos que son bastante efectivos para generar gráficas impresionantes. Este libro hace referencia frecuente a STATDISK, Minitab, Excel y a la calculadora TI-83/84 Plus, por lo que listamos las gráficas (que ya comentamos en esta sección y en la anterior) que es posible elaborar. (Para información detallada sobre los procedimientos, véase los manuales que complementan este libro).

STATDISK genera histogramas y diagramas de dispersión.

MINITAB genera histogramas, polígonos de frecuencias, gráficas de puntos, gráficas de tallo y hojas, gráficas de Pareto, gráficas circulares, diagramas de dispersión y gráficas de series de tiempo.

EXCEL genera histogramas, polígonos de frecuencias, gráficas circulares y diagramas de dispersión.

TI-83/84 PLUS genera histogramas y diagramas de dispersión.

A continuación se muestra un diagrama de dispersión generado con la TI-83/84 Plus, que es similar al primer diagrama de dispersión de Minitab que mostramos en esta sección.

TI-83/84 Plus



2-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **¿Por qué usar gráficas?** ¿Cuál es el principal objetivo de los datos gráficos?
2. **Diagrama de dispersión.** ¿Qué tipo de datos se requieren para la construcción de un diagrama de dispersión, y qué revela este tipo de gráfica acerca de los datos?
3. **Gráfica de series de tiempo.** ¿Qué tipo de datos se requieren para la construcción de una gráfica de series de tiempo y qué revela este tipo de gráfica acerca de los datos?
4. **Gráfica circular contra gráfica de Pareto.** ¿Por qué en general es mejor utilizar una gráfica de Pareto en vez de una gráfica circular?

En los ejercicios 5 a 8, utilice las 35 temperaturas máximas reales del conjunto de datos 8 en el apéndice B.

5. **Gráfica de puntos.** Construya una gráfica de puntos de las temperaturas máximas reales. ¿Qué sugiere la gráfica de puntos acerca de la distribución de las temperaturas máximas?
6. **Gráfica de tallo y hojas.** Utilice las 35 temperaturas máximas reales para construir una gráfica de tallo y hojas. ¿Qué sugiere esta gráfica acerca de la distribución de las temperaturas?
7. **Polígono de frecuencias.** Utilice las 35 temperaturas máximas reales para construir un polígono de frecuencias. En el eje horizontal, use los valores de marca de clase obtenidos de estos intervalos de clase: 50-59, 60-69, 70-79, 80-89.
8. **Ojiva.** Utilice las 35 temperaturas máximas reales para construir una ojiva. En el eje horizontal, use las siguientes fronteras de clase: 49.5, 59.5, 69.5, 79.5, 89.5. ¿Durante cuántos días la temperatura máxima real fue menor que 80°F?

En los ejercicios 9 a 12, utilice las 40 alturas de las erupciones del géiser Old Faithful listadas en el conjunto de datos 11 en el apéndice B.

9. **Gráfica de tallo y hojas.** Utilice las alturas para construir una gráfica de tallo y hojas. ¿Qué sugiere esta gráfica acerca de la distribución de las alturas?
10. **Gráfica de puntos.** Construya una gráfica de puntos con las alturas. ¿Qué sugiere esta gráfica acerca de la distribución de las alturas?

- 11. Ojiva.** Utilice las alturas para construir una ojiva. En el eje horizontal, use las siguientes fronteras de clase: 89.5, 99.5, 109.5, 119.5, 129.5, 139.5, 149.5, 159.5. ¿Cuántas erupciones tuvieron una altura menor de 120 pies?
- 12. Polígono de frecuencias.** Utilice las alturas para construir un polígono de frecuencias. En el eje horizontal, use las marcas de clase obtenidas de los siguientes intervalos de clase: 90-99, 100-109, 110-119, 120-129, 130-139, 140-149, 150-159.
- 13. Empleos.** Se realizó un estudio para determinar la cantidad de personas que obtienen un empleo. La siguiente tabla incluye los datos de 400 sujetos seleccionados al azar. Los datos se basan en resultados del National Center for Career Strategies. Construya una gráfica de Pareto que corresponda a los datos. Si alguien quiere conseguir un empleo, ¿cuál parece ser la técnica más efectiva?

Fuentes de empleo de los sujetos encuestados	Frecuencia
Anuncios clasificados	56
Empresas de búsqueda de ejecutivos	44
Contactos profesionales	280
Correo masivo	20

- 14. Empleos.** Remítase a los datos del ejercicio 13 y construya una gráfica circular. Compare esta gráfica con la gráfica de Pareto. ¿Puede usted determinar qué gráfica es más efectiva para mostrar la importancia relativa de las fuentes de empleo?
- 15. Lesiones laborales mortales.** En un año reciente, 5524 personas murieron mientras trabajaban. Las causas fueron las siguientes: transporte (2375), contacto con objetos o equipo (884), asaltos o actos violentos (829), caídas (718), exposición a sustancias dañinas o a un ambiente nocivo (552), incendios y explosiones (166). (Los datos son del Bureau of Labor Statistics). Construya una gráfica circular que represente esos datos.
- 16. Lesiones laborales mortales.** Remítase a los datos del ejercicio 15 y construya una gráfica de Pareto. Compare esta gráfica con la gráfica circular. ¿Qué gráfica es más efectiva para mostrar la importancia relativa de las causas de muerte en el trabajo?

En los ejercicios 17 y 18, utilice los datos apareados del apéndice B para construir un diagrama de dispersión.

- 17. Alquitrán/CO en cigarrillos.** Para el conjunto de datos 3, utilice el alquitrán en la escala horizontal y el monóxido de carbono (CO) en la escala vertical. Determine si hay una relación entre el alquitrán y el CO de los cigarrillos. Si es así, describa la relación.
- 18. Consumo de energía y temperatura.** Para el conjunto de datos 9, utilice las 10 temperaturas diarias promedio y las correspondientes 10 cantidades de consumo de energía (kWh). (Emplee las temperaturas en la escala horizontal). Con base en el resultado, ¿hay una relación entre las temperaturas diarias promedio y las cantidades de energía consumida? Trate de identificar al menos una razón por la que existe (o no) una relación.

En los ejercicios 19 y 20, utilice los datos que se le dan para construir una gráfica de series de tiempo.

- 19. Fallas de aterrizaje.** A continuación se muestra el número de fallas en el aterrizaje de aeronaves, listadas en orden por año, comenzando en 1990 (según datos de la Federal Aviation Administration). ¿Existe alguna tendencia? Si es así, ¿cuál?

281 242 219 186 200 240 275 292 325 321 421

- 20. Salas de cine techadas.** A continuación se muestra el número de salas de cine techadas, listadas en orden por año, comenzando en 1987 (según datos de la National Association of Theater Owners). ¿Cuál es la tendencia? ¿En qué difiere esta tendencia de la tendencia de los autocinemas? (En esta sección se presenta una gráfica de series de tiempo de los autocinemas).

20,595	21,632	21,907	22,904	23,740	24,344	24,789	25,830	26,995
28,905	31,050	33,418	36,448	35,567	34,490	35,170	35,361	

En los ejercicios 21 a 24, remítase a la figura de esta sección que describe la ruta de ida y vuelta que utilizó Napoleón en 1812, durante su campaña en Moscú (vea la página 64). La banda ancha a la izquierda describe el tamaño del ejército cuando inició su invasión a Rusia desde Polonia; en tanto que la banda inferior describe su retirada.

- 21.** Se indica que el número de hombres que iniciaron la campaña era de 422,000. Calcule el número y el porcentaje de hombres que sobrevivieron toda la campaña.
- 22.** Calcule el número y el porcentaje de hombres que murieron al cruzar el río Berezina.
- 23.** De los 320,000 hombres que marcharon de Vilna a Moscú, ¿cuántos llegaron a Moscú? ¿Aproximadamente qué distancia recorrieron desde Vilna hasta Moscú?
- 24.** ¿Cuál fue la temperatura más fría experimentada por cualquiera de los hombres, y cuándo se llegó a la temperatura más baja?

2-4 MÁS ALLÁ DE LO BÁSICO

- 25. Gráficas de tallo y hojas contiguas.** Remítase a las edades de las mejores actrices y de los mejores actores, que se muestran en la tabla 2-1 del problema del capítulo. Al margen encontrará un formato para las *gráficas de tallo y hojas contiguas*, donde aparecen las primeras dos edades de cada grupo. Complete los datos y luego compare los resultados.

- 26. Gráficas de tallo y hojas expandidas y condensadas.** En esta sección se incluyó una gráfica de tallo y hojas con las edades de las mejores actrices (tabla 2-1). Remítase a esa gráfica para lo siguiente:

- a. La gráfica de tallo y hojas se puede expandir subdividiendo los renglones en hojas que tengan dígitos del 0 al 4 y dígitos del 5 al 9. A continuación se muestran los primeros dos renglones de la gráfica después de su expansión. Complete los siguientes dos renglones de la gráfica de tallo y hojas expandida.

Tallo	Hojas	
2	1244	←Para las hojas del 0 al 4.
2	5555666677778888999999	←Para las hojas del 5 al 9.

- b. La gráfica de tallo y hojas se puede condensar combinando renglones adyacentes. A continuación se muestra el primer renglón de la gráfica condensada. Observe que insertamos un asterisco para separar los dígitos de las hojas asociadas con los números de cada tallo. Cada renglón de la gráfica condensada debe incluir exactamente un asterisco, para que la forma de la gráfica reducida no se distorsione. Complete la gráfica de tallo y hojas condensada identificando los datos restantes.

Tallo	Hojas
2 - 3	12445555666677778888999999*0011223333444555555677888899

Repaso

En este capítulo estudiamos procedimientos para resumir y graficar datos. Cuando se investiga un conjunto de datos, suelen ser muy importantes las características del centro, la variación, la distribución, los valores extremos y el cambio de patrones con el paso del tiempo. Además, este capítulo incluye una variedad de herramientas para investigar la distribución de los datos. Después de estudiar este capítulo, usted deberá ser capaz de hacer lo siguiente:

- Resumir datos al construir una distribución de frecuencias o una distribución de frecuencias relativas (sección 2-2).
- Mostrar visualmente la naturaleza de la distribución al construir un histograma (sección 2-3) o un histograma de frecuencias relativas.
- Investigar características importantes del conjunto de datos al crear presentaciones visuales como el polígono de frecuencias, la gráfica de puntos, la gráfica de tallo y hojas, la gráfica de Pareto, la gráfica circular, el diagrama de dispersión (para datos apareados) o la gráfica de series de tiempo (sección 2-4).

Además de crear tablas de distribuciones de frecuencias y gráficas, usted debería ser capaz de *entender e interpretar* dichos resultados. Por ejemplo, en el problema del capítulo se incluye la tabla 2-1, con las edades de las mejores actrices y los mejores actores ganadores del premio Óscar. Es probable que el simple hecho de examinar las dos listas de edades no revele mucha información significativa; sin embargo, las distribuciones de frecuencias y las gráficas nos permiten ver si existe una diferencia importante. Observamos que las actrices suelen ser mucho más jóvenes que los actores. Es posible explorar aún más tal diferencia considerando factores culturales relevantes; pero los métodos de la estadística nos ofrecen un excelente punto de partida al indicarnos la dirección correcta.

Conocimientos estadísticos y pensamiento crítico

1. **Exploración de datos.** Cuando investigamos la distribución de un conjunto de datos, ¿cuál es más eficaz: la distribución de frecuencias o el histograma? ¿Por qué?
2. **Comparación de datos.** Cuando comparamos dos conjuntos de datos, ¿cuál es mejor: la distribución de frecuencias o la distribución de frecuencias relativas? ¿Por qué?
3. **Bienes raíces.** Un agente de bienes raíces investiga los precios de venta que han tenido las casas de su región durante los últimos 50 años. ¿Qué gráfica sería mejor: un histograma a una gráfica de series de tiempo? ¿Por qué?
4. **Distribución normal.** Un histograma se construye a partir de un conjunto de valores muestrales. ¿Cuáles son dos características básicas del histograma que podrían sugerirnos que los datos tienen una distribución normal?

Ejercicios de repaso

1. **Distribución de frecuencias con las edades de los mejores actores.** Construya una distribución de frecuencias con las edades de los actores ganadores del Óscar que se incluyen en la tabla 2-1. Utilice los mismos intervalos de clase que utilizamos para la distribución de frecuencias de las actrices ganadoras, como se indica en la tabla 2-2. ¿Qué diferencias existen con la distribución de frecuencias de las actrices?
2. **Histograma de las edades de los mejores actores.** Construya el histograma que corresponde a la distribución de frecuencias del ejercicio 1. ¿Qué diferencias hay con el histograma de las actrices (figura 2-2)?

- 3. Gráfica de puntos con las edades de los mejores actores.** Construya una gráfica de puntos con las edades de los actores ganadores del Óscar, a partir de la tabla 2-1. ¿Qué diferencias existen con la gráfica de puntos para las edades de las actrices? La gráfica de puntos con las edades de las mejores actrices está incluida en la sección 2-4 (vea la página 58).
- 4. Gráfica de tallo y hojas con las edades de los mejores actores.** Construya una gráfica de tallo y hojas con las edades de los actores ganadores del Óscar, a partir de la tabla 2-1. ¿Qué diferencias existen con la gráfica de tallo y hojas para las edades de las actrices? La gráfica de tallo y hojas con las edades de las mejores actrices está incluida en la sección 2-4 (véase la página 59).
- 5. Diagrama de dispersión con las edades de las actrices y de los actores.** Remítase a la tabla 2-1 y utilice únicamente las primeras 10 edades de las actrices y las primeras 10 edades de los actores. Construya un diagrama de dispersión. Con base en los resultados, ¿al parecer existe una asociación entre las edades de las actrices y las edades de los actores?
- 6. Gráfica de series de tiempo.** Remítase a la tabla 2-1 y utilice las edades de las actrices ganadoras del Óscar. Dichas edades están acomodadas en orden. Construya una gráfica de series de tiempo. ¿Se percibe alguna tendencia? ¿Las edades cambian sistemáticamente con el paso del tiempo?

Tabla del ejercicio 1-4

Resultado	Frecuencia
1–5	43
6–10	44
11–15	59
16–20	47
21–25	57
26–30	56
31–35	49
36 o 0 o 00	25

Ejercicios de repaso acumulativo

En los ejercicios 1 a 4, remítase a la distribución de frecuencias que aparece al margen, la cual resume los resultados de 380 giros de una ruleta en el hotel y casino Belaggio de Las Vegas. Las ruletas estadounidenses tienen 38 ranuras. Una de las ranuras tiene el 0, otra el 00 y las ranuras restantes están numeradas del 1 al 36.

1. Considere los números que se obtuvieron con los giros. ¿Estos números miden o cuentan algo?
2. ¿Cuál es el nivel de medición de los resultados?
3. Examine la distribución de los resultados en la tabla. Dado que la última clase resume los resultados de las tres ranuras, ¿su frecuencia de 25 es aproximadamente consistente con los resultados que se esperarían de una ruleta legal? En general, ¿las frecuencias sugieren que la ruleta está cargada o que no lo está?
4. Si un jugador se da cuenta de que los últimos 500 giros de una ruleta específica produjeron números con un promedio (media) de 5, ¿sería útil esa información para ganar?
5. **Encuesta de consumidores.** La Consumer Advocacy Union envía una encuesta por correo a 500 propietarios de automóviles elegidos al azar, y se reciben 185 respuestas. Un reactivo pregunta la cantidad de dinero que se gastó en la compra de los automóviles. Se construyen una distribución de frecuencias y un histograma a partir de tales cantidades. ¿Se pueden utilizar esos resultados para obtener conclusiones válidas sobre la población de todos los propietarios de automóviles?

Actividades de cooperación en equipo

- 1. Actividad en clase** Remítase a la figura 2-10, donde aparece la gráfica que construyó Florence Nightingale hace aproximadamente 150 años. Esa gráfica ilustra el número de soldados que murió por heridas de combate, enfermedades prevenibles y otras causas. No es muy fácil entender la figura 2-10. Construya una nueva gráfica que describa los mismos datos; pero de tal manera que facilite mucho su comprensión.
- 2. Actividad en clase** A continuación se muestran las edades de motociclistas en el momento en que resultaron mortalmente heridos en accidentes de tránsito (según datos del U. S. Department of Transportation). Si su objetivo es resaltar el peligro que representan las motocicletas para la gente joven, ¿cuál sería más efectivo: un histograma, una gráfica de Pareto, una gráfica circular, una gráfica de puntos, una gráfica de tallo y hojas, etcétera? Construya la gráfica que cumpla mejor el objetivo de resaltar los peligros de conducir motocicletas. ¿Es aceptable distorsionar los datos de manera intencional si se tiene como finalidad salvar la vida de los motociclistas?

17	38	27	14	18	34	16	42	28
24	40	20	23	31	37	21	30	25
17	28	33	25	23	19	51	18	29

- 3. Actividad fuera de clase** Cada equipo de tres o cuatro estudiantes debe construir una gráfica que sirva para responder esta pregunta: ¿Existen una diferencia entre los valores del índice de masa corporal (IMC) de hombres y mujeres? (Véase el conjunto de datos 1 en el apéndice B).

Proyecto tecnológico

Aunque las gráficas construidas manualmente tienen cierto encanto primitivo, a menudo se consideran inadecuadas para publicaciones y presentaciones. Las gráficas generadas por computadora son mucho mejores para tales propósitos. Utilice un paquete de cómputo estadístico, como STATDISK, Minitab o Excel para generar tres histogramas: **1.** un histograma de la frecuencia cardiaca de los hombres, que se incluyen en el conjunto de datos 1 del apéndice B; **2.** un histograma de la

frecuencia cardiaca de las mujeres, incluidos en el conjunto de datos 1 del apéndice B; **3.** un histograma de la lista combinada de los pulsos cardiacos de hombres y mujeres. Después de obtener copias impresas de los histogramas, compárelos. ¿Al parecer las frecuencias cardíacas de hombres y mujeres tienen características similares? (Posteriormente estudiaremos más procedimientos formales para realizar este tipo de comparaciones. Véase, por ejemplo, la sección 9-4).

De los datos a la decisión

Pensamiento crítico

Bondad de ajuste Un tema importante en la estadística es determinar si ciertos resultados se ajustan a una distribución en particular. Por ejemplo, podríamos lanzar un dado 60 veces para determinar si los resultados se ajustan a la distribución que se esperaría obtener con un dado legal (donde todos los resultados aparecerían aproximadamente el mismo número de veces). En la sección 11-2 se presenta un método formal para una prueba de la *bondad de ajuste*. Este proyecto incluye un procedimiento informal basado en una comparación subjetiva. Aquí consideraremos el importante tema de las muertes por accidentes automovilísticos. Las muertes en este tipo de accidentes son devastadoras para los familiares, y con frecuencia provocan demandas y cuantiosas indemnizaciones por parte de las aseguradoras. A continuación se muestran las edades de 100 conductores elegidos al azar, que murieron en accidentes automovilísticos. También se incluye una distribución de frecuencias de conductores con licencia según sus edades.

Edades (en años) de conductores muertos en accidentes automovilísticos

37	76	18	81	28	29	18	18	27	20
18	17	70	87	45	32	88	20	18	28
17	51	24	37	24	21	18	18	17	40
25	16	45	31	74	38	16	30	17	34
34	27	87	24	45	24	44	73	18	44
16	16	73	17	16	51	24	16	31	38
86	19	52	35	18	18	69	17	28	38
69	65	57	45	23	18	56	16	20	22
77	18	73	26	58	24	21	21	29	51
17	30	16	17	36	42	18	76	53	27

Análisis

Convierta la distribución de frecuencias en una distribución de frecuencias relativas, y luego elabore una distribución de frecuencias relativas para las edades de los conductores muertos en accidentes automovilísticos. Compare las dos distribuciones de frecuencias relativas. ¿Qué categorías de edad parecen tener proporciones mucho mayores de decesos de conductores con licencia? Si usted fuera el responsable de establecer las tarifas de seguros para automóviles, ¿a qué categorías de edad les asignaría las tarifas más elevadas? Construya una gráfica que sirva para identificar las categorías de edad que son más proclives a los accidentes automovilísticos mortales.

Edad	Conductores con licencia (en millones)
16–19	9.2
20–29	33.6
30–39	40.8
40–49	37.0
50–59	24.2
60–69	17.5
70–79	12.7
80–89	4.3



Proyecto de Internet

Datos en Internet

Internet brinda un enorme cantidad de información, de la cual gran parte se presenta en forma de datos en bruto (sin analizar) que se reunieron u observaron. Muchas páginas de Internet resumen este tipo de datos con el uso de los métodos gráficos estudiados en este capítulo. Por ejemplo, encontramos la siguiente información con tan sólo unos cuantos clics:

- Una gráfica de barras en el sitio de U.S. Bureau of Labor Statistics nos indica que la tasa de desempleo del 3% es la más baja entre los individuos con un título universitario, en comparación de los grupos con menor nivel académico.
- Una gráfica circular proporcionada por la National Collegiate Athletic Association

(NCAA) muestra que un estimado 89.67% de las ganancias de la NCAA en 2004-05 provienen de tarifas por venta de derechos de televisión y marketing.

El proyecto de Internet para este capítulo, el cual se encuentra en el sitio de Internet de *Estadística*, explora con más detalle representaciones gráficas de conjuntos de datos localizados en Internet. En el proceso, usted observará y reunirá conjuntos de datos sobre deportes, datos demográficos de poblaciones y finanzas, y también realizará sus propios análisis gráficos.

El sitio Web para este capítulo se encuentra en

La estadística en el trabajo

“Las aplicaciones de la estadística son herramientas que pueden ser útiles casi en cualquier área de trabajo”.



Bob Sehlinger

Editor, Menasha Ridge Press

Menasha, Ridge Press edita, entre otras publicaciones, la serie *Unofficial Guide* para John Wesley & Sons (Wiley, Inc), que utiliza ampliamente la estadística para investigar las experiencias que enfrentarían los viajeros, y para ayudarlos a tomar decisiones informadas y a disfrutar de unas maravillosas vacaciones.

¿Cómo utiliza la estadística en su trabajo y qué conceptos específicos de la estadística emplea?

Usamos la estadística en cada faceta del negocio: El análisis del valor esperado para pronósticos de ventas; el análisis de regresión para determinar los libros que debemos publicar en una serie, etcétera; no obstante, somos mejor conocidos por nuestra investigación en las áreas sobre las colas y los cálculos evolutivos.

Las metodologías de investigación que se usan en la serie *Unofficial Guide* nos están conduciendo a un modelo verdaderamente innovador sobre la creación de guías para viajeros. Nuestros diseños de investigación y el uso de la tecnología del campo de la investigación de operaciones han sido citados en el mundo académico y comentados en revistas durante bastante tiempo.

Estamos utilizando un enfoque de equipo revolucionario y ciencia de vanguardia, para brindar a los lectores información sumamente valiosa que no está disponible en otras series de viajes. Toda nuestra organización está dirigida por individuos con amplio entrenamiento y experiencia en diseño de investigación, así como en recolección y análisis de datos.

Desde la primera edición de *Unofficial Guide* hasta nuestra investigación en Walt Disney World, disminuir las filas de espera para nuestros lectores ha sido una de nuestras grandes prioridades. Hemos desarrollado y ofrecido a nuestros lectores planes de viajes con pruebas de campo que les permiten disfrutar de la mayor cantidad de atracciones posibles, con el menor tiempo de espera. Probamos nuestro método en el parque; el grupo que viajaba sin nuestros planes pasó en promedio tres horas y media más esperando en la cola, y disfrutó 37% menos atracciones, que quienes usaron nuestros planes de viaje.

Conforme añadimos atracciones a nuestra lista, el número de planes de viaje posibles crece con rapidez. Las 44 atracciones del plan de viaje de un día en

el Magic Kingdom para adultos representan 51,090,942,171,709,440,000 planes de viaje posibles. ¿Qué tan buenos son nuestros nuevos planes de viaje en la *Unofficial Guide*? Nuestro programa de cómputo generalmente logra ubicarse dentro del 2% del plan de viaje óptimo. Para poner esto en perspectiva, si el plan hipotético “perfecto” de viaje de un día para un adulto requiere de aproximadamente 10 horas, el plan de viaje no oficial tomaría alrededor de 10 horas y 12 minutos. Dado que en una computadora poderosa le llevaría aproximadamente 30 años encontrar el plan “perfecto” los 12 minutos adicionales constituyen una ventaja razonable.

¿Qué conocimientos de estadística se necesitan para obtener un trabajo como el que usted desempeña?

Yo trabajo con estadísticos que tienen doctorados y programadores para el desarrollo y la aplicación de diseños de investigación. Yo obtuve una maestría y tenía mucha experiencia práctica en investigación de operaciones antes de trabajar en la editorial; no obstante, el principal requisito para hacer investigación consiste en saber lo suficiente de estadística para identificar oportunidades para aplicarla, y desarrollar información útil para nuestros lectores.

¿Les recomienda a los estudiantes universitarios actuales estudiar estadística? ¿Por qué?

Definitivamente sí. En un contexto de negocios, la estadística junto con la contabilidad y unas buenas bases de matemáticas financieras son los fundamentos cuantitativos. Además, la estadística es importante prácticamente en todos los aspectos de la vida.

¿Qué otras habilidades son importantes para los estudiantes universitarios actuales?

Buenas habilidades de expresión oral y escrita.



Estadísticos para describir, explorar y comparar datos

3



- 3-1 Panorama general**
- 3-2 Medidas de tendencia central**
- 3-3 Medidas de variación**
- 3-4 Medidas de posición relativa**
- 3-5 Análisis exploratorio de datos (AED)**

Continuación del capítulo 2: ¿Los Premios de la Academia discriminan por la edad?

El problema que abre el capítulo 2 incluye las edades de los ganadores del Óscar a la mejor actriz y al mejor actor. En el capítulo 2 utilizamos distribuciones de frecuencias y gráficas para investigar si las edades de las actrices eran significativamente diferentes de las edades de los actores. Con base en los resultados obtenidos en el capítulo 2, parece que las actrices ganadoras del Óscar son más jóvenes que los actores ganadores de este premio. En el presente capítulo continuamos investigando si existe una discrepancia en las edades, pero incluimos nuevas herramientas que servirán para comparar los dos conjuntos de datos.

Las distribuciones de frecuencias y las gráficas del capítulo 2 no resultan afectadas si los datos corresponden a una muestra o a una población completa. Sin embargo, esta diferencia sí afecta a algunas de las herramientas que se presentan en este capítulo. Se podría decir que los datos constituyen una población, ya que incluyen la edad de *cada* ganador del Óscar como mejor actor y mejor actriz desde la primera ceremonia de Premios de la Academia, celebrada en 1928, hasta los últimos resultados disponibles en el momento en que se escribe este libro. En vez de considerar que las edades son datos poblacionales, los manejaremos como datos muestrales que se obtuvieron de una población más grande. Algunos puristas podrían manifestarse en contra

de esto, pero es un enfoque común que nos permite enfrentar preguntas importantes como ésta: ¿Existe una diferencia significativa entre la edad promedio (media) de las mejores actrices y la edad promedio (media) de los mejores actores?

Los métodos que se estudiaron en el capítulo 2 nos permitieron construir distribuciones de frecuencias y gráficas que resumen y presentan visualmente la distribución de los datos. Los métodos que se presentan en este capítulo nos permitirán calcular valores numéricos de estadísticos importantes. (En el capítulo 1 aprendimos que un estadístico es una medición numérica que describe alguna característica de una muestra, en tanto que un parámetro es una medición numérica que describe alguna característica de una población). En vez de basarnos únicamente en distribuciones de frecuencias y gráficas, ahora empezaremos a incluir estadísticos importantes al comparar las edades de las mejores actrices y los mejores actores. Después de calcular los valores de estadísticos importantes, estaremos más preparados para comparar los dos conjuntos de datos y para responder la siguiente pregunta fundamental: ¿Existen diferencias sustanciales e importantes entre las edades de las mejores actrices y las edades de los mejores actores? Seguiremos utilizando los mismos datos de la tabla 2-1, incluidos en el problema del capítulo 2.

3-1 Panorama general

Este capítulo es sumamente importante porque presenta estadísticos básicos que describen las características fundamentales de un conjunto de datos. En el panorama general del capítulo 2 (sección 2-1) señalamos que al describir, explorar y comparar conjuntos de datos, las siguientes características suelen ser de suma importancia: **1.** el centro, **2.** la variación, **3.** la distribución, **4.** los valores extremos y **5.** las características de los datos que cambian con el tiempo.

Pensamiento crítico e interpretación: más allá de las fórmulas

La tecnología nos permite disfrutar del siguiente principio del uso de la estadística moderna: no es tan importante memorizar fórmulas o realizar cálculos aritméticos complejos a mano; en cambio, nos podemos concentrar en obtener resultados con algún tipo de tecnología (calculadoras o programas de cómputo), para luego dar un sentido práctico a los resultados a través del pensamiento crítico. Tenga esto en mente mientras estudia el presente capítulo. Por ejemplo, cuando estudie la *desviación estándar* en la sección 3-3, trate de observar cómo funciona la fórmula como medida de variación, luego aprenda a calcular valores de desviaciones estándar, pero realmente haga un esfuerzo por *comprender e interpretar* los valores de las desviaciones estándar.

Este capítulo incluye algunos pasos detallados de procedimientos relevantes, pero no es necesario dominar esos pasos en todos los casos. Sin embargo, le recomendamos que realice algunos cálculos manuales antes de utilizar su calculadora o computadora, ya que esto mejorará su comprensión y así podrá apreciar mejor los resultados obtenidos a partir de la tecnología.

Los métodos de este capítulo y del anterior se conocen como métodos de **estadística descriptiva**, ya que su objetivo es resumir o *describir* las características importantes de un conjunto de datos. Más adelante en este libro emplearemos métodos de **estadística inferencial**, cuando utilicemos datos muestrales para hacer inferencias (o generalizaciones) acerca de una población. Con la estadística inferencial hacemos una inferencia más allá de los datos conocidos. La estadística descriptiva y la estadística inferencial son dos divisiones generales de esta materia, y el presente capítulo, junto con el 2, se ocupa de principios fundamentales de la estadística descriptiva.



3-2 Medidas de tendencia central

Concepto clave Cuando describimos, exploramos y comparamos conjuntos de datos, las siguientes características suelen ser sumamente importantes: centro, variación, distribución, valores extremos y cambios a través del tiempo. [Recuerde que las siglas CVDVT (Cuidado con los Virus que Destruyen el Valioso Trabajo) sirven para no olvidar esas características]. En esta sección nos enfocamos en las características del *centro*. Buscamos obtener de alguna manera un número que represente el valor central de un conjunto de datos. Los conceptos de *media* y *mediana* deben quedar completamente claros. En específico, habrá que conocer muy bien los métodos para el cálculo de los valores de la media y la mediana. Asimismo, es importante saber que el valor de la media se puede ver muy afectado por la presencia de un valor extremo, aunque la mediana no es tan sensible a esto. (Un valor extremo es un valor que

está muy alejado de la gran mayoría de los datos). La parte 1 de esta sección incluye conceptos básicos que deben comprenderse muy bien antes de pasar a la parte 2.



Maniquíes ≠ realidad

La revista *Health* comparó las medidas de los maniquíes con las medidas de las mujeres. Los siguientes resultados se reportaron como “promedios”, que tal vez representan medias. Estatura de los maniquíes: 6 pies; estatura de las mujeres 5 pies 4 pulgadas. Tamaño de la cintura de los maniquíes: 23 pulgadas; cintura de las mujeres: 29 pulgadas. Tamaño de la cadera de los maniquíes: 34 pulgadas; cadera de las mujeres: 40 pulgadas. Talla de vestido de los maniquíes: 6; talla de vestido de las mujeres: 11. Cuando se comparan las medias, es evidente que los maniquíes y las mujeres reales son muy diferentes.

Parte 1: Conceptos básicos de las medidas de tendencia central

Definición

Una **medida de tendencia central** es un valor que se encuentra en el centro o a la mitad de un conjunto de datos.

Hay muchas formas distintas de determinar el centro, por lo que tenemos diferentes definiciones de las medidas de tendencia central, que incluyen la media, la mediana, la moda y la mitad del rango. Comenzaremos con la media.

Media

La media (aritmética), por lo general, es la medida numérica más importante que se utiliza para describir datos; comúnmente se le conoce como *promedio*.

Definición

La **media aritmética** de un conjunto de valores es la medida de tendencia central que se calcula al sumar los valores y dividir el total entre el número de valores. Esta medida de tendencia central se utilizará con frecuencia a lo largo del libro y nos referiremos a ella simplemente como la **media**.

Esta definición puede expresarse con la fórmula 3-1, que utiliza la letra griega Σ (sigma mayúscula) para indicar que los valores de los datos deben sumarse. Es decir, Σx representa la suma de todos los valores de los datos. El símbolo n denota el **tamaño de la muestra**, que es el número de valores en el conjunto de datos.

$$\text{Fórmula 3-1} \quad \text{Media} = \frac{\Sigma x}{n} \quad \begin{array}{l} \leftarrow \text{suma de todos los valores de la muestra} \\ \leftarrow \text{número de valores muestrales} \end{array}$$

Si el conjunto de datos es una *muestra* de una población más grande, la media se simboliza \bar{x} (y se lee “x barra”); cuando se usan todos los valores de la población, la media se simboliza por medio de μ (la letra griega mu minúscula). (Por lo general, los estadísticos muestrales se representan con letras del abecedario latino como \bar{x} , y los parámetros poblacionales con letras del alfabeto griego como μ .)

Notación

Σ	representa la <i>suma</i> de un conjunto de valores.
x	es la <i>variable</i> que generalmente se usa para representar los datos individuales.
n	representa el <i>número de valores</i> en una <i>muestra</i> .
N	representa el <i>número de valores</i> en una <i>población</i> .
$\bar{x} = \frac{\Sigma x}{n}$	es la <i>media</i> de un conjunto de valores <i>muestrales</i> .
$\mu = \frac{\Sigma x}{N}$	es la <i>media</i> de todos los valores de una <i>población</i> .



Poblaciones cambiantes

Una de las cinco características más importantes que se mencionaron en el capítulo 2 es el patrón de cambio de los datos con el paso del tiempo. Algunas poblaciones cambian, y sus estadísticos importantes cambian también. Los estándares de los cinturones de seguridad no han cambiado en 40 años, aun cuando el peso de los estadounidenses se ha incrementado de manera considerable desde entonces. En 1960 el 12.8% de los estadounidenses se consideraban obesos, en comparación con el 22.6% en 1994.

Según la Agencia Nacional de Carreteras de Estados Unidos (Highway Traffic Safety Administration), los cinturones de seguridad deben ajustarse a un maniquí estándar de choque (diseñado con datos de 1960) sentado en el asiento delantero, con un espacio libre de 4 pulgadas. En teoría, el cinturón se ajusta al 95% de los hombres y al 99% de las mujeres, pero estos porcentajes ahora son bajos ante el aumento de peso durante la segunda mitad del siglo pasado. Algunas empresas de automóviles ofrecen cinturones de seguridad con extensiones, pero otras no.

EJEMPLO Verificación del plomo en el aire Se sabe que el plomo tiene algunos efectos dañinos graves en la salud. A continuación se presentan cantidades de plomo (medidas en microgramos por metro cúbico o $\mu\text{g}/\text{m}^3$) en el aire. La Agencia de Protección Ambiental de Estados Unidos estableció un estándar de calidad del aire con respecto al plomo: un nivel máximo de $1.5 \mu\text{g}/\text{m}^3$. Las mediciones que se presentan a continuación se registraron en el edificio 5 del World Trade Center en distintos días, inmediatamente después de la destrucción causada por los ataques terroristas del 11 de septiembre de 2001. Calcule la media de esta muestra de niveles de plomo en el aire.

5.40 1.10 0.42 0.73 0.48 1.10

SOLUCIÓN La media se calcula empleando la fórmula 3-1. Primero se suman las puntuaciones y después se dividen entre el número de éstas:

$$\bar{x} = \frac{\sum x}{n} = \frac{5.40 + 1.10 + 0.42 + 0.73 + 0.48 + 1.10}{6} = \frac{9.23}{6} = 1.538$$

La media del nivel de plomo es $1.538 \mu\text{g}/\text{m}^3$. Además del valor de la media, también es notable que el conjunto de datos incluye un valor (5.40) que está muy distante de los demás. Sería importante investigar un “valor extremo” como éste. En tal caso, el nivel de plomo de $5.40 \mu\text{g}/\text{m}^3$ se midió después del colapso de las torres del World Trade Center, cuando había niveles excesivos de polvo y humo. Asimismo, es probable que parte del plomo fuera producto de las emisiones de la gran cantidad de vehículos que llegaron al lugar. Estos factores ofrecen una explicación razonable para un valor tan extremo.

Mediana

Una desventaja de la media es su sensibilidad a cada valor, de tal forma que una puntuación excepcional puede afectarla de manera drástica. La mediana resuelve, en gran medida, esa desventaja. La mediana es un “valor intermedio”, ya que la mitad de los valores de los datos están por debajo de la mediana y la otra mitad por arriba de ella. La siguiente definición es más precisa.

Definición

La **mediana** de un conjunto de datos es la medida de tendencia central que implica el *valor intermedio*, cuando los valores de los datos originales se presentan en orden de magnitud creciente (o decreciente). La mediana suele de denotarse con \tilde{x} (y se lee “x con tilde”).

Para calcular la mediana, primero se *ordenan* los valores (se acomodan en orden) y luego se sigue uno de los siguientes dos procedimientos:

1. Si el número de valores es impar, la mediana es el número que se localiza exactamente a la mitad de la lista.
2. Si el número de valores es par, la mediana se obtiene calculando la media de los dos números que están a la mitad.

EJEMPLO Verificación del plomo en el aire A continuación se presenta una lista de cantidades de plomo (medidas en $\mu\text{g}/\text{m}^3$) en el aire. Calcule la mediana de esta muestra.

5.40 1.10 0.42 0.73 0.48 1.10

SOLUCIÓN Primero ordene los valores:

0.42 0.48 0.73 1.10 1.10 5.40

Puesto que el número de valores es par (6), la mediana se obtiene calculando la media de los dos valores intermedios: 0.73 y 1.10.

$$\text{Mediana} = \frac{0.73 + 1.10}{2} = \frac{1.83}{2} = 0.915$$

Como el número de valores es par (6), la mediana en la media de los dos valores intermedios, es decir, $0.915 \mu\text{g}/\text{m}^3$. Observe que la mediana es muy diferente a la media de $1.538 \mu\text{g}/\text{m}^3$ que se obtuvo con el mismo conjunto de datos muestrales en el ejemplo anterior. La razón de esta gran discrepancia es el efecto que la puntuación 5.40 tuvo en la media. Si este valor extremo se redujera a 1.20, la media disminuiría de $1.538 \mu\text{g}/\text{m}^3$ a $0.838 \mu\text{g}/\text{m}^3$, pero la mediana no cambiaría.

EJEMPLO Verificación del plomo en el aire Repita el ejemplo anterior, después de incluir la medición de $0.66 \mu\text{g}/\text{m}^3$ que se registró otro día. Es decir, calcule la mediana de las siguientes mediciones de plomo:

5.40 1.10 0.42 0.73 0.48 1.10 0.66

SOLUCIÓN Primero ordene los valores:

0.42 0.48 0.66 0.73 1.10 1.10 5.40

Puesto que el número de valores es impar (7), la mediana es el valor que está exactamente a la mitad de la lista ordenada: $0.73 \mu\text{g}/\text{m}^3$.

Después de estudiar los dos ejemplos anteriores, debe quedar claro el procedimiento para obtener la mediana, así como el hecho de que la media se ve muy afectada por valores extremos, mientras que la mediana no. Como la mediana no es tan sensible a los valores extremos, a menudo se utiliza para conjuntos de datos que tienen un número relativamente pequeño de valores extremos. Por ejemplo, recientemente el U.S. Census Bureau reportó que la *mediana* del ingreso familiar es de \$36,078 anuales. Se usó la mediana porque existen pocas familias con ingresos realmente altos.



Paradoja del tamaño de la clase

Existen por lo menos dos formas de obtener el tamaño de una clase promedio, y ambas pueden dar resultados muy diferentes. En una universidad, si tomamos la cantidad de estudiantes de 737 clases, obtenemos una media de 40 estudiantes. Sin embargo, si reunimos una lista del tamaño de las clases para cada estudiante y utilizamos esta lista, obtendríamos una media de 147. Esta gran discrepancia se debe al hecho de que existen muchos estudiantes en clases grandes, en tanto que hay pocos estudiantes en clases pequeñas. Sin cambiar el número de clases o de profesores, podríamos reducir el tamaño de clase promedio para los estudiantes haciendo que todas las clases tengan un tamaño similar. Esto también aumentaría la asistencia, que es más alta en las clases más pequeñas.

Moda

Definición

La **moda** de un conjunto de datos es el valor que se presenta con *mayor frecuencia*.

- Cuando dos valores se presentan con la misma frecuencia y ésta es la más alta, ambos valores son modas, por lo que el conjunto de datos es **bimodal**.
- Cuando más de dos valores se presentan con la misma frecuencia y ésta es la más alta, todos los valores son modas, por lo que el conjunto de datos es **multimodal**.
- Cuando ningún valor se repite, se dice que no hay moda.

EJEMPLO Calcule las modas de los siguientes conjuntos de datos:

- 5.40 1.10 0.42 0.73 0.48 1.10
- 27 27 27 55 55 55 88 88 99
- 1 2 3 6 7 8 9 10

SOLUCIÓN

- El número 1.10 es la moda porque es el valor que se presenta con mayor frecuencia.
- Los números 27 y 55 son modas, ya que ambos se presentan con la misma frecuencia, que es la más alta. Este conjunto de datos es bimodal porque tiene dos modas.
- No hay moda porque ningún valor se repite.

En realidad, la moda no se utiliza mucho con los datos numéricos. Sin embargo, entre las diferentes medidas de tendencia central que consideramos, la moda es la única que puede usarse con datos de nivel de medición nominal. (Recuerde que el nivel de medición nominal se refiere a datos que consisten únicamente en nombres, etiquetas o categorías).

Mitad del rango

Definición

La **mitad del rango** es la medida de tendencia central que constituye el valor que está a la mitad, entre la puntuación más alta y la más baja, en el conjunto original de datos. Se calcula sumando el valor máximo con el valor mínimo y luego dividiendo la suma entre 2, de acuerdo con la siguiente fórmula:

$$\text{mitad del rango} = \frac{\text{valor máximo} + \text{valor mínimo}}{2}$$

La mitad del rango se emplea pocas veces. Puesto que utiliza sólo los valores máximo y mínimo, es demasiado sensible a esos extremos. Sin embargo, la mitad del rango posee tres características valiosas: **1.** es fácil de calcular; **2.** ayuda a reforzar la importante idea de que hay varias maneras de definir el centro de un conjunto de datos; **3.** en ocasiones se le utiliza incorrectamente en vez de la mediana, de manera que la confusión se reduce si se define claramente tanto la mitad del rango como la mediana.

Tabla 3-1 Comparación de las edades de las mejores actrices y los mejores actores		
	Mejores actrices	Mejores actores
Media	35.7	43.9
Mediana	33.5	42.0
Moda	35	41 y 42
Mitad del rango	50.5	52.5

EJEMPLO Verificación del plomo en el aire A continuación se presentan medidas de las cantidades de plomo (en $\mu\text{g}/\text{m}^3$) en el aire, en el World Trade Center en diferentes días después del 11 de septiembre de 2001. Calcule la mitad del rango de esta muestra.

5.40 1.10 0.42 0.73 0.48 1.10

SOLUCIÓN La mitad del rango se calcula de la siguiente manera:

$$\frac{\text{valor máximo} + \text{valor mínimo}}{2} = \frac{5.40 + 0.42}{2} = 2.910$$

La mitad del rango es $2.910 \mu\text{g}/\text{m}^3$.

Por desgracia, el término *promedio* en ocasiones se utiliza para cualquier medida de tendencia central y en otras para implicar la media. Debido a esta ambigüedad, el término *promedio* no se debería usar cuando nos referimos a una medida de tendencia central en particular. En vez de ello, es necesario usar el término específico, como media, mediana, moda o mitad del rango.

Con la idea de describir, explorar y comparar datos, incluimos la tabla 3-1, que resume las distintas medidas de tendencia central para las edades de las actrices y los actores ganadores del Óscar de la tabla 2-1, del problema del capítulo 2. Una comparación de las medidas de tendencia central sugiere que las mejores actrices son más jóvenes que los mejores actores. Existen métodos para determinar si estas diferencias aparentes son estadísticamente significativas; estudiaremos algunos de esos métodos más adelante en este libro (véase la sección 9-3).

Regla de redondeo

Una sencilla regla para redondear las respuestas es la siguiente:

Aumente una posición decimal a las que hay en el conjunto original de datos.

Cuando aplique esta regla, redondee sólo la respuesta final y *no los valores intermedios que aparecen durante los cálculos*. Así, la media de 2, 3, 5, es 3.33333..., que se redondea a 3.3. Como los valores originales son números enteros, redondeamos al decimal más cercano. Otro ejemplo sería la media de 80.4 y 80.6, que es igual a 80.50 (una posición decimal más de la que se empleó para los valores originales).

Pensamiento crítico

A partir de un conjunto de datos muestrales, es relativamente fácil calcular las medidas de tendencia central, como la media y la mediana. Sin embargo, siempre debemos preguntarnos si los resultados tienen sentido. En la sección 1-2 señala-

mos que no tiene sentido hacer cálculos numéricos con datos que están en el nivel de medición nominal. Puesto que los datos a nivel nominal consisten únicamente de nombres, etiquetas o categorías, no tiene sentido calcular el valor de estadísticos como la media y la mediana. También debemos pensar en el método que se utilizó para reunir los datos muestrales. Si el método no es sensato, es probable que los estadísticos que obtengamos sean engañosos.

EJEMPLO Pensamiento crítico y medidas de tendencia central

Para cada uno de los siguientes ejercicios podemos calcular medidas de tendencia central como la media y la mediana. Identifique una razón importante por la que, en estos casos, la media y la mediana *no* son estadísticos que pueden servir de manera precisa y efectiva como medidas de tendencia central.

- a. Códigos postales: 12601 90210 02116 76177 19102
- b. Clasificaciones de los niveles de estrés de distintos empleos: 2 3 1 7 9
- c. Los sujetos encuestados se codifican de la siguiente manera: 1 (demócratas), 2 (republicanos), 3 (liberales), 4 (conservadores) o 5 (cualquier otro partido político).

SOLUCIÓN

- a. Los códigos postales no miden ni cuentan algo. Los números en realidad son etiquetas de ubicaciones geográficas. En este caso, la media y la mediana son estadísticos carentes de significado.
- b. Las clasificaciones reflejan un orden, pero no miden ni cuentan algo. La clasificación 1 podría indicar un empleo que tiene un nivel de estrés mucho mayor que el nivel de un empleo con una clasificación 2, por lo que los distintos números no corresponden a las magnitudes de los niveles de estrés. En este caso, la media y la mediana son estadísticos carentes de significado.
- c. Los resultados codificados son números, pero no miden ni cuentan algo. Estos números son simplemente distintas maneras de expresar nombres. Como consecuencia, en este caso, la media y la mediana son estadísticos carentes de significado.

El ejemplo anterior incluyó datos con niveles de medición que no justifican el uso de estadísticos como la media o la mediana. El siguiente ejemplo implica un tema más sutil.

EJEMPLO Media del salario de maestros En cada una de las 50 entidades de Estados Unidos, un investigador conoce el salario promedio de maestros de escuela secundaria (datos de la National Education Association):

\$37,200 \$49,400 \$40,000 . . . \$37,800

La media de las 50 cantidades es \$42,210, pero ¿es ésta la media del salario de todos los maestros de secundaria de Estados Unidos? ¿Por qué?

SOLUCIÓN No, \$42,210 no es necesariamente la media del salario de todos los maestros de secundaria en Estados Unidos. El problema aquí es que algunos estados tienen más maestros de secundaria que otros. El cálculo de la media de ese país debe tomar en cuenta el número de maestros de secundaria que hay en cada estado. (Después de este ejemplo analizamos la media de una distribución de frecuencias). La media de todos los maestros de secundaria de Estados Unidos es \$45,200 y no \$42,210.

El ejemplo anterior ilustra que la media de una población no es necesariamente igual a la media de las medias calculadas en distintos subconjuntos de la población.

En vez de realizar mecánicamente y sin reflexión el cálculo de medias y medianas, los ejercicios de esta sección fomentan la *reflexión* acerca de los resultados calculados.

Parte 2: Más allá de los aspectos básicos de las medidas de tendencia central

La media de una distribución de frecuencias

Cuando trabajamos con datos resumidos en una distribución de frecuencias, no sabemos con exactitud los valores que caen dentro de una clase en particular. Para poder hacer cálculos, suponga que en cada clase todos los valores muestrales son iguales a las marcas de clase. Por ejemplo, considere el intervalo de clase de 21-30, con una frecuencia de 28 (como en la tabla 2-2). Supongamos que los 28 valores son iguales a 25.5 (la marca de clase). Si repetimos 28 veces el valor 25.5, obtenemos un total de $25.5 \cdot 28 = 714$. Luego, podemos añadir estos productos de cada clase para calcular el total de todos los valores muestrales, los cuales dividimos después entre el número total de valores muestrales. (El número de valores muestrales es la sumatoria de las frecuencias Σf .) La fórmula 3-2 se utiliza para calcular la media cuando los datos muestrales están resumidos en una distribución de frecuencias. La fórmula 3-2 no es un nuevo concepto, sino simplemente una variante de la fórmula 3-1.

Primero multiplique cada frecuencia y marca de clase; luego sume los productos.

↓

Fórmula 3-2
$$\bar{x} = \frac{\sum(f \cdot x)}{\sum f}$$
 (media de la distribución de frecuencias)

↑

Sumatoria de las frecuencias

Por ejemplo, observe la tabla 3-2. Las primeras dos columnas son iguales a la distribución de frecuencias (tabla 2-2) de las edades de las actrices ganadoras del

Tabla 3-2 Cálculo de la media de una distribución de frecuencias			
Edad de las actrices	Frecuencia f	Marca de clase x	$f \cdot x$
21-30	28	25.5	714
31-40	30	35.5	1065
41-50	12	45.5	546
51-60	2	55.5	111
61-70	2	65.5	131
71-80	2	75.5	151
TOTALES:	$\sum f = 76$		$\sum(f \cdot x) = 2718$
		$\bar{x} = \frac{\sum(f \cdot x)}{\sum f} = \frac{2718}{76} = 35.8$	

Óscar. La tabla 3-2 ilustra el procedimiento que se sigue para aplicar la fórmula 3-2 cuando se calcula la media de datos resumidos en una distribución de frecuencias. La tabla 3-2 produce el resultado $\bar{x} = 35.8$, pero si usamos la lista original con 76 valores individuales obtenemos $\bar{x} = 35.7$. Recuerde, la distribución de frecuencias produce una aproximación de \bar{x} , ya que no se basa en la lista original exacta de los valores muestrales.

Media ponderada

En algunos casos, los valores varían de acuerdo con su grado de importancia, por lo que podemos ponderarlos y calcular la **media ponderada** de los valores x , una media que se obtiene asignando distintos pesos (w) a los valores, tal como se muestra en la fórmula 3-3.

$$\text{Fórmula 3-3} \quad \text{media ponderada: } \bar{x} = \frac{\sum(w \cdot x)}{\sum w}$$

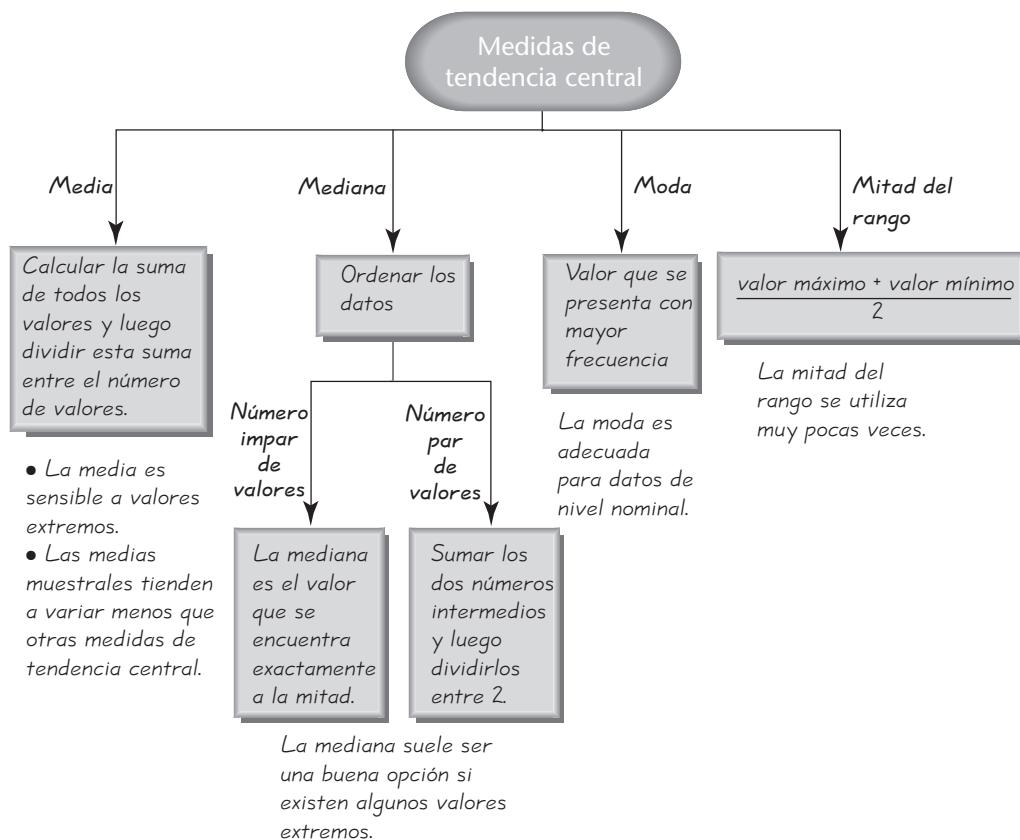
La fórmula 3-3 nos indica que debemos multiplicar cada peso w por el valor x correspondiente, luego sumar los productos y después dividir el total entre la sumatoria de los pesos w . Por ejemplo, suponga que necesitamos una media de tres calificaciones de una prueba (85, 90, 75), donde la primera prueba cuenta el 20%, la segunda el 30% y la tercera el 50% de la calificación final. Podemos asignar pesos de 20, 30 y 50 a las calificaciones de la prueba y luego calcular la media aplicando la fórmula 3-3 con $x = 85, 90, 75$ y los pesos correspondientes $w = 20, 30, 50$, de la siguiente manera:

$$\begin{aligned}\bar{x} &= \frac{\sum(w \cdot x)}{\sum w} \\ &= \frac{(20 \times 85) + (30 \times 90) + (50 \times 75)}{20 + 30 + 50} \\ &= \frac{8150}{100} = 81.5\end{aligned}$$

La mejor medida de tendencia central

Hasta ahora hemos considerado la media, mediana, moda y mitad del rango como medidas de tendencia central. ¿Cuál de ellas es la mejor? Por desgracia, no existe una respuesta única a esa pregunta, porque no hay criterios objetivos para determinar la medida más representativa para todos los conjuntos de datos. Las diferentes medidas de tendencia central ofrecen diversas ventajas y desventajas, algunas de las cuales se presentan en la figura 3-1. En lo que resta de este libro, la media se usará con frecuencia, la mediana ocasionalmente, en tanto que la moda y la mitad del rango no se utilizarán.

La media es relativamente *confiable*, de manera que cuando se seleccionan muestras de la misma población, las medias muestrales tienden a ser más consistentes que otras medidas de tendencia central. Es decir, las medias de muestras obtenidas de la misma población no varían tanto como las otras medidas de tendencia central. Otra ventaja importante de la media es que toma en cuenta a cada valor, y una desventaja importante es que en ocasiones se ve afectada de manera drástica por unos cuantos valores extremos. Esta desventaja se supera cuando empleamos una media recortada, como se describe en el ejercicio 29.

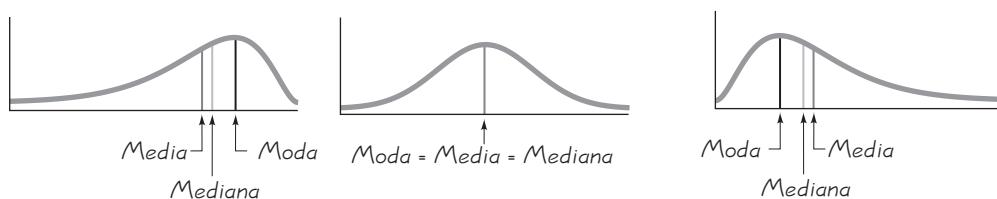


Sesgo

Una comparación de la media, la mediana y la moda puede revelar información acerca de las características de sesgo, que se define a continuación y se ilustra en la figura 3-2.

Definición

Una distribución de datos está **sesgada** si no es simétrica y se extiende más hacia un lado que hacia el otro. (Una distribución de datos es **simétrica** si la mitad izquierda de su histograma es aproximadamente una imagen en espejo de su mitad derecha).



a) Sesgo a la izquierda (sesgo negativo): la media y la mediana están a la izquierda de la moda.

b) Simétrica (sesgo cero): la media, la mediana y la moda son iguales.

c) Sesgo a la derecha (sesgo positivo): la media y la mediana están a la derecha de la moda.

Figura 3-2
Sesgo

Los datos **sesgados a la izquierda** (lo que también se conoce como *sesgo negativo*) poseen una cola izquierda más larga, y la media y la mediana se encuentran a la izquierda de la moda. [Aunque no siempre es posible predecirlo, los datos sesgados a la izquierda suelen tener una media menor a la mediana, como sucede en la figura 3-2a)]. Los datos **sesgados a la derecha** (lo que también se denomina *sesgo positivo*) poseen una cola derecha más larga, y la media y la mediana se encuentran a la derecha de la moda. [Una vez más, aunque no siempre es posible predecirlo, en los datos sesgados a la derecha la media suele estar a la derecha de la mediana, como ocurre en la figura 3-2c)].

En la práctica, muchas distribuciones de datos son simétricas y carecen de sesgo. Las distribuciones sesgadas hacia la derecha son más comunes que las sesgadas hacia la izquierda, ya que con frecuencia es más fácil obtener valores excepcionalmente grandes que valores excepcionalmente pequeños. Por ejemplo, en el caso de los ingresos anuales, es imposible obtener valores por debajo del límite inferior de cero, pero hay algunas personas que ganan millones de dólares en un año. Por lo tanto, el ingreso anual tiende a mostrar un sesgo hacia la derecha, como se observa en la figura 3-2c).

Uso de la tecnología

Los cálculos en esta sección son bastante sencillos, pero algunos de la siguiente sección requieren mayor esfuerzo. Muchos programas de cómputo le permiten introducir un conjunto de datos y utilizar una operación para obtener diversos estadísticos a partir de muestras, los cuales se engloban en la *estadística descriptiva*. (Véase la sección 3-4, donde se incluyen representaciones visuales de muestras que se obtienen con STATDISK,

Minitab, Excel y la calculadora TI-83/84 Plus). A continuación se incluyen algunos de los procedimientos para la obtención de tales representaciones visuales.

STATDISK Ingrese los datos en la ventana de datos. Haga clic en **Data** y elija **Descriptive Statistics**. Ahora haga clic en **Evaluate** para obtener los diversos estadísticos descriptivos, incluyendo la media, la mediana, la mitad del intervalo y otros que se discutirán en las secciones por venir.

MINITAB Ingrese los datos en la columna que tiene el encabezado C1. Haga clic en **Stat**, seleccione **Basic Statistics** y luego **Descriptive Statistics**. Los resultados incluirán la media y la mediana, así como otros estadísticos.

EXCEL Ingrese los datos de la muestra en la columna A. Seleccione **Tools**, después

Data Analysis y luego **Descriptive Statistics**; haga clic en **OK**. En el cuadro de diálogo, ingrese el rango de entrada de datos (tal como A1:A76 para 76 valores en la columna A), haga clic en **Summary Statistics** y después en **OK**. (Si Data Analysis no aparece en el menú Tools, deberá instalarlo haciendo clic en **Tools** y seleccionando **Add-Ins**).

TI-83/84 PLUS Primero ingrese los datos en la lista L1 presionando **STAT**, luego **Edit** y finalmente la tecla **ENTER**. Una vez que ingresó los datos, presione **STAT** y seleccione **CALC**, después seleccione **1-Var Stats** y finalmente presione **ENTER** dos veces. La representación visual incluirá la media \bar{x} , la mediana, el valor mínimo y el valor máximo. Utilice la flecha que va hacia abajo ↓ para ver los resultados que no aparecen en la primera representación visual.

3-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Medidas de tendencia central.** ¿En qué sentido la media, mediana, moda y mitad del rango son medidas de “tendencia central”?
- 2. Media y mediana.** Una clase de estadística se compone de 24 estudiantes, los cuales son desempleados o cuentan con un empleo de tiempo parcial con un salario bajo. La clase también incluye a un profesor que recibe un salario muy alto. ¿Cuál medida es mejor para describir el ingreso de una persona típica en la clase de 25 personas (incluyendo al profesor): la media o la mediana? ¿Por qué?



- 3. Datos nominales.** En el capítulo 1 señalamos que los datos están en un nivel de medición nominal si únicamente consisten en nombres y etiquetas. Un fanático de los Patriotas de Nueva Inglaterra registra el número de la playera de cada jugador del equipo en un juego del Súper Bowl. ¿Tiene algún sentido calcular la media de estos números? ¿Por qué?
- 4. Tiempo de transportación.** Una socióloga desea calcular la media del tiempo que tardan en transportarse al trabajo todos los residentes de Estados Unidos que tienen un empleo. Ella sabe que no es práctico encuestar a cada uno de los millones de individuos que trabajan, por lo que realiza una búsqueda en Internet y encuentra la media del tiempo de transporte en cada una de las 50 entidades de Estados Unidos. La socióloga suma los 50 tiempos y los divide entre 50. ¿Es probable que el resultado sea un buen estimado de la media del tiempo de transporte de todos los trabajadores? ¿Por qué?

En los ejercicios 5 a 12, calcule a) la media, b) la mediana, c) la moda y d) la mitad del rango de los datos muestrales listados. También responda las preguntas que se plantean.

- 5. Percepción del tiempo.** Algunos estudiantes de estadística participaron en un experimento con el fin de probar su capacidad para determinar el transcurso de 1 minuto (o 60 segundos). A continuación se presentan los resultados en segundos. Identifique al menos una buena razón por la que la media de esta muestra no sería un buen estimado del promedio de la población de adultos.

53 52 75 62 68 58 49 49

- 6. Cereal.** Un nutriólogo obtiene las cantidades de azúcar (en centigramos) de 100 centigramos (o 1 gramo) de 10 cereales diferentes, incluyendo Cheerios, Corn Flakes, Fruit Loops y otros siete. Las cantidades se presentan abajo. ¿Es probable que la media de esos valores sea un buen estimado de la cantidad de azúcar promedio en cada gramo de cereal que consume la población de todos los estadounidenses que comen cereal? ¿Por qué?

3 24 30 47 43 7 47 13 44 39

- 7. Fenotipos de guisantes.** Se realizó un experimento para determinar si una deficiencia de dióxido de carbono en la tierra afecta los fenotipos de los guisantes (chícharos). A continuación se indican los códigos de los fenotipos: 1 = amarillo claro, 2 = verde claro, 3 = amarillo rugoso y 4 = verde rugoso. ¿Se pueden obtener medidas de tendencia central para estos valores? ¿Los resultados tienen algún sentido?

2 1 1 1 1 1 4 1 2 2 1 2 3 3 2 3 1 3 1 3 1 3 2 2

- 8. Géiser Old Faithful.** Abajo se incluyen los intervalos (en minutos) entre las erupciones del géiser Old Faithful en el Parque Nacional Yellowstone. Después de cada erupción, el servicio de parques nacionales ofrece un estimado del tiempo que pasará hasta la siguiente erupción. Con base en estos valores, ¿cuál de éstos podría usarse como un estimado del tiempo que falta para la siguiente erupción? ¿Es probable que ese estimado tenga un intervalo de exactitud de 5 minutos?

98 92 95 87 96 90 65 92 95 93 98 94

- 9. Mediciones de la presión sanguínea.** Catorce estudiantes del segundo año de medicina del Bellevue Hospital midieron la presión sanguínea de la misma persona. A continuación se listan las lecturas sistólicas (en mmHg). ¿Qué es notorio acerca de este conjunto de datos?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

- 10. Temperaturas corporales.** Investigadores de la Universidad de Maryland reunieron lecturas de la temperatura corporal de una muestra de adultos, y a continuación se presentan ocho de esas medidas (en grados Fahrenheit). ¿La media de esta muestra es igual a 98.6, la cifra que generalmente se considera la temperatura corporal media de los adultos?

98.6 98.6 98.0 98.0 99.0 98.4 98.4 98.4

- 11. Moscas de la fruta.** Más adelante se indican las longitudes (en milímetros) del tórax de una muestra de moscas de la fruta machos. Las moscas de la fruta (*Drosophila*) son algunos de los sujetos favoritos de experimentación de los investigadores porque poseen una composición cromosómica sencilla, se reproducen con rapidez, tienen una gran cantidad de descendientes y son fáciles de cuidar. (Éstas son características de las moscas de la fruta y no necesariamente de los investigadores). Si nos enteramos de que las medidas se obtuvieron de moscas que volaban sobre una manzana que se encontraba en una mesa de cocina en Pocatello, ¿la media sirve como un estimado razonable de todas las moscas de la fruta que residen en Estados Unidos?

0.72 0.90 0.84 0.68 0.84 0.90 0.92 0.84 0.64 0.84 0.76

- 12. Ingreso personal.** A continuación se muestran las cantidades de ingreso personal per cápita (en dólares) de las primeras cinco entidades de Estados Unidos, ordenadas alfabéticamente: Alabama, Alaska, Arizona, Arkansas y California (datos del U.S. Bureau of Economic Analysis). Cuando se incluyen las 45 cantidades de los otros estados, la media de los 50 estados es de \$29,672.52. ¿La cantidad promedio del ingreso personal per cápita de todos los individuos en Estados Unidos es de \$29,672.52? En general, ¿podemos calcular la media de algunas características de todos los residentes de Estados Unidos al obtener la media de cada estado y después calcular la media de esos 50 resultados?

\$25,128 \$32,151 \$26,183 \$23,512 \$32,996

En los ejercicios 13 a 20, calcule la media y la mediana de cada una de las dos muestras y luego compare los dos conjuntos de resultados.

- 13. Llueven gatos.** A veces se utiliza la estadística para comparar o identificar autores de distintos trabajos. A continuación se incluye una lista con la longitud de las primeras 20 palabras que escribió Tennessee Williams en el prólogo de *The cat on a Hot Tin Roof*, junto con las primeras 20 palabras de *The Cat in the Hat*, del doctor Seuss. ¿Existe alguna diferencia?

Cat on a Hot Tin Roof: 2 6 2 2 1 4 4 2 4 2 3 8 4 2 2 7 7 2 3 11

The cat in the Hat: 3 3 3 3 5 2 3 3 3 2 4 2 2 3 2 3 5 3 4 4

- 14. Edades de polizones.** El *Queen Mary* navegaba entre Inglaterra y Estados Unidos; en ocasiones se encontraron polizones a bordo. A continuación se listan las edades (en años) de los polizones que iban con rumbo al este y de los que iban al oeste (datos de Cunard Steamship Co., Ltd.). Compare los dos conjuntos de datos.

Hacia el este: 24 24 34 15 19 22 18 20 20 17

Hacia el oeste: 41 24 32 26 39 45 24 21 22 21

- 15. Exactitud del pronóstico del clima.** En un análisis de la exactitud del pronóstico del clima se comparan las temperaturas máximas reales con las temperaturas máximas pronosticadas un día anterior y las temperaturas máximas pronosticadas cinco días antes. Más abajo se indican los errores entre las temperaturas pronosticadas y las temperaturas máximas reales para días consecutivos en el condado Dutchess, Nueva York. ¿Las medias y las medianas indican que las temperaturas pronosticadas un día antes son más exactas que las pronosticadas cinco días antes, como se esperaría?

(máxima real) – (máxima pronosticada un día antes): 2 2 0 0 -3 -2 1
-2 8 1 0 -1 0 1

(máxima real) – (máxima pronosticada cinco días antes): 0 -3 2 5 -6 -9 4
-1 6 -2 -2 -1 6 -4

- 16. Efecto de tratamiento.** Investigadores de la Universidad de Pennsylvania realizaron experimentos con álamos. A continuación se muestran los pesos (en kg) de álamos

que no recibieron tratamiento y de álamos tratados con fertilizantes y riego. ¿Parece existir una diferencia entre las dos medias? ¿Parece que el tratamiento con fertilizantes y riego es efectivo para incrementar el peso de los álamos?

Sin tratamiento:	0.15	0.02	0.16	0.37	0.22
Con fertilizantes y riego:	2.03	0.27	0.92	1.07	2.38

- 17. Tiempos de espera de clientes.** A continuación se presentan los tiempos de espera (en minutos) de los clientes del banco Jefferson Valley (donde todos los clientes forman una sola fila) y del banco Providence (donde los clientes esperan en filas individuales, en tres ventanillas diferentes). Determine si existe una diferencia entre los dos conjuntos de datos, que no sea aparente cuando se comparan las medidas de tendencia central. Si tal diferencia existe, ¿cuál es?

Jefferson Valley (una sola fila): 6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7

Providence (filas individuales): 4.2 5.4 5.8 6.2 6.7 7.7 7.7 8.5 9.3 10.0

- 18. Coca-Cola regular/Coca-Cola dietética.** Los siguientes son los pesos (en libras) de muestras del contenido de latas de Coca-Cola regular y dietética. ¿Parece existir una diferencia significativa entre los dos conjuntos de datos? ¿Cómo se podría explicar una diferencia como ésta?

Regular: 0.8192 0.8150 0.8163 0.8211 0.8181 0.8247

Dietética: 0.7773 0.7758 0.7896 0.7868 0.7844 0.7861

- 19. Reflexiones sobre peniques.** Los peniques o centavos estadounidenses acuñados antes de 1983 tienen un 97% de cobre y un 3% de zinc, en tanto que los acuñados después de 1983 tienen un 3% de cobre y un 97% de zinc. A continuación se indican los pesos (en gramos) de peniques acuñados antes y después de 1983. (El autor obtuvo estos datos). ¿Parece existir una diferencia considerable en las medias?

Antes de 1983: 3.1582 3.0406 3.0762 3.0398 3.1043 3.1274
3.0775 3.1038 3.1086 3.0586

Después de 1983: 2.5113 2.4907 2.5024 2.5298 2.4950 2.5127
2.4998 2.4848 2.4823 2.5163

- 20. IMC y género.** Es bien sabido que los hombres tienden a pesar más y a ser más altos que las mujeres. El índice de masa corporal (IMC) es una medida que se basa en el peso y la estatura. A continuación se listan los valores de IMC de hombres y mujeres elegidos de manera aleatoria. ¿Parece existir una diferencia notable?

Hombres: 23.8 23.2 24.6 26.2 23.5 24.5 21.5 31.4 26.4 22.7 27.8 28.1

Mujeres: 19.6 23.8 19.6 29.1 25.2 21.4 22.0 27.5 33.5 20.6 29.9 17.7

Conjuntos grandes de datos. Para los ejercicios 21 a 24, remítase al conjunto de datos en el apéndice B. Con un programa de cómputo o una calculadora obtenga las medias y las medianas; luego, compare los resultados tal como se indica.

- 21. Exactitud del pronóstico del clima.** El ejercicio 15 incluye los datos de temperaturas pronosticadas para 14 días. Repita el ejercicio 15 después de remitirse al conjunto de datos 8 del apéndice B. Use los datos de los 35 días para expandir la lista de las diferencias que se incluye en el ejercicio 15. (Los datos del ejercicio 15 se basan en los primeros 14 días del conjunto de datos 8). ¿Sus conclusiones cambian cuando usa el conjunto grande de datos? ¿Tiene mayor confianza en los resultados basados en el conjunto de datos más grande?

- 22. Coca-Cola regular/Coca-Cola dietética.** El ejercicio 18 incluye los pesos de Coca-Cola regular y Coca-Cola dietética del conjunto de datos 12 en el apéndice B. (Los datos del ejercicio 18 incluyen las primeras seis latas de Coca-Cola regular y las primeras

seis latas de Coca-Cola dietética). Repita el ejercicio 18 utilizando la lista completa de pesos de Coca-Cola regular y dietética. ¿Sus conclusiones cambian cuando se usa el conjunto grande de datos?

- 23. Peniques revisados.** El ejercicio 19 incluye los pesos de peniques acuñados antes y después de 1983. Repita el ejercicio 19 con la lista completa de los pesos de los peniques acuñados antes de 1983 y la lista completa de los peniques acuñados después de 1983, tal como aparecen en el conjunto de datos 14 del apéndice B. ¿Sus conclusiones cambian cuando se usa el conjunto grande de datos?
- 24. IMC y género.** El ejercicio 20 incluye los valores del IMC de 12 hombres y 12 mujeres. Repita el ejercicio 20 con la lista completa de los valores del IMC, tal como aparece en el conjunto de datos 1 del apéndice B. ¿Sus conclusiones cambian cuando se usa el conjunto grande de datos?

En los ejercicios 25 a 28, calcule la media de los datos que se resumen en la distribución de frecuencias dada. Además, compare las medias calculadas con las medias reales que se obtuvieron al utilizar la lista original de datos, que es la siguiente: (ejercicio 25) 53.8; (ejercicio 26) 0.230; (ejercicio 27) 46.7; (ejercicio 28) 98.20.

25. Temperatura mínima diaria (en °F)	Frecuencia	26. Precipitación diaria en (pulgadas)	Frecuencia
35–39	1	0.00–0.49	31
40–44	3	0.50–0.99	1
45–49	5	1.00–1.49	0
50–54	11	1.50–1.99	2
55–59	7	2.00–2.49	0
60–64	7	2.50–2.99	1
65–69	1		

Tabla del Ejercicio 27

Velocidad	Frecuencia
42–45	25
46–49	14
50–53	7
54–57	3
58–61	1

Tabla del Ejercicio 28

Temperatura	Frecuencia
96.5–96.8	1
96.9–97.2	8
97.3–97.6	14
97.7–98.0	22
98.1–98.4	19
98.5–98.8	32
98.9–99.2	6
99.3–99.6	4

- 27. Multas por exceso de velocidad.** La distribución de frecuencias dada describe la velocidad de los conductores multados por la policía de la ciudad en Poughkeepsie. Estos conductores viajaban por una zona de Creek Road, que pasa por la universidad del autor y tiene un límite de velocidad de 30 mi/h. ¿Qué diferencia existe entre la media y la velocidad límite de 30 mi/h?

- 28. Temperaturas corporales.** La distribución de frecuencias dada resume una muestra de temperaturas corporales humanas. (Véase las temperaturas de la medianoche el segundo día, tal como se muestran en el conjunto de datos 2 del apéndice B). ¿Qué diferencia existe entre la media y el valor de 98.6°F, que la mayoría de las personas consideran la temperatura promedio?

3-2 MÁS ALLÁ DE LO BÁSICO

- 29. Media recortada.** Ya que la media es muy sensible a los valores extremos, decimos que no es una medida de tendencia central *resistente*. La **media recortada** es más resistente. Para calcular la media recortada del 10% de un conjunto de datos, primero se acomodan los datos en orden, después se elimina el 10% de los valores inferiores y el 10% de los valores superiores y luego se calcula la media de los valores restantes. Para los pesos de los osos en el conjunto de datos 6 del apéndice B, calcule *a*) la media, *b*) la media recortada del 10%, *c*) la media recortada del 20%. ¿Qué diferencias hay en los resultados?

- 30. Grados de libertad.** Diez valores tienen una media de 75.0. Nueve de los valores son 62, 78, 90, 87, 56, 92, 70, 70 y 93.

- Calcule el décimo valor.
- Necesitamos crear una lista de n valores que tenga una media específica conocida.

Tenemos la libertad de seleccionar cualesquiera valores que deseemos para algunos de los n valores. ¿Cuántos de los n valores pueden asignarse libremente antes de determinar los valores restantes? (El resultado a menudo se conoce como el *número de grados de libertad*).

- 31. Datos censurados.** Se realizó un experimento para probar la longevidad de árboles tratados con fertilizantes. El experimento se lleva a cabo durante un periodo fijo de cinco años. (Se dice que la prueba se *censura* a los cinco años). Los resultados muestrales (en años) son 2.5, 3.4, 1.2, 5 +, 5 + (donde 5 + indica que el árbol aún estaba vivo al final del experimento). ¿Qué puede concluir acerca de la vida media de los árboles?

- 32. Datos transformados.** En cada uno de los siguientes casos, describa cómo se ven afectadas la media, la mediana, la moda y la mitad del rango.

- La misma constante k se suma a cada valor del conjunto de datos.
- Cada valor del conjunto de datos se multiplica por la misma constante k .

- 33. La media armónica** con frecuencia se utiliza como una medida de tendencia central para conjuntos de datos que consisten en tasas de cambios, como la rapidez. Para calcularla se divide el número de valores n entre la suma de los *recíprocos* de todos los valores, de la siguiente forma:

$$\frac{n}{\sum \frac{1}{x}}$$

(Ningún valor puede ser cero). Cuatro estudiantes conducen desde Nueva York hasta Florida (1200 millas), con una velocidad de 40 mi/h (¡sí, de verdad!). Como necesitan llegar a su clase de estadística a tiempo, viajan de regreso con una velocidad de 60 mi/h. ¿Cuál es la velocidad promedio del viaje completo? (La media armónica se utiliza para promediar velocidades).

- 34. La media geométrica** suele utilizarse en negocios y economía para calcular las tasas de cambio promedio, las tasas de crecimiento promedio o tasas promedio. Dados n valores (todos positivos), la media geométrica es la n -ésima raíz de su producto. El *factor de crecimiento promedio* de dinero compuesto con tasas de interés anual del 10, 8, 9, 12 y 7% puede obtenerse calculando la media geométrica de 1.10, 1.08, 1.09, 1.12 y 1.07. Calcule el factor de crecimiento promedio.

- 35. La media cuadrática** (o **cuadrado medio de raíz**, o CMR) suele utilizarse en aplicaciones de física. Por ejemplo, en los sistemas de distribución de energía, los voltajes y las corrientes suelen expresarse en términos de sus valores de CMR. La media cuadrática de un conjunto de valores se obtiene elevando al cuadrado cada valor, sumando los resultados, dividiendo entre el número de valores n y después sacando la raíz cuadrada del resultado, como se indica a continuación:

$$\text{Media cuadrática} = \sqrt{\frac{\sum x^2}{n}}$$

Calcule el CMR de las siguientes fuentes de energía (en volts) obtenidas de la casa del autor: 111.2, 108.7, 109.3, 104.8, 112.6.

36. Mediana. Cuando los datos se resumen en una distribución de frecuencias, la mediana puede calcularse identificando primero la *clase de la mediana* (la clase que contiene a la mediana). Luego suponemos que los valores en esa clase se distribuyen de manera uniforme y podemos interpolar. Este proceso se describe de la siguiente forma:

$$\text{(límite inferior de clase de la mediana)} \times \text{(anchura de clase)} \left(\frac{\left(\frac{n+1}{2} \right) - (m+1)}{\text{frecuencia de la clase de la mediana}} \right)$$

donde n es la suma de todas las frecuencias de clase y m es la suma de las frecuencias de clase que *preceden* la clase de la mediana. Utilice este procedimiento para calcular la mediana del conjunto de datos que se resume en la tabla 2-2. ¿Qué diferencias existen entre este resultado y la mediana de la lista original de datos, que es de 33.5? ¿Cuál valor de la mediana es mejor: el que se calculó para la tabla de frecuencias o el valor de 33.5?

3-3

Medidas de variación

Concepto clave Esta sección es una de las más importantes de todo el libro, pues presenta el tema de la variación, un concepto muy relevante en estadística. Primero lea la parte 1 de forma rápida y obtenga una comprensión general de las características de la variación. Luego, aprenda a usar el conjunto de datos para calcular los valores del rango y la desviación estándar. Después lea la parte 2 y trate de comprender la regla práctica del intervalo para interpretar valores de desviación estándar; también trate de comprender el razonamiento que subyace en la fórmula de la desviación estándar, pero no dedique mucho tiempo a memorizar fórmulas o a hacer cálculos aritméticos. En vez de ello, dé mayor importancia a *interpretar* los valores de la desviación estándar.

Parte 1: Conceptos básicos de la variación

En la figura 3-3 se presenta un ejemplo visual de variación, el cual muestra gráficas de barras de los tiempos de espera de los clientes en tres bancos diferentes. En el primer banco, el gerente controla de forma muy cuidadosa los tiempos de espera modificando el número de cajeros según sea necesario. En el segundo banco, todos los clientes esperan en una sola fila y son atendidos por los cajeros disponibles. En el tercer banco hay una fila para cada ventanilla. A continuación se muestran los tiempos de espera (en minutos) específicos de los clientes, que se describen en la figura 3-3:

Banco 1: filas variables	6	6	6
Banco 2: una sola fila	4	7	7
Banco 3: múltiples filas	1	3	14

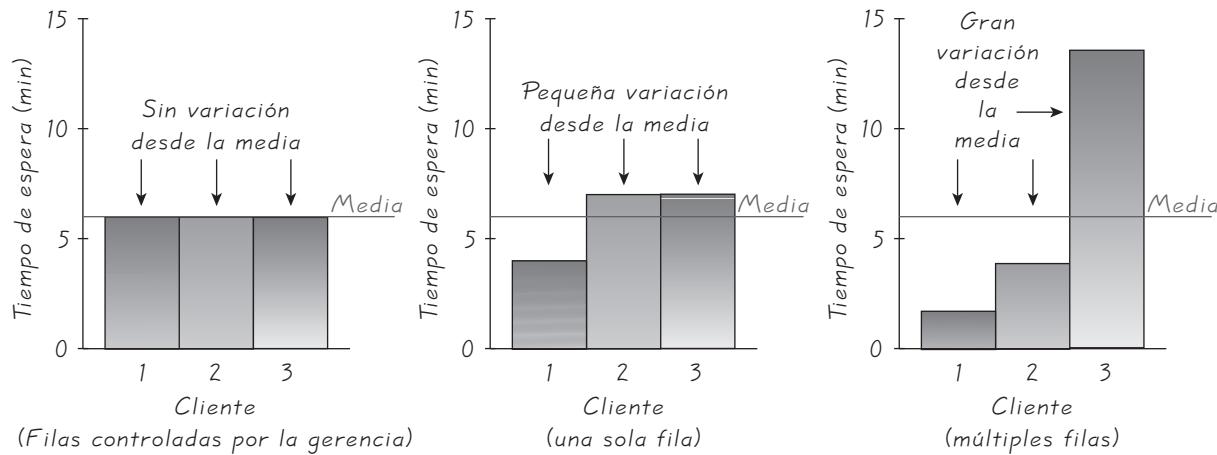


Figura 3-3 Tiempos de espera (en minutos) de clientes de bancos

Si sólo consideramos la media, no podremos reconocer ninguna diferencia entre las tres muestras, ya que todos tienen una media $\bar{x} = 6.0$ min. Sin embargo, debe ser evidente que las muestras difieren mucho con respecto a la *variación* de los tiempos de espera. En el primer banco todos los tiempos de espera son de 6 minutos, por lo que no hay variación. Los tiempos de espera de los clientes en múltiples filas varían mucho más que los tiempos en una sola fila. En esta sección queremos desarrollar la habilidad de medir y comprender este tipo de variaciones.

Ahora estudiaremos algunas formas específicas para medir la variación, de manera que podamos utilizar números específicos en vez de juicios subjetivos. Comenzaremos con el rango.

Rango

Definición

El **rango** de un conjunto de datos es la diferencia entre el valor máximo y el valor mínimo.

$$\text{rango} = (\text{valor máximo}) - (\text{valor mínimo})$$

Para calcular el rango, sólo se resta el valor mínimo del valor máximo. Para los clientes del primer banco, el rango es $6 - 6 = 0$ min. Para los tiempos de espera de la fila única, el rango es $7 - 4 = 3$ min. Los tiempos de espera en múltiples filas tienen un rango de 13 min; este valor más grande sugiere una mayor variación.

Es muy fácil calcular el rango, pero como depende únicamente de los valores máximo y mínimo, no es tan útil como otras medidas de variación que incluyen cada valor.



Confiabilidad y validez

La confiabilidad de los datos se refiere a la consistencia con que éstos se presentan, en tanto que la validez de los datos se refiere a qué tan bien miden lo que se supone que deben medir. La confiabilidad de una prueba de cociente intelectual (CI) se puede evaluar al comparar las puntuaciones de una aplicación con las puntuaciones de una aplicación posterior. Para probar la validez de una prueba de CI, podríamos comparar las puntuaciones de la prueba con otro indicador de la inteligencia, como el rendimiento académico. Muchos críticos consideran que las pruebas de CI son confiables, pero no válidas: arrojan resultados consistentes, pero en realidad no miden la inteligencia.

Desviación estándar de una muestra

La desviación estándar es, por lo general, la medida de variación más importante y útil. Definimos ahora la desviación estándar, pero para comprenderla por completo es necesario estudiar el apartado “Interpretación y comprensión de la desviación estándar”, que aparece posteriormente en esta sección.

Definición

La **desviación estándar** de un conjunto de valores muestrales, es la medida de variación de los valores con respecto a la media. Es un tipo de desviación promedio de los valores con respecto a la media, que se calcula utilizando las fórmulas 3-4 o 3-5. La fórmula 3-5 es sólo una versión diferente de la 3-4, pero algebraicamente son iguales.

$$\textbf{Fórmula 3-4} \quad s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}} \quad \text{desviación estándar de una muestra}$$

$$\textbf{Fórmula 3-5} \quad s = \sqrt{\frac{n\sum(x^2) - (\sum x)^2}{n(n - 1)}} \quad \text{versión abreviada de la fórmula de la desviación estándar de una muestra}$$

Más adelante, en esta sección, analizaremos los fundamentos de esas fórmulas; por ahora le recomendamos que utilice la fórmula 3-4 para resolver algunos ejemplos y después aprenda a calcular los valores de la desviación estándar por medio de su calculadora o de un programa de cómputo. (La mayoría de las calculadoras científicas están diseñadas para introducir una lista de valores y obtener la desviación estándar de manera automática). Mientras tanto, citamos propiedades importantes que son consecuencia de la forma en que se define la desviación estándar.

- La desviación estándar es una medida de variación de todos los valores con respecto a la *media*.
- El valor de la desviación estándar s generalmente es positivo. Sólo es igual a cero cuando todos los valores de los datos son el mismo número. (Nunca es negativa). Además, valores grandes de s implican mayores cantidades de variación.
- El valor de la desviación estándar s puede aumentar de manera drástica con la inclusión de uno o más valores extremos (valores de datos que se encuentran muy lejos de los demás).
- Las unidades de la desviación estándar s (como minutos, pies, libras, etcétera) son las mismas de los datos originales.

Procedimiento para calcular la desviación estándar con la fórmula 3-4

Paso 1: Calcule la media \bar{x} .

Paso 2: Reste la media de cada valor individual para obtener una lista de desviaciones de la forma $(x - \bar{x})$.

Paso 3: Eleve al cuadrado cada una de las diferencias obtenidas en el paso 2. Esto produce números de la forma $(x - \bar{x})^2$.

Paso 4: Sume todos los cuadrados obtenidos en el paso 3. Éste es el valor de $\sum(x - \bar{x})^2$.

continúa

Paso 5: Divida el total del paso 4 entre el número ($n - 1$); es decir, 1 menos que el total de valores presentes.

Paso 6: Calcule la raíz cuadrada del resultado del paso 5.

EJEMPLO Uso de la fórmula 3-4 Utilice la fórmula 3-4 para calcular la desviación estándar de los tiempos de espera de las múltiples filas. Esos tiempos (en minutos) son 1, 3 y 14.

SOLUCIÓN Seguiremos los seis pasos de este proceso. Remítase a los pasos y a la tabla 3-3, que presenta los cálculos detallados.

Paso 1: Obtenga la media de 6.0 sumando los valores y después dividiendo entre el número de valores:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{18}{3} = 6.0 \text{ min}$$

Paso 2: Reste la media de 6.0 de cada valor para obtener los valores de $(x - \bar{x})$: $-5, -3, 8$.

Paso 3: Eleve al cuadrado cada valor obtenido en el paso 2 para obtener los valores de $(x - \bar{x})^2$: $25, 9, 64$.

Paso 4: Sume todos los valores anteriores para obtener el valor de

$$\Sigma(x - \bar{x})^2 = 98.$$

Paso 5: Con $n = 3$ valores, divida entre 1 menos que 3:

$$\frac{98}{2} = 49.0$$

Paso 6: Calcule la raíz cuadrada de 49.0. La desviación estándar es

$$\sqrt{49.0} = 7.0 \text{ min}$$

De manera ideal ahora interpretaríamos el significado de los resultados, pero esas interpretaciones se discutirán más adelante en esta sección.

Tabla 3-3 Cálculo de la desviación estándar

x	$x - \bar{x}$	$(x - \bar{x})^2$
1	-5	25
3	-3	9
14	8	64
Totales: 18		98
$\bar{x} = \frac{18}{3} = 6.0 \text{ min}$		$s = \sqrt{\frac{98}{3 - 1}} = \sqrt{49} = 7.0 \text{ min}$



¿Dónde están los bateadores de 0.400?

El último beisbolista que bateó más de 0.400 fue Ted Williams, quien promedió 0.406 en 1941. Hubo promedios por arriba de 0.400 en 1876, 1879, 1887, 1894, 1895, 1896, 1897, 1899, 1901, 1911, 1920, 1922, 1924, 1925 y 1930, pero ninguno desde 1941. ¿Será que ya no existen grandes bateadores? Stephen Jay Gould, de la Universidad de Harvard, señaló que el promedio de bateo medio se ha mantenido estable en 0.260 durante aproximadamente 100 años, pero la desviación estándar disminuyó de 0.049 en la década de 1870 hasta 0.031 en la actualidad. Él argumenta que las estrellas de hoy son tan buenas como las del pasado, pero que los mejores lanzadores actuales mantienen promedios por debajo de 0.400.

EJEMPLO Uso de la fórmula 3-5 En el ejemplo anterior se utilizó la fórmula 3-4 para calcular la desviación estándar de los tiempos de espera en el banco con múltiples filas. Utilice el mismo conjunto de datos (1, 3, 14) y calcule la desviación estándar con la fórmula 3-5.

SOLUCIÓN La fórmula 2.5 requiere que primero encontremos valores para n , Σx , and Σx^2 .

$$n = 3 \quad (\text{ya que existen tres valores en la muestra})$$

$$\Sigma x = 18 \quad (\text{se obtiene sumando los tres valores muestrales: } 1 + 3 + 14 = 18)$$

$$\Sigma x^2 = 206 \quad (\text{se obtiene al sumar los cuadrados de los valores muestrales, } 1^2 + 3^2 + 14^2 = 206)$$

Si usamos la fórmula 3-5, obtenemos

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}} = \sqrt{\frac{3(206) - (18)^2}{3(3 - 1)}} = \sqrt{\frac{294}{6}} = 7.0 \text{ min}$$

Una actividad adecuada es detenerse aquí y calcular la desviación estándar de los tiempos de espera de 4, 7 y 7 min (para la fila única). Siga los mismos procedimientos utilizados en los dos ejemplos anteriores y verifique que $s = 1.7$ min. (También será importante desarrollar la habilidad para obtener valores de desviaciones estándar con el uso de calculadoras y de programas de cómputo). Aun cuando las interpretaciones de tales desviaciones estándar se analizarán posteriormente, ahora podemos compararlas para observar que la desviación estándar de los tiempos de espera de una sola fila (1.7 min) es mucho menor que la desviación estándar de las múltiples filas (7.0 min). Esto apoya nuestra conclusión subjetiva de que los tiempos de espera con la fila única tienen mucho menor variación que los tiempos de espera con múltiples filas. El banco que cuenta con un gerente preocupado, quien controla cuidadosamente los tiempos de espera, tiene una desviación estándar de 0 min, que es la más baja.

Desviación estándar de una población

En nuestra definición de la desviación estándar nos referimos a datos *muestrales*. Para calcular la desviación estándar σ (sigma minúscula) de una población, se utiliza una fórmula ligeramente diferente: en vez de dividir entre $n - 1$, se divide entre el tamaño N de la población, como en la siguiente expresión:

$$\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}} \quad \text{desviación estándar de la población}$$

Como generalmente usamos datos muestrales, con frecuencia utilizaremos la fórmula 3-4, en la que dividimos entre $n - 1$. Muchas calculadoras dan tanto la desviación estándar muestral como la desviación estándar poblacional, pero usan una gran variedad de notaciones diferentes. Asegúrese de identificar la notación que utiliza su calculadora, de manera que pueda obtener el resultado correcto.

Varianza de una muestra y una población

Usamos el término *variación* como una descripción general de la cantidad que varían los valores entre sí. (En ocasiones se aplica el término *dispersión* en vez de *variación*). El término *varianza* se refiere a una definición específica.

Definición

La **varianza** de un conjunto de valores es una medida de variación igual al cuadrado de la desviación estándar.

Varianza muestral: s^2 el cuadrado de la desviación estándar s .

Varianza poblacional: σ^2 el cuadrado de la desviación estándar poblacional σ .

Se dice que la varianza muestral s^2 es un **estimador sin sesgo** de la varianza poblacional σ^2 , lo que significa que los valores de s^2 tienden a igualar el valor de σ^2 , en lugar de tender, de manera sistemática, a sobreestimar o subestimar σ^2 . Por ejemplo, considere una prueba de cociente intelectual (CI) diseñada de tal forma que tiene una varianza de 225. Si usted repite el proceso de elegir aleatoriamente 100 sujetos, aplicarles la prueba y calcular la varianza muestral s^2 en cada caso, las varianzas muestrales que obtendrá tenderán a concentrarse alrededor de 225, que es la varianza de la población.

EJEMPLO Cálculo de la varianza En el ejemplo anterior empleamos los tiempos de espera de 1, 3 y 14 min, y encontramos una desviación estándar $s = 7.0$ min. Calcule la varianza de esa misma muestra.

SOLUCIÓN Ya que la varianza es el cuadrado de la desviación estándar, obtenemos los resultados que se presentan abajo. Observe que las unidades de los valores de los datos están dadas en minutos y que la desviación estándar es de 7.0 minutos, pero la varianza está dada en unidades de min^2 .

$$\text{Varianza muestral} = s^2 = 7.0^2 = 49.0 \text{ min}^2$$

La varianza es un estadístico importante que se utiliza en algunos métodos estadísticos relevantes, como el análisis de varianza, que se explica el capítulo 12. Para nuestros propósitos, la varianza tiene una gran desventaja: *las unidades de la varianza son diferentes de las unidades del conjunto original de datos*. Por ejemplo, si los tiempos de espera originales de los clientes están dados en minutos, las unidades de varianza están dadas en minutos cuadrados (min^2). ¿Qué es un minuto cuadrado? Como la varianza utiliza unidades distintas, es sumamente difícil comprenderla si la relacionamos con el conjunto original de datos. Por esta propiedad, nos enfocaremos en la desviación estándar cuando tratemos de comprender la variación más adelante en este capítulo.

Ahora presentamos la notación y la regla de redondeo que utilizamos.

Notación

s = desviación estándar *muestral*

s^2 = varianza *muestral*

σ = desviación estándar *poblacional*

σ^2 = varianza *poblacional*

Nota: Los artículos de las revistas y los reportes científicos suelen usar DE (o bien, SD, por *standard deviation* en inglés) para la desviación estándar y VAR para la varianza.

Regla del redondeo

Usamos la misma regla de redondeo que se empleó en la sección 3-2:

Aumentamos una posición decimal a las que había en los datos originales.

Redondee únicamente la respuesta final, no los valores a la mitad de un cálculo. (Si se vuelve absolutamente necesario redondear a la mitad, deberemos llevar al menos el doble de posiciones decimales de las que se utilizarán en la respuesta final).

Parte 2: Más allá de los aspectos básicos de la variación

Interpretación y comprensión de la desviación estándar

Este apartado es sumamente importante, pues ahora trataremos de darle sentido a la desviación estándar. Primero, debemos comprender con claridad que la desviación estándar mide la *variación* entre los valores. Los valores cercanos producirán una desviación estándar pequeña, mientras que los valores muy dispersos producirán una desviación estándar más grande.

Una herramienta rudimentaria pero sencilla para comprender la desviación estándar es la **regla práctica del intervalo**, que se basa en el principio de que, para muchos conjuntos de datos, la vasta mayoría (tanto como el 95%) de los valores muestrales se ubican dentro de dos desviaciones estándar a partir de la media. (Es posible mejorar la precisión de esta regla si tomamos en cuenta factores como el tamaño de la muestra y la naturaleza de la distribución, pero preferimos sacrificar precisión en aras de la sencillez. Además, podríamos usar tres o incluso cuatro desviaciones estándar en vez de 2, lo cual constituye una decisión un tanto arbitraria. Sin embargo, deseamos una regla sencilla que nos ayude a interpretar los valores de las desviaciones estándar; métodos posteriores producirán resultados más precisos).

Regla práctica del intervalo

Para estimar el valor de la desviación estándar s : Para obtener un estimado de la desviación estándar, utilice

$$s \approx \frac{\text{rango}}{4}$$

donde el rango = (valor máximo) – (valor mínimo).

Para interpretar un valor conocido de la desviación estándar: Si se conoce la desviación estándar s , utilícela para calcular estimados de los valores muestrales mínimos y máximos “comunes” por medio de

valor mínimo “común” = (media) – $2 \times$ (desviación estándar)

valor máximo “común” = (media) + $2 \times$ (desviación estándar)

Cuando calcule una desviación estándar por medio de las fórmulas 3-4 o 3-5, la regla práctica del intervalo resulta útil para verificar el resultado, pero debe estar

consciente de que, si bien la aproximación nos acerca a la respuesta, puede tener un error considerable.

EJEMPLO Edades de las mejores actrices Utilice la regla práctica del intervalo para calcular un estimado de la desviación estándar con la muestra de las 76 edades de las actrices que ganaron un Óscar en la categoría de mejor actriz. Las edades se presentan en la tabla 2-1, que viene incluida en el problema del capítulo 2.

SOLUCIÓN Al emplear la regla práctica del intervalo para estimar la desviación estándar de datos muestrales, calculamos el rango y lo dividimos entre 4. Si observamos la lista de las edades de las actrices, notaremos que el valor máximo es de 80 y el valor mínimo de 21; por lo tanto, la regla práctica del intervalo para estimar la desviación estándar s se utiliza de la siguiente manera:

$$s \approx \frac{\text{rango}}{4} = \frac{80 - 21}{4} = \frac{59}{4} = 14.75 \approx 15$$

INTERPRETACIÓN Este resultado se acerca al valor correcto de 11.1, que se obtiene al calcular el valor exacto de la desviación estándar con las fórmulas 3-4 o 3-5, aunque el resultado de 15 se aleja de la desviación estándar real de forma considerable. Esto demuestra que la regla práctica del intervalo produce un estimado “burdo” que puede alejarse mucho del resultado real.

El siguiente ejemplo es particularmente importante como ilustración de una forma de *interpretar* el valor de una desviación estándar.

EJEMPLO Pulso cardiaco de mujeres Resultados anteriores de la encuesta sobre salud National Health Survey sugieren que el pulso cardiaco (latidos por minuto) tiene una media de 76.0 y una desviación estándar de 12.5. Utilice la regla práctica del intervalo para calcular las frecuencias máxima y mínima “comunes”. (Estos resultados podrían ayudar a un médico a identificar pulsos cardíacos “poco comunes” asociados con alguna enfermedad). Luego determine si un pulso cardiaco de 110 sería considerado “poco común”.

SOLUCIÓN Con una media de 76.0 y una desviación estándar de 12.5, utilizamos la regla práctica del intervalo para calcular los pulsos cardíacos mínimo y máximo comunes de la siguiente manera:

$$\begin{aligned}\text{valor mínimo “común”} &= (\text{media}) - 2 \times (\text{desviación estándar}) \\ &= 76.0 - 2(12.5) = 51 \text{ latidos por minuto}\end{aligned}$$

$$\begin{aligned}\text{valor máximo “común”} &= (\text{media}) + 2 \times (\text{desviación estándar}) \\ &= 76.0 + 2(12.5) = 101 \text{ latidos por minuto}\end{aligned}$$

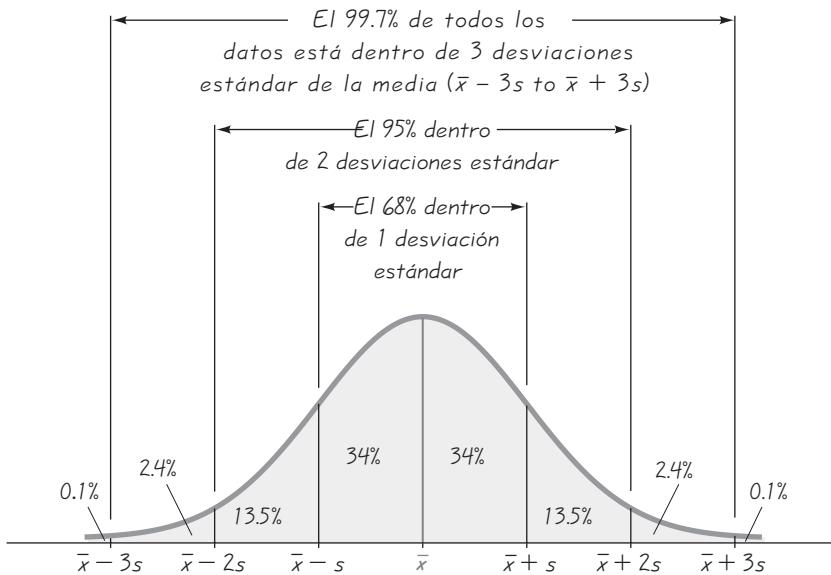
INTERPRETACIÓN Con base en estos resultados, esperamos que la mujer común tenga un pulso cardiaco de entre 51 y 101 latidos por minuto. Puesto que 110 latidos por minuto no cae dentro de esos límites, ese valor sería considerado poco común. Con un pulso cardiaco de 110, un médico trataría de encontrar la razón de esta lectura poco común.



Más acciones, menos riesgo

En su libro *Investments*, los autores Zvi Bodie, Alex Kane y Alan Marcus afirman que “la desviación estándar promedio de los rendimientos de carteras compuestas por un solo tipo de acciones fue de 0.554. El riesgo promedio disminuye rápidamente cuando aumenta el número de acciones incluidas en la cartera”. También señalan que con 32 acciones la desviación estándar es de 0.325, lo que indica mucho menos variación y riesgo. Los autores destacan que con sólo unas cuantas acciones, una cartera tiene alto grado de riesgo “específico de una empresa”, lo que significa que el riesgo puede atribuirse a la escasa cantidad de acciones implicadas. Con más de 30 acciones hay muy poco riesgo específico de una empresa; en esa situación, casi todo el riesgo es “riesgo de mercado”, atribuible al mercado global de acciones. Además, señalan que estos principios son “sólo una aplicación de la bien conocida ley de promedios”.

Figura 3-4
La regla empírica



Regla empírica para datos con distribución normal (o 68-95-99.7)

Otra regla útil para interpretar los valores de una desviación estándar es la **regla empírica**. Esta regla establece que las siguientes propiedades se aplican a *conjuntos de datos con una distribución aproximadamente normal*. (Véase la figura 3-4).

- Aproximadamente el 68% de todos los valores están dentro de una desviación estándar de la media.
- Aproximadamente el 95% de todos los valores están dentro de 2 desviaciones estándar de la media.
- Aproximadamente el 99.7% de todos los valores están dentro de 3 desviaciones estándar de la media.

EJEMPLO Puntuaciones de CI Las puntuaciones de CI tienen una distribución normal, con una media de 100 y una desviación estándar de 15. ¿Qué porcentaje de las puntuaciones se ubican entre 70 y 130?

SOLUCIÓN La clave para resolver este problema consiste en reconocer que 70 y 130 están exactamente a 2 desviaciones estándar de la media de 100, como se indica abajo.

$$2 \text{ desviaciones estándar} = 2s = 2(15) = 30$$

Por lo tanto, 2 desviaciones estándar de la media equivalen a

$$\begin{aligned} 100 - 30 &= 70 \\ 100 + 30 &= 130 \end{aligned}$$

La regla empírica nos indica que aproximadamente el 95% de todos los valores están dentro de dos desviaciones estándar de la media, de manera que el 95% de todas las puntuaciones de CI se encuentran entre 70 y 130.

Sugerencia: Las dificultades para aplicar la regla empírica suelen surgir de la confusión al interpretar frases tales como “dentro de 2 desviaciones estándar de la media”. Deténgase aquí y revise el ejemplo anterior hasta que el significado de la frase quede claro. Además, observe las siguientes interpretaciones generales de ese tipo de frases.

Frase	Significado
Dentro de 1 desviación estándar de la media	Entre $(\bar{x} - s)$ y $(\bar{x} + s)$
Dentro de 2 desviaciones estándar de la media	Entre $(\bar{x} - 2s)$ y $(\bar{x} + 2s)$
Dentro de 3 desviaciones estándar de la media	Entre $(\bar{x} - 3s)$ y $(\bar{x} + 3s)$

Un tercer concepto útil para comprender o interpretar el valor de una desviación estándar es el **teorema de Chebyshev**. La regla empírica anterior se aplica sólo a conjuntos de datos con una distribución normal. El teorema de Chebyshev, en vez de limitarse a conjuntos de datos con distribuciones normales, se aplica a *cualquier* conjunto de datos, pero sus resultados son muy aproximados. Como los resultados son límites inferiores (“al menos”), este teorema tiene una utilidad limitada.

Teorema de Chebyshev

La proporción (o fracción) de cualquier conjunto de datos que está dentro de K desviaciones estándar a partir de la media siempre es *al menos* $1 - 1/K^2$, donde K es cualquier número positivo mayor que 1. Para $K = 2$ y $K = 3$ tenemos las siguientes aseveraciones:

- Al menos $3/4$ (o el 75%) de todos los valores están dentro de 2 desviaciones estándar de la media.
- Al menos $8/9$ (o el 89%) de todos los valores están dentro de 3 desviaciones estándar de la media.

EJEMPLO Puntuaciones de CI Las puntuaciones de CI tienen una media de 100 y una desviación estándar de 15. ¿Qué podemos concluir a partir del teorema de Chebyshev?

SOLUCIÓN Si aplicamos el teorema de Chebyshev con una media de 100 y una desviación estándar de 15, podemos llegar a las siguientes conclusiones:

- Al menos $3/4$ (o el 75%) de las puntuaciones de CI están dentro de 2 desviaciones estándar de la media (entre 70 y 130).
- Al menos $8/9$ (o el 89%) de las puntuaciones de CI están a 3 desviaciones estándar de la media (entre 55 y 145).

Cuando intentemos darle un significado al valor de una desviación estándar, debemos usar uno o más de los tres conceptos anteriores. Para comprender aún mejor la naturaleza de la desviación estándar, consideraremos los fundamentos subyacentes que conducen a la fórmula 3-4, que es la base de su definición. (La fórmula 3-5 es sencillamente otra versión de la fórmula 3-4, derivada de tal manera que los cálculos aritméticos pueden simplificarse).

Fundamentos de la desviación estándar

La desviación estándar de un conjunto de datos muestrales se define con las fórmulas 3-4 y 3-5, las cuales son equivalentes en el sentido de que siempre producen el mismo resultado. La fórmula 3-4 tiene la ventaja de reforzar el concepto de que la desviación estándar es un tipo de desviación promedio. La fórmula 3-5 tiene la ventaja de ser más fácil de usar cuando hay que calcular desviaciones estándar por

nuestra cuenta. La fórmula 3-5 también elimina los errores de redondeo intermedios que se introducen en la fórmula 3-4, cuando no se utiliza el valor exacto de la media. La fórmula 3-5 se aplica en calculadoras y programas, ya que requiere sólo de tres lugares de memoria (para n , Σx y Σx^2) en vez de un lugar de memoria para cada valor del conjunto de datos.

¿Para qué definir una medida de variación en la forma descrita por la fórmula 3-4? Al medir la variación en un conjunto de datos muestrales, parece lógico iniciar con las cantidades individuales con las que los valores se desvían de la media. Para un valor particular x , la cantidad de **desviación** es $x - \bar{x}$, que es la diferencia entre el valor individual x y la media. Para los tiempos de espera de 1, 3 y 14, la media es 6.0, de manera que las desviaciones de la media son -5 , -3 y 8 . Sería bueno combinar de alguna forma esas desviaciones en un solo valor colectivo. La simple suma de las desviaciones no funciona, ya que la suma siempre será cero. Para obtener un estadístico que mida la variación (en vez de que siempre sea cero), necesitamos evitar la cancelación de números positivos y negativos. Un método consiste en sumar valores absolutos, como en $\sum |x - \bar{x}|$. Si calculamos la media de esta suma, obtendremos la **desviación media absoluta (DMA)**, que es la distancia media de los datos con respecto a la media.

$$\text{desviación media absoluta} = \frac{\sum |x - \bar{x}|}{n}$$

Puesto que los tiempos de espera de 1, 3 y 14 tienen desviaciones de -5 , -3 y 8 , la desviación media absoluta es $(5 + 3 + 8)/3 = 16/3 = 5.3$.

¿Por qué no utilizar la desviación media absoluta? Como la desviación media absoluta requiere que usemos valores absolutos, emplea una operación que no es algebraica. (Las operaciones algebraicas incluyen la suma, la multiplicación, la raíz cuadrada y la elevación a potencias enteras o fraccionarias, pero el valor absoluto no está incluido). El uso de valores absolutos crea problemas algebraicos en los métodos inferenciales de la estadística. Por ejemplo, la sección 9-3 presenta un método para hacer inferencias acerca de las medias de dos poblaciones, y ese método se construye alrededor de una propiedad de adición de las varianzas, pero la desviación media absoluta no posee tal propiedad de adición. (He aquí una versión simplificada de la propiedad de adición de la varianza: si se tienen dos poblaciones independientes y si selecciona aleatoriamente un valor de cada población y se suman, esas sumas tendrán una varianza que es igual a la suma de las varianzas de las dos poblaciones). La misma propiedad de adición subyace en los fundamentos de la regresión, que se presenta en el capítulo 10, y el análisis de varianza, que se estudia en el capítulo 12. La desviación media absoluta carece de esta importante propiedad de adición. Además, el valor de la media absoluta está *sesgado*, lo que significa que cuando se calculan valores de media absoluta de muestras, no se tiende a igualar el valor medio absoluto de la población. En contraste, la desviación estándar utiliza únicamente operaciones algebraicas. Puesto que se basa en la raíz cuadrada de una suma de cuadrados, la desviación estándar se asemeja a las fórmulas de distancia que se emplean en álgebra. Existen muchos ejemplos en los que un procedimiento estadístico se basa en una suma de cuadrados similar. Por lo tanto, en vez de emplear valores absolutos, obtenemos una mejor medida de variación si logramos que ninguna de las desviaciones ($x - \bar{x}$) sea negativa, elevando todas al cuadrado; este método conduce a la desviación estándar. Por esas razones, las calculadoras científicas suelen incluir

una función para la desviación estándar, pero casi nunca para la desviación media absoluta.

¿Por qué dividir entre $n - 1$? Despues de obtener todos los valores individuales de $(x - \bar{x})^2$, los combinamos calculando su suma y luego obtenemos un promedio dividiéndola entre $n - 1$. Dividimos entre $n - 1$ porque existen solamente $n - 1$ valores independientes. Es decir, con una media dada, sólo a $n - 1$ valores se les puede asignar un número con libertad, antes que se determine el último valor. Vea el ejercicio 38, que ofrece números concretos, los cuales ilustran cómo la división entre $n - 1$ es mejor que la división entre n . Este ejercicio demuestra que si s^2 se definiera con la división entre n , de forma sistemática subestimaría el valor de σ^2 , por lo que lo compensamos al incrementar su valor general haciendo que su denominador sea más pequeño (usando $n - 1$ en vez de n). El ejercicio 38 demuestra cómo la división entre $n - 1$ provoca que la varianza muestral s^2 iguale el valor de la varianza poblacional σ^2 , mientras que la división entre n causa que la varianza muestral s^2 subestime el valor de la varianza poblacional σ^2 .

El paso 6 del procedimiento para calcular la desviación estándar implica sacar una raíz cuadrada. Esto se hace para compensar la elevación al cuadrado que se realizó en el paso 3. Una consecuencia importante de la obtención de la raíz cuadrada es que *la desviación estándar tiene las mismas unidades de medición que los valores originales*. Por ejemplo, si el tiempo de espera de los clientes está dado en minutos, la desviación estándar de tales tiempos también estará dada en minutos. Si nos detuviéramos en el paso 5, el resultado estaría dado en unidades de "minutos cuadrados", que es un concepto abstracto sin relación directa con la realidad.

Comparación de la variación en diferentes poblaciones

Anteriormente afirmamos que, como las unidades de la desviación estándar son las mismas que las unidades de los datos originales, es más fácil comprender la desviación estándar que la varianza. Sin embargo, esta misma propiedad dificulta comparar la variación de valores tomados de distintas poblaciones. Como el resultado es un valor libre de unidades de medida específicas, el *coeficiente de variación* resuelve esta desventaja.

Definición

El **coeficiente de variación (CV)** de un conjunto de datos muestrales o poblacionales, expresado como porcentaje, describe la desviación estándar en relación con la media. El coeficiente de variación está dado de la siguiente forma:

$$\begin{array}{ll} \text{Muestra} & \text{Población} \\ CV = \frac{s}{\bar{x}} \cdot 100\% & CV = \frac{\sigma}{\mu} \cdot 100\% \end{array}$$

EJEMPLO Estatura y peso de hombres Si utilizamos los datos de la muestra de estaturas y pesos de los 40 hombres del conjunto de datos 1 del apéndice B, obtendremos los estadísticos que aparecen en la siguiente tabla. Calcule el coeficiente de variación de las estaturas, después calcule el coeficiente de variación de los pesos; finalmente, compare los dos resultados.

continúa

	Media (\bar{x})	Desviación estándar (s)
Estatura	68.34 in.	3.02 in.
Peso	172.55 lb	26.33 lb

SOLUCIÓN Tenemos estadísticos muestrales, así que los dos coeficientes de variación se obtienen de la siguiente manera:

$$\text{Estaturas: } CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{3.02 \text{ in.}}{68.34 \text{ in.}} \cdot 100\% = 4.42\%$$

$$\text{Pesos: } CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{26.33 \text{ lb}}{172.55 \text{ lb}} \cdot 100\% = 15.26\%$$

Aun cuando la diferencia en unidades imposibilita la comparación de la desviación estándar de 3.02 pulgadas con la desviación estándar de 26.33 libras, es posible comparar los coeficientes de variación, que carecen de unidades. Podemos ver que las estaturas (con $CV = 4.42\%$) tienen una variación considerablemente menor que los pesos (con $CV = 15.26\%$). Lo anterior tiene sentido, ya que por lo general vemos que los pesos de los hombres varían mucho más que sus estaturas. Por ejemplo, es muy raro encontrar un hombre adulto que mida el doble que otro, pero es mucho más común ver a un hombre que pese el doble que otro.

Después de estudiar esta sección, debería quedar claro que la desviación estándar es una medida de variación entre valores. A partir de datos muestrales, usted debería ser capaz de calcular el valor de la desviación estándar, así como de interpretar los valores de las desviaciones estándar que calcule. También debería saber que para conjuntos de datos comunes, es inusual que un valor difiera de la media por más de dos o tres desviaciones estándar.

Uso de la tecnología

STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus pueden usarse para hacer los

importantes cálculos de esta sección. Use los mismos procedimientos que se describen al final de la sección 3-2.



3-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Variación.** ¿Por qué la desviación estándar se considera una medida de *variación*? Describa con sus propias palabras las características de un conjunto de datos medido con la desviación estándar.
- 2. Comparación de la variación.** ¿Cuáles datos cree usted que tengan mayor variación: las puntuaciones de CI de 30 estudiantes de un curso de estadística o las puntuaciones de CI de 30 individuos que ven una película? ¿Por qué?

3. ¿Valor infrecuente? Una profesora de estadística aplica un examen que tiene una media de 50 y una desviación estándar de 10. (Ella no está utilizando un examen típico con una puntuación máxima de 100 y promete hacer una curva con las calificaciones). Un estudiante obtiene una calificación de 85 en el examen. En este contexto, ¿la calificación de 85 es “poco común”? ¿Por qué?

4. ¿Enunciado correcto? En el libro *How to Lie with Charts*, Gerald E. Jones escribe que “la desviación estándar suele definirse como más o menos la diferencia entre la puntuación más alta y la media, y la puntuación más baja y la media. Por ejemplo, si la media es 1, el valor más alto es 3 y el valor más bajo es –1, la desviación estándar es ± 2 .”. ¿Es correcto este enunciado? ¿Por qué?

En los ejercicios 5 a 12, calcule el rango, la varianza y la desviación estándar de los datos muestrales. (Se usarán los mismos datos de la sección 3-2, donde se calcularon medidas de tendencia central. Aquí se calculan las medidas de variación). Asimismo, responda las preguntas que se plantean.

5. Percepción del tiempo. Algunos estudiantes de estadística participaron en un experimento que intentaba probar su capacidad para determinar el transcurso de 1 minuto (o 60 segundos). A continuación se presentan los resultados en segundos. Identifique al menos una buena razón por la que la desviación estándar de esta muestra no sería un buen estimado de la desviación estándar de la población de adultos.

53 52 75 62 68 58 49 49

6. Cereal. Un nutriólogo obtiene las cantidades de azúcar (en centigramos) de 100 centigramos (o 1 gramo) de 10 cereales diferentes, incluyendo Cheerios, Corn Flakes, Fruit Loops y otros siete. Las cantidades se listan abajo. ¿Es probable que la desviación estándar de esos valores sea un buen estimado de la desviación estándar de la cantidad de azúcar en cada gramo de cereal que consume la población de todos los estadounidenses que comen cereal? ¿Por qué?

3 24 30 47 43 7 47 13 44 39

7. Fenotipos de guisantes. Se realizó un experimento para determinar si una deficiencia de dióxido de carbono en la tierra afecta los fenotipos de los guisantes (o chícharos). A continuación se indican los códigos de los fenotipos: 1 = amarillo claro, 2 = verde claro, 3 = amarillo rugoso y 4 = verde rugoso. ¿Es posible obtener medidas de variación para estos valores? ¿Los resultados tienen algún sentido?

2 1 1 1 1 1 4 1 2 2 1 2 3 3 2 3 1 3 1 3 1 3 2 2

8. Géiser Old Faithful. Abajo se indican los intervalos (en minutos) entre las erupciones del géiser Old Faithful en el Parque Nacional Yellowstone. Con base en los resultados, ¿es poco común un intervalo de 100 minutos?

98 92 95 87 96 90 65 92 95 93 98 94

9. Mediciones de la presión sanguínea. Catorce estudiantes del segundo año de medicina de Bellevue Hospital midieron la presión sanguínea de la misma persona. A continuación se listan las lecturas sistólicas (en mmHg). Si la presión sanguínea del sujeto permanece constante y los estudiantes de medicina aplican correctamente la misma técnica de medición, ¿cuál debería ser el valor de la desviación estándar?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

10. Temperaturas corporales. Investigadores de la Universidad de Maryland reunieron lecturas de la temperatura corporal de una muestra de adultos; a continuación se listan ocho de esas medidas (en grados Fahrenheit). Con base en esos resultados, ¿es poco

común una temperatura corporal de 104.0°F? ¿Por qué? ¿Qué debería hacerse con alguien que tiene una temperatura corporal de 104.0°F?

98.6 98.6 98.0 98.0 99.0 98.4 98.4 98.4

- 11. Moscas de la fruta.** Más adelante se listan las longitudes (en milímetros) del tórax de una muestra de moscas de la fruta machos. Si nos enteramos de que las medidas se obtuvieron de moscas que volaban sobre una manzana que se encontraba en una mesa en Pocattelo, ¿la desviación estándar sirve como un estimado razonable de la desviación estándar de todas las moscas de la fruta que residen en Estados Unidos?

0.72 0.90 0.84 0.68 0.84 0.90 0.92 0.84 0.64 0.84 0.76

- 12. Ingreso personal.** A continuación se indican las cantidades de ingreso personal per cápita (en dólares) de las primeras cinco entidades de Estados Unidos, ordenadas alfabéticamente: Alabama, Alaska, Arizona, Arkansas y California (datos del U.S. Bureau of Economic Analysis). Suponga que se incluyen las 45 cantidades de los otros estados y que la desviación estándar de las 50 entidades es de \$4337. ¿La desviación estándar de las cantidades de ingreso personal per cápita para todos los individuos de Estados Unidos es necesariamente de \$4337? ¿Por qué?

\$25,128 \$32,151 \$26,183 \$23,512 \$32,996

En los ejercicios 13 a 20, calcule el rango, la varianza y la desviación estándar de cada una de las dos muestras y luego compare los dos conjuntos de resultados. (En la sección 3-2 se emplearon los mismos datos).

- 13. Llueven gatos.** A veces se utiliza la estadística para comparar o identificar autores de distintos trabajos. A continuación se incluye una lista con la longitud de las primeras 20 palabras que escribió Tennessee Williams en el prefacio de *The Cat on a Hot Tin Roof*, junto con las primeras 20 palabras de *The Cat in the Hat*, del doctor Seuss. ¿Parece haber una diferencia en la variación?

Cat on a Hot Tin Roof: 2 6 2 2 1 4 4 2 4 2 3 8 4 2 2 7 7 2 3 11

The Cat in the Hat: 3 3 3 3 5 2 3 3 3 2 4 2 2 3 2 3 5 3 4 4

- 14. Edades de polizones.** El *Queen Mary* navegaba entre Inglaterra y Estados Unidos; en ocasiones se encontraron polizones a bordo. A continuación se indican las edades (en años) de los polizones que iban con rumbo al este y los que iban al oeste (datos de Cunard Steamship Co., Ltd.). Compare la variación de los dos conjuntos de datos.

Hacia el este: 24 24 34 15 19 22 18 20 20 17

Hacia el oeste: 41 24 32 26 39 45 24 21 22 21

- 15. Exactitud del pronóstico del clima.** En un análisis de la exactitud del pronóstico del clima se comparan las temperaturas máximas reales con las temperaturas máximas pronosticadas un día anterior y con las temperaturas máximas pronosticadas cinco días antes. Más abajo se señalan los errores entre las temperaturas pronosticadas y las temperaturas máximas reales para días consecutivos en el condado Dutchess, Nueva York. ¿La desviación estándar sugiere que las temperaturas pronosticadas un día antes son más exactas que las pronosticadas cinco días antes, como se esperaría?

(máxima real) – (máxima pronosticada un día antes):	2 2 0 0 -3 -2 1 -2 8 1 0 -1 0 1
--	--

(máxima real) – (máxima pronosticada cinco días antes):	0 -3 2 5 -6 -9 4 -1 6 -2 -2 -1 6 -4
--	--

- 16. Efecto de tratamiento.** Investigadores de la Universidad de Pennsylvania realizaron experimentos con álamos. A continuación se indican los pesos (en kg) de los álamos

que no recibieron tratamiento y de los álamos tratados con fertilizantes y riego. ¿Parece que existe una diferencia entre las dos desviaciones estándar?

Sin tratamiento: 0.15 0.02 0.16 0.37 0.22

Con fertilizantes y riego: 2.03 0.27 0.92 1.07 2.38

- 17. Tiempos de espera de clientes.** A continuación se presentan los tiempos de espera (en minutos) de los clientes del banco Jefferson Valley (donde todos los clientes se forman en una sola fila) y del banco Providence (donde los clientes esperan en filas individuales, en tres ventanillas diferentes). Compare la variación de los dos conjuntos de datos.

Jefferson Valley (una sola fila): 6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7

Providence (filas individuales): 4.2 5.4 5.8 6.2 6.7 7.7 7.7 8.5 9.3 10.0

- 18. Coca-Cola regular/Coca-Cola dietética.** Los siguientes son los pesos (en libras) de muestras del contenido de latas de Coca-Cola regular y Coca-Cola dietética. Compare la variación de los dos conjuntos de datos.

Regular: 0.8192 0.8150 0.8163 0.8211 0.8181 0.8247

Dietética: 0.7773 0.7758 0.7896 0.7868 0.7844 0.7861

- 19. Reflexiones sobre peniques.** Los peniques o centavos estadounidenses acuñados antes de 1983 tienen un 97% de cobre y un 3% de zinc, en tanto que los acuñados después de 1983 tienen un 3% de cobre y un 97% de zinc. A continuación se indican los pesos (en gramos) de peniques acuñados antes y después de 1983. (El autor obtuvo estos datos). ¿Parece que existe una diferencia considerable en las desviaciones estándar?

Antes de 1983:	3.1582	3.0406	3.0762	3.0398	3.1043	3.1274	3.0775
	3.1038	3.1086	3.0586				

Después de 1983:	2.5113	2.4907	2.5024	2.5298	2.4950	2.5127	2.4998
	2.4848	2.4823	2.5163				

- 20. IMC y género.** Es bien sabido que los hombres tienden a pesar más y a ser más altos que las mujeres. El índice de masa corporal (IMC) es una medida que se basa en el peso y en la estatura. A continuación se muestran los valores de IMC de hombres y mujeres elegidos de manera aleatoria. ¿Parece existir una diferencia en la variación entre los dos conjuntos de datos?

Hombres: 23.8 23.2 24.6 26.2 23.5 24.5 21.5 31.4 26.4 22.7 27.8 28.1

Mujeres: 19.6 23.8 19.6 29.1 25.2 21.4 22.0 27.5 33.5 20.6 29.9 17.7

Conjuntos grandes de datos. En los ejercicios 21 a 24, remítase al conjunto de datos en el apéndice B. Con un programa de cómputo o una calculadora obtenga el rango, la varianza y la desviación estándar de cada muestra y luego compare los resultados.

- 21. Exactitud del pronóstico del clima.** El ejercicio 15 incluye los datos de temperaturas pronosticadas para 14 días. Repita el ejercicio 15 después de remitirse al conjunto de datos 8 del apéndice B. Utilice los datos de los 35 días para expandir la lista de las diferencias que se incluye en el ejercicio 15. (Los datos del ejercicio 15 se basan en los primeros 14 días del conjunto de datos 8). ¿Sus conclusiones cambian cuando usa el conjunto grande de datos? ¿Tiene mayor confianza en los resultados basados en el conjunto de datos más grande?

- 22. Coca-Cola regular/Coca-Cola dietética.** El ejercicio 18 incluye los pesos de Coca-Cola regular y Coca-Cola dietética del conjunto de datos 12 en el apéndice B. (Los datos del ejercicio 18 incluyen las primeras seis latas de Coca-Cola regular y las primeras seis latas de Coca-Cola dietética). Repita el ejercicio 18 utilizando la lista completa de pesos de Coca-Cola regular y dietética. ¿Sus conclusiones cambian cuando se usa el conjunto grande de datos?

- 23. Peniques revisados.** El ejercicio 19 incluye los pesos de peniques acuñados antes y después de 1983. Repita el ejercicio 19 con la lista completa de los pesos de los peniques acuñados antes de 1983 y la lista completa de los peniques acuñados después de 1983, tal como aparecen en el conjunto de datos 14 del apéndice B. ¿Sus conclusiones cambian cuando se usa el conjunto grande de datos?
- 24. IMC y género.** El ejercicio 20 incluye los valores del IMC de 12 hombres y 12 mujeres. Repita el ejercicio 20 con la lista completa de los valores del IMC, tal como aparece en el conjunto de datos 1 del apéndice B. ¿Sus conclusiones cambian cuando se usa el conjunto grande de datos?

Cálculo de la desviación estándar a partir de una distribución de frecuencias. En los ejercicios 25 a 28, calcule la desviación estándar de los datos muestrales que se resumen en una tabla de distribución de frecuencias, utilizando la siguiente fórmula, donde x representa la marca de clase y f representa la frecuencia de clase. También, compare las desviaciones estándar calculadas con las desviaciones estándar obtenidas por medio de la fórmula 3-4 con la lista original de los valores de los datos: (ejercicio 25) 6.9, (ejercicio 26) 0.630, (ejercicio 27) 4.0, (ejercicio 28) 0.62.

$$s = \sqrt{\frac{n[\sum(f \cdot x^2)] - [\sum(f \cdot x)]^2}{n(n-1)}} \quad \text{desviación estándar para una distribución de frecuencias}$$

25. Temperatura mínima diaria (en °F)	Frecuencia	26. Precipitación mínima diaria (en pulgadas)	Frecuencia
35–39	1	0.00–0.49	31
40–44	3	0.50–0.99	1
45–49	5	1.00–1.49	0
50–54	11	1.50–1.99	2
55–59	7	2.00–2.49	0
60–64	7	2.50–2.99	1
65–69	1		

Tabla del Ejercicio 27

Velocidad	Frecuencia
42–45	25
46–49	14
50–53	7
54–57	3
58–61	1

Tabla del Ejercicio 28

Temperatura	Frecuencia
96.5–96.8	1
96.9–97.2	8
97.3–97.6	14
97.7–98.0	22
98.1–98.4	19
98.5–98.8	32
98.9–99.2	6
99.3–99.6	4

- 27. Multas por exceso de velocidad.** La distribución de frecuencias dada describe las velocidades de conductores multados por la policía de la ciudad de Poughkeepsie. Estos conductores circulaban por una zona con velocidad límite de 30 mi/h en Creek Road, que pasa por la universidad del autor.
- 28. Temperaturas corporales.** La distribución de frecuencias al margen resume una muestra de temperaturas corporales humanas. (Consulte las temperaturas de medianoche del segundo día, incluidas en el conjunto de datos 2 del apéndice B).
- 29. Regla práctica del intervalo.** Utilice la regla práctica del intervalo para estimar la desviación estándar de las edades de los profesores de su universidad.
- 30. Regla práctica del intervalo.** Utilice los datos muestrales del conjunto de datos 1 del apéndice B, sobre una muestra de 40 mujeres que tienen muslos con una longitud media de 38.86 cm y una desviación estándar de 3.78 cm. Aplique la regla práctica del intervalo para identificar las longitudes mínima y máxima “comunes” del muslo. ¿Una longitud de 47.0 cm se consideraría inusual en este contexto?
- 31. Regla práctica del intervalo.** La media de las cantidades de consumo de energía eléctrica de la casa del autor durante un periodo de dos meses es de 2838 kWh, mientras que la desviación estándar es de 504 kWh. (El autor realmente registra esta información). Utilice la regla práctica del intervalo para identificar las cantidades mínima y máxima “comunes” de consumo de energía eléctrica. Para un periodo de dos meses en particular, la empresa que suministra la electricidad registró un consumo de 578 kWh. ¿Es poco común esa cantidad?

32. Regla práctica del intervalo. Latas de aluminio con un grosor de 0.0111 pulgadas tienen cargas axiales con una media de 281.8 lb y una desviación estándar de 27.8 lb. La carga axial se mide aplicando presión sobre la superficie de la lata hasta que ésta colapsa. (Véase el conjunto de datos 15 en el apéndice B). Utilice la regla práctica del intervalo para calcular las cargas axiales mínima y máxima “comunes”. Una lata específica tuvo una carga axial de 504 lb. ¿Es esto poco común?

33. Regla empírica. Las estaturas de un grupo de hombres tienen una distribución normal, con una media de 176 cm y una desviación estándar de 7 cm. Por medio de la regla empírica, ¿cuál es el porcentaje aproximado de hombres entre

- 169 cm y 183 cm?
- 155 cm y 197 cm?

34. Coeficiente de variación. Remítase a los datos del ejercicio 17 y calcule el coeficiente de variación de cada una de las dos muestras; después, compare los resultados.

35. Coeficiente de variación. Utilice los siguientes datos muestrales para calcular el coeficiente de variación de cada muestra; después, compare los resultados.

Estaturas (en pulgadas) de hombres: 71 66 72 69 68 69

Largo (en mm) de huevos de ave: 19.7 21.7 21.9 22.1 22.1 22.3
22.7 22.9 23.9

36. Comprensión de las unidades de medición. Si un conjunto de datos indica la longevidad (en días) de moscas de la fruta, ¿qué unidades se utilizan para la desviación estándar? ¿Qué unidades se utilizan para la varianza?

3-3 MÁS ALLÁ DE LO BÁSICO

37. Interpretación de valores extremos. Un conjunto de datos consta de 20 valores bastante cercanos. Se incluye otro valor, pero este nuevo dato es un valor extremo (muy alejado de los otros valores). ¿Cómo afecta el valor extremo a la desviación estándar? ¿No hay ningún efecto? ¿Hay un pequeño efecto? ¿Hay un gran efecto?

38. ¿Por qué dividir entre $n - 1$? Permita que una población consista en los valores 3, 6 y 9. Suponga que las muestras de 2 valores se eligen aleatoriamente *con reemplazo*.

- Calcule la varianza σ^2 de la población $\{3, 6, 9\}$.
- Liste las nueve muestras diferentes posibles de 2 valores seleccionados con reemplazo, luego calcule la varianza muestral s^2 (que incluye la división entre $n - 1$) para cada una. Si usted selecciona de manera repetida 2 valores muestrales, ¿cuál es el valor promedio de las varianzas muestrales s^2 ?
- Para cada una de las nueve muestras, calcule la varianza al tratar cada muestra como si fuera una población. (Asegúrese de usar la fórmula para la varianza poblacional, que incluye la división entre n). Si usted selecciona de manera repetida 2 valores muestrales, ¿cuál es el valor promedio de las varianzas poblacionales?
- ¿Qué método produce valores que son mejores estimados de σ^2 : el inciso b) o el inciso c)? Cuando se calculan las varianzas de muestras, ¿se debe dividir entre n o entre $n - 1$?
- Los incisos anteriores demuestran que s^2 es un estimador sin sesgo de σ^2 . ¿Es s un estimador sin sesgo de σ ?



Índice del costo de la risa

En realidad hay un Índice del Costo de la Risa (ICR), que busca los costos de artículos como pollos de plástico, anteojos de Groucho Marx, entradas a clubes de comediantes y otros 13 indicadores principales del humor. Éste es el mismo método básico que se utiliza en la creación del Índice de Precios al Consumidor (IPC), que se basa en un promedio ponderado de bienes y servicios adquiridos por consumidores comunes. Mientras que las puntuaciones estándar y los percentiles nos permiten comparar valores diferentes ignorando cualquier elemento del tiempo, los números índice, como el ICR y el IPC, nos permiten comparar el valor de alguna variable con su valor en un periodo tomado como base. El valor de un número índice es el valor real dividido entre el valor base, multiplicado por 100.

3-4 Medidas de posición relativa

Concepto clave Esta sección presenta medidas que resultan útiles para comparar valores de diferentes conjuntos de datos o para comparar valores dentro del mismo conjunto de datos. El concepto más importante de esta sección es la puntuación z , por lo que debemos comprender el papel que desempeña (para comparar valores de distintos conjuntos de datos) y desarrollar nuestra capacidad para convertir valores de datos en puntuaciones z . También estudiaremos los cuartiles y percentiles, que se utilizan para comparar valores dentro del mismo conjunto de datos.

Puntuaciones z

Una puntuación z (o valor estandarizado) se calcula convirtiendo un valor a una escala estandarizada, como se establece en la siguiente definición. Utilizaremos ampliamente las puntuaciones z en el capítulo 6 y en capítulos posteriores, ya que son muy importantes.

Definición

Una **puntuación z (o valor estandarizado)** es el número de desviaciones estándar que un valor x se encuentra por arriba o por debajo de la media. Se calcula utilizando las siguientes expresiones:

$$\begin{array}{ll} \text{Muestra} & \text{Población} \\ z = \frac{x - \bar{x}}{s} & \text{o} \quad z = \frac{x - \mu}{\sigma} \end{array}$$

(Redondear z a dos espacios decimales).

El siguiente ejemplo ilustra la forma en que se utilizan las puntuaciones z para comparar valores, aun cuando provengan de distintas poblaciones.

EJEMPLO Comparación de estaturas Con una estatura de 75 pulgadas, Lyndon Johnson fue el presidente de Estados Unidos más alto del siglo pasado. Con una estatura de 85 pulgadas, Shaquille O'Neal es el jugador más alto del equipo de básquetbol Miami Heat. ¿Quién es *relativamente* más alto: Lyndon Johnson entre los presidentes del siglo pasado o Shaquille O'Neal entre los jugadores de su equipo Miami Heat? La estatura media de los presidentes del siglo pasado era de 71.5 pulgadas, con una desviación estándar de 2.1 pulgadas. Los jugadores de básquetbol del equipo Miami Heat tienen una estatura media de 80.0 pulgadas, con una desviación estándar de 3.3 pulgadas.

SOLUCIÓN Las estaturas de los presidentes y de los jugadores de básquetbol provienen de poblaciones muy diferentes; para compararlas es necesario estandarizar las estaturas convirtiéndolas en puntuaciones z .

$$\text{Lyndon Johnson: } z = \frac{x - \mu}{\sigma} = \frac{75 - 71.5}{2.1} = 1.67$$

$$\text{Shaquille O'Neal: } z = \frac{x - \mu}{\sigma} = \frac{85 - 80.0}{3.3} = 1.52$$

INTERPRETACIÓN La estatura de Lyndon Johnson está a 1.67 desviaciones estándar por arriba de la media, mientras que la estatura de Shaquille O'Neal está a 1.52 desviaciones estándar por arriba de la media. La estatura de Lyndon

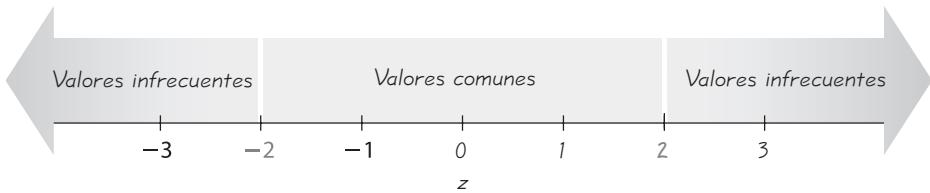


Figura 3-5 Interpretación de las puntuaciones z

Los valores infrecuentes son aquellos con puntuaciones z menores que -2.00 o mayores que 2.00 .

Johnson, entre los presidentes del siglo pasado, es relativamente mayor que la estatura de Shaquille O’Neal entre los jugadores de básquetbol del equipo de Miami Heat. Shaquille O’Neal es mucho más alto que Lyndon Johnson, pero este último es relativamente más alto cuando lo comparamos con sus colegas.

Puntuaciones z y valores infrecuentes

En la sección 3-3 utilizamos la regla práctica del intervalo para concluir que un valor es “infrecuente” si está a más de 2 desviaciones estándar de la media. Por lo tanto, los valores infrecuentes tienen puntuaciones z menores que -2 y mayores que $+2$. (Véase la figura 3-5). Si aplicamos este criterio, Lyndon Johnson no tiene una estatura infrecuente cuando lo comparamos con los presidentes del siglo pasado, y Shaquille O’Neal no tiene una estatura infrecuente cuando lo comparamos con sus compañeros de equipo, ya que ninguno de los dos tiene una estatura con una puntuación z mayor que 2 .

Valores comunes: $-2 \leq z$ puntuación ≤ 2

Valores infrecuentes: puntuación $z < -2$ o puntuación $z > 2$

Si consideramos a los jugadores de básquetbol del equipo de Miami Heat, el jugador más bajo es Damon Jones, quien mide 75 pulgadas. Su puntuación z es -1.52 , como se muestra en el siguiente cálculo. (Nuevamente, usamos $\mu = 80.0$ pulgadas y $\sigma = 3.3$ pulgadas del equipo de Miami Heat).

$$\text{Damon Jones: } z = \frac{x - \mu}{\sigma} = \frac{75 - 80.0}{3.3} = -1.52$$

La estatura de Damon Jones ilustra el siguiente principio sobre los valores que están por debajo de la media:

Siempre que un valor sea menor que la media, su puntuación z correspondiente será negativa.

Las puntuaciones z son medidas de posición, en el sentido de que describen la localización de un valor (en términos de desviaciones estándar) en relación con la media. Una puntuación z de 2 indica que un valor está a dos desviaciones estándar *por arriba* de la media, en tanto que una puntuación z de -3 indica que un valor está a tres desviaciones estándar *por debajo* de la media. Los cuartiles y los percentiles también son medidas de posición, pero se definen de forma distinta que las puntuaciones z y son útiles para comparar valores dentro del mismo conjunto de datos o entre distintos conjuntos de datos.

Cuartiles y percentiles

En la sección 3-2 aprendimos que la mediana de un conjunto de datos es el valor que está a la mitad, de manera que el 50% de los valores son iguales o menores

que la mediana y el 50% de los valores son mayores o iguales que la mediana. Así como la mediana divide los datos en dos partes iguales, los tres **cuartiles**, denotados por Q_1 , Q_2 y Q_3 , dividen los valores ordenados en cuatro partes iguales. (Los valores están *ordenados* cuando están acomodados en orden). He aquí las descripciones de los tres cuartiles:

Q_1 (primer cuartil): Separa el 25% inferior de los valores ordenados del 75% superior. (Para ser más precisos, al menos el 25% de los valores ordenados son menores o iguales que Q_1 , y al menos el 75% de los valores son mayores o iguales que Q_1).

Q_2 (segundo cuartil): Igual a la mediana; separa el 50% inferior de los valores ordenados del 50% superior.

Q_3 (tercer cuartil): Separa el 75% inferior de los valores ordenados del 25% superior. (Para ser más precisos, al menos el 75% de los valores ordenados son menores o iguales que Q_3 , y al menos el 25% de los valores son mayores o iguales que Q_3).

Describiremos un procedimiento para el cálculo de cuartiles después de analizar los percentiles. No existe un acuerdo universal respecto de un procedimiento único para el cálculo de cuartiles; por esa razón, con frecuencia los distintos programas de cómputo producen resultados diferentes. Por ejemplo, si usted utiliza el conjunto de datos 1, 3, 6, 10, 15, 21, 28 y 36, obtendrá los siguientes resultados:

	Q_1	Q_2	Q_3
STATDISK	4.5	12.5	24.5
Minitab	3.75	12.5	26.25
SPSS	3.75	12.5	26.25
Excel	5.25	12.5	22.75
SAS	4.5	12.5	24.5
TI-83/84 Plus	4.5	12.5	24.5

Si usted utiliza una calculadora o un programa de cómputo para resolver ejercicios que comprenden cuartiles, es posible que obtenga resultados que difieran ligeramente de las respuestas que vienen al final del libro.

Así como existen tres cuartiles que separan un conjunto de datos en cuatro partes, también existen 99 **percentiles**, que se denotan P_1 , P_2 , ..., P_{99} , los cuales separan los datos en 100 grupos, con aproximadamente el 1% de los valores en cada grupo. (Los cuartiles y percentiles son ejemplos de *cuantiles* o *fractiles*, que separan los datos en grupos con aproximadamente el mismo número de valores).

El proceso para calcular el percentil que corresponde a un valor x específico es bastante sencillo, tal como se indica en la siguiente expresión:

$$\text{Percentil del valor } x = \frac{\text{número de valores menores que } x}{\text{número total de valores}} \cdot 100$$

(el resultado se redondea al entero más cercano)

EJEMPLO Edades de las mejores actrices La tabla 3-4 lista las 76 edades ordenadas de las mejores actrices ganadoras del Óscar que se incluyen en la tabla 2-1. (La tabla 2-1 aparece en el problema del capítulo 2). Calcule el percentil correspondiente a una edad de 30 años.

Tabla 3-4 Edades ordenadas de las 76 mejores actrices									
21	22	24	24	25	25	25	25	26	26
26	26	27	27	27	27	28	28	28	28
29	29	29	29	29	29	30	30	31	31
31	32	32	33	33	33	33	33	34	34
34	35	35	35	35	35	35	35	36	37
37	38	38	38	38	39	39	40	41	41
41	41	41	42	42	43	45	46	49	50
54	60	61	63	74	80				

SOLUCIÓN En la tabla 3-4 observamos que existen 26 edades menores que 30; por lo tanto,

$$\text{percentil de } 30 = \frac{26}{76} \cdot 100 = 34$$

INTERPRETACIÓN La edad de 30 años es el percentil 34.

El ejemplo anterior muestra cómo convertir un valor muestral dado a su percentil correspondiente. Existen diversos métodos para el procedimiento inverso de convertir un percentil en el valor correspondiente del conjunto de datos. El procedimiento que utilizamos se resume en la figura 3-6, que emplea la siguiente notación.

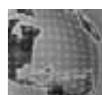
Notación

- n número total de valores en el conjunto de datos
- k percentil utilizado (ejemplo: para el percentil 25, $k = 25$)
- L localizador que da la *posición* de un valor (ejemplo: para el valor en el lugar 12 en la lista ordenada, $L = 12$)
- P_k percentil k -ésimo (ejemplo: P_{25} es el percentil 25)

EJEMPLO Edades de las mejores actrices Remítase a las edades ordenadas de las mejores actrices en la tabla 3-4 y utilice la figura 3-6 para calcular el valor del percentil 20, P_{20} .

SOLUCIÓN Al consultar la figura 3-6, observamos que los datos muestrales ya están ordenados, de manera que podemos proceder a calcular el valor del localizador L . En este cálculo utilizamos $k = 20$ porque estamos tratando de calcular el valor del percentil 20. Usamos $n = 76$ porque tenemos 76 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{20}{100} \cdot 76 = 15.2$$



LA ESTADÍSTICA EN LAS NOTICIAS

El crecimiento de la estadística

El reportero Richard Rothstein escribió en el *New York Times* que el estudio del álgebra, la trigonometría y la geometría en preparatoria “dejan muy poco espacio para el estudio de la estadística y la probabilidad. No obstante, los estudiantes necesitan fundamentos sobre el análisis de datos”. El reportero indicó que el cálculo tiene un papel relevante en los estudios universitarios, a pesar de que “sólo pocos trabajos, principalmente en el campo técnico, realmente lo utilizan”. Rothstein citó un estudio realizado por el profesor Clifford Konold, de la Universidad de Massachusetts, en el cual se contó el número de presentaciones de datos en el *New York Times*. En los ejemplares del *Times* de 1972, el doctor Konold encontró cuatro gráficas o tablas en cada una de 10 ediciones de entre semana (sin incluir las secciones de deportes y negocios), pero en 1982 había 8 y en 1992 había 44; “al año siguiente él (el doctor Konold) encontró más de 100”. El crecimiento de la estadística como disciplina se ha visto fomentado, en parte, por un mayor uso de este tipo de presentaciones de datos en los medios de comunicación masiva.

continúa

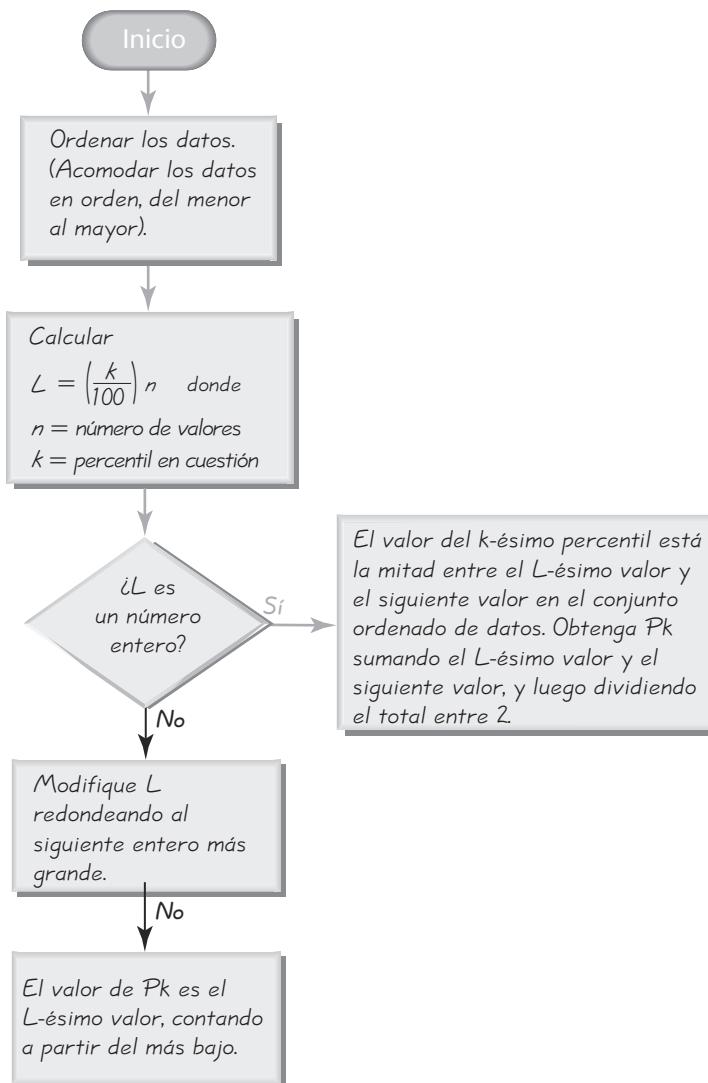


Figura 3-6 Conversión del k -ésimo percentil al valor del dato correspondiente

NDespués, nos preguntamos si L es un número entero. La respuesta es no, por lo que procedemos al siguiente recuadro inferior, donde modificamos L redondeando su valor al entero más alto, de 15.2 a 16. (En este libro generalmente redondeamos de la forma acostumbrada, pero éste es uno de los dos casos en que redondeamos *hacia arriba* en vez de redondear *hacia el entero más cercano*). Por último, el recuadro inferior muestra que el valor de P_{20} es el decimosexto valor, contando del más bajo al más alto. En la tabla 3-4 el valor 16º es 27. Es decir, $P_{20} = 27$ años de edad.

EJEMPLO Edades de las mejores actrices Remítase a las edades de las mejores actrices en la tabla 3-4. Utilice la figura 3-6 para calcular el valor de Q_3 , que es el tercer cuartil.

SOLUCIÓN Primero señalamos que Q_3 es igual a P_{75} , de manera que procedemos con el objetivo de calcular el valor del percentil 75. Si nos remitimos a la figura 3-6, observamos que los datos de la muestra ya están ordenados, así que procedemos a calcular el valor del localizador L . Para este cálculo utilizamos $k = 75$, ya que estamos tratando de obtener el valor del percentil 75, y usamos $n = 76$ porque son 76 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{75}{100} \cdot 76 = 57$$

Luego, nos preguntamos si L es un número entero y respondemos que sí, de manera que vamos al siguiente recuadro a la derecha. Ahora podemos ver que el valor del percentil k -ésimo (75°) se encuentra a la mitad entre el valor L -ésimo (57°) y el siguiente valor del conjunto original de datos. Es decir, el valor del percentil 75 se encuentra a la mitad entre el quincuagésimo séptimo valor (57°) y el quincuagésimo octavo valor (58°). El valor 57° es 39 años y el valor 58° es 40 años; por lo tanto, el valor a la mitad de ellos es 39.5 años. Concluimos que el percentil 75 es $P_{75} = 39.5$ años. El valor del tercer cuartil Q_3 también es de 39.5 años.

El ejemplo anterior demuestra que al calcular un valor cuartilar (como Q_3), es posible utilizar el valor del percentil equivalente (como P_{75}) en su lugar. Al margen se indican las relaciones equivalentes entre cuartiles y percentiles.

En secciones anteriores de este capítulo, describimos diversos estadísticos, incluyendo media, mediana, moda, rango y desviación estándar. Algunos otros estadísticos se definen con el uso de cuartiles y percentiles, como los siguientes:

$$\begin{aligned} Q_1 &= P_{25} \\ Q_2 &= P_{50} \\ Q_3 &= P_{75} \end{aligned}$$

$$\text{rango intercuartilar (o RIC)} = Q_3 - Q_1$$

$$\text{rango semiintercuartilar} = \frac{Q_3 - Q_1}{2}$$

$$\text{cuartil medio} = \frac{Q_3 + Q_1}{2}$$

$$\text{rango de percentiles 10-90} = P_{90} - P_{10}$$

Después de completar esta sección, usted podrá convertir un valor en su puntuación z (o puntuación estándar) correspondiente, de forma que sea posible compararlo con otros valores que provienen de diferentes conjuntos de datos. También podrá convertir un valor en su valor percentilar correspondiente, de manera que pueda compararlo con otros valores en algún conjunto de datos. También sabrá convertir un percentil en su valor de dato correspondiente. Finalmente, comprenderá el significado de los cuartiles y podrá relacionarlos con sus valores percentiles correspondientes (como en $Q_3 = P_{75}$).

Uso de la tecnología

STATDISK

Column 1 Descriptive Statistics	
Sample Size, n:	76
Mean:	35.68421
Median:	33.5
Midrange:	50.5
RMS:	37.33842
Variance, s^2 :	122.4056
St Dev, s :	11.06371
Mean Abs Dev:	7.759003
Range:	59
Minimum:	21
1st Quartile:	28
2nd Quartile:	33.5
3rd Quartile:	39.5
Maximum:	80
Sum:	2712
Sum Sq:	105956

Existe una variedad de programas de cómputo y calculadoras para calcular muchos de los estadísticos estudiados hasta ahora en este capítulo. En la sección 3-2 dimos instrucciones específicas para el uso de STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus. Señalamos que en ocasiones es posible introducir un conjunto de datos y utilizar

Minitab

Descriptive Statistics: Actresses

Variable	Mean	StDev	Variance	Sum	Minimum	Q1	Median	Q3
Actresses	35.68	11.06	122.41	2712.00	21.00	28.00	33.50	39.75

Variable	Maximum	Range
Actresses	80.00	59.00

Excel

Column1	
Mean	35.68421053
Standard Error	1.269094238
Median	33.5
Mode	35
Standard Deviation	11.06370707
Sample Variance	122.405614
Kurtosis	4.399199941
Skewness	1.869827875
Range	59
Minimum	21
Maximum	80
Sum	2712
Count	76

TI-83/84 Plus

1-Var Stats
 $\bar{x}=35.68421053$
 $\Sigma x=2712$
 $\Sigma x^2=105956$
 $Sx=11.06370707$
 $\sigma x=10.9906785$
 $n=76$

TI-83/84 Plus

1-Var Stats
 $n=76$
 $\min X=21$
 $Q_1=28$
 $Med=33.5$
 $Q_3=39.5$
 $\max X=80$

una operación para obtener diversos estadísticos muestrales, frecuentemente llamados *estadísticos descriptivos*. En las siguientes representaciones visuales se mencionan ejemplos de tales resultados, los cuales provienen de las edades de las mejores actrices

que aparecen en la tabla 2-1, en el problema del capítulo 2. Los resultados de la calculadora TI-83/84 Plus se presentan en dos plantillas, ya que no caben en una sola.

3-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Puntuaciones z.** Se descubre que un valor de un conjunto de datos grande tiene una puntuación z de -2 . ¿El valor está por arriba o por debajo de la media? ¿A cuántas desviaciones estándar de la media se encuentra este valor?
- Puntuaciones z.** Una muestra consiste en las longitudes de águilas calvas americanas, medidas en centímetros. Si la longitud de un águila en específico se convierte en una puntuación z , ¿qué unidades se utilizan para la puntuación z ? ¿Centímetros?
- Cuartiles.** Para un conjunto grande de datos, se determina que el primer cuartil es 15. ¿Qué significa cuando decimos que 15 es el primer cuartil?
- Códigos postales.** Los códigos postales en Estados Unidos están ordenados de este a oeste. Los estados del este, como Maine, tienen códigos postales que inician con 0, en tanto que los estados del oeste, como Hawai, Alaska y California, tienen códigos postales que inician con 9. Si se calcula el primer cuartil Q_1 de una lista de todos los códigos postales, ¿el resultado corresponde al lugar que se localiza a un 25%

de la distancia del lugar que está más al este con respecto al lugar que está más al oeste?

En los ejercicios 5 a 8, exprese todas las puntuaciones z con dos decimales.

5. **Estatura de Darwin.** La media de la estatura de los hombres es de 176 cm y su desviación estándar de 7 cm. Charles Darwin tenía una estatura de 182 cm.
 - a. ¿Qué diferencia hay entre la estatura de Darwin y la media?
 - b. ¿A cuántas desviaciones estándar corresponde [la diferencia obtenida en el inciso a]]?
 - c. Convierta la estatura de Darwin a una puntuación z.
 - d. Si consideramos que las estaturas “comunes” son aquellas que corresponden a puntuaciones z entre –2 y 2, ¿la estatura de Darwin es común o infrecuente?
6. **CI de Einstein.** Las puntuaciones de CI de la prueba Stanford Binet tienen una media de 100 y una desviación estándar de 16. Se dice que Albert Einstein tenía un CI de 160.
 - a. ¿Qué diferencia hay entre el CI de Einstein y la media?
 - b. ¿A cuántas desviaciones estándar corresponde [la diferencia obtenida en el inciso a]]?
 - c. Convierta el CI de Einstein a una puntuación z.
 - d. Si consideramos que las puntuaciones de CI “comunes” son aquellas que corresponden a puntuaciones z entre –2 y 2, ¿el CI de Einstein es común o infrecuente?
7. **Estatura de presidentes.** Con una estatura de 67 pulgadas, William McKinley fue el presidente más bajo del siglo pasado. Los presidentes del siglo pasado tenían una estatura media de 71.5 pulgadas y una desviación estándar de 2.1 pulgadas.
 - a. ¿Qué diferencia hay entre la estatura de McKinley y la media de la estatura de los presidentes del siglo pasado?
 - b. ¿A cuántas desviaciones estándar corresponde [la diferencia obtenida en el inciso a]]?
 - c. Convierta la estatura de McKinley a una puntuación z.
 - d. Si consideramos que las estaturas “comunes” son aquellas que corresponden a puntuaciones z entre –2 y 2, ¿la estatura de McKinley es común o infrecuente?
8. **La mujer más alta del mundo.** Sandy Allen es la mujer más alta del mundo; mide 91.25 pulgadas (o 7 pies, 7.25 pulgadas). La media de la estatura de las mujeres es de 63.6 pulgadas, con una desviación estándar de 2.5 pulgadas.
 - a. ¿Qué diferencia hay entre la estatura de Sandy Allen y la media de la estatura de las mujeres?
 - b. ¿A cuántas desviaciones estándar corresponde [la diferencia obtenida en el inciso a]]?
 - c. Convierta la estatura de Sandy Allen a una puntuación z.
 - d. ¿La estatura de Sandy Allen cumple el criterio de ser infrecuente porque corresponde a una puntuación z que no cae entre –2 y 2?
9. **Temperaturas corporales.** Las temperaturas corporales humanas tienen una media de 98.20°F y una desviación estándar de 0.62°F. Convierta las siguientes temperaturas a puntuaciones z?
 - a. 97.5°F
 - b. 98.60°F
 - c. 98.20°F
10. **Estatura de mujeres.** El Beanstalk Club es exclusivo para mujeres y hombres muy altos. La estatura mínima requerida para las mujeres es de 70 pulgadas. La estatura de las mujeres tiene una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas. Calcule la puntuación z correspondiente a una mujer con una estatura de 70 pulgadas, y determine si esa estatura es infrecuente.
11. **Duración del embarazo.** Una mujer escribió a *Dear Abby* y afirmó haber dado a luz 308 días después de una visita de su esposo, quien estaba en la Marina. La duración de los embarazos tiene una media de 268 días y una desviación estándar de 15 días. Calcule la puntuación z para 308 días. ¿Esta duración es infrecuente? ¿Qué concluye usted?

- 12. Niveles de colesterol.** Para los hombres de entre 18 y 24 años de edad, los niveles de colesterol sérico tienen una media de 178.1 (mg/100 mL) y una desviación estándar de 40.7 (basado en datos de la National Health Survey). Calcule la puntuación z correspondiente a un hombre de entre 18 y 24 años, que tiene un nivel de colesterol sérico de 259.0 mg/100 mL. ¿Este nivel es excepcionalmente alto?
- 13. Comparación de calificaciones de exámenes.** ¿Cuál es relativamente mejor: una calificación de 85 en un examen de psicología o una calificación de 45 en un examen de economía? Las calificaciones del examen de psicología tienen una media de 90 y una desviación estándar de 10. Las calificaciones del examen de economía tienen una media de 55 y una desviación estándar de 5.
- 14. Comparaciones de pruebas.** Tres estudiantes toman pruebas de estrés equivalentes. ¿Cuál es la puntuación relativa más alta?
- Una puntuación de 144 en una prueba con una media de 128 y una desviación estándar de 34.
 - Una puntuación de 90 en una prueba con una media de 86 y una desviación estándar de 18.
 - Una puntuación de 18 en una prueba con una media de 15 y una desviación estándar de 5.

En los ejercicios 15 a 18, utilice las 76 edades ordenadas de las mejores actrices que se incluyen en la tabla 3-4. Calcule el percentil correspondiente a cada edad.

15. 25 16. 35 17. 40 18. 50

En los ejercicios 19 a 26, utilice las 76 edades ordenadas de las mejores actrices que se incluyen en la tabla 3-4. Calcule el percentil o cuartil indicado.

19. P_{10} 20. Q_1 21. P_{25} 22. Q_2
 23. P_{33} 24. P_{66} 25. P_1 26. P_{85}

3-4 MÁS ALLÁ DE LO BÁSICO

- 27. Edades de las mejores actrices.** Utilice las 76 edades ordenadas de las mejores actrices que se incluyen en la tabla 3-4.
- Calcule el rango intercuartilar.
 - Obtenga el cuartil medio.
 - Calcule el rango de percentiles 10-90.
 - ¿ $P_{50} = Q_2$? Si es así, ¿ P_{50} es siempre igual a Q_2 ?
 - ¿ $Q_2 = (Q_1 + Q_3)/2$? Si es así, ¿ Q_2 es siempre igual a $(Q_1 + Q_3)/2$?
- 28. Interpolación.** Cuando se calculan percentiles usando la figura 3-6, si el localizador L no es un número entero, se debe redondear al siguiente entero más grande. Una alternativa a este método es la *interpolación*. Por ejemplo, el uso de la interpolación con un localizador $L = 23.75$ nos conduce a un valor que está a 0.75 (o 3/4) de distancia entre los valores 23° y 24°. Utilice el método de interpolación para calcular P_{35} para las edades de las mejores actrices de la tabla 3-4. ¿Qué diferencia hay entre el resultado y el valor que se obtendría utilizando la figura 3-6 sin interpolación?
- 29. Deciles y quintiles.** Para un conjunto de datos existen nueve **deciles**, designados como D_1, D_2, \dots, D_9 que separan los datos ordenados en 10 grupos, con aproximadamente el 10% de los valores en cada uno. También existen cuatro **quintiles**, que dividen los datos ordenados en 5 grupos, con aproximadamente el 20% de los valores en cada uno. (Note la diferencia entre los quintiles y los cuantiles, que describimos anteriormente en esta sección).
- ¿Qué percentil es equivalente a D_1 ? ¿A D_5 ? ¿A D_8 ?
 - Calcule los nueve deciles con las edades ordenadas de las mejores actrices de la tabla 3-4.
 - Calcule los cuatro quintiles con las edades ordenadas de las mejores actrices de la tabla 3-4.

3-5 Análisis exploratorio de datos (AED)

Concepto clave En esta sección se estudian los valores extremos; luego, se presenta una nueva gráfica estadística llamada gráfica del cuadro, que sirve para visualizar la distribución de datos. (Las gráficas de cuadro no se incluyeron en la sección 2-4 porque utilizan cuartiles, y éstos se presentaron en la sección 3-14). Debemos saber construir una gráfica de cuadro sencilla. Esta sección también se enfoca en el principio importante de *explorar* los datos antes de utilizar métodos estadísticos específicos.

Comenzamos esta sección definiendo el análisis exploratorio de datos; luego presentaremos los valores extremos, el resumen de 5 números y las gráficas de cuadro. Las gráficas de cuadro modificadas, que se estudian casi al final de esta sección, son un poco más complicadas, pero ofrecen información más específica sobre los valores extremos.

Definición

El **análisis exploratorio de datos** es el proceso de utilizar herramientas estadísticas (como gráficas, medidas de tendencia central y medidas de variación) con la finalidad de investigar conjuntos de datos para comprender sus características importantes.

Recuerde que en la sección 2-1 mencionamos cinco características importantes de los datos: centro, variación, distribución, valores extremos y patrones de cambio con el paso del tiempo. Las características centrales se pueden investigar con medidas como la media y la mediana; la variación, con medidas como la desviación estándar y el rango; y la distribución de los datos por medio de herramientas como distribuciones de frecuencias e histogramas. Hemos visto que algunos estadísticos importantes (como la media y la desviación estándar) pueden resultar muy afectados por la presencia de un valor extremo. Por lo general es importante investigar con mayor detalle el conjunto de datos para identificar aspectos sobresalientes, especialmente aquellos que puedan afectar los resultados y las conclusiones. Uno de esos aspectos es la presencia de valores extremos, los cuales estudiaremos con mayor detalle.



Valores extremos

A lo largo de este libro consideraremos un **valor extremo** como un valor que está muy alejado de la mayor parte de los demás valores. (Existe una definición alternativa más específica en la sección “Gráficas de cuadro modificadas”, casi al final de esta sección). En relación con los otros datos, un valor es *extremo* cuando está muy alejado del patrón general de la mayoría de los datos. Cuando se explora un conjunto de datos, se deben considerar los valores extremos, ya que pueden revelar información importante y afectar en gran medida el valor de la media y de la desviación estándar, así como distorsionar gravemente un histograma. El siguiente ejemplo utiliza un valor incorrecto para ejemplificar un valor extremo, pero no todos los valores extremos son errores; algunos son valores correctos.

EJEMPLO Edades de las mejores actrices Cuando se utiliza un programa de cómputo o una calculadora, es fácil cometer errores al presionar las teclas. Remítase a las edades de las mejores actrices de la tabla 2-1. (La tabla 2-1 se presenta en el problema del capítulo 2). Suponga que al registrar las edades,



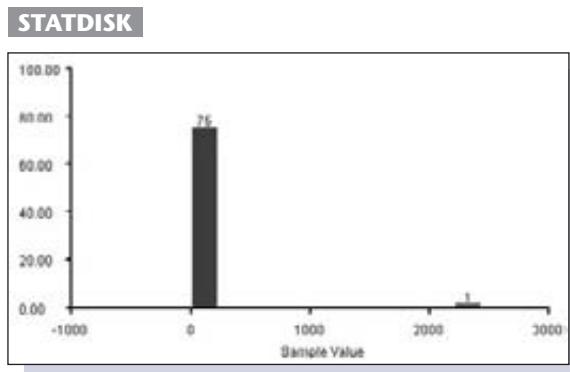
Minería de datos

El término *minería de datos* generalmente se utiliza para describir la práctica ahora común de analizar un gran conjunto de datos existente con el propósito de encontrar relaciones, patrones o cualquier resultado interesante que no se descubrió en los estudios originales de tal conjunto de datos. Algunos especialistas en estadística se muestran preocupados por la inferencia *ad hoc*, una práctica que consiste en que un investigador revisa con detalle datos antiguos, encuentra algo significativo y después identifica una pregunta importante que ya se respondió. Robert Gentleman, editor de una columna de la revista *Chance*, escribe que “existen algunos temas estadísticos interesantes y fundamentales de los que se puede ocupar la minería de datos. Simplemente esperamos que su éxito y popularidad actuales no dañen demasiado nuestra disciplina (la estadística) antes de que se determinen sus limitaciones”.

continúa

usted introduce el primer dato de 22 de manera incorrecta como 2222 porque presionó la tecla demasiado tiempo cuando se distrajo por un meteorito que aterrizó en su jardín. El dato incorrecto de 2222 es un valor extremo porque se localiza muy lejos de los demás valores. ¿Cómo afecta el valor extremo a la media, la desviación estándar y el histograma?

SOLUCIÓN Cuando el dato 22 es reemplazado por el valor extremo 2222, la media cambia de 35.7 años a 64.6 años, de manera que su efecto es muy grande. El dato incorrecto 2222 provoca que la desviación estándar cambie de 11.1 a 251.0, por lo que el efecto del valor extremo aquí también es muy importante. La figura 2-2 de la sección 2-3 describe el histograma de las edades correctas de las actrices en la tabla 2-1, pero la siguiente pantalla del STATDISK muestra el histograma que resulta al utilizar los mismos datos con el valor 22 reemplazado por el dato incorrecto 2222. Compare este histograma de STATDISK con la figura 2-2 y podrá ver con facilidad que la presencia del valor extremo afecta de manera drástica la forma de la distribución.



El ejemplo anterior ilustra estos principios importantes:

1. **Un valor extremo puede tener un efecto importante sobre la media.**
2. **Un valor extremo puede tener un efecto importante sobre la desviación estándar.**
3. **Un valor extremo puede tener un efecto importante sobre la escala del histograma, de forma que la verdadera naturaleza de la distribución se oculte por completo.**

Un procedimiento sencillo para encontrar valores extremos es el examen de una lista *ordenada* de los datos. En particular, observe los valores mínimo y máximo muestrales y determine si se alejan mucho de los demás valores. Algunos valores extremos son valores correctos y algunos son erróneos, como en el ejemplo anterior. Si estamos seguros de que un valor extremo es un error, debemos corregirlo o eliminarlo. Si incluimos un valor extremo porque sabemos que es correcto, podríamos estudiar sus efectos por medio de la construcción de gráficas y el cálculo de estadísticos que incluyan y que descarten los valores extremos.

Gráficas de cuadro

Además de las gráficas presentadas en las secciones 2-3 y 2-4, una gráfica de cuadro es otro tipo de gráfica que se utiliza a menudo. Las gráficas de cuadro son útiles para revelar la tendencia central de los datos, su dispersión, su distribución y la

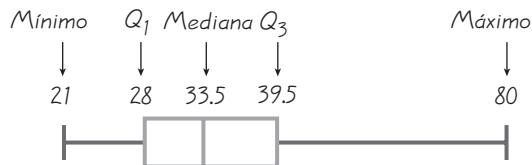


Figura 3-7 Gráfica de cuadro de las edades de las mejores actrices

presencia de valores extremos. La construcción de una gráfica de cuadro requiere que primero obtengamos el valor mínimo, el valor máximo y los cuartiles, tal como se define en el *resumen de los 5 números*.

Definiciones

Para un conjunto de datos, el **resumen de los 5 números** consiste en el valor mínimo; el primer cuartil, Q_1 ; la mediana (o segundo cuartil, Q_2); el tercer cuartil, Q_3 ; y el valor máximo.

Una **gráfica de cuadro** (o **diagrama de cuadro y bigotes**) es una gráfica de un conjunto de datos que consiste en una línea, que se extiende desde el valor mínimo hasta el valor máximo, y una caja con líneas trazadas en el primer cuartil, Q_1 , la mediana y el tercer cuartil, Q_3 . (Véase la figura 3-7).

Procedimiento para construir una gráfica de cuadro

1. Elabore el resumen de los 5 números, que consiste en el valor mínimo, Q_1 , la mediana, Q_3 , y el valor máximo.
2. Construya una escala con valores que incluyan el valor mínimo y el valor máximo.
3. Construya un cuadro (un rectángulo) que se extienda desde Q_1 hasta Q_3 y dibuje una línea en la caja, en el valor de la mediana.
4. Dibuje líneas que se extiendan hacia fuera del cuadro hasta los valores mínimo y máximo.

Las gráficas de cuadro no muestran tanta información detallada como los histogramas o las gráficas de tallo y hojas, por lo que quizás no sean la mejor elección cuando se maneja un solo conjunto de datos. Suelen ser muy útiles para comparar dos o más conjuntos de datos. Cuando se utilicen dos o más gráficas de cuadro para comparar distintos conjuntos de datos, es importante emplear la misma escala, de forma que puedan realizarse comparaciones correctas.

EJEMPLO Edades de las actrices Remítase a las 76 edades de las mejores actrices en la tabla 2-1 (sin el error 2222 utilizado en vez del 22 en el ejemplo anterior).

- a. Elabore el resumen de los 5 números.
- b. Construya una gráfica de cuadro.



Una propina extrema

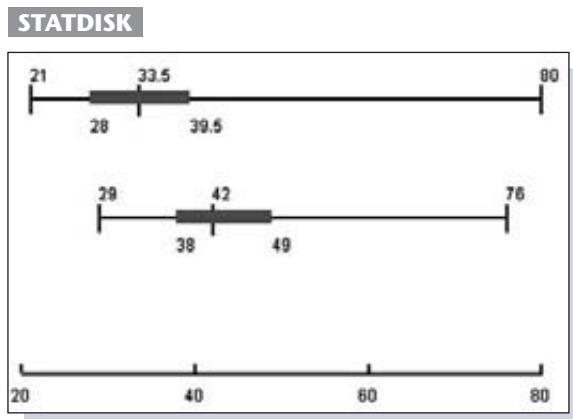
Es importante tomar en cuenta los valores extremos ya que, en muchos casos, un valor extremo podría tener un efecto muy importante en los estadísticos y en las conclusiones que se derivan de ellos. En algunos casos un valor extremo es un error que debe corregirse o eliminarse. En otros, un valor extremo es un valor válido que debe investigarse para obtener información importante. Algunos alumnos del autor reunieron datos consistentes en facturas y propinas de restaurantes, pero no encontraron valores extremos sobresalientes en esos valores muestrales. Sin embargo, un valor extremo es la propina de \$16,000 que se dio por una cuenta de \$8,899.78 en un restaurante. Esta propina la dio un ejecutivo de Londres no identificado al mesero Lenny Lorando, en el restaurante Nello's, ubicado en la ciudad de Nueva York. Lorando dijo que había atendido al cliente antes y que "él siempre es generoso, pero nunca antes de esta forma". Y agregó: "Tengo que hablarle a mi hermana acerca de él".

continúa

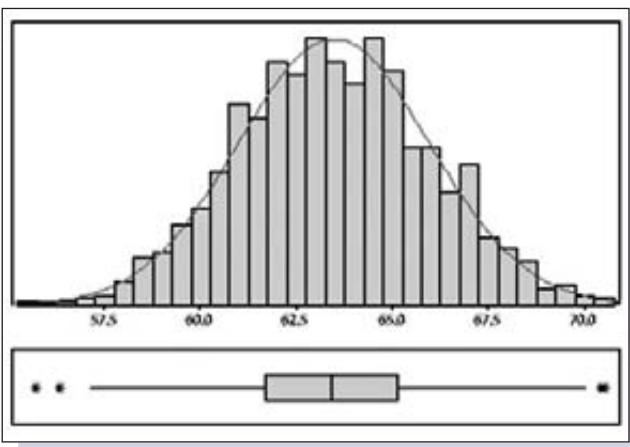
SOLUCIÓN

- El resumen de los 5 números consiste en el mínimo, Q_1 , la mediana, Q_3 y el valor máximo. Para calcular estos valores, primero ordene los datos (del más bajo al más alto). Es fácil identificar el mínimo de 21 y el máximo de 80 en la lista ordenada. Ahora proceda a calcular los cuartiles. Para el primer cuartil tenemos $Q_1 = P_{25} = 28$. [Si se utiliza el diagrama de flujo de la figura 3-6, el localizador es $L = (25/100)76 = 19$, que es un número entero, de manera que la figura 3-6 indica que Q_1 es el valor que está la mitad entre el valor 19° y el valor 20° en la lista ordenada]. La mediana es 33.5; es decir, el valor que está a la mitad entre los valores 38° y 39°. También encontramos que $Q_3 = 39.5$ al utilizar la figura 3-6 para calcular el percentil 75. Por lo tanto, el resumen de los 5 números es 21, 28, 33.5, 39.5 y 80.
- En la figura 3-7 se muestra la gráfica de cuadro para los datos. Primero usamos el mínimo (21) y el máximo (80) para determinar una escala de valores; después, graficamos los valores con el resumen de los 5 números, tal como se muestra.

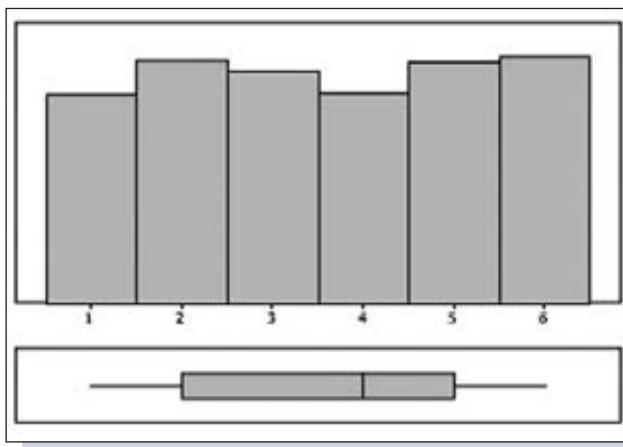
Las gráficas de cuadro son especialmente útiles para comparar conjuntos de datos, en particular cuando pertenecen a la misma escala. A continuación se muestran las gráficas de cuadro generadas con STATDISK para las edades de las mejores actrices (gráfica de cuadro superior) y los mejores actores, presentadas en la misma escala. En las dos gráficas de cuadro podemos ver que las edades de las actrices tienden a variar más y a ser menores que las edades de los actores.



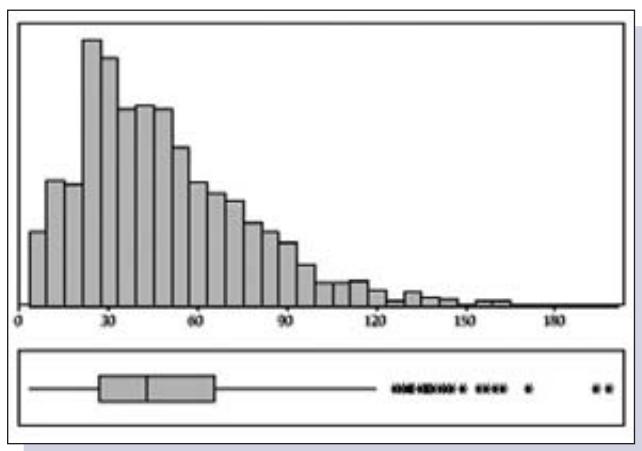
Observe las pantallas generadas por Minitab en la siguiente página, para descubrir la manera en que las gráficas de cuadro se relacionan con los histogramas al mostrar la distribución de los datos. En cada caso, la gráfica de cuadro se muestra debajo del histograma correspondiente. Dos de las gráficas de cuadro incluyen asteriscos para identificar valores extremos. Estos símbolos especiales forman parte de una gráfica de caja modificada, la cual se describe en el siguiente apartado.

Minitab

a) Distribución normal (con forma de campana)
1000 estaturas (en pulgadas) de mujeres

Minitab

b) Distribución uniforme
1000 lanzamientos de un dado

Minitab

c) Distribución sesgada
Ingresos (en miles de dólares) de 1000 profesores
de estadística

Gráficas de cuadro modificadas

Hasta ahora describimos **gráficas de cuadro esqueléticas** (o **regulares**), pero algunos paquetes de programas estadísticos elaboran gráficas de cuadro modificadas, las cuales representan los valores extremos como puntos especiales (como la gráfica de cuadro de la estatura de mujeres, generada con Minitab). Por ejemplo, Minitab usa asteriscos para mostrar valores extremos que se identifican utilizando un criterio específico que emplea el *rango intercuartilar* (RIC). En la sección 3-4 se señaló que el rango intercuartilar es: Rango intercuartilar (RIC) = $Q_3 - Q_1$. Luego, el valor del RIC se utiliza para identificar valores extremos de la siguiente manera:

Un valor es un valor extremo si está . . .

por arriba de Q_3 en una cantidad mayor que $1.5 \times \text{RIC}$

o

por debajo de Q_1 en una cantidad mayor que $1.5 \times \text{RIC}$

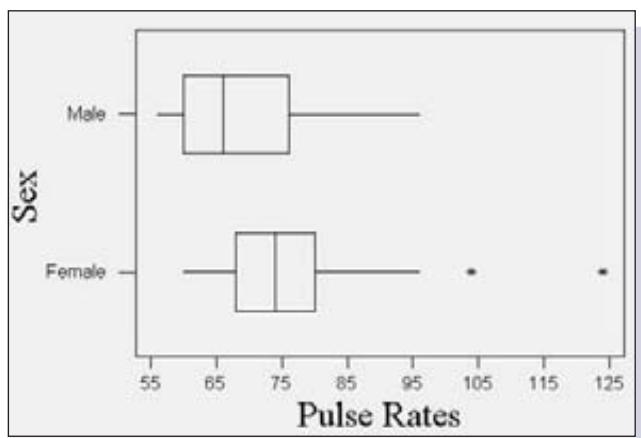
Una **gráfica de cuadro modificada** es aquella que se construye con las siguientes modificaciones: **1.** Se usa un símbolo especial (como un asterisco) para identificar valores extremos, como se definieron aquí, y **2.** la línea horizontal sólida se extiende únicamente hasta el valor del dato mínimo que no es un valor extremo y hasta el valor del dato máximo que no es un valor extremo. (*Nota: Los ejercicios sobre gráficas de cuadro modificadas sólo se encuentran en la sección “Más allá de lo básico”*).

EJEMPLO ¿Los hombres y las mujeres tienen los mismos pulsos cardíacos? Está bien documentado que existen diferencias fisiológicas importantes entre hombres y mujeres. Los hombres tienden a pesar más y a ser más altos que las mujeres. ¿Pero existe una diferencia en los pulsos cardíacos de hombres y mujeres? El conjunto de datos 1, en el apéndice B, lista los pulsos cardíacos de una muestra de 40 hombres y de una muestra de 40 mujeres. Más adelante en este libro, describiremos métodos estadísticos importantes que se pueden utilizar para probar diferencias de manera formal, pero por ahora exploraremos los datos para ver qué información podemos obtener. (Incluso si ya sabemos aplicar estos métodos estadísticos formales, es mejor explorar primero los datos antes de realizar el análisis formal).

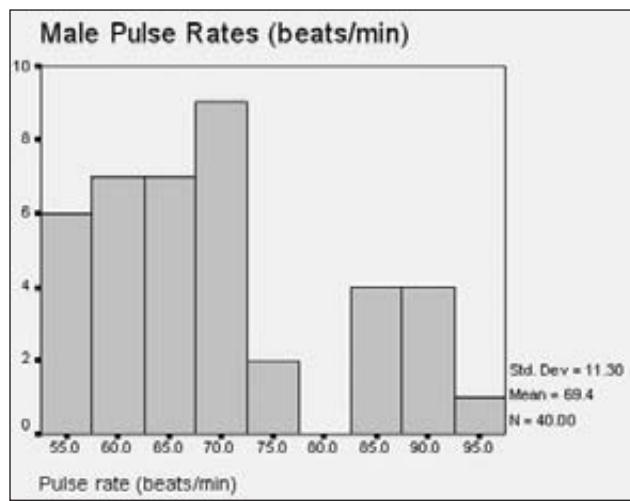
SOLUCIÓN Comencemos investigando los principales elementos de centralidad, variación, distribución, valores extremos y cambios con el paso del tiempo (la misma lista “CVDVT” que se presentó en la sección 2-1). A continuación se presentan las medidas de tendencia central (media), de variación (desviación estándar) y el resumen de los 5 números para los pulsos cardíacos del conjunto de datos 1. También se muestran las gráficas de caja de Minitab para cada muestra (donde los asteriscos indican valores que corresponden a valores extremos), un histograma generado en SPSS de los pulsos cardíacos de los hombres y un histograma de frecuencias relativas generado con SAS de los pulsos cardíacos de las mujeres.

	Media	Desviación estándar	Mínimo	Q ₁	Mediana	Q ₃	Máximo
Hombres	69.4	11.3	56	60.0	66.0	76.0	96
Mujeres	76.3	12.5	60	68.0	74.0	80.0	124

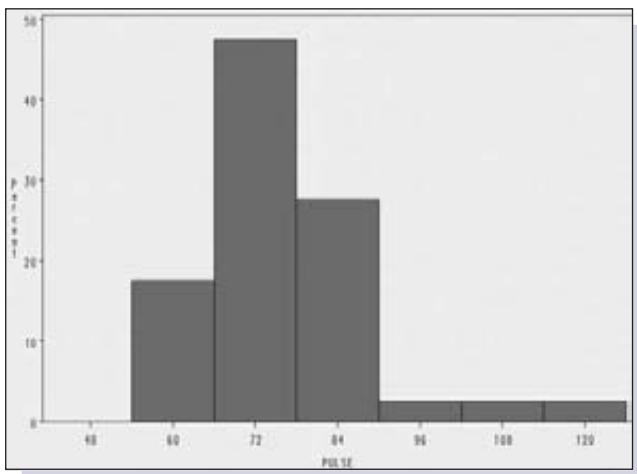
Minitab: Gráfica de caja de los pulsos cardíacos de hombres y mujeres



SPSS: Histograma de los pulsos cardíacos de hombres



SAS: Histograma de frecuencias relativas de los pulsos cardiacos de mujeres



INTERPRETACIÓN Al examinar y comparar los estadísticos y las gráficas, hicimos las siguientes observaciones importantes.

- **Centro:** Al parecer, las medias de los pulsos cardiacos de los hombres (69.4) y de las mujeres (76.3) son muy diferentes. Las gráficas de caja se basan en los valores mínimo y máximo, junto con los cuartiles, por lo que no describen los valores de las dos medias, aunque sí sugieren que los pulsos cardiacos de los hombres son un poco más bajos que los pulsos cardiacos de las mujeres, por lo que se infiere que la media de los pulsos cardiacos de los hombres es menor que la media de las mujeres.
- **Variación:** Las desviaciones estándar de 11.3 y 12.5 no son muy diferentes. Asimismo, las gráficas de cuadro describen la dispersión de los datos. Al parecer, las anchuras de las gráficas de caja no difieren mucho, lo que apoya la observación de que las desviaciones estándar no se ven demasiado diferentes.
- Los valores dados para el mínimo, el primer cuartil, la mediana, el tercer cuartil y el máximo sugieren que los valores de los hombres son más bajos en cada caso, de manera que los pulsos cardiacos de los hombres parecen ser más bajos que los de las mujeres. Existe una diferencia muy grande entre los valores máximos de 96 (para los hombres) y 124 (para las mujeres), pero el valor máximo de 124 es un valor extremo porque difiere mucho de los otros pulsos cardiacos. Ahora debemos preguntarnos si el pulso cardiaco de 124 es correcto o si se trata de un error, y debemos investigar el efecto de ese valor extremo en los resultados generales. Por ejemplo, si eliminamos el valor extremo de 124, ¿los pulsos cardiacos de los hombres continúan siendo menores que los de las mujeres? Lea los siguientes comentarios sobre los valores extremos.
- **Valores extremos:** Las gráficas de cuadro generadas en Minitab incluyen dos asteriscos que corresponden a los pulsos cardiacos de mujeres que se consideran valores extremos. El pulso cardiaco más elevado de 124 para las mujeres es un valor extremo que no afecta de manera drástica los resultados. Si eliminamos el valor de 124, la media del pulso cardiaco de las mu-

continúa



Detectores de mentiras humanos

Investigadores probaron la capacidad de 13,000 personas para determinar si alguien miente. Encontraron a 31 individuos con habilidades excepcionales para identificar mentiras. Estos detectores de mentiras humanos tenían tasas de éxito de alrededor del 90%. Además, descubrieron que los agentes federales y los alguaciles eran bastante hábiles para detectar mentiras, con tasas de éxito de alrededor del 80%. La profesora de psicología, Maureen O'Sullivan entrevistó a todas las personas hábiles para identificar mentiras, y comentó que "todas ellas ponen atención a indicios no verbales y a los matices en el uso de palabras; además, los aplican de forma diferente a distintas personas. Estos individuos son capaces de decir ocho cosas sobre una persona después de observar un video de dos segundos. Es increíble las cosas que estas personas pueden observar". Es posible usar métodos estadísticos para distinguir entre las personas que son capaces de detectar mentiras y aquellas que no lo son.

res cambia de 76.3 a 75.1, de manera que el pulso cardíaco de los hombres continúa siendo mucho más bajo. Incluso si eliminamos los valores extremos de 104 y 124, la media del pulso cardíaco de las mujeres cambia de 76.3 a 74.3, y la media de 69.4 de los hombres continúa siendo baja.

- **Distribuciones:** El histograma de frecuencias relativas generado con SAS para los pulsos cardíacos de mujeres y el histograma generado con SPSS para los pulsos cardíacos de hombres muestran distribuciones que no son radicalmente diferentes. Ambas gráficas parecen tener una distribución aproximadamente normal, como se podría esperar. Si el uso de un método estadístico específico requiere poblaciones con distribución normal (en forma de campana), este requisito queda aproximadamente satisfecho en ambos conjuntos de datos muestrales.

Ahora tenemos bastante información sobre la naturaleza de los pulsos cardíacos de hombres y mujeres. Con base en nuestra exploración, podemos concluir que, al parecer, los hombres tienen pulsos cardíacos con una media menor que las mujeres. Existen métodos más avanzados que podremos utilizar más adelante (como los métodos de la sección 9-3), pero las herramientas que se estudiaron en este capítulo nos ofrecen bastante información.

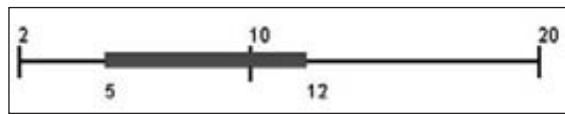
Pensamiento crítico

Armados con una lista de herramientas para investigar las características centrales, de variación, de distribución, de valores extremos y de cambio con el tiempo, podríamos sentirnos tentados a usar un procedimiento irreflexivo y de memorización, pero el pensamiento crítico es sumamente importante. Además de utilizar las herramientas presentadas en este capítulo, debemos tomar en cuenta otros factores relevantes que resultan cruciales para nuestras conclusiones. Podríamos plantear preguntas como éstas: ¿Es probable que la muestra sea representativa de la población, o la muestra tiene algún sesgo? ¿Cuál es la fuente de los datos y esa fuente podría ser alguien con intereses que afecten la calidad de los datos? Suponga, por ejemplo, que deseamos estimar el promedio de los ingresos de profesores de estadística. Además, suponga que enviamos cuestionarios por correo a 200 profesores de estadística y que recibimos 20 respuestas. Con estos datos podríamos calcular la media, la desviación estándar, construir gráficas, identificar valores extremos, etcétera, pero los resultados serán lo que los estadísticos denominan "basura". Se trata de una muestra de respuesta voluntaria, y lo más probable es que no sea representativa de la población de todos los profesores de estadística. Además de las herramientas estadísticas presentadas en este capítulo, ¡también debemos pensar!

3-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Gráfica de cuadro.** Remítase a la gráfica de cuadro generada con STATDISK que se presenta abajo. ¿Qué nos dicen los valores de 2, 5, 10, 12 y 20 acerca del conjunto de datos con el que se construyó la gráfica de cuadro?



Uso de la tecnología

Esta sección presentó los valores extremos, los resúmenes de los 5 números y las gráficas de cuadro. Para encontrar valores extremos, se acomodan los datos en orden de menor a mayor, después se examinan los valores máximo y mínimo para determinar si se encuentran muy lejos de los otros valores muestrales. El STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus pueden dar valores de cuartiles, de manera que es fácil elaborar el resumen de los 5 números. STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus pueden utilizarse para crear gráficas de cuadro. Ahora describimos los distintos procedimientos. (*Precaución:* Recuerde que los valores cuartilares calculados por medio de Minitab y la calculadora TI-83/84 Plus pueden diferir ligeramente de aquellos calculados a partir de la figura 3-6, por lo que tal vez las gráficas de cuadro también difieran ligeramente).

STATDISK Ingrese los datos en la ventana correspondiente, luego haga clic en **Data** y después en **Boxplot**. Haga clic en las columnas que desea incluir y luego haga clic en **Plot**.

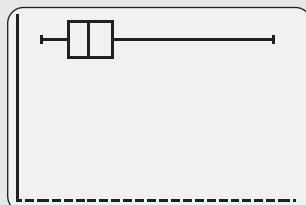
MINITAB Introduzca los datos en la columna C1, luego seleccione **Graph** y **Boxplot**. Con Release 14 seleccione un formato de la primera columna. Introduzca C1 en la celda superior izquierda y luego haga clic en **OK**.

EXCEL Aunque Excel no está diseñado para generar gráficas de cuadro, éstas pueden crearse utilizando la función Data Desk XL que complementa este libro. Primero introduzca los datos en la columna A, haga clic en **DDXL** y seleccione **Graphs and Plots**. Estando en la función **Type**, seleccione la opción de **Boxplot**. En el cuadro de diálogo haga clic en el ícono del lápiz e introduzca el rango de datos, como A1:A76, si usted tiene 76 valores listados en la columna A. Haga clic en **OK**. El resultado es una gráfica de cuadro modificada, con indicaciones de los datos ligeramente y notablemente extremos, tal como se describe en el ejercicio 15. También se muestran los valores del resumen de los 5 números.

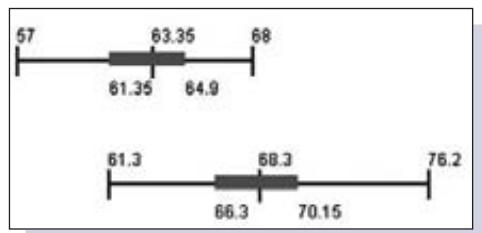
TI-83/84 PLUS Introduzca los datos muestrales en la lista L1. Ahora seleccione **STAT PLOT** presionando la segunda tecla

después de la tecla denominada Y =. Presione la tecla **ENTER**, después seleccione la opción de **ON** y seleccione el tipo de gráfica de cuadro que se ubica a la mitad del segundo renglón. Xlist debe indicar L1 y el valor Freq debe ser 1. Ahora presione la tecla **ZOOM** y seleccione la opción 9 para **ZoomStat**. Presione la tecla **ENTER** y debe aparecer la gráfica de cuadro. Puede utilizar las teclas con flechas para moverse hacia la derecha o izquierda, de manera que pueda leer los valores desde la escala horizontal. La siguiente pantalla corresponde a la figura 3-7, de la página 121.

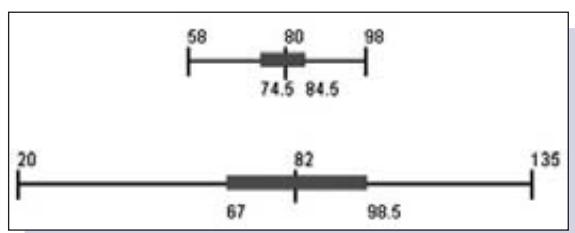
TI-83/84 Plus



- 2. Comparaciones entre gráficas de cuadro.** Remítase a las dos gráficas de cuadro generadas con STATDISK que se presentan a continuación y que se basan en la misma escala. Una de las gráficas representa las estaturas de hombres elegidos al azar y la otra, las estaturas de mujeres elegidas al azar. ¿Cuál gráfica de cuadro representa a las mujeres? ¿Cómo lo sabe?



- 3. Variación.** Las dos gráficas de cuadro que se muestran a continuación corresponden a los tiempos de servicio de dos compañías diferentes que reparan unidades de aire acondi-



cionado. Ambas gráficas utilizan la misma escala. La gráfica de cuadro de arriba corresponde a la empresa Sigma Air Conditioning Company, mientras que la gráfica de abajo corresponde a la empresa Newport Repair Company. ¿Qué empresa tiene tiempos de reparación menos variables? ¿Qué empresa debe tener costos más predecibles? ¿Por qué?

- 4. Valores extremos.** Un conjunto de 20 valores muestrales incluye un valor extremo que está muy alejado de los 19 valores restantes. ¿Qué efecto tiene ese valor extremo en cada uno de los siguientes estadísticos: media, mediana, desviación estándar y rango?
- 5. Prueba de semillas de maíz.** En 1908 William Gosset publicó el artículo “The Probable Error of a Mean” bajo el seudónimo de “Student” (*Biometrika*, vol. 6, núm. 1). El autor incluyó los datos que se muestran a continuación, acerca de dos diferentes tipos de semillas de maíz (normales y secadas en horno) que se utilizaron en parcelas de tierra adyacentes. Los valores listados corresponden a la producción de cabezas (o mazorcas) de maíz en libras por acre. Utilizando la producción de las *semillas regulares*, calcule el resumen de los 5 números y construya una gráfica de cuadro.

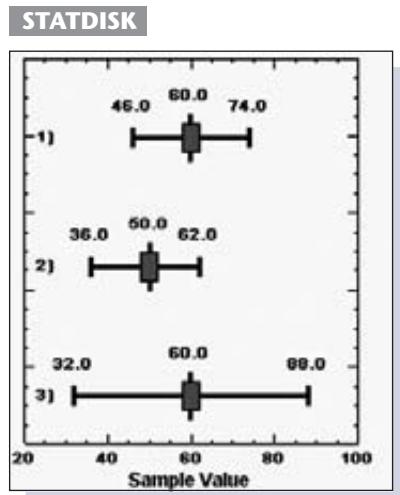
Regulares	1903	1935	1910	2496	2108	1961	2060	1444	1612	1316	1511
Secadas en horno	2009	1915	2011	2463	2180	1925	2122	1482	1542	1443	1535

- 6. Prueba de semillas de maíz.** Utilice los datos de la producción de *semillas secadas en horno* del ejercicio 5 para elaborar el resumen de los 5 números y construir una gráfica de cuadro. ¿Los resultados parecen ser muy diferentes de los obtenidos en el ejercicio 5?
- 7. Pesos de monedas de 25 centavos.** Remítase al conjunto de datos 14 y utilice los pesos de las monedas de plata de 25 centavos (acuñadas antes de 1964) para elaborar el resumen de los cinco números y construir una gráfica de cuadro.
- 8. Pesos de monedas de 25 centavos.** Remítase al conjunto de datos 14 y utilice los pesos de las monedas de 25 centavos (acuñadas después de 1964) para elaborar el resumen de los 5 números y construir una gráfica de cuadro. ¿Los resultados parecen ser muy diferentes de los obtenidos en el ejercicio 7?
- 9. Estaturas de osos.** Remítase al conjunto 6 de datos, que incluye las estaturas (en pulgadas) de 54 osos que fueron anestesiados y medidos. Elabore el resumen de los 5 números y construya una gráfica de cuadro. ¿La distribución de las estaturas parece ser simétrica o sesgada?
- 10. Temperaturas corporales.** Remítase al conjunto 2 de datos en el apéndice B, que incluye las 106 temperaturas corporales de las 12 a. m. el día 2. Elabore el resumen de los 5 números y construya una gráfica de cuadro; luego determine si los valores de la muestra sustentan la creencia generalizada de que la media de la temperatura corporal es de 98.6°F.
- 11. Valores del IMC.** Remítase al conjunto de datos 1 en el apéndice B y utilice los valores del índice de masa corporal (IMC) de los hombres para elaborar el resumen de los 5 números y construir una gráfica de cuadro.
- 12. Valores del IMC.** Remítase al conjunto de datos 1 en el apéndice B y utilice los valores del índice de masa corporal (IMC) de las mujeres para elaborar el resumen de los 5 números y construir una gráfica de cuadro. ¿Los resultados parecen ser muy diferentes de los obtenidos en el ejercicio 11?

3-5 MÁS ALLÁ DE LO BÁSICO

- 13.** Observe la siguiente pantalla de STATDISK con tres gráficas de cuadro que representan la medida de la duración (en meses) de las muestras de tres acumuladores para au-

tomóvil diferentes. Si usted fuera el gerente de una flotilla de automóviles y tuviera que elegir una de las tres marcas, ¿cuál gráfica de cuadro representa la marca que debe seleccionar? ¿Por qué?



- 14. Valores extremos.** En vez de considerar que un valor es extremo si “está muy alejado de casi todos los demás valores de los datos”, considere un dato como extremo si está por arriba de Q_3 en una cantidad mayor que $1.5 \times \text{RIC}$ o por debajo de Q_1 en una cantidad mayor que $1.5 \times \text{RIC}$. Utilice el siguiente conjunto de datos y calcule lo que se le pide:
- El resumen de los 5 números
 - El rango intercuartilar (RIC)
 - Los valores extremos

2 3 4 5 7 9 14 15 15 16 19 32 50

- 15. Datos ligeramente extremos y notablemente extremos.** Algunos paquetes estadísticos de cómputo construyen gráficas de cuadro que identifican datos ligeramente extremos (que a menudo se grafican como puntos sólidos) y datos notablemente extremos (que a menudo se grafican como círculos huecos). Los **datos ligeramente extremos** están por debajo de Q_1 o por arriba de Q_3 en una cantidad mayor que $1.5 \times \text{RIC}$, pero no mayor que $3 \times \text{RIC}$. Los datos notablemente extremos están por debajo de Q_1 en más de $3 \times \text{RIC}$ o por arriba de Q_3 en más de $3 \times \text{RIC}$. Remítase al conjunto de datos 10 en el apéndice B y utilice las cantidades de lluvia del lunes para identificar cualquier dato ligeramente o notablemente extremo.

Repaso

En este capítulo consideramos las medidas de tendencia central, las medidas de variabilidad, las medidas de posición relativa y los métodos generales para describir, explorar y comparar conjuntos de datos. Cuando se investiga un conjunto de datos, por lo general las siguientes características son muy importantes:

1. *Centro*: un valor representativo o promedio.
2. *Variación*: una medida de la cantidad en que varían los valores.

3. **Distribución:** la naturaleza o forma de la distribución de los datos (como normal, uniforme o sesgada).
4. **Valores extremos:** valores muestrales que se ubican muy lejos de la gran mayoría del resto de valores muestrales.
5. **Tiempo:** características cambiantes de los datos a través del tiempo.

Es especialmente importante desarrollar las siguientes habilidades y comprender los siguientes conceptos:

- Calcular medidas de tendencia central como la media y la mediana (sección 3-2).
- Calcular medidas de variación como la desviación estándar, la varianza y el rango (sección 3-3).
- *Comprender e interpretar* la desviación estándar utilizando herramientas como la regla práctica del intervalo (sección 3-3).
- Comparar valores individuales utilizando puntuaciones z , cuartiles o percentiles (sección 3-4).
- Investigar y explorar la dispersión de los datos, el centro de los datos y el rango de los valores por medio de la construcción de una gráfica de cuadro (sección 3-5).

Además de calcular valores de estadísticos y crear gráficas, es sumamente importante *comprender e interpretar* estos resultados. Por ejemplo, debe quedar claro que la desviación estándar es una medida de la variación de datos; usted debe ser capaz de usar la desviación estándar para distinguir entre valores frecuentes e infrecuentes.

Conocimientos estadísticos y pensamiento crítico

1. **Centro y variación.** Un ingeniero de control de calidad diseña procedimientos de reparación para reducir la desviación estándar de los tiempos de reparación. ¿Esto implica que las reparaciones se hacen en menos tiempo? ¿Por qué?
2. **Especificaciones de producción.** Al diseñar el procedimiento de producción de baterías utilizadas en marcapasos cardíacos, un ingeniero especifica que “las baterías deben tener una vida media mayor de 10 años, y se puede ignorar la desviación estándar de la vida de las baterías”. Si la vida media de la batería es mayor de 10 años, ¿se puede ignorar la desviación estándar? ¿Por qué?
3. **Valor extremo.** Después de elegir al azar a 50 poseedores de tarjetas de crédito, se calculan las cantidades que deben en la actualidad. Luego se determinan los valores de la media, la mediana y la desviación estándar. Después se incluye una cantidad adicional de \$1,000,000. ¿Qué efecto tendrá esta cantidad adicional sobre la media, la mediana y la desviación estándar?
4. **Encuesta de Internet.** Un proveedor de servicios de Internet realiza una encuesta anónima por ese medio a sus suscriptores y 2500 de ellos responden reportando los valores de los automóviles que poseen en la actualidad. Como el tamaño de la muestra es tan grande, ¿es probable que los resultados produzcan una media que sea muy cercana a la media del valor de todos los automóviles que son propiedad de los estadounidenses? ¿Por qué?

Ejercicios de repaso

- 1. Altura de árboles.** En un estudio sobre la relación entre las alturas y los diámetros del tronco de árboles, estudiantes de botánica reunieron datos muestrales. A continuación se presentan las circunferencias de los árboles (en pies). Los datos se basan en los resultados de “Tree measurements”, por Stanely Rice, *American Biology Teacher*, vol. 61, núm. 9. Utilice las circunferencias y calcule *a*) la media, *b*) la mediana, *c*) la moda, *d*) la mitad del rango, *e*) el rango, *f*) la desviación estándar, *g*) la varianza, *h*) Q_1 , *i*) Q_3 , *j*) P_{10} .

1.8 1.9 1.8 2.4 5.1 3.1 5.5 5.1 8.3 13.7 5.3 4.9 3.7 3.8 4.0 3.4 5.2 4.1 3.7 3.9

2. **a.** Con los resultados del ejercicio 1, convierta la circunferencia de 13.7 pies a una puntuación *z*.
b. En el contexto de estos datos muestrales, ¿la circunferencia de 13.7 pies es “infrecuente”? ¿Por qué?
c. Utilice la regla práctica del intervalo e identifique cualquier otra circunferencia que sea infrecuente.
3. **Distribución de frecuencias.** Utilice las mismas circunferencias de los árboles del ejercicio 1 para construir una distribución de frecuencias. Utilice siete clases, donde la primera clase tenga un límite inferior de 1.0, con una anchura de clase de 2.0.
4. **Histograma.** Utilice la distribución de frecuencias del ejercicio 3 y construya un histograma e identifique la naturaleza general de la distribución (como uniforme, normal o sesgada).
5. **Gráfica de cuadro.** Con las mismas circunferencias del ejercicio 1, construya una gráfica de caja e identifique los valores que conforman el resumen de los 5 números.
6. **Comparación de puntuaciones.** Un psicólogo industrial de Citation Corporation desarrolla dos pruebas diferentes para medir la satisfacción en el trabajo. ¿Cuál puntuación es mejor: una puntuación de 72 en la prueba de gerentes, que tiene una media de 80 y una desviación estándar de 12, o una puntuación de 19 en la prueba de empleados de producción, que tiene una media de 20 y una desviación estándar de 5? Explique.
7. **Estimación de la media y la desviación estándar**
 - a.** Calcule la media de la antigüedad de los automóviles que conducen los estudiantes de su universidad.
 - b.** Utilice la regla práctica del intervalo para hacer una estimación de la desviación estándar de la antigüedad de los automóviles que conducen los estudiantes de su universidad.
8. **Interpretación de la desviación estándar.** La estatura media de los hombres es de 69.0 pulgadas, con una desviación estándar de 2.5 pulgadas. Utilice la regla práctica del intervalo para estimar la estatura mínima “común” y la estatura máxima “común”. En este contexto, ¿una estatura de 72 pulgadas (o 6 pies) es infrecuente? ¿Por qué?

Ejercicios de repaso acumulativo

1. **Medidas de árboles.** Remítase a la muestra de circunferencia de árboles (en pies) del ejercicio 1 de repaso.
 - a.** ¿Los valores dados pertenecen a una población que es discreta o continua?
 - b.** ¿Cuál es el nivel de medición de sus valores? (Nominal, ordinal, de intervalo o de razón).

- 2. a.** Un conjunto de datos tiene un nivel de medición nominal; usted desea obtener el valor de un dato representativo. ¿Cuál de los siguientes es más apropiado: la media, la mediana, la moda o la mitad del rango? ¿Por qué?
- b.** Un botánico desea obtener datos sobre las plantas que crecen en las casas. Se obtiene una muestra al llamar por teléfono a las primeras 250 personas listadas en el directorio telefónico local. ¿Qué tipo de muestreo se está utilizando? (Aleatorio, estratificado, sistemático, por conglomerados, de conveniencia).
- c.** Se realiza una encuesta de salida al entrevistar a todas las personas que salen de las casillas de 50 distritos de votación elegidos al azar. ¿Qué tipo de muestreo se está utilizando? (Aleatorio, estratificado, sistemático, por conglomerados, de conveniencia).
- d.** Un fabricante hace varas fertilizadas para el cultivo de plantas. Un gerente descubre que las cantidades de fertilizante que se colocan en las varas no son muy consistentes, de manera que algunos procesos de fertilización duran más tiempo del anunciado, pero otros no cubren ese periodo. El gerente desea mejorar la calidad logrando que las cantidades de fertilizante sean más consistentes. Al analizar las cantidades de fertilizante, ¿cuál de los siguientes estadísticos es más relevante: la media, la mediana, la moda, la mitad del rango, la desviación estándar, el primer cuartil o el tercer cuartil? ¿El valor de ese estadístico debe incrementarse, disminuirse o dejarse sin cambios?

Actividades de cooperación en equipo

1. Actividad fuera de clase ¿Influyen las cifras de anclaje en los estimados? En el artículo “Weighing Anchors”, de la revista *Omni*, el autor John Rubin observó que cuando la gente estima un valor, su estimación suele estar “anclada a” (o influida por) un número anterior, aun cuando ese número no tenga ninguna relación con la cantidad que se estima. Para demostrar esto, pidió a varias personas que le dieran un estimado rápido del valor de $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$. La respuesta promedio fue 2250, pero cuando se invirtió el orden de los números el promedio fue 512. Rubin explicó que cuando iniciamos un cálculo con números grandes (como en $8 \times 7 \times 6$), nuestros estimados tienden a ser grandes. Señaló que tanto 2250 como 512 son mucho menores que el producto correcto, 40,320. El artículo sugiere que números irrelevantes podrían desempeñar un papel importante en los avalúos de bienes raíces, así como en las estimaciones del valor de un automóvil o de la posibilidad de una guerra nuclear.

Realice un experimento para probar este argumento. Seleccione a algunos sujetos y pídaleles que estimen rápidamente el valor de

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

Después, seleccione a otros sujetos y pídaleles que estimen rápidamente el valor de

$$1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$$

Registre los estimados junto con el orden utilizado. Diseñe con cuidado el experimento, de forma que las condiciones sean uniformes y que los dos grupos muestrales se seleccionen de forma que se minimice cualquier sesgo. No describa la posible explicación a los sujetos sino hasta después de que hayan dado sus estimaciones. Compare los dos conjuntos de resultados muestrales a través del uso de los métodos de este capítulo. Prepare un reporte impreso que incluya los datos reunidos, los métodos utilizados explicados con detalle, el método de análisis, las gráficas y/o estadísticos relevantes, así como las conclusiones. Incluya una crítica dando razones por las que los resultados podrían ser incorrectos y describa formas de mejorar el experimento.

2. Actividad fuera de clase Cada equipo, formado por tres o cuatro estudiantes, debe reunir un conjunto original de datos que estén a un nivel de medición de intervalo o de razón. Proporcionen lo siguiente: 1. una lista de valores muestrales, 2. resultados de computadora impresos con estadísticos descriptivos y gráficas, y 3. una descripción por escrito de la naturaleza de los datos, el método de recolección y las características importantes.

Proyecto tecnológico

Cuando se manejan conjuntos grandes de datos, la introducción manual de éstos suele ser tediosa y requiere de mucho tiempo. Hay mejores cosas que hacer con su tiempo, como aprender los principios aerodinámicos de un “Frisbee”. Remítase al conjunto de datos 17 del apéndice B, que incluye las distancias de los *home runs* de tres jugadores de béisbol excepcionales: Barry Bonds (temporada 2001), Mark McGwire (temporada 1998) y Sammy Sosa (temporada 1998). En vez de introducir manualmente las 209 distancias de los tres conjuntos de datos, utilice la calculadora TI-83/84 Plus,

STATDISK, Minitab o Excel para cargar el conjunto de datos, los cuales están disponibles en el CD incluido en este libro. Proceda a generar histogramas y obtenga estadísticos apropiados que le permitan comparar los tres conjuntos de datos. ¿Existen algunas diferencias significativas? ¿Existen algunos valores extremos? ¿Parece que los jugadores que golpean más lejos hacen más *home runs*? ¿Por qué? Analice los últimos dígitos de las distancias y determine si los valores parecen ser estimaciones o mediciones. Escriba un reporte breve que incluya sus conclusiones y gráficas de apoyo.

De los datos a la decisión

Pensamiento crítico

¿Existe alguna configuración del teclado que sea más eficiente que la que usa la mayoría de nosotros? La configuración tradicional del teclado se denomina *QWERTY* por la posición de las letras *QWERTY* en la fila superior. Desarrollada en 1872, la configuración *QWERTY* fue diseñada para obligar a las personas a escribir más despacio para que las primeras máquinas de escribir no se trabaran. En 1936 se desarrolló el teclado Dvorak, el cual, se supone, es un arreglo más eficiente al presentar las teclas más utilizadas en la fila intermedia (o fila “base”), donde son más accesibles.

Un artículo de la revista *Discover* sugirió que uno puede medir la facilidad de la escritura utilizando el siguiente sistema de calificación: a cada letra de la fila intermedia asigne un 0, a cada letra de la fila superior asigne un 1 y a cada letra de la fila inferior asigne un 2. (Véase, “Typecasting”, por Scott Kim, *Discover*). Con el uso de este sistema de calificación en cada una de las 52 palabras del preámbulo de la Constitución estadounidense, se obtienen las calificaciones que se muestran a continuación.

Interpretación de los resultados

Como es probable que la comparación visual de los dos conjuntos de datos no revele

Calificaciones de las palabras del teclado QWERTY para el preámbulo de la Constitución estadounidense

2	2	5	1	2	6	3	3	4	2
4	0	5	7	7	5	6	6	8	10
7	2	2	10	5	8	2	5	4	2
6	2	6	1	7	2	7	2	3	8
1	5	2	5	2	14	2	2	6	3
1	7								

Calificaciones de las palabras del teclado Dvorak para el preámbulo de la Constitución estadounidense

2	0	3	1	0	0	0	0	2	0
4	0	3	4	0	3	3	1	3	5
4	2	0	5	1	4	0	3	5	0
2	0	4	1	5	0	4	0	1	3
0	1	0	3	0	1	2	0	0	0
1	4								

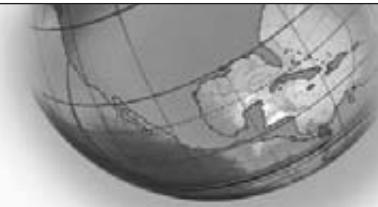
mucho, utilice los métodos del capítulo 2 y de éste para describirlos, explorarlos y compararlos. Responda al menos una de las siguientes preguntas:

- a. ¿Al parecer hay una diferencia significativa en la facilidad de escritura entre las dos configuraciones del teclado? ¿La configuración Dvorak parece ser más eficiente?
- b. Si existe una diferencia significativa y la configuración Dvorak es más eficiente, ¿por qué no se ha adoptado?
- c. ¿Son representativas las palabras del preámbulo? ¿Se llegaría a la misma con-

clusión con una muestra de palabras de este libro, que fue escrito más recientemente?

- d. Cada valor muestral es la calificación total de las letras de una *palabra*, pero ¿debenmos utilizar palabras o simplemente letras individuales?
- e. ¿Es apropiado el sistema de calificación? (A las letras de la fila superior se les asigna 1, a las letras de la fila intermedia se les asigna 0 y a las letras de la fila inferior se les asigna 2). ¿Existe un sistema de calificación diferente que refleje mejor la dificultad para escribir?





Proyecto de Internet

Uso de estadísticos para resumir datos

No podemos subestimar la importancia de la estadística como herramienta para resumir datos. Por ejemplo, considere conjuntos de datos como las edades de todos los estudiantes de su escuela o los ingresos anuales de cada habitante de Estados Unidos. En papel, estos conjuntos de datos serían listas de números tan largas que sería difícil entenderlos e interpretarlos por sí mismos. En el capítulo anterior aprendimos diversas herramientas gráficas que se utilizan para representar este tipo de conjuntos de datos. El presente capítulo se enfocó en el uso de números estadísticos para resumir varios aspectos de los datos.

La capacidad para resumir datos con estadísticos es tan importante como la capacidad de

interpretar tales estadísticos cuando se presenten. Dado un número como la media aritmética, usted no sólo debe comprender lo que ésta le indica acerca de los datos subyacentes, sino también los estadísticos adicionales que necesita para poner el valor de la media en contexto.

El proyecto de Internet de este capítulo le ayudará a desarrollar estas habilidades en el uso de datos de diversos campos como la meteorología, el entretenimiento y la salud. Además, descubrirá usos para estadísticos como la media geométrica, que jamás había imaginado.

El sitio Web de este capítulo se encuentra en

<http://www.pearsoneducacion.net/triola>

La estadística en el trabajo

“Los conocimientos básicos sobre probabilidad, resumen de datos y los principios de la estadística inferencial son esenciales para comprender el método científico y para evaluar los informes de los estudios científicos”.



Robert S. Holzman, MD

Profesor de medicina y de medicina ambiental, Escuela de Medicina de NYU; epidemiólogo en Bellevue Hospital Center, Ciudad de Nueva York.

El doctor Holzman es internista especializado en enfermedades infecciosas. Es responsable del Programa de control de infecciones en el Bellevue Hospital. Además, es profesor de estudiantes de medicina y de posdoctorado en las materias de Enfermedades clínicas infecciosas y Epidemiología.

¿Cómo utiliza la estadística en su trabajo y qué conceptos estadísticos específicos emplea?

Muchas de las tareas que realizo todos los días implican un análisis estadístico aplicado, incluyendo la determinación del tamaño de muestras, el análisis de ensayos clínicos y experimentos de laboratorio, así como el desarrollo de modelos de regresión para estudios retrospectivos, principalmente con el uso de la regresión logística. También hago un seguimiento de las tasas de infección hospitalarias por medio de gráficas de control.

Por favor, describa un ejemplo específico donde el uso de la estadística le haya servido para mejorar una práctica o servicio.

Hace dos meses, una enfermera supervisora detectó un aumento del aislamiento de cierto tipo de bacteria en los pacientes de una unidad de cuidados intensivos. Tomamos medidas para solucionar ese incremento y utilizamos gráficas de control para verificar que los niveles de la bacteria estuvieran regresando a sus niveles de línea base.

¿Qué estudios de estadística se necesitan para realizar un trabajo como el suyo? ¿Qué otros requisitos educativos se necesitan?

Para ser un médico académico es necesario asistir a la escuela de medicina y después realizar estudios de posgrado al menos durante cinco años, a veces más. En muchos casos, ahora los estudiantes están

combinando su educación médica con un programa de estudios de posgrado orientado hacia la investigación. Yo adquirí mis conocimientos estadísticos y epidemiológicos al combinar una relación con profesionales de la estadística en el trabajo, la lectura de textos estadísticos y el trabajo en cursos de posgrado. En la actualidad, estos conocimientos se obtienen en programas educativos, pero se necesita experiencia para dominar el campo.

¿Cree usted que, en el lugar que en que trabaja, las personas que solicitan un empleo se ven favorecidas si tienen estudios de estadística?

Definitivamente, la capacidad de aplicar conocimientos estadísticos y epidemiológicos para evaluar la literatura médica se considera una ventaja para los médicos, incluso para aquellos que dan atención clínica. El trabajo como epidemiólogo requiere de estudios estadísticos adicionales.

¿Recomendaría a los estudiantes universitarios de hoy que estudien estadística? ¿Por qué?

Los conocimientos básicos sobre probabilidad, resumen de datos y los principios de la estadística inferencial son esenciales para comprender el método científico y para evaluar los informes de los estudios científicos. Tales informes aparecen todos los días en los periódicos, y el hecho de conocer las aportaciones del informe sirve para disminuir la aceptación irreflexiva de nuevos “hechos”.



Probabilidad

4



- 4-1 Panorama general**
- 4-2 Fundamentos**
- 4-3 Regla de la suma**
- 4-4 Regla de la multiplicación: Fundamentos**
- 4-5 Regla de la multiplicación: Complementos y probabilidad condicional**
- 4-6 Probabilidades por medio de simulaciones**
- 4-7 Conteo**
- 4-8 Teorema de Bayes (en CD-ROM)**

¿Debe preocuparse de que le realicen una prueba de detección de drogas cuando solicite un trabajo?

Según la American Management Association, alrededor del 70% de las empresas estadounidenses realizan pruebas de detección de drogas al menos a algunos empleados y aspirantes. El U.S. National Institute on Drug Abuse afirma que aproximadamente el 15% de los jóvenes entre 18 y 25 años consumen drogas ilegales. Quest Diagnostic estima que el 3% de la fuerza laboral general de Estados Unidos consume marihuana. Supongamos que usted solicitó un empleo, tiene excelentes aptitudes (las cuales incluyen la aprobación exitosa de un curso de estadística), le hicieron una prueba de consumo de marihuana y no le dieron el empleo. Usted podría sospechar que no pasó el examen de marihuana, aun cuando no consume esta droga.

Análisis de los resultados

La tabla 4-1 muestra los resultados de la prueba “1-Panel-THC” para identificar el consumo de marihuana. Este dispositivo de prueba cuesta \$5.95 y la empresa Drug Test Success lo distribuye. Los resultados de la prueba fueron confirmados con cromatografía de gases y espectrometría de masas, que la empresa describe como “el método de confirmación preferido”. (Esos resultados se basan en el uso de 50 ng/mL como nivel de corte para determinar la presencia de marihuana). Con base en los resultados de la tabla 4-1, ¿qué probabilidades hay de que la prueba indique que usted consumió marihuana, aunque

no sea así? Cuando una prueba muestra la presencia de alguna condición, como una enfermedad o los residuos de alguna droga, se dice que el resultado de la prueba es *positivo*. Cuando la prueba indica un resultado positivo, pero la condición en realidad no está presente, el resultado es un *falso positivo*. Es decir, un falso positivo es un error en el que la prueba indica la presencia de una condición, cuando en realidad esta última no se presenta. En este caso, el aspirante al empleo podría sentirse angustiado por la probabilidad de un resultado falso positivo, ya que sería un error que provocaría de manera injusta la negación del empleo. (El contratante podría sentirse preocupado por otro tipo de error, un *falso negativo*, que se presenta cuando una prueba indica que el aspirante no consume marihuana, cuando en realidad sí lo hace. Este falso negativo podría causar la contratación de un individuo que consume marihuana, y este error puede ser grave para algunos trabajos, como los que realizan los pilotos, los cirujanos o los ingenieros de trenes).

En este capítulo analizaremos preguntas relevantes como éstas: Dados los resultados muestrales de la tabla 4-1, ¿cuál es la probabilidad de un resultado falso positivo? ¿Cuál es la probabilidad de un resultado falso negativo? ¿Esas probabilidades son lo suficientemente bajas como para que los aspirantes y los contratantes no se preocupen por tomar decisiones incorrectas motivadas por resultados erróneos de las pruebas?

Tabla 4-1 Resultados de exámenes sobre el consumo de marihuana

¿Los sujetos realmente consumen marihuana?		
	Sí	No
Resultado de prueba positivo (La prueba indica que la marihuana está <i>presente</i>).	119 (verdadero positivo)	24 (falso positivo)
Resultado de prueba negativo (La prueba indica que la marihuana está <i>ausente</i>).	3 (falso negativo)	154 (verdadero negativo)

4-1 Panorama general

La probabilidad conforma los cimientos sobre los cuales se construyen los métodos importantes de la estadística inferencial. Como un sencillo ejemplo, suponga que usted ha creado un procedimiento de selección del género y afirma que éste incrementa en gran medida la probabilidad de que un bebé sea niña. Suponga que los resultados de pruebas independientes con 100 parejas demuestran que su procedimiento dio por resultado 98 niñas y sólo 2 niños. Aunque existe la probabilidad de que nazcan 98 niñas en 100 nacimientos sin ningún tratamiento especial, tal probabilidad es tan baja que se rechazaría como una explicación razonable. En cambio, se reconocería de manera general que los resultados indican fuertes evidencias para afirmar que la técnica de selección del género es efectiva. Ésta es precisamente la forma de pensar de los especialistas en estadística: rechazan las explicaciones basadas en probabilidades muy bajas y utilizan la *regla del suceso infrecuente para la estadística inferencial*.

Regla del suceso infrecuente para estadística inferencial

Si, bajo un supuesto dado, la probabilidad de un suceso particular observado es extremadamente pequeña, concluimos que el supuesto probablemente es incorrecto.

El objetivo principal de este capítulo es desarrollar una buena comprensión de los valores de probabilidad, los cuales se usarán en los capítulos siguientes. Un objetivo secundario es desarrollar las habilidades básicas necesarias para determinar los valores de probabilidad en una variedad de circunstancias importantes.



4-2 Fundamentos

Concepto clave En esta sección se presenta el concepto básico de la *probabilidad* de un suceso. Se presentarán tres métodos diferentes para calcular valores de probabilidad. Veremos que los valores de probabilidad se expresan en números entre 0 y 1, inclusive. Sin embargo, el objetivo más importante de esta sección consiste en aprender a *interpretar* valores de probabilidad. Por ejemplo, debemos comprender que una pequeña probabilidad, como 0.001, corresponde a un suceso que es *infrecuente*, en el sentido de que ocurre en pocas ocasiones. En capítulos posteriores hablaremos de valores específicos llamados “valores *P*” y veremos que éstos tienen un papel sumamente importante en diversos métodos de estadística inferencial. Sin embargo, tales valores *P* sólo son valores de probabilidad, como se describe en esta sección. Concéntrese en desarrollar una intuición para interpretar valores de probabilidad, en especial los que son relativamente pequeños.

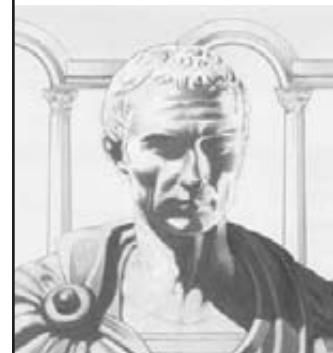
Al considerar la probabilidad, tratamos con procedimientos (como tirar un dado, contestar una pregunta de opción múltiple en un examen, o ser sometido a una prueba de consumo de drogas) que producen resultados.

Definiciones

Un **suceso** es cualquier conjunto de resultados o consecuencias de un procedimiento.

Un **suceso simple** es un resultado o un suceso que ya no puede desglosarse en componentes más simples.

El **espacio muestral** de un procedimiento se compone de todos los sucesos *simples* posibles. Es decir, el espacio muestral está formado por todos los resultados que ya no pueden desglosarse más.



EJEMPLOS En la siguiente presentación usamos f para indicar que se trata de una niña y una m para indicar que se trata de un varón.

Procedimiento	Ejemplo de suceso	Espacio muestral completo
Un solo nacimiento	niña (suceso simple)	{f, m}
3 nacimientos	2 niñas y un niño (ffm, fmf, mff son sucesos simples que dan como resultado 2 niñas y un niño)	{fff, ffm, fmf, fmm, mff, mfm, mmf, mmm}

Con un solo nacimiento, el hecho de que resulte niña es un *suceso simple*, porque no se puede desglosar en eventos más simples. Con tres nacimientos, el suceso de “2 niñas y un niño” no es un *suceso simple*, porque se puede desglosar en sucesos más simples, como ffm, fmf o mff. Con tres nacimientos, el *espacio muestral* consiste en los 8 sucesos simples nombrados antes. Con tres nacimientos, el resultado ffm se considera un suceso simple, porque no se puede desglosar en eventos más simples. Podríamos pensar de manera incorrecta que ffm se puede desglosar aún más en los resultados individuales f, f y m, pero f, f y m no son los resultados individuales de tres nacimientos. Con tres nacimientos, existen exactamente 8 resultados que son sucesos simples: fff, ffm, fmf, fmm, mff, mfm, mmf y mmm.

Existen diferentes formas para definir la probabilidad de un suceso; por ahora expondremos tres enfoques. Para iniciar, presentamos una lista de algunas notaciones básicas.

Notación de probabilidades

P denota una probabilidad.

A, B y C denotan sucesos específicos.

$P(A)$ denota la probabilidad de que ocurra el suceso A.

Regla 1: Aproximación de la probabilidad por frecuencias relativas

Realice (u observe) un procedimiento un gran número de veces y cuente las veces que el suceso A ocurre en realidad. Con base en estos resultados reales, $P(A)$ se *estima* de la siguiente forma:

$$P(A) = \frac{\text{número de veces que ocurrió } A}{\text{número de veces que se repitió el ensayo}}$$

Probabilidades que desafían la intuición

En ciertos casos, nuestros estimados subjetivos de valores de probabilidad son drásticamente distintos de las probabilidades reales. He aquí un ejemplo clásico: si usted inhala profundamente, hay una probabilidad mayor al 99% de que inhale una molécula que haya sido exhalada en el último aliento de César al morir. Con este mismo ánimo morboso y poco intuitivo, si la fatal taza con cicuta que mató a Sócrates era en su mayor parte agua, el siguiente vaso de agua que usted beba muy probablemente contendrá una de esas mismas moléculas. He aquí un ejemplo más, pero menos morboso, que puede verificarse: en grupos de 25 estudiantes, la probabilidad de que al menos dos estudiantes cumplan años el mismo día es mayor del 50%.



Puedes apostarlo

En una lotería estatal típica, la “casa” tiene una ventaja del 65 o 70%, ya que sólo entre el 30 y 35% del dinero apostado se devuelve en forma de premios. La ventaja de la casa en los hipódromos suele ser de alrededor del 15%. En los casinos, la ventaja de la casa es del 5.26% para la ruleta, 5.9% para el veintiuno, 1.4% para los dados y del 3 al 22% para las máquinas tragamonedas. Algunos jugadores profesionales pueden ganar sistemáticamente en el veintiuno, usando complicadas técnicas de conteo de cartas. Ellos saben cuando un mazo tiene una proporción elevada de cartas altas; es entonces cuando hacen apuestas cuantiosas. Muchos casinos reaccionan expulsando a los contadores de cartas o revolviendo los mazos con más frecuencia.

Regla 2: Método clásico de la probabilidad (requiere resultados igualmente probables)

Suponga que un procedimiento dado tiene n sucesos simples distintos y que *cada uno de esos sucesos simples tiene la misma posibilidad de ocurrir*. Si el suceso A puede ocurrir en s de estas n formas, entonces

$$P(A) = \frac{\text{número de formas en que puede ocurrir } A}{\text{número de sucesos simples diferentes}} = \frac{s}{n}$$

Regla 3: Probabilidades subjetivas

$P(A)$, la probabilidad del suceso A , se *estima* con base en el conocimiento de las circunstancias relevantes.

Es muy importante señalar que el método clásico (*regla 2*) requiere *resultados igualmente probables*. Si los resultados no son igualmente probables, debemos usar el estimado de frecuencias relativas o confiar en nuestro conocimiento de las circunstancias para hacer una *conjetura entrenada*. La figura 4-1 ilustra los tres métodos.

Al calcular probabilidades con el método de frecuencias relativas (*regla 1*), obtenemos una *aproximación* en vez de un valor exacto. Conforme el número total de observaciones se incrementa, los estimados correspondientes tienden a acercarse

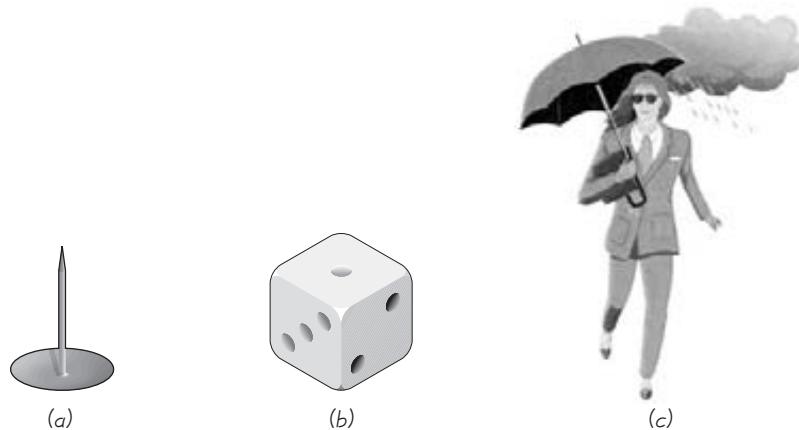


Figura 4-1 Tres métodos para calcular la probabilidad

(a) Método de las frecuencias relativas (*regla 1*): Cuando se trata de determinar $P(\text{tachuela cae con la punta hacia arriba})$, debemos repetir muchas veces el procedimiento de lanzar la tachuela y después calcular el cociente del número de veces que la tachuela cae con la punta hacia arriba entre el número de lanzamientos.

(b) Método clásico (*regla 2*): Cuando se trata de determinar $P(2)$ con un dado balanceado, cada una de las seis caras tiene la misma probabilidad de ocurrir.

$$P(2) = \frac{\text{número de formas en que 2 puede ocurrir}}{\text{número de sucesos simples}} = \frac{1}{6}$$

(c) Probabilidad subjetiva (*regla 3*): Cuando se trata de estimar la probabilidad de que mañana llueva, los meteorólogos usan su conocimiento experto de las condiciones del tiempo para desarrollar un estimado de la probabilidad.

a la probabilidad real. Esta propiedad se enuncia en forma de teorema, que se conoce comúnmente como la *ley de los números grandes*.

Ley de los números grandes

Conforme un procedimiento se repite una y otra vez, la probabilidad de frecuencias relativas (a partir de la regla 1) de un suceso, tiende a aproximarse a la probabilidad real.

La ley de los números grandes indica que los estimados por frecuencias relativas de la regla 1 tienden a mejorar si se hacen más observaciones. Esta ley refleja una simple noción fundamentada en el sentido común: una estimación de probabilidad basada sólo en unos cuantos ensayos puede desviarse en cantidades sustanciales, pero con un número muy grande de ensayos, la estimación tiende a ser mucho más exacta. Por ejemplo, suponga que deseamos realizar una encuesta para estimar la probabilidad de que alguien pueda darse una palmada en la cabeza y frotarse el estómago al mismo tiempo. Si sólo encuestamos a cinco personas, el estimado podría tener un nivel elevado de error; pero si encuestamos a miles de individuos elegidos aleatoriamente, hay muchas más probabilidades de que se acerque al verdadero valor de la población.

Probabilidad y resultados que no son igualmente probables Un error común consiste en suponer que los resultados son igualmente probables sólo porque no sabemos nada acerca de la probabilidad de cada resultado. Por ejemplo, la probabilidad de que un republicano gane la siguiente elección presidencial de Estados Unidos *no* es $1/2$. Con frecuencia se plantea el valor de $1/2$ como la probabilidad, con base en el razonamiento incorrecto de que un republicano ganará o no ganará, y que estos dos resultados son igualmente probables, pero no es así. La probabilidad de que un republicano gane la siguiente elección presidencial depende de factores tales como la habilidad de los candidatos, las cantidades de dinero recaudadas, la situación de la economía, la situación de la seguridad nacional e incluso del clima que impere el día de las elecciones. De manera similar, es incorrecto concluir que existe una probabilidad de $1/2$ de aprobar su próximo examen de estadística. Con la preparación adecuada, la probabilidad debe ser mucho mayor que $1/2$. Cuando no se sabe nada acerca de la probabilidad de distintos resultados posibles, no debe suponerse que son igualmente probables.

EJEMPLO Anotación de un tiro libre Calcule la probabilidad que tiene el jugador de básquetbol de la NBA, Reggie Miller, de anotar un tiro libre después de recibir una falta. En cierto momento de su carrera, anotó 5915 tiros libres entre 6679 intentos (de acuerdo con datos de la NBA).

SOLUCIÓN El espacio muestral consiste en dos sucesos simples: Miller anota el tiro libre o no lo hace. Puesto que el espacio muestral consiste en sucesos que no tienen la misma probabilidad, no podemos utilizar el método clásico (regla 2), sino el método de las frecuencias relativas (regla 1) con sus resultados anteriores, y obtenemos lo siguiente

$$P(\text{Miller anota un tiro libre}) = \frac{5915}{6679} = 0.886$$



¿Qué tan probable?

Cómo interpretamos términos como probable, improbable o extremadamente probable? La FFA interpreta estos términos de la siguiente manera. *Probable*: Una probabilidad del orden de 0.00001 o mayor para cada hora de vuelo. Se espera que ocurran sucesos de este tipo varias veces durante la vida operacional de cada nave aérea.

Improbable: Una probabilidad del orden de 0.00001 o menor. No se espera que ocurran sucesos de esta clase durante toda la vida operacional de un solo avión de un tipo en particular, aunque pueden ocurrir durante toda la vida operacional de todos los aviones de un tipo en particular. *Extremadamente improbable*: Una probabilidad en el orden de 0.00000001 o menor. Eventos como éste son tan improbables que no es necesario considerar su ocurrencia.



Probabilidades subjetivas en el hipódromo

Algunos investigadores han estudiado la capacidad para estimar probabilidades subjetivas realistas de los apostadores en los hipódromos. (Véase “Racetrack Betting: Do Bettors Understand the Odds?”, de Brown, D’Amato y Gertner, revista *Chance*, vol. 7, núm. 3). Después de analizar los resultados de 4400 carreras, los autores concluyeron que aunque los apostadores sobreestiman un poco las probabilidades de ganar de los que no son favoritos y subestiman ligeramente las probabilidades de ganar de los favoritos, su desempeño general es muy bueno. Las probabilidades subjetivas se calcularon a partir de las ganancias de los apostadores, con base en las cantidades apostadas, en tanto que las probabilidades reales se calcularon a partir de los resultados reales de las carreras.

EJEMPLO Genotipo Como parte de un estudio sobre los genotipos AA, Aa, aA y aa, anote cada genotipo individual en una ficha, luego mezcle las cuatro fichas y elija una al azar. ¿Qué probabilidad tiene de elegir el genotipo Aa?

SOLUCIÓN Puesto que el espacio muestral (AA, Aa, aA, aa) en este caso incluye resultados igualmente posibles, empleamos el método clásico (regla 2) para obtener

$$P(Aa) = \frac{1}{4}$$

EJEMPLO Choque de meteoritos ¿Cuál es la probabilidad de que su automóvil sea impactado por un meteorito este año?

SOLUCIÓN En ausencia de datos históricos de meteoritos que chocan contra automóviles, no podemos usar el método de frecuencias relativas de la regla 1. Hay dos posibles resultados (chocar o no chocar), pero no son igualmente probables, por lo que no podemos usar el método clásico de la regla 2. Esto nos deja con la regla 3, por medio de la cual hacemos un estimado subjetivo. En este caso, todos sabemos que la probabilidad en cuestión es muy, muy pequeña. Estimemos que sea, digamos, 0.00000000001 (equivalente a una en un billón). Este estimado subjetivo, basado en nuestro conocimiento general, puede encontrarse en el campo general de la probabilidad real.

En problemas de probabilidad básica del tipo que estamos considerando ahora, es muy importante examinar cuidadosamente la información disponible para identificar de manera correcta el número total de resultados posibles. En algunos casos se cuenta con el total de resultados posibles, pero en otros debe calcularse, como en el siguiente ejemplo en el cual debemos calcular el número total de resultados posibles.

EJEMPLO Clonación de seres humanos Se seleccionan adultos al azar para una encuesta Gallup y se les pregunta si consideran que la clonación de seres humanos debe permitirse o no. Entre los adultos elegidos al azar y encuestados, 91 dijeron que se debe permitir la clonación de seres humanos, 901 que no se debe permitir y 20 se abstuvieron de opinar. Con base en estos resultados, estime la probabilidad de que una persona elegida al azar considere que se debe permitir la clonación de seres humanos.

SOLUCIÓN Sugerencia: En vez de tratar de formular una respuesta directamente a partir de la declaración escrita, resuma la información dada en un formato que le permita comprenderla mejor. Por ejemplo, utilice el siguiente formato:

91	se debe permitir la clonación de seres humanos
901	no se debe permitir la clonación de seres humanos
20	sin opinión
1012	respuestas totales

Ahora podemos utilizar el método de las frecuencias relativas (regla 1) de la siguiente manera:

$P(\text{se debe permitir la clonación de seres humanos})$

$$= \frac{\text{número de individuos que consideran que se debe permitir la clonación de seres humanos}}{\text{número total de personas encuestadas}} = \frac{91}{1012}$$

$$= 0.0899$$

Estimamos que, cuando se elige al azar a un adulto, existe una probabilidad de 0.0899 de que considere que debe permitirse la clonación de seres humanos. Como ocurre con todas las encuestas, la exactitud de este resultado depende de la calidad del método de muestreo y del procedimiento de encuesta. Como la encuesta fue realizada por la organización Gallup, es probable que los resultados tengan una exactitud razonable. En el capítulo 7 se incluirán procedimientos más avanzados para analizar este tipo de resultados de encuesta.

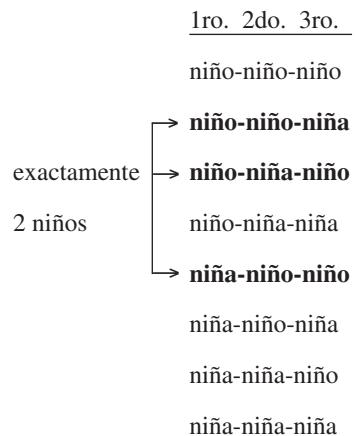
EJEMPLO Género de los hijos Determine la probabilidad de que exactamente dos de los tres hijos de una pareja sean varones. Suponga que es igualmente probable dar a luz un niño que una niña, y que el género de cualquier hijo no influye en el género de otro.

SOLUCIÓN El mayor obstáculo aquí es identificar correctamente el espacio muestral. Esto implica más que trabajar sólo con los números 2 y 3 que se dieron en el planteamiento del problema. El espacio muestral consiste en 8 diferentes formas en que 3 hijos pueden presentarse, y las listamos al margen. Como los 8 resultados son igualmente probables, usamos la regla 2. De los ocho posibles resultados, tres corresponden a exactamente dos varones, así que

$$P(2 \text{ varones en } 3 \text{ nacimientos}) = \frac{3}{8} = 0.375$$

INTERPRETACIÓN Existe una probabilidad de 0.375 de que si una pareja tiene tres hijos, exactamente dos de ellos sean niños.

Los enunciados de las tres reglas para calcular probabilidades y los ejemplos anteriores parecen sugerir que siempre debemos usar la regla 2 cuando un procedimiento tiene resultados igualmente probables. En realidad, muchos procedimientos son tan complicados que el uso del método clásico (regla 2) no es práctico. En el juego de solitario, por ejemplo, todos los resultados (del reparto) son igualmente probables, pero es extremadamente frustrante tratar de usar la regla 2 para calcular la probabilidad de ganar. En estos casos podemos obtener buenas estimaciones con mayor facilidad utilizando el método de frecuencias relativas (regla 1). Es muy común que las simulaciones sean útiles cuando se usa este método. (Una **simulación** de un procedimiento es un proceso que se comporta de la misma forma que el procedimiento mismo; por lo tanto, produce resultados similares. Véase la sección 4-6 y el Proyecto tecnológico casi al final de este capítulo). Por ejemplo, al estimar la probabilidad de ganar en el solitario, es mucho más fácil usar la regla 1 y repetir el juego muchas veces (o correr una simulación por computadora) que realizar los cálculos extremadamente complejos que se requieren con la regla 2.



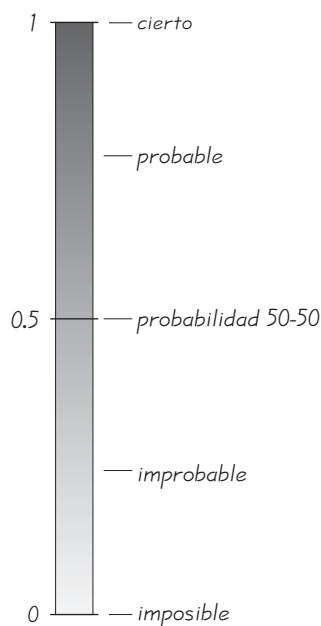


Figura 4-2 Valores posibles para probabilidades

EJEMPLO Día de Acción de Gracias Si se elige un año de manera aleatoria, calcule la probabilidad de que el Día de Acción de Gracias sea un *a)* miércoles o *b)* jueves.

SOLUCIÓN

- El Día de Acción de Gracias se celebra siempre el cuarto jueves de noviembre. Por lo tanto, es imposible que sea un miércoles. Cuando un suceso es imposible, decimos que su probabilidad es 0.
- Es cierto que el Día de Acción de Gracias será un jueves. Cuando es seguro que un suceso ocurra, decimos que su probabilidad es 1.

Ya que cualquier suceso imaginable es imposible, cierto, o algún punto intermedio, se deduce que la probabilidad matemática de cualquier suceso es 0, 1 o un número entre 0 y 1 (véase la figura 4-2).

- **La probabilidad de un suceso imposible es 0.**
- **La probabilidad de un suceso que ocurrirá con certeza es 1.**
- **Para cualquier suceso A, la probabilidad de A se encuentra entre 0 y 1, inclusive. Es decir, $0 \leq P(A) \leq 1$.**

En la figura 4-2, se muestra la escala de 0 hasta 1 y se incluyen las expresiones más comunes y familiares de probabilidad.

Sucesos complementarios

Algunas veces necesitamos calcular la probabilidad de que un suceso *A no* ocurra.

Definición

El **complemento** de un suceso *A*, denotado por \bar{A} , consiste en todos los resultados en los cuales el suceso *A no* ocurre.

EJEMPLO Género al nacer En realidad nacen más niños que niñas. En un grupo típico, hay 205 bebés recién nacidos y 105 de ellos son niños. Si un bebé del grupo es seleccionado al azar, ¿cuál es la probabilidad de *no* sea un niño?

SOLUCIÓN Ya que 105 de los 205 bebés son niños, se deduce que 100 de ellos son niñas, entonces

$$P(\text{no seleccionar a un niño}) = P(\bar{\text{nino}}) = P(\text{niña}) = \frac{100}{205} = 0.488$$

Aun cuando es difícil desarrollar una regla universal para el redondeo de probabilidades, el siguiente lineamiento se aplicará a la mayoría de los problemas en este libro.

Redondeo de probabilidades

Cuando se expresa el valor de una probabilidad, hay que dar la fracción o el número decimal *exacto*, o bien, redondear los resultados decimales finales a tres cifras significativas. (*Sugerencia:* Cuando una probabilidad no sea una fracción simple como 2/3 o 5/9, exprésela como un decimal para que el número resulte más claro).

Todos los dígitos en un número son significativos, excepto los ceros que se incluyen para la colocación apropiada del punto decimal.

EJEMPLOS

- La probabilidad de 0.021491 tiene cinco dígitos significativos (21491) y puede redondearse a 0.0215, con tres dígitos relevantes.
- La probabilidad de $1/3$ puede permanecer como fracción o redondearse a 0.333 (*no* a 0.3).
- La probabilidad de obtener cara en el lanzamiento de una moneda puede expresarse como $1/2$ o 0.5; como 0.5 es exacto, no hay necesidad de expresarlo como 0.500.
- La fracción $432/7842$ es exacta, pero su valor no es obvio. Exprésela como el decimal 0.0551.

La expresión matemática de la probabilidad como un número entre 0 y 1 es un concepto importante en esta sección. Esta forma de expresión es fundamental y común en los procedimientos estadísticos, y la usaremos en el resto de este libro. Por ejemplo, un resultado de computadora común puede incluir una expresión “valor P ” como “nivel de significancia menor que 0.001”. Más tarde analizaremos el significado de los valores P , los cuales son esencialmente probabilidades del tipo que se considera en esta sección. Por ahora, usted debe reconocer que una probabilidad de 0.001 (equivalente a $1/1000$) corresponde a un suceso tan infrecuente que puede ocurrir, en promedio, una sola vez en mil ensayos.

Posibilidades

Las expresiones de probabilidad a veces se proponen como *posibilidades*, por ejemplo, 50:1 (o “50 a 1”). Una grave desventaja de las posibilidades es que hacen que muchos cálculos sean extremadamente difíciles. Por ello, los matemáticos y los especialistas en estadística así como en otros campos científicos prefieren usar probabilidades. La ventaja de las posibilidades es que facilitan el manejo de las transacciones de dinero asociadas a los juegos de azar, por lo que tienden a usarse en casinos, loterías e hipódromos. Note que en las tres definiciones siguientes, las posibilidades reales en contra y las posibilidades reales a favor describen la probabilidad real de algún suceso, pero las posibilidades de pago describen la relación entre la apuesta y la cantidad del pago. Las posibilidades reales corresponden a los resultados reales, pero las posibilidades de pago están establecidas por los operadores de los casinos y los hipódromos. Las pistas de carreras y los casinos están en el negocio con el fin de lograr su propio beneficio. Por ello, las posibilidades de pago no serán las mismas que las posibilidades reales.

Definición

Las **posibilidades reales en contra** de que ocurra un suceso A están indicadas por el cociente $P(\bar{A})/P(A)$, casi siempre expresado en la forma $a:b$ (o “ a a b ”), donde a y b son enteros que no tienen factores comunes.

Las **posibilidades reales a favor** del suceso A son el recíproco de las posibilidades reales en contra de ese suceso. Si las posibilidades en contra de A son $a:b$, entonces las posibilidades a favor de A son $b:a$.

Las **posibilidades de pago** contra el suceso A representan la proporción de la ganancia neta (si usted gana) con respecto a la cantidad de la apuesta.

posibilidades de pago en contra del suceso A = (ganancia neta) : (cantidad apostada)



El cumpleaños más común: el 5 de octubre

Un sitio Web indicó que “una investigación reciente en bases de datos realizada por Anybirthday.com sugiere que el 5 de octubre es la fecha de nacimiento más popular en Estados Unidos”. Se señaló que una concepción en la noche de Año Nuevo podría dar por resultado probablemente un nacimiento el 5 de octubre. La fecha de nacimiento menos común se identificó como el 22 de mayo.

Al parecer, el 18 de agosto no tiene el mismo encanto que la noche de Año Nuevo.

EJEMPLO Si usted apuesta \$5 al número 13 de la ruleta, su probabilidad de ganar es $1/38$ y las posibilidades de pago están dadas por el casino como $35:1$.

- Calcule las posibilidades reales en contra del resultado de 13.
- ¿Cuánta ganancia neta podría obtener si gana apostando al 13?
- Si el casino estuviera funcionando solamente por diversión y las posibilidades de pago se modificaran para igualar las posibilidades reales en contra del 13, ¿cuánto ganaría usted si el resultado fuera 13?

SOLUCIÓN

- Con $P(13) = 1/38$ y $P(\text{no } 13) = 37/38$, tenemos

$$\text{posibilidades reales en contra del } 13 = \frac{P(\text{no } 13)}{P(13)} = \frac{37/38}{1/38} = \frac{37}{1} \text{ o } 37:1$$

- Puesto que las posibilidades de pago en contra del 13 son $35:1$, tenemos

$$35:1 = (\text{ganancia neta}):(\text{monto apostado})$$

entonces, hay una ganancia de \$35 por cada \$1 apostado. Para una apuesta de \$5, la ganancia neta es de \$175. El apostador que gane podría recoger \$175 más la apuesta original de \$5. La cantidad total obtenida debería ser \$180, con una ganancia neta de \$175.

- Si el casino estuviera funcionando por diversión y no por ganancia, las posibilidades de pago serían iguales a las posibilidades reales en contra del resultado de 13. Si las posibilidades de pago cambiaron de $35:1$ a $37:1$, usted obtendría una ganancia neta de \$37 por cada \$1 apostado. Si usted apuesta \$5, su ganancia neta sería de \$185. (El casino logra su ganancia pagando sólo \$175 en vez de los \$185 que se pagarían en un juego de ruleta justo que no favoreciera al casino).

4-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Interpretación de probabilidad.** ¿Qué significa cuando decimos que “la probabilidad de ganar el Gran Premio de la lotería de Illinois es $1/20,358,520$ ”? ¿Un triunfo como éste es *infrecuente*?
- Probabilidad de lluvia.** Al hablar acerca de la probabilidad de que llueva en Boston el 4 de julio del próximo año, el reportero de un periódico afirma que la probabilidad es de $1/2$, ya que lloverá o no lloverá. ¿Este razonamiento es correcto? ¿Por qué?
- Probabilidad y sucesos infrecuentes.** Un reportero de noticias afirma que un suceso particular es *infrecuente* porque su probabilidad es de sólo 0.001. ¿Es correcta esta afirmación? ¿Por qué?
- Probabilidad subjetiva.** Utilice el juicio subjetivo para estimar la probabilidad de que la próxima vez que suba a un elevador, éste se quede atorado entre dos pisos.

En los ejercicios 5 y 6, exprese el grado indicado de probabilidad como un valor de probabilidad entre 0 y 1.

5. Identificación de valores de probabilidad

- a. “Como estudió a conciencia y comprendió los conceptos, seguramente aprobará el examen de estadística”.
- b. “El pronóstico de mañana indica un 10% de probabilidad de lluvia”.
- c. “Usted tiene la probabilidad de una bola de nieve en el infierno de casarse con mi hija”.

6. Identificación de valores de probabilidad

- a. “Al lanzar una moneda de 25 centavos, existe la probabilidad de 50-50 de que el resultado sea una cara”.
- b. “Usted tiene una probabilidad en cinco de adivinar la respuesta correcta”.
- c. “Usted tiene una probabilidad del 1% de tener una cita con la persona que acaba entrar en la habitación”.

7. Identificación de valores de probabilidad. ¿Cuáles de los siguientes valores *no* pueden ser probabilidades?

$$0, \quad 1, \quad -1, \quad 2, \quad 0.0123, \quad 3/5, \quad 5/3, \quad \sqrt{2}$$

8. Identificación de valores de probabilidad

- a. ¿Cuál es la probabilidad de que ocurra un suceso inevitable?
- b. ¿Cuál es la probabilidad de un suceso imposible?
- c. Un espacio muestral consiste en 10 sucesos separados que son igualmente probables. ¿Cuál es la probabilidad de cada uno?
- d. En un examen de verdadero/falso, ¿cuál es la probabilidad de responder una pregunta correctamente si usted elige al azar?
- e. En un examen de opción múltiple, con cinco posibles respuestas para cada pregunta, ¿cuál es la probabilidad de responder una pregunta correctamente si usted elige al azar?

9. Género de los hijos. En esta sección dimos un ejemplo que incluía una lista de los ocho resultados que son posibles cuando una pareja tiene tres hijos. Remítase a esa lista y calcule la probabilidad de cada suceso.

- a. De entre tres hijos, hay exactamente una niña.
- b. De entre tres hijos, hay exactamente dos niñas.
- c. De entre tres hijos, todos son niñas.

10. Teléfonos celulares y cáncer cerebral. En un estudio de 420,095 usuarios de teléfono celular en Dinamarca, se encontró que 135 desarrollaron cáncer cerebral o del sistema nervioso. Estime la probabilidad de que un usuario de teléfono celular seleccionado al azar desarrolle un cáncer de este tipo. ¿El resultado es muy diferente de 0.000340, que fue el que se encontró para la población general? ¿Qué sugiere el resultado acerca de los teléfonos celulares como causantes de cáncer de este tipo, como se afirma?

11. Genética mendeliana. Cuando Mendel realizó sus famosos experimentos genéticos con guisantes, una muestra de vástagos consistió en 428 plantas de guisantes verdes y 152 de guisantes amarillos. Con base en esos resultados, estime la probabilidad de obtener un vástago de guisantes verdes. ¿El resultado es lo suficientemente cercano al valor de $3/4$ que se espera?

12. Ser alcanzado por un rayo. En un año reciente, de los 281,421,906 habitantes de Estados Unidos, 389 fueron alcanzados por un rayo. Calcule la probabilidad de que una persona seleccionada al azar en Estados Unidos sea alcanzada por un rayo este año.

13. Selección del género. En una prueba de la técnica de selección de género MicroSort, los resultados consistieron en 295 niñas y 30 niños (según datos del Genetics & IVF Institute). A partir de este resultado, ¿cuál es la probabilidad de que una pareja que utiliza el método MicroSort tenga una niña? ¿La técnica parece ser útil para incrementar la probabilidad de que un bebé sea niña?

14. Reconocimiento de marca

- a. En un estudio de reconocimiento de marcas, 831 consumidores conocían la sopa Campbell's y 18 no (según datos de Total Research Corporation). Use estos resultados para estimar la probabilidad de que un consumidor seleccionado al azar reconozca la sopa Campbell's.
- b. *Estime* la probabilidad subjetiva de que un consumidor adulto estadounidense, seleccionado al azar, conozca el nombre de la marca McDonald's, la cadena de restaurantes de comida rápida.
- c. *Estime* la probabilidad subjetiva de que un consumidor adulto estadounidense seleccionado al azar reconozca el nombre de la marca Veeco Instruments, un fabricante de productos de microelectrónica.

15. Dulces azules sencillos M&M

- a. Remítase a los 100 dulces M&M listados en el conjunto de datos 13 del apéndice B y estime la probabilidad de obtener un dulce azul al elegir al azar un dulce M&M sencillo.
- b. The Mars Company afirma que el 24% de sus dulces M&M sencillos son azules. ¿La estimación del inciso a) coincide aproximadamente con esta afirmación o al parecer existe una gran diferencia?

16. Botones para el paso de peatones. La ciudad de Nueva York tiene 750 botones para peatones que funcionan y otros 2500 que no funcionan (de acuerdo con datos de "For Exercise in New York Futility, Push Button" por Michael Luo, *New York Times*). Si se elige un botón para peatones al azar en la ciudad de Nueva York, ¿cuál es la probabilidad de que funcione? ¿Es posible que la misma probabilidad sea un buen estimado para otra ciudad como Chicago?

Uso de la probabilidad para identificar sucesos infrecuentes. En los ejercicios 17 a 24, considere que un suceso es "infrecuente" si la probabilidad de que ocurra es menor o igual que 0.05. (Esto es equivalente al mismo criterio que se utiliza comúnmente en la estadística inferencial, aunque el valor de 0.05 no es absolutamente rígido; en ocasiones se utilizan otros valores como 0.01).

17. Adivinación de fechas del nacimiento. En su primera cita, Kelly le pide a Mike que adivine su fecha de nacimiento, omitiendo el año.

- a. ¿Cuál es la probabilidad de que Mike adivine correctamente? (Ignore los años bisiestos).
- b. Sería "infrecuente" que él adivinara con acierto en el primer intento?
- c. Si usted fuera Kelly, y Mike adivinara correctamente en su primer intento, ¿creería que él tuvo un golpe de suerte? ¿O tendría la seguridad de que él ya sabía la fecha en que usted nació?
- d. Si Kelly le pide a Mike que adivine su edad, y la respuesta de Mike es más alta por 15 años, ¿cuál es la probabilidad de que Mike y Kelly tengan una segunda cita?

18. Exactitud del IRS. La U.S. General Accounting Office puso a prueba la exactitud de las respuestas del Internal Revenue Service (IRS) con respecto a preguntas sobre los contribuyentes. En 1733 ensayos, el IRS acertó 1107 veces y se equivocó 626.

- a. Estime la probabilidad de que una pregunta de un contribuyente seleccionado al azar tenga una respuesta incorrecta.
- b. ¿Es infrecuente que el IRS dé una respuesta incorrecta a una pregunta de un contribuyente? ¿Debería ser infrecuente?

19. Probabilidad de un accidente automovilístico. Entre 400 conductores seleccionados al azar, en el rango de edades de 20 a 24 años, 136 sufrieron un accidente automovilístico durante el año anterior (según datos del National Safety Council). Si se selecciona al azar a un conductor en ese rango de edad, ¿cuál es la probabilidad aproximada de que él o ella sufra un accidente automovilístico durante el siguiente año? ¿Es infrecuente que un conductor en ese rango de edad sufra un accidente automovilístico en el transcurso de un año? Será el valor resultante lo bastante alto para preocupar a los individuos entre 20 y 24 años de edad?

- 20. Efecto negativo del Lipitor.** En una prueba clínica de Lipitor (atorvastatin), un fármaco que se utiliza con frecuencia para disminuir los niveles de colesterol, un grupo de pacientes recibió tratamiento de tabletas de 10 mg del medicamento. En ese grupo, 19 pacientes experimentaron síntomas de resfriado y 844 pacientes no (según datos de Pfizer, Inc.).
- Estime la probabilidad de que un paciente que tome el fármaco experimente síntomas de resfriado.
 - ¿Es “infrecuente” que un paciente que toma el fármaco experimente síntomas de resfriado?
- 21. Efecto negativo del Viagra.** Cuando el fármaco Viagra se probó clínicamente, 117 pacientes reportaron dolor de cabeza y 617 no (de acuerdo con datos de Pfizer, Inc.). Utilice esta muestra para estimar la probabilidad de que un usuario de Viagra sufra dolor de cabeza. ¿Es infrecuente que un usuario de Viagra sufra dolor de cabeza? ¿Es la probabilidad lo bastante alta como para preocupar a los usuarios de Viagra?
- 22. Interpretación de la efectividad de un tratamiento.** Se diseña un experimento doble ciego para probar la eficacia del fármaco Statisticzene como tratamiento para la ceguera a los números. Los sujetos muestran una mejoría cuando son tratados con Statisticzene. Los investigadores calculan que existe una probabilidad de 0.04 de que el grupo de tratamiento muestre una mejoría si el fármaco no tiene efecto alguno. ¿Es infrecuente que una persona tratada con un fármaco ineficaz muestre mejoría? ¿Qué debemos concluir acerca de la eficacia del Statisticzene?
- 23. Probabilidad de un resultado incorrecto.** La tabla 4-1 indica que, de los 178 sujetos que *no* consumieron marihuana, el resultado de la prueba del consumo de esta droga fue incorrecto en 24 ocasiones.
- Con base en los resultados disponibles, calcule la probabilidad de obtener un resultado de prueba incorrecto para una persona que no consume marihuana.
 - ¿Es “infrecuente” que el resultado de la prueba sea incorrecto en las personas que no utilizan marihuana?
- 24. Probabilidad de resultado incorrecto.** La tabla 4-1 indica que, de los 122 sujetos que consumen marihuana, el resultado de la prueba del consumo de esta droga fue incorrecto en 3 ocasiones.
- Con base en los resultados disponibles, calcule la probabilidad de obtener un resultado de prueba incorrecto para una persona que consume marihuana.
 - ¿Es “infrecuente” que el resultado de la prueba sea incorrecto en las personas que sí consumen marihuana?
- Construcción de espacios muestrales.* En los ejercicios 25 a 28, construya el espacio muestral que se le indica y responda las preguntas.
- 25. Género de los hijos: Construcción de espacio muestral.** Esta sección incluyó una tabla que resume los resultados de género para una pareja que planea tener tres hijos.
- Construya una tabla similar para una pareja que planea tener *dos* hijos.
 - Suponiendo que los resultados listados en el inciso a) sean igualmente probables, calcule la probabilidad de tener dos niñas.
 - Calcule la probabilidad de tener exactamente un hijo de cada género.
- 26. Genética: Construcción del espacio muestral.** Ambos progenitores tienen los genes de color de ojos café/azul, y cada uno contribuye con un gen para su hijo. Suponga que si el hijo tiene al menos un gen café, ese color dominará y los ojos serán cafés. (La determinación real del color de los ojos es un tanto más complicada).
- Haga una lista de los posibles resultados diferentes. Suponga que estos resultados son igualmente probables.
 - ¿Cuál es la probabilidad de que un hijo de estos padres tenga el par de genes azul/azul?
 - ¿Cuál es la probabilidad de que el hijo tenga ojos cafés?
- 27. Genética: Construcción del espacio muestral.** Repita el ejercicio 26 suponiendo que uno de los padres tiene genes de color de ojos café/café, en tanto que el otro tiene genes de color de ojos café/azul.

- 28. Genética: construcción del espacio muestral.** Repita el ejercicio 26, suponiendo que uno de los padres tiene genes de color de ojos café/café, en tanto que el otro tiene genes de color de ojos azul/azul.

Posibilidades. En los ejercicios 29 a 32, responda las preguntas que implican posibilidades.

- 29. Posibilidades en el solitario.** Puesto que los cálculos en el solitario son tan complejos, se jugó el juego 500 veces para estimar la probabilidad de ganar. (Los resultados son del juego de solitario de Microsoft y se utilizaron las reglas de Vegas de “tomar 3”, con una apuesta de \$52 y una devolución de \$5 por carta). De los 500 ensayos, se ganaron 77 juegos. Con base en estos resultados, calcule las posibilidades en contra de ganar.
- 30. Posibilidades en el Derby de Kentucky.** La probabilidad de que el caballo Outta Here ganara el 129º Derby de Kentucky era 1/50. ¿Cuáles eran las posibilidades reales en contra de que Outta Here ganara esa carrera?
- 31. Cálculo de posibilidades en la ruleta.** Una rueda de ruleta tiene 38 ranuras, una ranura es 0, otra es 00 y cada una de las demás están numeradas del 1 al 36. Usted está apostando a un número impar.
- ¿Cuál es su probabilidad de ganar?
 - ¿Cuáles son las posibilidades reales en contra?
 - Cuando se apuesta a número impar, las posibilidades de pago son 1:1. ¿Qué ganancia podría obtener al apostar \$18 si gana?
 - ¿Qué ganancia podría obtener al apostar \$18, si de alguna manera pudiera convencer al casino de modificar sus posibilidades de pago para que fueran las mismas que las posibilidades reales en contra? (*Recomendación:* No trate realmente de convencer a ningún casino de esto; carecen totalmente de sentido del humor cuando se trata de asuntos de este tipo).
- 32. Posibilidades en el Derby de Kentucky.** Cuando el caballo Funny Cide ganó el 129º Derby de Kentucky, una apuesta de \$2 a que Funny Cide ganaría dio por resultado un reintegro de \$27.60.
- ¿Qué ganancia neta hubo al ganar con una apuesta de \$2 a Funny Cide?
 - ¿Cuáles fueron las posibilidades de pago en contra de que Funny Cide ganara?
 - Con base en el paseo preliminar a la carrera, los apostadores colectivamente creyeron que Funny Cide tenía una probabilidad de ganar de 2/33. Suponiendo que 2/33 era la probabilidad real de la victoria de Funny Cide, ¿cuáles fueron las posibilidades reales en contra?
 - Si las posibilidades de pago fueran iguales a las posibilidades reales calculadas en el inciso c), ¿cuánto valdría un boleto de \$2 después de que Funny Cide ganara?

4-2 MÁS ALLÁ DE LO BÁSICO

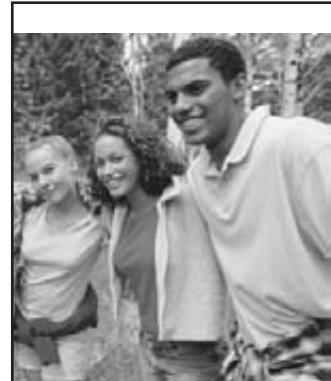
- 33. Cálculo de probabilidad a partir de posibilidades.** Si las posibilidades reales en contra de un suceso A son $a:b$, entonces $P(A) = b/(a + b)$. Calcule la probabilidad de que el caballo Buddy Gil ganara el 129º Derby de Kentucky, dado que las posibilidades reales en contra eran 9:1.
- 34. Riesgo relativo y razón de probabilidad.** En un ensayo clínico con 734 sujetos tratados con Viagra, 117 reportaron dolor de cabeza. En un grupo de control de 725 sujetos no tratados con Viagra, 29 reportaron sufrir dolor de cabeza. Si la proporción de dolores de cabeza en el grupo de tratamiento se denota como p_t y la proporción de dolores de cabeza en el grupo de control como p_c , el **riesgo relativo** es p_t/p_c . El riesgo relativo es una medida de la fuerza del efecto del tratamiento con Viagra. Otra medida como ésta es la **razón de probabilidad**, que es el cociente de las posibilidades a favor de un

dolor de cabeza en el grupo de tratamiento entre las posibilidades a favor de un dolor de cabeza en el grupo de control, el cual se calcula evaluando lo siguiente:

$$\frac{p_t/(1 - p_t)}{p_c/(1 - p_c)}$$

El riesgo relativo y la razón de probabilidad se utilizan comúnmente en estudios médicos y epidemiológicos. Calcule el riesgo relativo y la razón de probabilidad de los datos del dolor de cabeza.

35. **Moscas en una naranja.** Si dos moscas se posan sobre una naranja, calcule la probabilidad de que ambas se localicen en puntos pertenecientes al mismo hemisferio.
36. **Puntos en un palo.** Se seleccionan al azar dos puntos a lo largo de un palo recto. Después se rompe el palo en esos dos puntos. Calcule la probabilidad de que los tres pedazos que quedan se puedan acomodar para formar un triángulo. (Quizá éste sea el ejercicio más difícil del libro).



Los niños y las niñas no son igualmente probables

4-3 Regla de la suma

Concepto clave El objetivo principal de esta sección es presentar la *regla de la suma* como un método para calcular probabilidades que pueden expresarse de la forma $P(A \text{ o } B)$, es decir, la probabilidad de que ocurra el suceso A o de que ocurra el suceso B (o de que ambos ocurran), como único resultado de un procedimiento. Para calcular la probabilidad de que ocurra el suceso A o el suceso B , primero debemos obtener el número total de maneras en que puede ocurrir A y de maneras en que puede ocurrir B , pero calculamos ese total sin contar cada resultado más de una vez.

La palabra clave en esta sección es “o”. A lo largo de este texto usaremos el *o inclusive*, que significa: uno o el otro o ambos. (Con excepción del ejercicio 26, no consideramos el *o exclusivo*, que significa uno o el otro, pero no ambos).

En la sección anterior presentamos aspectos fundamentales de la probabilidad y estudiámos sucesos calificados como *simples*. En esta sección y en la siguiente nos ocuparemos de *sucesos compuestos*.



Definición

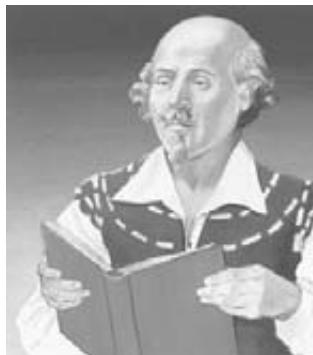
Un **suceso compuesto** es cualquier suceso que combine dos o más sucesos simples.

Notación de la regla de la suma

$$P(A \text{ o } B) = P(\text{en un solo ensayo, ocurre el suceso } A \text{ u ocurre el suceso } B \text{ o ambos ocurren})$$

Comprensión de la notación En esta sección, $P(A \text{ y } B)$ denota la probabilidad de que tanto A como B ocurran en el mismo ensayo, pero en la siguiente sección utilizamos $P(A \text{ y } B)$ para denotar la probabilidad de que el evento A ocurra en un ensayo, seguido por el suceso B en otro ensayo. Por lo tanto, el verdadero significado de $P(A \text{ y } B)$ sólo se determina sabiendo si nos referimos a un ensayo que puede tener resultados de A y B , o dos ensayos en donde el suceso A ocurra en el primer ensayo y el suceso B ocurrira en el segundo. Así pues, el significado de $P(A \text{ y } B)$ depende del contexto.

En muchos cálculos de probabilidad, se obtienen buenos resultados al suponer que los niños y las niñas tienen las mismas probabilidades de nacer. En realidad, es más probable que nazca un niño (con una probabilidad de 0.512) que una niña (con una probabilidad de 0.488). Estos resultados se basan en datos recientes del National Center for Health Statistics, según los cuales de los 4,058,814 nacimientos en un año, 2,076,969 fueron niños y 1,981,845 fueron niñas. Los investigadores revisan estas probabilidades para descubrir cambios que podrían sugerir factores como modificaciones en el ambiente y la exposición a sustancias químicas.



El vocabulario de Shakespeare

Según Bradley Efron y Ronald Thisted, los escritos de Shakespeare incluyen 31,534 palabras diferentes. Ellos usaron la teoría de la probabilidad para concluir que Shakespeare probablemente conocía al menos otras 35,000 palabras que no usó en sus escritos. Estimar el tamaño de una población es un problema importante que se encuentra con frecuencia en estudios ecológicos, pero el resultado que aquí se presenta es otra aplicación interesante. (Véase “Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?” en *Biometrika*, vol. 63, núm. 3).

Tabla 4-1 Resultados de exámenes sobre el consumo de marihuana

	¿Los sujetos realmente consumen marihuana?	
	Sí	No
Resultado de prueba positivo (La prueba indica que la marihuana está <i>presente</i>).	119 (verdadero positivo)	24 (falso positivo)
Resultado de prueba negativo (La prueba indica que la marihuana está <i>ausente</i>).	3 (falso negativo)	154 (verdadero negativo)

Remítase a la tabla 4-1 que se reproduce aquí para su comodidad. En la muestra de 300 sujetos representados en la tabla, ¿cuántos de ellos resultaron positivos o consumían marihuana? (Recuerde, “resultaron positivos o consumían marihuana” realmente significa “resultaron positivos, consumían marihuana o ambos”). Un examen de la tabla debe indicarle que un total de 146 sujetos resultaron positivos o consumían marihuana. (*Nota importante:* Es incorrecto sumar los 143 sujetos que resultaron positivos con los 122 sujetos que consumían marihuana, ya que este total de 265 contaría dos veces a 119 de los sujetos, que sólo deben contarse una vez). Vea el papel que desempeña el total correcto de 146 en el siguiente ejemplo.

EJEMPLO Prueba de drogas Remítase a la tabla 4-1 que se reproduce aquí para su comodidad. Suponiendo que se elige al azar a una de las 300 personas que fueron examinadas, calcule la probabilidad de seleccionar a un sujeto que haya resultado positivo o que consumía marihuana.

SOLUCIÓN En la tabla 4-1 observamos que hay 146 sujetos que tuvieron un resultado de prueba positivo o consumían marihuana. El total de 146 se obtuvo al sumar a los sujetos que resultaron positivos con los sujetos que consumían marihuana, teniendo cuidado de contar a cada uno sólo una vez. Al dividir el total de 146 entre el total general de 300, obtenemos el siguiente resultado: $P(\text{resultado positivo de la prueba o consumo de marihuana}) = 146/300$ o 0.487.

En el ejemplo anterior existen varias estrategias que usted podría utilizar para contar a los sujetos que resultaron positivos o consumían marihuana. Cualquiera de los siguientes podría funcionar:

- Coloree las celdas que representan a los sujetos que resultaron positivos o consumían marihuana, luego sume los números de las celdas coloreadas, teniendo cuidado de sumar cada número sólo una vez. Este método da por resultado

$$119 + 24 + 3 = 146$$

- Sume los 143 sujetos que resultaron positivos con los 122 sujetos que consumían marihuana, pero el total de 265 incluye un doble conteo de 119 sujetos, de manera que para compensar esto se resta el traslape que consiste

en los 119 sujetos que resultaron positivos y consumían marihuana. Este método produce el siguiente resultado

$$143 + 122 - 119 = 146$$

- Comience con el total de 143 sujetos que resultaron positivos, luego sume los sujetos que consumían marihuana y que aún no habían sido incluidos en el total, para obtener un resultado de

$$143 + 3 = 146$$

Estudie cuidadosamente el ejemplo anterior para comprender esta característica fundamental del cálculo de la probabilidad de un suceso A o de un suceso B : el uso la palabra “o” sugiere una suma, y la suma se debe realizar sin un conteo doble.

Este ejemplo sugiere una regla general por medio de la cual sumamos el número de resultados que corresponden a cada uno de los sucesos en cuestión:

Para calcular la probabilidad de que un suceso A ocurra o un suceso B ocurra, calcule el número total de formas en que A puede ocurrir y el número de formas en que B puede ocurrir, pero calcule ese total de tal manera que ningún resultado se cuente más de una vez.

Un modo de formalizar la regla consiste en combinar el número de formas en que un suceso A puede ocurrir con el número de formas en que un suceso B puede ocurrir y, si hay algún traslape, se debe compensar restando el número de resultados que se contaron dos veces, como se hace en la siguiente regla.

Regla formal de la suma

$$P(A \text{ o } B) = P(A) + P(B) - P(A \text{ y } B)$$

donde $P(A \text{ y } B)$ denota la probabilidad de que A y B ocurran al mismo tiempo, como resultado en un ensayo de un procedimiento.

La regla formal de la suma se presenta como una fórmula, pero no se recomienda el uso irreflexivo de fórmulas. En general, es mejor *comprender* el espíritu de la regla y utilizar esa comprensión de la siguiente forma.

Regla intuitiva de la suma

Para obtener $P(A \text{ o } B)$, calcule la suma del número de formas en que puede ocurrir el suceso A y el número de formas en que puede ocurrir el suceso B , *sumando de tal manera que cada resultado se cuente sólo una vez*. $P(A \text{ o } B)$ es igual a esa suma, dividida entre el número total de resultados en el espacio muestral.

Puesto que el traslape de sucesos es un aspecto esencial en la regla de la suma, existe un término especial para describirlo:

Definición

Los sucesos A y B son **disjuntos** (o **mutuamente excluyentes**) cuando ambos no pueden ocurrir al mismo tiempo. (Es decir, los sucesos disjuntos no se traslanan).

EJEMPLO Examen de drogas De nuevo, remítase a la tabla 4-1.

- Considere el procedimiento de elegir al azar a uno de los 300 sujetos incluidos en la tabla 4-1. Determine si los siguientes sucesos son disjuntos: A: elegir a un sujeto con un resultado de prueba negativo; B: elegir a un sujeto que no consumía marihuana.
- Suponiendo que se elige al azar a una de las 300 personas que fueron sometidas a la prueba, calcule la probabilidad de elegir a un sujeto con un resultado de prueba negativo o que no consumía marihuana.

SOLUCIÓN

- En la tabla 4-1 observamos que hay 157 sujetos con resultados de prueba negativos y 178 sujetos que no consumían marihuana. El suceso de elegir a un sujeto con un resultado de prueba negativo y elegir a un sujeto que no consumía marihuana pueden ocurrir al mismo tiempo (ya que existen 154 sujetos con resultados de prueba negativos y que no consumían marihuana). Como esos eventos se traslanan, pueden ocurrir al mismo tiempo y decimos que los sucesos *no son* disjuntos.
- En la tabla 4-1 debemos calcular el número total de sujetos que tuvieron resultados de prueba negativos y que no consumían marihuana, pero debemos hacerlo sin contar dos veces a cada uno. Obtenemos un total de 181.

Puesto que 181 sujetos tuvieron resultados de prueba negativos o no consumían marihuana, y como hay un total de 300 sujetos, obtenemos

$$P(\text{resultados de prueba negativos o no consumían marihuana}) = \frac{181}{300} = 0.603$$

En la figura 4-3 se muestra un diagrama de Venn que nos ofrece una ilustración de la regla formal de la suma. En esta figura podemos ver que la probabilidad de A o B es igual a la probabilidad de A (círculo izquierdo) más la probabilidad de B (círculo derecho) menos la probabilidad de A y B (región media con forma de balón de fútbol americano). Esta figura nos muestra que la suma de las áreas de los dos círculos haría que se contara dos veces la región media. Éste es el concepto básico que subyace en la regla de la suma. Debido a la relación entre la regla de la suma y el diagrama de Venn que se muestra en la figura 4-3, es común el uso de la notación $P(A \cup B)$ en vez de $P(A$ o B). De manera similar, se usa con frecuencia la notación $P(A \cap B)$ en vez de $P(A$ y B), de manera que la regla formal de la suma se expresa como

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

La regla de la suma se simplifica cuando A y B son disjuntos (no pueden ocurrir simultáneamente), de manera que $P(A$ y B) se convierte en cero. La figura 4-4 indica que si A y B son disjuntos, tenemos $P(A$ o B) = $P(A) + P(B)$.

Podemos resumir los puntos clave de esta sección de la siguiente manera:

- Para calcular $P(A$ o B), primero debemos asociar el uso de la palabra “o” con la suma.
- Considere si los sucesos A y B son disjuntos; es decir, ¿pueden ocurrir al mismo tiempo? Si no son disjuntos (es decir, si pueden ocurrir al mismo tiempo),

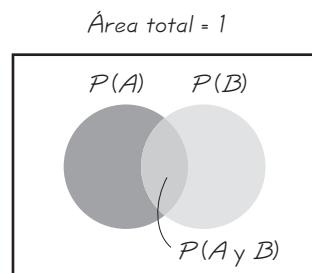


Figura 4-3 Diagrama de Venn de sucesos que no son disjuntos

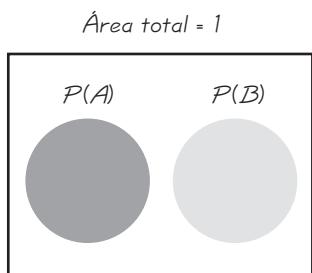


Figura 4-4 Diagrama de Venn de sucesos disjuntos

asegúrese de evitar (o al menos compensar) un conteo doble cuando se suman las probabilidades relevantes. Si usted comprende la importancia de no realizar un conteo doble cuando calcule $P(A \text{ o } B)$, no necesariamente debe calcular el valor de $P(A) + P(B) - P(A \text{ y } B)$.

Los errores que se cometan al aplicar la regla de la suma a menudo implican un conteo doble; es decir, tratar a los sucesos que no son disjuntos como si lo fueran. Una indicación de este tipo de error es una probabilidad total mayor que 1; sin embargo, no siempre los errores que implican a la regla de la suma hacen que la probabilidad total sea mayor que 1.

Sucesos complementarios

En la sección 4-2 definimos el complemento del suceso A y lo denotamos como \bar{A} . Dijimos que \bar{A} consiste en todos los resultados en los que el suceso A *no* ocurre. Los sucesos A y \bar{A} deben ser disjuntos, porque es imposible que un suceso y su complemento ocurran al mismo tiempo. Además, podemos estar absolutamente seguros de que A ocurre, o bien, de que no ocurre, lo que implica que debe ocurrir A o \bar{A} . Estas observaciones nos permiten aplicar la regla de la suma para sucesos mutuamente excluyentes de la siguiente manera:

$$P(A \text{ o } \bar{A}) = P(A) + P(\bar{A}) = 1$$

Justificamos $P(A \text{ o } \bar{A}) = P(A) + P(\bar{A})$ señalando que A y \bar{A} son disjuntos; justificamos el total de 1 por nuestra certeza absoluta de que A ocurre, o bien, no ocurre. Este resultado de la regla de la suma da lugar a las siguientes tres formas equivalentes.

Regla de los sucesos complementarios

$$P(A) + P(\bar{A}) = 1$$

$$P(\bar{A}) = 1 - P(A)$$

$$P(A) = 1 - P(\bar{A})$$

Área total = 1

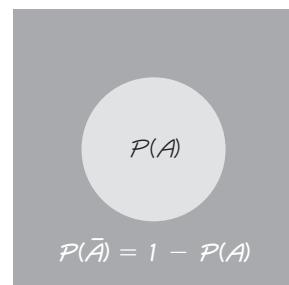


Figura 4-5 Diagrama de Venn del complemento del suceso A

La figura 4-5 es una representación visual de la relación entre $P(A)$ y $P(\bar{A})$.

EJEMPLO En realidad, cuando nace un bebé, $P(\text{n}\tilde{\i}n\tilde{o}) = 0.512$. Calcule $P(\bar{\text{n}\tilde{\i}n\tilde{o}})$.

SOLUCIÓN Usando la regla de los sucesos complementarios, tenemos

$$P(\bar{\text{n}\tilde{\i}n\tilde{o}}) = 1 - P(\text{n}\tilde{\i}n\tilde{o}) = 1 - 0.512 = 0.488$$

Es decir, la probabilidad de no tener un niño, que es la misma que la de tener una niña, es de 0.488.

La principal ventaja de la *regla de los sucesos complementarios* es que puede simplificar mucho ciertos problemas. Ilustraremos esta ventaja en la sección 4-5.

4-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Sucesos disjuntos.** Con sus palabras, describa qué significa que dos sucesos sean *disjuntos*.
- 2. Regla de la suma.** Con sus palabras, describa cómo se aplica la regla de la suma para calcular la probabilidad de que ocurra el suceso *A* o de que ocurra el suceso *B*.
- 3. Encuesta.** Para un proyecto de investigación, usted necesita calcular la probabilidad de que una persona sea zurda o conduzca un automóvil. ¿En qué error incurría si encuestara a un grupo de 500 personas, formado por sus amigos más cercanos y parientes?
- 4. Sucesos disjuntos y complementos.** Si un suceso es el complemento de otro suceso, ¿los dos sucesos deben ser disjuntos? ¿Por qué?

Determinar si los sucesos son disjuntos. En cada uno de los incisos de los ejercicios 5 y 6, ¿los dos eventos son disjuntos para un solo ensayo? (Sugerencia: Considere que “disjunto” es equivalente a “separado” o “que no se traslapan”).

- 5. a.** Elección de un presidente de Estados Unidos
Elección de una candidata
- b.** Seleccionar al azar a una persona que fuma puro
Seleccionar al azar a un hombre
- c.** Seleccionar al azar a una persona tratada con el fármaco Lipitor que reduce los niveles de colesterol
Seleccionar al azar a una persona de un grupo de control que no recibe el medicamento
- 6. a.** Seleccionar al azar una mosca de la fruta con ojos rojos
Seleccionar al azar una mosca de la fruta con ojos sepia (café oscuro)
- b.** Recibir una llamada telefónica de un sujeto de encuesta voluntario que se opone a la clonación
Recibir una llamada telefónica de un sujeto de encuesta voluntario que aprueba la clonación de ovejas
- c.** Seleccionar al azar a un enfermero
Seleccionar al azar a un hombre
- 7. Cálculo de complementos**
 - a.** Si $P(A) = 0.05$, calcule $P(\bar{A})$.
 - b.** Las mujeres tienen una tasa del 0.25% de ceguera a los colores rojo y verde. Si se elige una mujer al azar, ¿cuál es la probabilidad de que no tenga ceguera a los colores rojo y verde? (Sugerencia: Considere que el equivalente decimal de 0.25% es 0.0025, no 0.25).
- 8. Cálculo de complementos**
 - a.** Calcule $P(\bar{A})$, dado que $P(A) = 0.01$.
 - b.** Una encuesta de Reuters/Zogby indicó que el 61% de los estadounidenses creen que existe vida en otros lugares de la galaxia. ¿Cuál es la probabilidad de elegir al azar a una persona que no tenga esta creencia?

En los ejercicios 9 a 12, utilice los datos de la siguiente tabla que resume los resultados de 985 muertes de peatones causadas por accidentes (según datos de la National Highway Traffic Safety Administration).

		¿El peatón estaba intoxicado?	
		Sí	No
¿El conductor estaba intoxicado?	Sí	59	79
	No	266	581

- 9. Muertes de peatones.** Si se elige al azar una de las muertes de peatones, calcule la probabilidad de que el peatón estuviera intoxicado o que el conductor estuviera intoxicado.
- 10. Muertes de peatones.** Si se elige al azar una de las muertes de peatones, calcule la probabilidad de que el peatón no estuviera intoxicado o que el conductor no estuviera intoxicado.
- 11. Muertes de peatones.** Si se elige al azar una de las muertes de peatones, calcule la probabilidad de que el peatón estuviera intoxicado o que el conductor no estuviera intoxicado.
- 12. Muertes de peatones.** Si se elige al azar una de las muertes de peatones, calcule la probabilidad de que el conductor estuviera intoxicado o que el peatón no estuviera intoxicado.

En los ejercicios 13 a 20, utilice los datos de la siguiente tabla que resume los grupos sanguíneos y los factores Rh de 100 personas comunes. Estos valores pueden variar en diferentes regiones de acuerdo con el origen étnico de la población.

		Grupo			
		O	A	B	AB
Tipo	Rh ⁺	39	35	8	4
	Rh ⁻	6	5	2	1

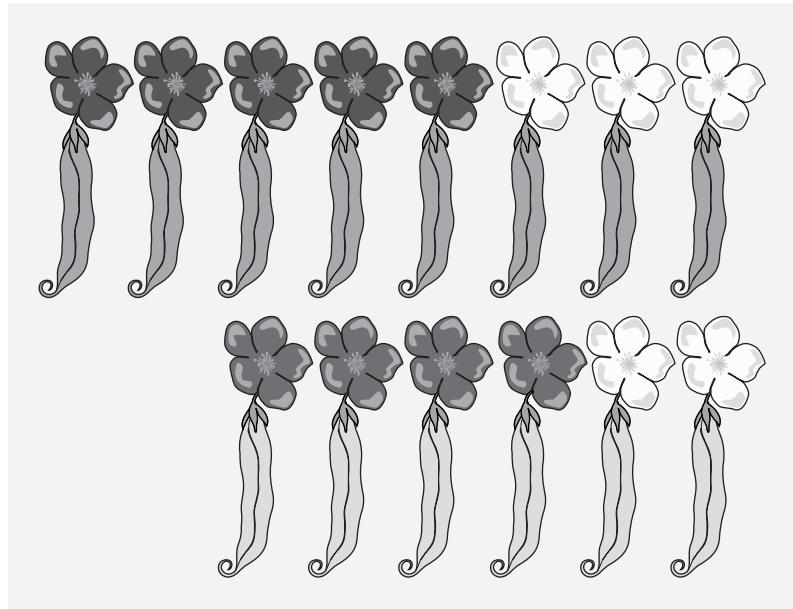
- 13. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule la probabilidad de seleccionar a alguien que no sea del grupo A.
- 14. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule la probabilidad de seleccionar a alguien del tipo Rh⁻.
- 15. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule la probabilidad de seleccionar a alguien que sea del grupo A o del tipo Rh⁻.
- 16. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule la probabilidad de seleccionar a alguien que sea del grupo A o del grupo B.
- 17. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule $P(\text{no del tipo Rh}^+)$.
- 18. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule $P(\text{grupo B o tipo Rh}^+)$.
- 19. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule $P(\text{grupo AB o tipo Rh}^+)$.
- 20. Grupos y tipos sanguíneos.** Si se elige a una persona al azar, calcule $P(\text{grupo A u O o tipo Rh}^+)$.

En los ejercicios 21 y 22, remítase la figura (en la parte superior de la siguiente página) que describe los guisantes utilizados en un estudio genético. (Las probabilidades tienen un papel importante en la genética, y Mendel realizó sus famosos experimentos de hibridación con guisantes, como los que se muestran en la figura).

- 21. Construcción de tabla.** Utilice la figura de la siguiente página para identificar las frecuencias en la tabla que aparece al margen. (Las flores son las porciones superiores y las vainas son las porciones inferiores. Para completar la tabla considere que el color morado está representado en la figura por gris oscuro, y el verde por gris medio. El color amarillo está representado por gris claro, en tanto que el blanco aparece como tal).
- 22. Experimento de hibridación.** Suponga que se elige al azar uno de los guisantes. (Recuerde que en la figura el color morado está representado por gris oscuro, y el verde por gris medio. El color amarillo está representado por gris claro, en tanto que el blanco aparece como tal).
- a. Remítase a la figura y calcule $P(\text{vaina verde o flor morada})$.
- b. Remítase a la tabla construida en el ejercicio 21 y calcule $P(\text{vaina verde o flor morada})$.
- c. ¿Qué formato es más fácil de usar: la figura o la tabla?

Tabla del ejercicio 21

	Flor		
	Morada	Blanca	
Vaina	Verde	?	?
	Amarilla	?	?



Guisantes utilizados en un experimento de hibridación

- 23. Resistencia a la encuesta.** Las empresas que realizan encuestas están preocupadas por los niveles decrecientes de cooperación de las personas que se eligen para ser encuestadas. Un encuestador se pone en contacto con 84 personas entre 18 y 21 años de edad; encuentra que 73 responden y 11 se rehúsan a contestar. Cuando se pone en contacto con 275 personas entre 22 y 29 años, 255 responden y 20 se rehúsan a responder (según datos de “I Hear You Knocking but You Can’t Come In”, de Fitzgerald y Fuller, *Sociological Methods and Research*, vol. 11, núm. 1). Suponga que se selecciona al azar a 1 de 359 personas. Calcule la probabilidad de que sea una persona en el rango de edad de 18 a 21 años o una persona que se rehúsa a responder.
- 24. Resistencia a la encuesta.** Remítase al mismo conjunto de datos utilizado en el ejercicio 23. Suponga que se selecciona al azar a 1 de las 359 personas y calcule la probabilidad de que sea una persona en el rango de 18 a 21 años o alguien que sí respondió.

4-3 MÁS ALLÁ DE LO BÁSICO

- 25. Sucesos disjuntos.** Si los sucesos A y B son disjuntos, y los sucesos B y C son disjuntos, ¿los sucesos A y C deben ser disjuntos? Dé un ejemplo que apoye su respuesta.
- 26. *O exclusivo*.** ¿En qué se modifica la regla de la suma si se utiliza *o exclusivo* en vez de *o inclusive*? En esta sección se señaló que *o exclusivo* significa uno o el otro, pero no ambos.
- 27. Extensión de la regla de la suma.** La regla formal de la suma, presentada en esta sección, expresa la probabilidad de A o B como sigue: $P(A \text{ o } B) = P(A) + P(B) - P(A \text{ y } B)$. Extienda esta regla formal para desarrollar una expresión aplicable a $P(A \text{ o } B \text{ o } C)$. (*Sugerencia:* Dibuje un diagrama de Venn).

Regla de la multiplicación: 4-4 Fundamentos

Concepto básico En la sección 4-3 presentamos la regla de la suma para calcular $P(A \text{ o } B)$, la probabilidad de que un solo ensayo tenga un resultado de A o B o ambos. Esta sección presenta la regla básica de la multiplicación, la cual se utiliza para calcular $P(A \text{ y } B)$, la probabilidad de que el suceso A ocurra en un primer ensayo y que el suceso B ocurra en un segundo ensayo. Si el resultado del primer suceso A afecta de alguna forma la probabilidad del segundo suceso B , es importante ajustar la probabilidad de B para que refleje la ocurrencia del suceso A . La regla para el cálculo de $P(A \text{ y } B)$ se denomina regla de la multiplicación porque implica multiplicar la probabilidad del suceso A por la probabilidad del suceso B (donde la probabilidad del suceso B se ajusta por el resultado del suceso A).

Notación

$P(A \text{ y } B) = P(\text{el suceso } A \text{ ocurre en un primer ensayo y el suceso } B \text{ ocurre en un segundo ensayo})$

En la sección 4-3 asociamos el uso de la palabra *o* con la suma; en esta sección asociaremos el uso de la palabra *y* con la multiplicación.

La teoría de la probabilidad se utiliza extensamente en el análisis y diseño de pruebas estandarizadas, como SAT, ACT, MCAT (para medicina) y LSAT (para derecho). Con el fin de facilitar el proceso de calificación, las pruebas de este tipo suelen usar preguntas de falso/verdadero o de opción múltiple. Supongamos que el primer reactivo de un examen es del tipo falso/verdadero y que el segundo es de opción múltiple con cinco respuestas posibles (a, b, c, d y e). Vamos a usar los dos reactivos siguientes. ¡Intente responderlos!

1. Verdadero o falso: Una libra de plumas es más pesada que una libra de oro.
2. ¿Cuál de las opciones ha tenido la mayor influencia en nuestra comprensión de la genética?
 - a. Gene Hackman
 - b. Gene Simmons
 - c. Gregor Mendel
 - d. Los jeans
 - e. Jean-Jacques Rousseau

Las respuestas a los dos reactivos son V (de “verdadero”) y c. (La primera pregunta es verdadera. Los pesos de las plumas se expresan en libras Avoirdupois, pero los pesos del oro se expresan en libras troy). Calculemos la probabilidad de que si alguna persona hace suposiciones al azar para ambas respuestas, la respuesta al primer reactivo sea correcta y la respuesta al segundo reactivo también sea correcta. Una forma de calcular esta probabilidad es elaborar una lista del espacio muestral, como sigue:

V,a	V,b	V,c	V,d	V,e
F,a	F,b	F,c	F,d	F,e

Si las respuestas son conjetas al azar, entonces los 10 posibles resultados son igualmente probables, por lo tanto



LA ESTADÍSTICA
EN LAS NOTICIAS

Redundancia

La confiabilidad de los sistemas puede mejorarse considerablemente con la redundancia de componentes esenciales. Los automóviles de carreras de las series de la Copa Winston NASCAR tienen dos sistemas de ignición para que, si uno falla, exista otro de reserva. Los aviones poseen dos sistemas eléctricos independientes, y los que se usan para vuelos instrumentales suelen tener dos radios distintos.

La siguiente cita se tomó de un artículo de *Popular Science* acerca de los aviones antiradar: “Un avión construido en buena parte con fibra de carbono fue el Lear Fan 2100, que tenía que llevar dos transpondedores de radar. La razón es que si fallaba una unidad de transpondedor, el avión seguiría siendo casi invisible para el radar”. Tal redundancia es una aplicación de la regla de la multiplicación de la teoría de probabilidad. Si un componente tiene una probabilidad de 0.001 de fallar, la probabilidad de que dos componentes independientes fallen es de sólo 0.000001.



Motores a reacción independientes

Poco después de salir de Miami, el vuelo 855 de Eastern Airlines tuvo que apagar un motor porque se encendió el indicador de baja presión de aceite. Cuando el jet L-1011 regresaba a Miami para aterrizar, los indicadores de baja presión de los otros dos motores también se encendieron. Entonces falló otro motor y después falló el último motor que estaba funcionando. El jet descendió sin propulsión desde 13,000 pies hasta 4000 pies; entonces la tripulación pudo arrancar un motor. La aeronave, con 172 personas a bordo, aterrizó con seguridad. Con motores a reacción independientes, la probabilidad de que los tres fallen es de sólo 0.0001³, es decir, alrededor de una en un billón. La FAA averiguó que el mismo mecánico que cambió el aceite de los tres motores se equivocó al reemplazar los anillos de sello del tapón de aceite. El empleo de un solo mecánico hizo que el funcionamiento de los motores se volviera dependiente, situación que se corrigió exigiendo que los motores reciban mantenimiento por mecánicos diferentes.

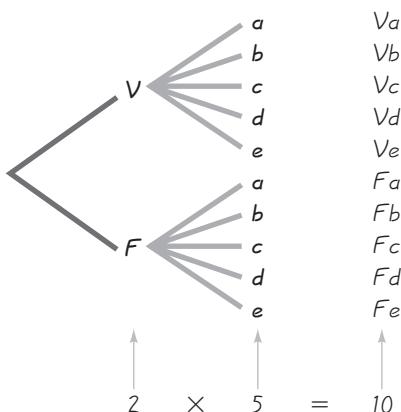


Figura 4-6 Diagrama de árbol de las respuestas del examen

$$P(\text{ambas correctas}) = P(V \text{ y } c) = \frac{1}{10} = 0.1$$

Ahora note que $P(V \text{ y } c) = 1/10$, $P(V) = 1/2$ y $P(c) = 1/5$; por lo tanto, vemos que

$$\frac{1}{10} = \frac{1}{2} \cdot \frac{1}{5}$$

de manera que

$$P(V \text{ y } c) = P(V) \times P(c)$$

Esto sugiere que, en términos generales, $P(A \text{ y } B) = P(A) \cdot P(B)$, pero antes de hacer esta generalización, consideremos otro ejemplo.

En primer lugar, observe que los diagramas de árbol en ocasiones sirven para determinar el número de resultados posibles en el espacio muestral. Un **diagrama de árbol** es una imagen gráfica de los resultados posibles de un procedimiento, los cuales se muestran como líneas queeman de un punto de partida. Estos diagramas a veces son útiles, si el número de posibilidades no es demasiado grande. El diagrama de árbol de la figura 4-6 resume los resultados de los reactivos de verdadero/falso y opción múltiple. En la figura 4-6 vemos que, si las dos respuestas son conjetas al azar, las 10 ramas son igualmente probables, y la probabilidad de obtener el par correcto (V,c) es de 1/10. Para cada respuesta a la primera pregunta, hay cinco respuestas a la segunda. *El número total de resultados es 5 dos veces, es decir, 10.* Por lo tanto, el diagrama de árbol de la figura 4-6 ilustra la razón del uso de la multiplicación.

Nuestro primer ejemplo, el de los reactivos de verdadero/falso y opción múltiple, sugiere que $P(A \text{ y } B) = P(A) \cdot P(B)$, pero el siguiente ejemplo incluye otro elemento importante.

PROBLEMA
4
DEL CAPÍTULO

EJEMPLO Examen de drogas. El problema del capítulo incluye la tabla 4-1, la cual se reproduce a continuación. Si se elige al azar a dos de los sujetos incluidos en la tabla, *sin reemplazo*, calcule la probabilidad de que la primera persona seleccionada tenga un resultado de prueba positivo y que la segunda tenga un resultado de prueba negativo.

Tabla 4-1

Resultados de exámenes sobre el consumo de marihuana

		¿Los sujetos realmente consumen marihuana?	
		Sí	No
Resultado de prueba positivo (La prueba indica que la marihuana está <i>presente</i>).		119 (verdadero positivo)	24 (falso positivo)
Resultado de prueba negativo (La prueba indica que la marihuana está <i>ausente</i>).		3 (falso negativo)	154 (verdadero negativo)

SOLUCIÓN

Primera selección:

$P(\text{resultado de prueba positivo}) = 143/300$ (porque hay 143 sujetos que resultaron positivos, y el número total de sujetos es de 300)

Segunda selección:

$P(\text{resultado de prueba negativo}) = 157/299$ (después de la primera selección de un sujeto con un resultado de prueba positivo, quedan 299 sujetos, 157 de los cuales tienen resultados negativos de la prueba)

Con $P(\text{el primer sujeto tiene un resultado de prueba positivo}) = 143/300$ y $P(\text{el segundo sujeto tiene un resultado de prueba negativo}) = 157/299$, tenemos

$$P(\text{el primer sujeto tiene un resultado de prueba positivo y el segundo sujeto tiene un resultado de prueba negativo}) = \frac{143}{300} \cdot \frac{157}{299} = 0.250$$

El punto clave es que *se debe ajustar la probabilidad del segundo suceso para reflejar el resultado del primer suceso*. Como la selección del segundo sujeto se realiza sin reemplazar al primer sujeto, la segunda probabilidad debe tomar en cuenta el hecho de que la primera selección eliminó a un sujeto que resultó positivo, de manera que en la segunda selección sólo existen 299 sujetos, y 157 de ellos tuvieron un resultado de prueba negativo.

Este ejemplo ilustra el importante principio de que *la probabilidad del segundo suceso B debe tomar en cuenta el hecho de que el primer suceso A ya ocurrió*. Este principio suele expresarse usando la siguiente notación.

Por ejemplo, jugar a la lotería de California y después jugar a la lotería de Nueva York son sucesos *independientes* porque el resultado de la lotería de California no



LA ESTADÍSTICA EN LAS NOTICIAS

Recomendación para la lotería

Un columnista del diario *New York Daily News*, Stephen Allensworth, hace poco dio recomendaciones para seleccionar números en el juego *Daily Numbers* del estado de Nueva York. En la descripción de un sistema para ganar, escribió que “comprende números dobles asociados con dígitos fríos. (Un dígito frío es uno que sale una vez o no sale nunca en un periodo de siete días)”. Allensworth procedió a identificar algunos números específicos que “tienen una excelente probabilidad de salir esta semana”.

Allensworth supone que algunos números están “rezagados”, pero la selección de números en la lotería es independiente de los resultados pasados. El sistema que describe no tiene bases reales y no funcionará. Los lectores que siguen una recomendación tan pobre como ésta, están siendo engañados y perderán más dinero, ya que creen erróneamente que sus probabilidades de ganar mejoran.

Notación para la probabilidad condicional

$P(B|A)$ representa la probabilidad de que un suceso *B* ocurra después de suponer que el suceso *A* ya ocurrió. (Podemos leer $B|A$ como “*B* dado *A*” o como “el suceso *B* ocurre después de que el suceso *A* ya ocurrió”).

Definiciones

Dos sucesos A y B son **independientes** cuando la ocurrencia de uno no afecta la probabilidad de la ocurrencia del otro. (De manera similar, muchos otros sucesos son independientes si la ocurrencia de cualquiera no afecta las probabilidades de la ocurrencia de los demás). Si A y B no son independientes, se dice que son **dependientes**.

tiene ningún efecto en las probabilidades de los resultados de la lotería de Nueva York. En contraste, el suceso de intentar arrancar su automóvil y el suceso de llegar a clase a tiempo son sucesos *dependientes*, porque el resultado del intento de arrancar su automóvil afecta la probabilidad de llegar a clase a tiempo.

Usando la notación y las definiciones anteriores, junto con los principios ilustrados en los ejemplos previos, podemos resumir el concepto clave de esta sección como la siguiente *regla formal de la multiplicación*, pero le recomendamos trabajar con la *regla intuitiva de la multiplicación*, la cual tiene más probabilidades de facilitar la comprensión que el uso a ciegas de una fórmula.

Regla formal de la multiplicación

$$P(A \text{ y } B) = P(A) \cdot (B|A)$$

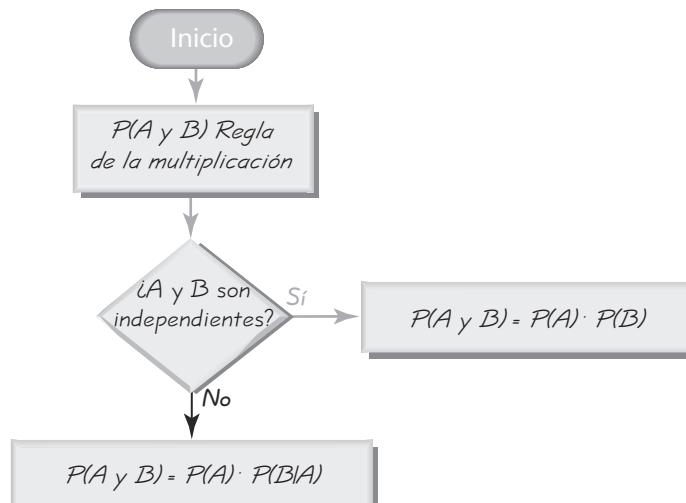
Si A y B son sucesos independientes, $P(B|A)$ es en realidad lo mismo que $P(B)$. Vea la siguiente *regla intuitiva de la multiplicación*. (Véase también la figura 4-7).

Regla intuitiva de la multiplicación

Cuando calcule la probabilidad de que el suceso A ocurra en un ensayo y el suceso B ocurra en el siguiente ensayo, multiplique la probabilidad del suceso A por la probabilidad del suceso B , pero asegúrese de que la probabilidad del suceso B toma en cuenta la ocurrencia previa del suceso A .

Figura 4-7

Aplicación de la regla de la multiplicación



EJEMPLO Plantas Una bióloga experimenta con una muestra de dos plantas vasculares (indicadas por V) y cuatro plantas no vasculares (indicadas por N). A continuación se listan los códigos de las seis plantas estudiadas. La bióloga desea elegir al azar dos de las plantas para estudiarlas con mayor profundidad. Calcule la probabilidad de que la primera planta seleccionada sea no vascular (N) y que la segunda planta también sea no vascular (N). Suponga que las selecciones se hacen *a)* con reemplazo; *b)* sin reemplazo.

V V N N N N

SOLUCIÓN

- a.** Si las dos plantas se seleccionan *con reemplazo*, las dos selecciones son independientes, ya que el segundo suceso no se ve afectado por el primer resultado. En cada selección, de las seis plantas, cuatro son no vasculares (N), de manera que obtenemos

$$P(\text{la primera planta es N y la segunda planta es N}) = \frac{4}{6} \cdot \frac{4}{6} = \frac{16}{36} \text{ o } 0.444$$

- b.** Si las dos plantas se seleccionan *sin reemplazo*, las dos selecciones son dependientes, ya que la probabilidad del segundo suceso se ve afectada por el primer resultado. En la primera selección, cuatro de las seis plantas son no vasculares (N). Después de elegir a una planta no vascular en la primera selección, quedan cinco plantas, incluyendo tres que son no vasculares. Por lo tanto,

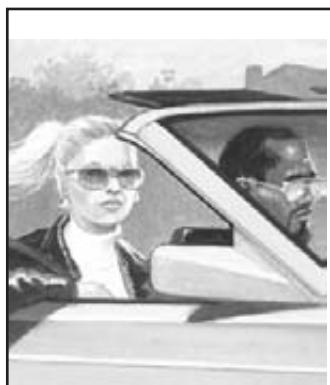
$$P(\text{la primera planta es N y la segunda planta es N}) = \frac{4}{6} \cdot \frac{3}{5} = \frac{12}{30} = \frac{2}{5} \text{ o } 0.4$$

Observe que en este caso ajustamos la segunda probabilidad para tomar en cuenta la selección de una planta no vascular (N) en el primer resultado. Después de seleccionar N la primera vez, de las cinco plantas restantes, tres serían no vasculares.

Hasta aquí hemos analizado dos sucesos, pero la regla de la multiplicación puede extenderse fácilmente a varios sucesos. En general, la probabilidad de cualquier secuencia de sucesos independientes es simplemente el producto de sus probabilidades correspondientes. Por ejemplo, la probabilidad de lanzar una moneda tres veces y obtener siempre cara es $0.5 \cdot 0.5 \cdot 0.5 = 0.125$. También podemos extender la regla de la multiplicación para que se aplique a varios eventos dependientes, simplemente ajuste las probabilidades conforme se va avanzando.

Tratar sucesos dependientes como si fueran independientes El inciso *b*) del último ejemplo implicó la selección de elementos sin reemplazo y, por consiguiente, consideramos los sucesos como dependientes. Sin embargo, es común tratar sucesos como independientes cuando se toman *muestras pequeñas de poblaciones grandes*. En estos casos, es raro que se seleccione el mismo elemento dos veces. He aquí un lineamiento común:

Si el tamaño de la muestra no es mayor que el 5% del tamaño de la población, trate las selecciones como si fueran *independientes* (si las selecciones se hacen sin reemplazo, de manera que sean técnicamente dependientes).



Sentenciados por probabilidad

Un testigo describió a una asaltante de Los Ángeles como una mujer caucásica de pelo rubio, peinado en cola de caballo, que escapó en un automóvil amarillo conducido por un hombre afroestadounidense que usaba barba y bigote. Janet y Malcom Collins se ajustaban a esta descripción y se les condenó con fundamento en el testimonio de que existe aproximadamente una posibilidad en 12 millones de que cualquier pareja tenga tales características. Se estimó que la probabilidad de tener un automóvil amarillo es de 1/10, en tanto que las demás probabilidades se estimaron en 1/4, 1/10, 1/3, 1/10 y 1/1,000. Más tarde, las condenas se anularon, cuando se señaló que no se presentó evidencia que apoyara las probabilidades que se estimaron o la independencia de los sucesos. Sin embargo, puesto que la pareja no se seleccionó al azar, se cometió un error grave al no considerar la probabilidad de que hubiera otras parejas en la misma región con las mismas características.



Calificación perfecta en el SAT

Si se selecciona al azar a un sujeto que responde el Scholastic Aptitude Test (SAT), ¿cuál es la probabilidad de elegir a una persona que obtenga una calificación perfecta? ¿Cuál es la probabilidad de obtener una calificación perfecta en el SAT, adivinando las respuestas? Éstas son dos preguntas muy diferentes.

La prueba del SAT cambió de dos a tres secciones en 2005, y entre los 300,000 estudiantes que tomaron la primera prueba en marzo de 2005, 107 lograron calificaciones perfectas de 2400 puntos al obtener 800 en cada una de las tres secciones correspondientes a escritura, lectura de comprensión y matemáticas. Con base en estos resultados, la probabilidad de seleccionar al azar a uno de los sujetos de prueba y elegir a una persona con calificación perfecta es $107/300,000$ o aproximadamente 0.000357. En un año reciente en el que aún se aplicaba el antiguo examen SAT, aproximadamente 1.3 millones de personas lo tomaron, de las cuales, 587 recibieron calificaciones perfectas de 1600, lo que representa una probabilidad de 0.000452. Sólo una parte del SAT consiste en 35 preguntas de opción múltiple, y la probabilidad de responder a todas ellas correctamente adivinando es de $(1/5)^{35}$, una cantidad tan pequeña que cuando se escribe como un decimal, deben escribirse 24 ceros entre el punto y el primer número diferente de cero.

Los encuestadores usan este lineamiento cuando encuestan apenas a 1000 adultos de poblaciones de millones. Ellos suponen independencia, aunque toman la muestra sin reemplazo.

El siguiente ejemplo nos da idea del importante procedimiento de *prueba de hipótesis* que se estudiará en el capítulo 8.

EJEMPLO Eficacia de la selección del género Un genetista desarrolla un procedimiento para aumentar la probabilidad de engendrar una niña. En una prueba inicial, 20 parejas utilizan el método, lo que da como resultado 20 niñas en 20 nacimientos. Suponiendo que el procedimiento de selección del género no tiene efecto, calcule la probabilidad de que nazcan 20 niñas en 20 nacimientos, debido al azar. Con base en los resultados, ¿existe una fuerte evidencia que apoye la afirmación del genetista de que el procedimiento es eficaz para incrementar la probabilidad de engendrar una niña?

SOLUCIÓN Deseamos calcular $P(\text{los 20 bebés son niñas})$, suponiendo que el procedimiento no tiene ningún efecto, de manera que la probabilidad de que cualquier descendiente sea mujer es 0.5. Como se utilizan pares separados de padres, trataremos los sucesos como si fueran independientes. Se obtiene el siguiente resultado:

$$\begin{aligned} P(\text{los 20 bebés son niñas}) &= P(\text{el primero es niña, y el segundo es niña, y el tercero es niña \dots} \\ &\quad \text{y el vigésimo es niña}) \\ &= P(\text{niña}) \cdot P(\text{niña}) \cdot \dots \cdot P(\text{niña}) \\ &= 0.5 \cdot 0.5 \cdot \dots \cdot 0.5 \\ &= 0.5^{20} = 0.000000954 \end{aligned}$$

La baja probabilidad de 0.000000954 indica que en vez de tener 20 niñas al azar, una explicación más razonable es que con el procedimiento de selección de género hay más probabilidades de que nazcan niñas. Como existe una probabilidad tan pequeña (0.000000954) de que resulten 20 niñas en 20 nacimientos, tenemos evidencia suficiente para concluir que el procedimiento de selección de género sirve para incrementar la probabilidad de que un bebé sea niña. Es decir, parece que el procedimiento es eficaz.

Podemos resumir los fundamentos de las reglas de la suma y de la multiplicación como sigue:

- En la regla de la suma, la palabra “o” en $P(A \text{ o } B)$ sugiere suma. Sume $P(A)$ y $P(B)$, teniendo cuidado de sumar en tal forma que cada resultado se cuente sólo una vez.
- En la regla de la multiplicación, la palabra “y” en $P(A \text{ y } B)$ sugiere multiplicación. Multiplique $P(A)$ y $P(B)$, pero asegúrese de que la probabilidad del suceso B tome en cuenta la ocurrencia previa del suceso A .

4-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Sucesos independientes.** Con sus propias palabras, explique qué significa que dos sucesos sean independientes.
2. **Muestreo con reemplazo.** El profesor de una clase de 25 estudiantes elige al azar a un estudiante y luego selecciona también al azar a un segundo estudiante. Si en la segunda selección los 25 estudiantes están disponibles, ¿se trata de un muestreo con reemplazo o de un muestreo sin reemplazo? ¿El segundo resultado es independiente del primero?
3. **Muestreo sin reemplazo.** El profesor de una clase de 25 estudiantes elige al azar a un estudiante y luego selecciona también al azar a un segundo estudiante. Si en la segunda selección están disponibles 24 estudiantes, ¿se trata de un muestreo con reemplazo o de un muestreo sin reemplazo? ¿El segundo resultado es independiente del primero?
4. **Notación.** ¿Qué representa la notación $P(B|A)$?

Identificación de sucesos como independientes o dependientes. En los ejercicios 5 y 6, clasifique cada par de sucesos como independientes o dependientes. (Si dos sucesos son técnicamente dependientes, pero se pueden tratar como si fueran independientes, considérelos independientes).

5. a. Elegir al azar una moneda de 25 centavos acuñada antes de 2001
Elegir al azar una segunda moneda de 25 centavos acuñada antes de 2001
- b. Elegir al azar a un televidente que está viendo *The Barry Manilow Biography*
Elegir al azar a un segundo televidente que está viendo *The Barry Manilow Biography*
- c. Usar pantalón corto de tela escocesa con calcetines negros y sandalias
Pedirle a alguien una cita y recibir una respuesta positiva
6. a. Descubrir que su calculadora funciona
Descubrir que su teléfono celular funciona
- b. Descubrir que su tostador de pan no funciona
Descubrir que su refrigerador no funciona
- c. Beber alcohol o consumir drogas hasta deteriorar su capacidad de conducir
Verse implicado en un accidente automovilístico
7. **Conjeturas.** Un examen rápido consiste en una pregunta de verdadero/falso, seguida de una pregunta de opción múltiple con cuatro respuestas posibles (a, b, c y d). Si ambas preguntas se responden con conjeturas, calcule la probabilidad de que las dos respuestas sean correctas. ¿Hacer conjeturas es una buena estrategia en este examen?
8. **Letra y dígito.** La dueña de una nueva computadora crea una contraseña de dos caracteres. Seleccionó al azar la letra del alfabeto para el primer carácter y un dígito (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) para el segundo carácter. ¿Cuál es la probabilidad de que su contraseña sea “K9”? ¿Servirá esta contraseña para evitar que otra persona tenga acceso a su computadora?
9. **Uso de ropa anaranjada de cazador.** Un estudio que pretendía descubrir la relación entre sufrir heridas de caza y usar ropa anaranjada “de cazador” mostró que de 123 cazadores heridos por ser confundidos con presas, 6 usaban ropa anaranjada (según datos de los Centers for Disease Control). Si un estudio de seguimiento comenzara con la selección aleatoria de cazadores de esta muestra de 123, calcule la probabilidad de que los primeros dos cazadores seleccionados usaran ropa anaranjada.

continúa

- a. Suponga que el primer cazador se reemplaza antes de que el siguiente se seleccione.
 - b. Suponga que el primer cazador no se reemplaza antes de que el segundo cazador se seleccione.
 - c. Dada la alternativa entre seleccionar con reemplazo y sin reemplazo, ¿cuál es más lógica para esta situación? ¿Por qué?
- 10. Selección de senadores en Estados Unidos.** En el 108º Congreso de Estados Unidos, el senado consta de 51 republicanos, 48 demócratas y 1 independiente. Si un cabildero de la industria del tabaco selecciona al azar a tres diferentes senadores, ¿cuál es la probabilidad de que todos sean republicanos? ¿Sería probable que un cabildero usara la selección aleatoria en esta situación?
- 11. Muestreo de aceptación.** Con cierto método de un procedimiento llamado *muestreo de aceptación*, se selecciona aleatoriamente y sin reemplazo una muestra de artículos; el lote completo se acepta si cada artículo en la muestra es aprobado. La Niko Electronics Company acaba de fabricar 5000 CD y 100 están defectuosos. Si se seleccionan al azar 4 de estos CD para probarlos, ¿cuál es la probabilidad de que se acepte el lote completo?
- 12. Nivel de confianza de encuesta.** En las encuestas de opinión pública, es común manejar un “nivel de confianza” del 95%, lo cual significa que hay un 0.95 de probabilidad de que los resultados de la encuesta sean precisos dentro de los márgenes de error establecidos. Si seis organizaciones diferentes realizan encuestas independientes, ¿cuál es la probabilidad de que las seis sean precisas dentro de los márgenes de error establecidos? ¿El resultado sugiere que con un nivel de confianza del 95% podemos esperar que casi todas las encuestas estén dentro del margen de error establecido?
- 13. Prueba de eficacia de un método de selección de género.** Descubrimientos recientes parecen hacer posible que las parejas aumenten, de forma considerable, la posibilidad de tener un hijo del género de su elección. En una prueba de un método de selección del género, 12 parejas tratan de concebir niñas. Si este método de selección del género no tuviera efecto, ¿cuál es la probabilidad de que los 12 bebés sean niñas? Si en realidad entre 12 hijos 12 resultan niñas, ¿parece ser efectivo este método de selección de género? ¿Por qué?
- 14. Identificación de la voz de un criminal.** En un caso legal en Riverhead, Nueva York, nueve víctimas diferentes de un crimen escucharon grabaciones de la voz de cinco hombres diferentes. Las nueve víctimas identificaron la misma voz como la del criminal. Si las identificaciones de voz se hubieran hecho al azar, calcule la probabilidad de que las nueve víctimas seleccionaran a la misma persona. ¿Constituye esto una duda razonable?
- 15. Redundancia.** El principio de redundancia se utiliza cuando la confiabilidad de un sistema se mejora por medio de componentes redundantes o de respaldo. Suponga que su reloj despertador tiene un 0.975 de probabilidad de funcionar en cualquier mañana dada.
- a. ¿Cuál es la probabilidad de que su reloj despertador *no* funcione en la mañana de un examen final importante?
 - b. Si usted tiene dos relojes despertadores como el descrito, ¿cuál es la probabilidad de que ambos fallen en la mañana de un examen final importante?
 - c. Con un reloj despertador, tenemos un 0.975 de probabilidad de ser despertados. ¿Cuál es la probabilidad de ser despertado si estamos usando dos relojes despertadores?
 - d. ¿Un segundo despertador representa una confiabilidad considerablemente mayor?
- 16. Habilidades sociales.** Bob considera que cuando le pide a una mujer una cita, ella puede aceptar o rechazar su propuesta, y supone que tiene una probabilidad de 0.05 de conseguir la cita. Si esta suposición es correcta, ¿cuál es la probabilidad de que Bob reciba cinco rechazos al pedir citas a cinco mujeres diferentes? ¿Este resultado es la probabilidad correcta de que Bob reciba cinco rechazos al pedir citas a cinco mujeres diferentes? ¿Por qué?

En los ejercicios 17 a 20, utilice los datos de la siguiente tabla, que resume los resultados de 985 muertes de peatones, causadas por accidentes (según datos de la National Highway Traffic Safety Administration).

		¿Peatón intoxicado?	
		Sí	No
¿Conductor intoxicado?	Sí	59	79
	No	266	581

17. Conductores intoxicados. Si se seleccionan al azar dos muertes *diferentes* de peatones, calcule la probabilidad de que ambas involucren a conductores intoxicados.

18. Peatones intoxicados. Si se seleccionan al azar dos muertes *diferentes* de peatones, calcule la probabilidad de que ambas involucren a peatones intoxicados.

19. Muertes de peatones

- a. Si se elige al azar una de las muertes de peatones, ¿cuál es la probabilidad de que sea un caso en donde ni el peatón ni el conductor estuvieran intoxicados?
- b. Si se eligen al azar dos muertes *diferentes* de peatones, ¿cuál es la probabilidad de que, en ambos casos, ni el peatón ni el conductor estuvieran intoxicados?
- c. Si se eligen al azar dos muertes de peatones con reemplazo, ¿cuál es la probabilidad de que en ambos casos, ni el peatón ni el conductor estuvieran intoxicados?
- d. Compare los resultados de los incisos b) y c).

20. Muertes de peatones

- a. Si se elige al azar una de las muertes de peatones, ¿cuál es la probabilidad de que involucre a un peatón intoxicado y a un conductor intoxicado?
- b. Si se eligen al azar dos muertes *diferentes* de peatones, ¿cuál es la probabilidad de que en ambos casos, tanto el peatón como el conductor estuvieran intoxicados?
- c. Si se eligen al azar dos muertes de peatones con reemplazo, ¿cuál es la probabilidad de que en ambos casos, tanto el peatón como el conductor estuvieran intoxicados?
- d. Compare los resultados de los incisos b) y c).

4-4 MÁS ALLÁ DE LO BÁSICO

21. Las mismas fechas de cumpleaños. Calcule la probabilidad de que dos personas no tengan la misma fecha de cumpleaños, cuando el número de personas seleccionadas al azar es

- a. 3
- b. 5
- c. 25

22. Género de hijos

- a. Si una pareja planea tener ocho hijos, calcule la probabilidad de que todos sean del mismo género.
- b. Suponiendo que los niños y las niñas fueran igualmente probables, calcule la probabilidad de que, de 1000 nacimientos, todos los bebés sean niñas. ¿El resultado indica que un suceso como éste es imposible?

23. Selección de cartas. Se van a seleccionar dos cartas de un mazo revuelto, al azar y sin reemplazo. Calcule la probabilidad de obtener un as en la primera carta y una espada en la segunda carta.

24. Complementos y la regla de la suma

- Desarrolle una fórmula para la probabilidad de no obtener A o B en un mismo ensayo. Esto es, calcule una expresión para $P(\overline{A} \text{ o } \overline{B})$.
- Desarrolle una fórmula para la probabilidad de no obtener A o no obtener B en un mismo ensayo. Esto es, calcule una expresión para $P(\overline{A} \text{ o } \overline{B})$.
- Compare los resultados de los incisos a) y b). ¿Es $P(\overline{A} \text{ o } \overline{B}) = P(\overline{A} \text{ o } \overline{B})$?

La regla de la multiplicación: Complementos y probabilidad

4-5 Complementos y probabilidad condicional

En la sección 4-4 se presentó el concepto básico de la regla de la multiplicación; en esta sección extenderemos el uso de esa regla a dos aplicaciones especiales. Primero, cuando queremos identificar la probabilidad de que, en varios ensayos, ocurra *al menos uno* de ciertos eventos específicos, un método sencillo consiste en calcular la probabilidad de que *ninguno* de los sucesos ocurra y luego calcular el complemento de tal suceso. En segundo lugar, consideraremos la probabilidad condicional, que es la probabilidad de un suceso, dada la información adicional de que algún otro suceso ya ocurrió.

Empecemos con situaciones en las que deseamos calcular la probabilidad de que, en varios ensayos, *al menos uno* produzca algún resultado especificado.

Complementos: la probabilidad de “uno al menos”

La regla de la multiplicación y la regla de los complementos pueden utilizarse en conjunto para simplificar en gran medida la solución a este tipo de problema: calcular la probabilidad de que entre varios ensayos, *al menos uno* dé algún resultado especificado. En casos como éste, es esencial que el significado del lenguaje se comprenda con claridad:

- “Al menos uno” equivale a “uno o más”.
- El complemento de obtener al menos uno de los elementos de un tipo en particular es que usted *no* obtenga elementos de ese tipo.

Suponga que una pareja planea tener tres hijos y quiere conocer la probabilidad de tener al menos una niña. Veamos las siguientes interpretaciones:

Al menos una niña de entre tres hijos = 1 o más niñas.

El complemento de “al menos una niña” = ninguna niña = los tres hijos son varones.

Podríamos calcular esta probabilidad con facilidad, a partir de una lista del espacio muestral completo de ocho resultados, pero queremos ilustrar el uso de los complementos, los cuales podrán usarse en muchos otros problemas que no se resuelven tan fácilmente.

EJEMPLO Género de hijos Calcule la probabilidad de que una pareja tenga al menos 1 niña entre tres hijos. Suponga que los niños y las niñas son igualmente probables y que el género de un hijo es independiente del género de cualquier hermano o hermana.

SOLUCIÓN

Paso 1: Use un símbolo para representar el suceso deseado. En este caso, sea A = al menos 1 de los tres hijos es una niña.

Paso 2: Identifique el suceso que es el complemento de A .

$$\begin{aligned}\bar{A} &= \text{no tener al menos una niña entre tres hijos} \\ &= \text{los tres hijos son niños} \\ &= \text{nino y niño y niño}\end{aligned}$$

Paso 3: Calcule la probabilidad del complemento.

$$\begin{aligned}P(\bar{A}) &= P(\text{nino y niño y niño}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}\end{aligned}$$

Paso 4: Calcule $P(A)$ evaluando $1 - P(\bar{A})$.

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{1}{8} = \frac{7}{8}$$

INTERPRETACIÓN Existe una probabilidad en $7/8$ de que si una pareja tiene tres hijos, al menos uno sea niña.

El principio utilizado en este ejemplo se resume como sigue:

Para calcular la probabilidad de *al menos uno* de algo, calcule la probabilidad de *ninguno* y reste este resultado de 1. Esto es,

$$P(\text{al menos uno}) = 1 - P(\text{ninguno})$$

Probabilidad condicional

Ahora consideraremos el siguiente punto principal de esta sección, el cual se basa en el principio de que la probabilidad de un suceso suele verse afectada por el conocimiento previo de las circunstancias. Por ejemplo, si usted selecciona aleatoriamente a una persona de la población general, la probabilidad de obtener un varón es de 0.5, pero si usted ya sabe que la persona a seleccionar fuma puro, entonces aumenta la probabilidad de manera drástica de que la persona sea hombre (porque el 85% de los fumadores de puros son hombres). Una *probabilidad condicional* de un suceso ocurre cuando la probabilidad se ve afectada por el conocimiento de otras circunstancias. La probabilidad condicional de que el suceso B ocurra, dado que el suceso A ya ocurrió, puede calcularse mediante la regla de la multiplicación [$P(A \text{ y } B) = P(A) \cdot P(B|A)$] y resolviendo para $P(B|A)$ al dividir ambos lados de la ecuación entre $P(A)$.

Definición

La **probabilidad condicional** de un suceso es una probabilidad obtenida con la información adicional de algún otro evento que ya ocurrió. $P(B|A)$ denota la probabilidad condicional de que el suceso B ocurra, dado que el suceso A ya ocurrió, y puede calcularse dividiendo la probabilidad de que ambos sucesos A y B ocurran entre la probabilidad del suceso A :

$$P(B|A) = \frac{P(A \text{ y } B)}{P(A)}$$



¿Coincidencias?

John Adams y Thomas Jefferson (el segundo y tercer presidentes de Estados Unidos) murieron el mismo día, el 4 de julio de 1826. El presidente Lincoln murió asesinado en el teatro Ford; el presidente Kennedy murió asesinado en un automóvil Lincoln fabricado por la Ford Motor Company. Los sucesores a la presidencia tanto de Lincoln como de Kennedy fueron vicepresidentes apellidados Johnson. Catorce años *antes* del naufragio del *Titanic*, una novela describió el hundimiento del *Titan*, un barco que chocó con un iceberg; véase *The Wreck of the Titanic Foretold?* de Martin Gardner. Este autor señala: “En casi todos los casos de coincidencias desconcertantes, es imposible hacer siquiera una estimación aproximada de su probabilidad”.

Esta fórmula es una expresión formal de la probabilidad condicional, pero no se recomienda el uso irreflexivo de las fórmulas. Recomendamos el siguiente método intuitivo.

Método intuitivo para la probabilidad condicional

La probabilidad condicional de B dado A puede calcularse suponiendo que el suceso A ya ocurrió y, bajo ese supuesto, calcular la probabilidad de que ocurra el suceso B .

EJEMPLO Prueba de drogas Remítase a la tabla 4-1, la cual se reproduce aquí para su comodidad. Calcule lo siguiente:

- Si se elige al azar a uno de los 300 sujetos de prueba, calcule la probabilidad de que la persona resulte positiva, dado que en realidad consumió marihuana.
- Si se elige al azar a uno de los 300 sujetos de prueba, calcule la probabilidad de que la persona realmente haya consumido marihuana, dado que tuvo un resultado de prueba positivo.

SOLUCIÓN

- Deseamos conocer $P(\text{positivo} | \text{consumo de marihuana})$, la probabilidad de elegir a una persona que resultó positiva, *dado que la persona seleccionada realmente consumió marihuana*. He aquí el punto clave: si suponemos que la persona seleccionada consume marihuana, sólo nos estamos refiriendo a los 122 sujetos de la primera columna en la tabla 4-1. De esos 122 sujetos, 119 resultaron positivos, de manera que

$$P(\text{positivo} | \text{consumo de marihuana}) = \frac{119}{122} = 0.975$$

Podemos obtener el mismo resultado utilizando la fórmula que incluimos con la definición de la probabilidad condicional. En el siguiente cálculo tomamos en cuenta el hecho de que 119 de los 300 sujetos eran consumidores de marihuana y resultaron positivos en la prueba. Asimismo, 122 de los 300

Tabla 4-1 Resultados de exámenes sobre el consumo de marihuana

	¿Los sujetos realmente consumen marihuana?	
	Sí	No
Resultado de prueba positivo (La prueba indica que la marihuana está <i>presente</i>).	119 (verdadero positivo)	24 (falso positivo)
Resultado de prueba negativo (La prueba indica que la marihuana está <i>ausente</i>).	3 (falso negativo)	154 (verdadero negativo)

sujetos consumieron marihuana. Por lo tanto,

$$\begin{aligned} P(\text{positivo} | \text{consumo de marihuana}) &= \frac{P(\text{consumo de marihuana y positivo})}{P(\text{consumo de marihuana})} \\ &= \frac{119/300}{122/300} = 0.975 \end{aligned}$$

- b. Deseamos conocer $P(\text{consumo de marihuana} | \text{positivo})$. Si suponemos que la persona elegida resultó positiva, nos estamos refiriendo a los 143 sujetos del primer renglón de la tabla 4-1. De esos 143 sujetos, 119 consumieron marihuana, de manera que

$$P(\text{consumo de marihuana} | \text{positivo}) = \frac{119}{143} = 0.832$$

De nuevo, el mismo resultado puede encontrarse aplicando la fórmula para la probabilidad condicional:

$$\begin{aligned} P(\text{consumo de marihuana} | \text{positivo}) &= \frac{P(\text{positivo y consumo de marihuana})}{P(\text{positivo})} \\ &= \frac{119/300}{143/300} = 0.832 \end{aligned}$$

Si comparamos los resultados de los incisos a) y b), vemos que $P(\text{positivo} | \text{consumo de marihuana})$ no es igual que $P(\text{consumo de marihuana} | \text{positivo})$.

INTERPRETACIÓN El primer resultado de $P(\text{positivo} | \text{consumo de marihuana}) = 0.975$ indica que un consumidor de marihuana tiene una probabilidad de 0.975 de obtener un resultado de prueba positivo. El segundo resultado de $P(\text{consumo de marihuana} | \text{positivo}) = 0.832$, indica que, en algunos resultados de prueba positivos, existe una probabilidad de 0.832 de que la persona realmente haya consumido marihuana.

Confusión del inverso

Observe que en el ejemplo anterior, $P(\text{positivo} | \text{consumo de marihuana})$ $P(\text{consumo de marihuana} | \text{positivo})$. El hecho de creer de manera incorrecta que $P(B|A)$ y $P(A|B)$ son iguales, o utilizar incorrectamente un valor por otro suele llamarse *confusión del inverso*. Diversos estudios demuestran que los médicos a menudo dan información incorrecta cuando confunden el inverso. Con base en estudios reales, tienden a confundir $P(\text{cáncer} | \text{prueba positiva})$ con $P(\text{prueba positiva} | \text{cáncer})$. Alrededor del 95% de los médicos estimaron que $P(\text{cáncer} | \text{prueba positiva})$ implica 10 veces más probabilidades; como consecuencia, esos pacientes recibieron diagnósticos muy confusos, sintiéndose innecesariamente alterados por la información incorrecta.



Muestreo compuesto

En una ocasión el ejército estadounidense hizo pruebas para determinar la presencia de sífilis en cada recluta, tomando una muestra de sangre individual que se analizaba por separado. Un investigador sugirió mezclar muestras de sangre por pares. Después de analizar los pares mezclados, los reclutas con sífilis podían identificarse volviendo a analizar las muestras de sangre individuales de los pocos pares que dieron resultados positivos en el análisis. El número total de análisis se redujo apareando muestras de sangre, así que ¿por qué no colocarlos en grupos de tres o cuatro o más? Se usó la teoría de probabilidad para determinar el tamaño de grupo más eficiente y se desarrolló una teoría general para detectar los defectos en cualquier población. Esta técnica se conoce como *muestreo compuesto*.

4-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Probabilidad de al menos uno.** Usted desea calcular la probabilidad de obtener al menos un defecto al elegir al azar y probar 10 marcapasos cardíacos. ¿Qué sabe usted acerca del número exacto de defectos si “al menos uno” de los 10 artículos es defectuoso?

2. **Probabilidad condicional.** Con sus palabras, describa la probabilidad condicional y dé un ejemplo.
3. **Cálculo de probabilidad.** Un investigador de mercados necesita calcular la probabilidad de que un comprador sea hombre, dado que utilizó una tarjeta de crédito para realizar una compra. El investigador razona que existen dos resultados (hombre, mujer), de manera que la probabilidad es $1/2$. ¿Está en lo correcto? ¿Qué información importante no está incluida en su proceso de razonamiento?
4. **Confusión del inverso.** ¿En qué consiste la confusión del inverso?

Descripción de complementos. En los ejercicios 5 a 8, realice una descripción escrita del complemento del suceso dado.

5. **Prueba de sangre.** Cuando a seis aspirantes a un empleo se les hace una prueba del consumo de marihuana, al menos uno resulta positivo.
6. **Control de calidad.** Cuando se envían 50 unidades de electrocardiógrafos, todas están libres de defectos.
7. **Trastorno relacionado con el cromosoma X.** Cuando a 12 hombres se les hace una prueba de un gen recesivo específico vinculado con el cromosoma X, los resultados revelan que ninguno de ellos posee el gen.
8. **Un éxito con las mujeres.** Cuando Brutus pide una cita a 12 mujeres diferentes, al menos una acepta.
9. **Probabilidad condicional subjetiva.** Utilice la probabilidad subjetiva para estimar la probabilidad de que una tarjeta de crédito sea utilizada de manera fraudulenta, dado que los cargos del día de hoy se realizaron en varios países diferentes.
10. **Probabilidad subjetiva.** Utilice la probabilidad subjetiva para estimar la probabilidad de seleccionar a un adulto al azar y obtener un varón, dado que la persona seleccionada posee una motocicleta. Si un investigador de crímenes encuentra que una motocicleta está registrada a nombre de Pat Ryan, ¿será razonable creer que Pat es hombre?
11. **Probabilidad de engendrar al menos una niña.** Si una pareja planea tener cuatro hijos, ¿cuál es la probabilidad de que tengan al menos una niña? ¿Esta probabilidad es lo suficientemente alta como para que la pareja tenga mucha confianza en que de sus cuatro hijos al menos uno sea niña?
12. **Probabilidad de concebir al menos una niña.** Si una pareja planea tener 10 hijos (puede suceder), ¿cuál es la probabilidad de que tengan al menos una niña? Si la pareja finalmente tuvo 10 hijos y todos fueron niños, ¿qué puede concluir la pareja?
13. **Probabilidad de una niña.** Calcule la probabilidad de que una pareja tenga una niña cuando nazca su tercer hijo, dado que los primeros dos hijos fueron niñas. ¿El resultado es igual a la probabilidad de que sus tres hijos sean niñas?
14. **Prueba de drogas.** Remítase a la tabla 4-1 y suponga que se elige al azar a uno de los 300 sujetos de prueba. Calcule la probabilidad de seleccionar a alguien que haya resultado positivo, dado que no consumió marihuana. ¿Por qué este caso es especialmente problemático para los sujetos de prueba?
15. **Prueba de drogas.** Remítase a la tabla 4-1 y suponga que se elige al azar a uno de los 300 sujetos de prueba. Calcule la probabilidad de seleccionar a alguien que haya resultado negativo, dado que no consumió marihuana.
16. **Prueba de drogas.** Remítase a la tabla 4-1 y suponga que se elige al azar a uno de los 300 sujetos de prueba. Calcule la probabilidad de seleccionar a alguien que no consumió marihuana, dado que resultó negativo en la prueba. Compare este resultado con el obtenido en el ejercicio 15.

- 17. Redundancia en relojes despertadores.** Un profesor de estadística quiere asegurarse de no llegar tarde a una clase temprano en la mañana como resultado del mal funcionamiento de un reloj despertador. En vez de usar un reloj despertador, decide usar tres. ¿Cuál es la probabilidad de que al menos uno de sus relojes despertadores funcione correctamente, si cada reloj despertador por separado tiene un 95% de probabilidad de funcionar correctamente? ¿Gana en realidad mucho el profesor al usar tres relojes despertadores en vez de uno solo?
- 18. Muestreo de aceptación.** Con un método del procedimiento llamado *muestreo de aceptación*, se selecciona al azar y sin reemplazo una muestra de artículos; el lote completo se rechaza si se encuentra al menos un defecto. La Medtyme Company acaba de fabricar 5000 medidores de presión sanguínea y el 4% de ellos están defectuosos. Si se seleccionan 3 y se prueban, ¿cuál es la probabilidad de que se rechace el lote completo?
- 19. Uso de muestras de sangre compuestas.** Cuando se hacen pruebas de sangre para detectar infecciones por VIH, el procedimiento puede hacerse de forma más eficiente y menos costosa mezclando muestras de sangre. Así, si las muestras de tres personas se combinan y la mezcla resulta negativa, sabemos que las tres muestras individuales son negativas. Calcule la probabilidad de un resultado positivo para tres muestras combinadas en una mezcla, suponiendo que la probabilidad de que una muestra de sangre individual dé resultado positivo es de 0.1 (la probabilidad de la población “en riesgo”, de acuerdo con datos del New York State Health Department).
- 20. Uso de muestras de agua compuestas.** El departamento de salud pública del condado de Orange realiza pruebas al agua para determinar contaminación por la presencia de la bacteria *E. coli* (*Escherichia coli*). Con la finalidad de reducir costos de laboratorio, se mezclan las muestras de agua de seis albercas públicas para realizar una sola prueba; sólo se hace una prueba posterior si la muestra mezclada falla. Con base en resultados previos, existe un 2% de probabilidad de encontrar la bacteria *E. coli* en una alberca pública. Calcule la probabilidad de que una muestra combinada de seis albercas públicas revele la presencia de la bacteria *E. coli*.

Probabilidades condicionales. En los ejercicios 21 a 24, utilice los siguientes datos sobre 100 senadores del 108º Congreso de Estados Unidos.

	Republicano	Demócrata	Independiente
Hombre	46	39	1
Mujer	5	9	0

21. Si seleccionamos al azar a un senador, ¿cuál es la probabilidad de elegir a un republicano, dado que seleccionamos a un hombre?
22. Si seleccionamos al azar a un senador, ¿cuál es la probabilidad de elegir a un hombre, dado que seleccionamos a un individuo republicano? ¿Es el mismo resultado obtenido en el ejercicio 21?
23. Si seleccionamos al azar a un senador, ¿cuál es la probabilidad de elegir a una mujer, dado que seleccionamos a un individuo independiente?
24. Si seleccionamos al azar a un senador, ¿cuál es la probabilidad de elegir un demócrata o un independiente, dado que seleccionamos a un hombre?

4-5 MÁS ALLÁ DE LO BÁSICO

- 25. Fecha de cumpleaños compartida.** Calcule la probabilidad de que, entre 25 personas seleccionadas al azar,
- no haya dos que compartan la misma fecha de cumpleaños.
 - al menos dos compartan la misma fecha de cumpleaños.



Probabilidad de un suceso que nunca ha ocurrido

Algunos sucesos son posibles, pero tan poco probables que nunca han ocurrido. He aquí un problema de este tipo, de gran interés para los científicos políticos: estime la probabilidad de que su voto determine el ganador de una elección presidencial de Estados Unidos. Andrew Gelman, Gary King y John Boscardin escribieron en el *Journal of the American Statistical Association* (vol. 93, núm. 441) que “el valor exacto de esta probabilidad es de interés menor, pero el número tiene implicaciones importantes para la comprensión de la ubicación óptima de los recursos de campaña, para saber si los estados y los grupos de votantes reciben su parte de atención de los candidatos presidenciales, y de qué manera los modelos formales de ‘elección racional’ del comportamiento del votante pueden ser capaces de explicar por qué las personas votan”. Los autores demuestran cómo se obtiene el valor de probabilidad de 1 en 10 millones para las elecciones cerradas.

26. **¿Quién fue?** La planta de Atlanta de la Medassist Pharmaceutical Company fabricó 400 marcapasos, de los cuales tres están defectuosos. La planta de Baltimore de la misma compañía fabricó 800 marcapasos, dos de los cuales están defectuosos. Si se selecciona al azar uno de los 1200 marcapasos y se encuentra que está defectuoso, ¿cuál es la probabilidad de que haya sido fabricado en Atlanta?
27. **Montaña rusa.** La montaña rusa Rock'n'Roller de los estudios Disney-MGM en Orlando tiene dos asientos en cada una de sus 12 filas. Los pasajeros se asignan a los asientos en el orden en que van llegando. Si usted se sube a esta montaña rusa una vez, ¿cuál es la probabilidad de obtener el codiciado lugar de la fila de adelante? ¿Cuántas veces se tiene que subir para tener un mínimo del 95% de probabilidad de que le toque el asiento delantero al menos una vez?
28. **Monedas ocultas.** Un profesor de estadística lanza dos monedas que ningún estudiante puede ver. Un estudiante pregunta si una de las monedas cayó en cara. Puesto que la respuesta del profesor es “sí”, calcule la probabilidad de que ambas monedas cayeran en cara.

Probabilidades por medio 4-6 de simulaciones

Concepto clave Hasta ahora en este capítulo hemos identificado varias reglas básicas e importantes, las cuales suelen usarse para calcular probabilidades; en esta sección presentamos un enfoque muy diferente, que puede vencer gran parte la dificultad encontrada en los métodos formales de las secciones anteriores de este capítulo. En vez de utilizar reglas formales para calcular probabilidades, el siguiente método alternativo consiste en realizar una simulación, en la que utilizamos un procedimiento hasta cierto punto diferente, el cual se comporta de la misma manera que el procedimiento que estamos considerando.

Definición

La **simulación** es un proceso que se comporta de la misma forma que el procedimiento sometido a consideración, de manera que produce resultados semejantes.

Considere los siguientes ejemplos para comprender mejor el uso de la simulación.

EJEMPLO Selección del género Cuando los investigadores médicos prueban técnicas de selección del género, necesitan conocer valores de probabilidad de diferentes resultados, como por ejemplo, la probabilidad de tener al menos 60 niñas entre 100 niños. Suponiendo que el nacimiento de un varón o de una niña es igualmente probable, describa una simulación que dé como resultado los géneros de 100 bebés recién nacidos.

SOLUCIÓN Una opción es simplemente lanzar una moneda al aire 100 veces; la cara representa a las niñas y la cruz a los varones. Otra opción es usar una calculadora o computadora para generar aleatoriamente ceros y unos (el 0 representa un niño y el 1 representa una niña). Los números deben generarse de forma que sean igualmente probables.

EJEMPLO Las mismas fechas de cumpleaños El ejercicio 25 de la sección 4-5 se refiere al clásico problema de la fecha de cumpleaños, en el que encontramos la probabilidad de que, en un grupo 25 personas seleccionadas al azar, al menos dos comparten la misma fecha de cumpleaños. La solución teórica es difícil, ya que resulta poco práctico encuestar a muchos grupos diferentes de 25 personas, por lo que desarrollamos una simulación. Describa una simulación que podría utilizarse para encontrar la probabilidad de que entre 25 personas seleccionadas al azar, al menos dos comparten la misma fecha de cumpleaños.

SOLUCIÓN Comience por representar fechas de cumpleaños con números enteros del 1 a 365, donde 1 = 1 de enero, 2 = 2 de enero, . . . , 365 = 31 de diciembre. Después, use una calculadora o un programa de cómputo para generar 25 números aleatorios entre 1 y 365. Estos números pueden ordenarse, ya que así será fácil estudiar la lista para determinar si 2 de las fechas de cumpleaños simuladas son iguales. Es posible repetir el proceso tantas veces como queramos hasta estar satisfechos de tener bases firmes para determinar la probabilidad. Nuestro estimado de la probabilidad es el número de veces que tuvimos al menos dos fechas de cumpleaños iguales, dividido entre el número total de grupos de 25 que se generaron.

Hay varias maneras de obtener números del 1 al 365 generados aleatoriamente, incluidas las siguientes:

- **Una tabla de números aleatorios:** Remítase, por ejemplo, al *CRC Standard Probability and Statistics Tables and Formulae*, que contiene una tabla de 14,000 dígitos. (Existen muchas formas de extraer números del 1 al 365 en este tipo de tablas. Una forma es tomar los dígitos en las primeras tres columnas, ignorando 000 y cualquier número mayor a 365).
- **STATDISK:** Seleccione **Data** de la barra del menú principal, luego seleccione **Uniform Generator** y proceda a introducir un tamaño muestral de 25, un mínimo de 1 y un máximo de 365; introduzca 0 para el número de lugares decimales. La pantalla que resulta en el STATDISK se muestra en la página siguiente. Use **copy/paste** y copie el conjunto de datos al **Sample Editor**, donde pueden ordenarse los valores. (Para ordenar los números, haga clic en **Data Tools** y seleccione la opción **Sort Data**). En la pantalla del STATDISK vemos que el 7º y 8º individuos tienen la misma fecha de cumpleaños: el sexagésimo octavo (68º) día del año.
- **Minitab:** Seleccione **Calc** en la barra del menú principal, después seleccione **Random Data** y después seleccione **Integer**. En el cuadro de diálogo, introduzca 25 para el número de renglones, guarde los resultados en la columna C1 e ingrese un mínimo de 1 y un máximo de 365. Entonces puede usar **Manip** y **Sort** para acomodar los datos en orden creciente. El resultado se presenta en la siguiente página, pero los números no serán los mismos. Este resultado del Minitab de 25 números indica que el 9º y el 10º son iguales.
- **Excel:** Haga clic en la celda que se encuentra en la esquina superior izquierda, después haga clic en el ícono de función **fx**. Seleccione **Math & Trig**, después seleccione **RANDBETWEEN**. En el cuadro de diálogo, escriba 1 para el límite inferior (*bottom*) y 365 como límite superior (*top*). Después de obtener el número aleatorio en la primera celda, haga clic, mantenga presionado el



Para ganar, apueste con audacia

El diario *New York Times* publicó una nota de Andrew Pollack en el que se reportó que el casino *Mirage* de Las Vegas tenía menores ganancias que las esperadas. El periodista afirmó que “las ganancias del *Mirage* resultan particularmente volátiles, ya que se favorece a los grandes apostadores, quienes pueden apostar \$100,000 o más en una mano de cartas. La ley de los promedios no funciona con tanta consistencia para unas cuantas apuestas grandes como lo hace para miles de pequeñas...”. Esto refleja el principio más fundamental al apostar: para ganar, ¡haga una apuesta grande en vez de muchas apuestas chicas! Con el juego adecuado, como puede ser el de los dados, usted tiene poco menos del 50% de posibilidades de duplicar su dinero si hace una apuesta grande. Al hacer muchas apuestas pequeñas, la probabilidad de duplicar su dinero disminuye sustancialmente.

STATDISK

Row	1 Ran...
1	7
2	8
3	16
4	38
5	42
6	46
7	68
8	68
9	104
10	117
11	140
12	195
13	204
14	244
15	271
16	274

Minitab

	C1	C2
1	38	
2	48	
3	59	
4	71	
5	101	
6	107	
7	122	
8	129	
9	153	
10	153	
11	163	

Excel

	A
1	15
2	3
3	15
4	362
5	164
6	184
7	158
8	59
9	143
10	85
11	134

TI-83/84 Plus

```
randInt(1,365,25
→L1
{79 206 340 133...
SortA(L1)      Done
L1
{17 34 46 70 79...
```

botón del ratón para arrastrar la esquina inferior derecha de la primera celda y continúe hacia abajo hasta resaltar 25 celdas. Cuando usted suelte el botón del ratón, deben aparecer los 25 números aleatorios. La pantalla que se reproduce aquí, muestra que el primero y el tercer número son iguales.

- **Calculadora TI-83/84 Plus:** Oprima la tecla **MATH**, seleccione **PRB**, luego escoja **randInt** y proceda a introducir el mínimo de 1, el máximo de 365 y 25 para el número de valores. Vea la pantalla de la TI-83/84 Plus, que indica que usamos **randInt** para generar los números, los cuales se guardaron en la lista L1, donde se ordenaron y mostraron. La imagen de la pantalla que se observa aquí indica que no hay números iguales entre los pocos que se pueden ver. Oprima **STAT** y seleccione **Edit** para ver la lista completa de números generados.

Es de suma importancia construir una simulación que se comporte precisamente como el procedimiento real. En el siguiente ejemplo demostramos la forma correcta y una forma incorrecta de construir una simulación.

EJEMPLO Simulación de dados Describa un procedimiento para simular el lanzamiento de un par de dados.

SOLUCIÓN En el procedimiento de tirar un par de dados, cada uno de los dos dados nos da un número entre 1 y 6 (inclusive); estos dos números se suman. Cualquier simulación debe hacer lo mismo. Hay una manera correcta y una incorrecta de simular un lanzamiento de dos dados.

La manera correcta: Generar aleatoriamente un número entre 1 y 6, generar aleatoriamente otro número entre 1 y 6, y luego sumar los dos resultados.

La manera incorrecta: Generar aleatoriamente números entre 2 y 12. Este procedimiento es similar a tirar dos dados en el sentido de que los resultados caen siempre entre 2 y 12, pero estos resultados entre 2 y 12 son igualmente probables. Con dados reales, los valores entre 2 y 12 *no son* igualmente probables. Esta simulación produciría muchos resultados confusos.

Algunos problemas de probabilidad se resuelven sólo por estimación de la probabilidad, utilizando observaciones reales o construyendo una simulación. La extensa disponibilidad de calculadoras y computadoras ha facilitado mucho el uso de métodos de simulación, tanto que ahora las simulaciones se usan con frecuencia para determinar valores de probabilidad.

4-6 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Simulación.** ¿Qué es una simulación? Si se utiliza un método de simulación para un problema de probabilidad, ¿el resultado es la respuesta correcta exacta?
- Simulación.** Si nacen tres bebés, pueden resultar 0 niñas, 1 niño, 2 niñas o 3 niñas. Un investigador simula tres nacimientos de la siguiente manera: en una tarjeta se anota el 0, en otra tarjeta se anota el 1, en otra se anota el 2 y en otra el 3. Luego, las cuatro tarjetas se mezclan en un recipiente y se elige una al azar. Si consideramos sólo el resultado del número de niñas en los tres nacimientos, ¿este proceso simula los tres nacimientos de forma similar a los nacimientos reales? ¿Por qué?
- Simulación.** Un estudiante desea simular 25 fechas de nacimiento, como se describió en esta sección, pero no dispone de una calculadora ni de un programa de cómputo, por lo que da 25 números entre 1 y 365. ¿Es correcto realizar la simulación de esta manera? ¿Por qué?
- Simulación.** Un estudiante desea simular tres nacimientos. Para ello anota “niño” en una tarjeta y “niña” en otra. El estudiante revuelve las tarjetas y luego selecciona una y anota el género. Luego, revuelve las tarjetas por segunda ocasión, elige una y registra el género. Por tercera vez revuelve las cartas, selecciona una y registra el género. ¿Se trata de un buen proceso para simular tres nacimientos?

En los ejercicios 5 a 8, describa el procedimiento de simulación. (Por ejemplo, para simular 10 nacimientos, utilice un generador de números aleatorios para obtener 10 enteros entre 0 y 1, inclusive, considerando que el 0 representa a un hombre y el 1 a una mujer).

- Simulación de un estudio de seguridad de motocicletas.** En un estudio sobre muertes causadas por accidente de motocicleta, se descubrió que el 95% de los conductores de motocicletas son hombres (según datos de “Motorcycle Rider Conspicuity and Crash Related Injury”, de Wells *et al.*, *BJM USA*). Describa un procedimiento para utilizar un programa de cómputo o la calculadora TI-83/84 Plus para simular la elección aleatoria de 20 conductores de motocicletas. Cada resultado individual debe indicar si el conductor es hombre o mujer.
- Simulación de hibridación.** Cuando Mendel realizó sus famosos experimentos de hibridación, utilizó plantas de guisantes con vainas verdes y amarillas. Un experimento incluyó la crusa de plantas de guisantes de tal forma que se esperaba que el 25% de los vástagos tuvieran vainas amarillas y que el 75% de los vástagos tuvieran vainas verdes. Describa un procedimiento para utilizar un programa de cómputo o una calculadora TI-83/84 Plus con el fin de simular 12 plantas de guisantes en un experimento de hibridación como éste.



Monos mecanógrafos

Una aseveración clásica dice que si un mono golpea al azar un teclado, tarde o temprano produciría las obras completas de Shakespeare, suponiendo que continuara tecleando un siglo tras otro. Se utilizó la regla de la multiplicación para probabilidades con la finalidad de obtener estimados de esta clase. Algunos consideran demasiado pequeño un resultado de 1,000,000,000, 000,000,000,000,000,000, 000,000 años. Con algo similar en mente, Sir Arthur Eddington escribió este poema: “Había una vez un sesudo babuino, que soplabía y soplabía un fagot. El babuino decía: ‘Estoy convencido de que, si sigo soplando, en miles de millones de años me saldrá una canción’”.

- 7. Simulación de fabricación.** Describa un procedimiento en el que se utilice un programa de cómputo o una calculadora TI-83/84 Plus para simular 500 teléfonos celulares manufacturados. Para cada teléfono celular, el resultado debe indicar si el teléfono celular funciona adecuadamente o tiene algún defecto. El proceso de fabricación tiene una tasa de defectos del 2%.
- 8. Simulación de zurdos.** El 15% de los hombres estadounidenses son zurdos (según datos de una encuesta de Scripps Survey Research Center). Describa un procedimiento en el que se utilice un programa de cómputo o una calculadora TI-83/84 Plus para simular la selección aleatoria de 200 hombres. Los resultados deben indicar si cada uno de los hombres es zurdo o no.

En los ejercicios 9 a 12, realice una simulación utilizando una calculadora TI-83/84 Plus, STATDISK, Minitab, Excel o cualquier otra calculadora o programa adecuado.

- 9. Simulación de un estudio de seguridad de motocicletas.** Remítase al ejercicio 5, en el que se le pidió la descripción de una simulación.
- Realice la simulación y registre el número de hombres conductores de motocicletas. Si es posible, obtenga una copia impresa de los resultados. ¿El porcentaje de hombres de la simulación se acerca razonablemente al valor de 95%?
 - Repita la simulación hasta que la realice un total de 10 veces. Registre el número de hombres en cada caso. Con base en los resultados, ¿el número de hombres es muy consistente? Con base en los resultados, ¿sería *infrecuente* seleccionar al azar a 20 conductores de motocicletas y descubrir que la mitad de ellos son mujeres?
- 10. Simulación de hibridación.** Remítase al ejercicio 6, en el que se le pidió la descripción de una simulación de hibridación.
- Realice la simulación y registre el número de vástagos de guisantes amarillos. Si es posible, obtenga una copia impresa de los resultados. ¿El porcentaje de vástagos de guisantes amarillos de la simulación se acerca razonablemente al valor de 25%?
 - Repita la simulación hasta que la realice un total de 10 veces. Registre el número de vástagos de guisantes con vainas amarillas en cada caso. Con base en los resultados, ¿el número de vástagos de guisantes con vainas amarillas es muy consistente? Con base en los resultados, ¿sería *infrecuente* seleccionar al azar 12 de los vástagos y encontrar que ninguno de ellos tiene vainas amarillas?
- 11. Simulación de fabricación.** Remítase al ejercicio 7, en el que se le pidió la descripción de una simulación de fabricación.
- Realice la simulación y registre el número de teléfonos celulares defectuosos. ¿El porcentaje de teléfonos celulares defectuosos de la simulación se acerca razonablemente al valor de 2%?
 - Repita la simulación hasta que la realice un total de 4 veces. Registre el número de teléfonos celulares defectuosos en cada caso. Con base en los resultados, ¿sería *infrecuente* seleccionar al azar 500 teléfonos celulares y encontrar que ninguno de ellos tiene defectos?
- 12. Simulación de zurdos.** Remítase al ejercicio 8, en el que se le pidió la descripción de una simulación.
- Realice la simulación y registre el número de hombres zurdos. ¿El porcentaje de hombres zurdos de la simulación se acerca razonablemente al valor de 15%?
 - Repita la simulación hasta que la realice un total de 5 veces. Registre el número de hombres zurdos en cada caso. Con base en los resultados, ¿sería *infrecuente* seleccionar al azar 200 hombres y encontrar que ninguno de ellos es zurdo?
- 13. Análisis de la eficacia de un fármaco.** Se ha descubierto que, cuando alguien trata de dejar de fumar en ciertas circunstancias, la tasa de éxito es del 20%. Existe un nuevo fármaco sustituto de la nicotina para ayudar a quienes desean dejar de fumar. En una prueba de 50 fumadores que utilizaron el fármaco cuando trataban de dejar de fumar,

se encontró que 12 de ellos lo lograron. El fabricante del fármaco argumenta que los 12 éxitos son mejores que los 10 que se esperarían sin el fármaco, por lo que éste es eficaz. Realice una simulación de 50 fumadores que tratan de abandonar el hábito y suponga que el fármaco no tiene efecto, por lo que la tasa de éxito continúa siendo del 20%. Repita la simulación varias veces y determine si podrían alcanzarse fácilmente 12 éxitos con un fármaco ineficaz. ¿Qué concluye acerca de la eficacia del fármaco?

14. **Análisis de la efectividad de un método de selección de género.** Para poner a prueba la efectividad de una técnica de selección de género, se realizó un ensayo con 20 parejas que deseaban tener una niña. De los 20 bebés que nacieron, 18 fueron niñas. Realice una simulación de 20 nacimientos, suponiendo que el método de selección de género no tiene efecto. Repita la simulación varias veces y determine si fácilmente podrían nacer 18 niñas con un método de selección de género ineficaz. ¿Qué concluye acerca de la efectividad del método?

4-6 MÁS ALLÁ DE LO BÁSICO

15. **Simulación del problema de Monty Hall.** Un problema que ha atraído gran atención en los últimos años es el *problema de Monty Hall*, inspirado en el antiguo programa de concurso de televisión “Let’s Make a Deal”, conducido por Monty Hall. Suponga que usted es un concursante que escogió una de tres puertas, después de saber que detrás de dos de ellas no hay nada, pero que detrás de una de las tres está un Corvette rojo último modelo. Luego, el presentador abre una de las puertas que usted no eligió y muestra que no hay nada detrás. Ahora le da la opción de quedarse con su primera selección o cambiar a otra puerta cerrada. ¿Le conviene quedarse con su primera elección o cambiarla? Realice una simulación de este juego y determine si debe mantener su decisión o modificarla. (De acuerdo con la revista *Chance*, las escuelas de negocios de instituciones como Harvard y Stanford usan este problema para ayudar a los estudiantes a familiarizarse con la toma de decisiones).

16. **Simulación de fechas de cumpleaños**

- Elabore una simulación para calcular la probabilidad de que, cuando se seleccionan 50 personas al azar, al menos dos tengan la misma fecha de cumpleaños. Describa la simulación y estime la probabilidad.
- Elabore una simulación para calcular la probabilidad de que, cuando se seleccionan 50 personas al azar, al menos tres tengan la misma fecha de cumpleaños. Describa la simulación y estime la probabilidad.

17. **Genética: simulación del control demográfico.** Un clásico problema de probabilidad cuenta la historia de un rey que quería incrementar la proporción de mujeres en su comarca decretando que, después de que una madre diera a luz a un hijo varón, se le prohibiera tener más hijos. El rey consideraba que algunas familias sólo tendrían un varón, mientras que otras familias tendrían pocas mujeres y un varón, de manera que la proporción de niñas se incrementaría. ¿Es correcto su razonamiento? ¿Se incrementará la proporción de niñas?

4-7 Conteo

Concepto clave En muchos problemas de probabilidad, el mayor obstáculo consiste en obtener el número total de resultados. En esta sección presentamos varios métodos diferentes para calcular tales números. Por ejemplo, la lotería Fantasy 5 de California implica la selección de cinco números (enteros) diferentes entre 1 y 39, inclusive. Como para ganar el premio mayor es necesario que usted elija los



Escasez de números telefónicos

Las compañías telefónicas con frecuencia dividen regiones que tienen un código de área en sectores con dos o más códigos de área, porque las nuevas líneas de fax e Internet están próximas a agotar los números que pueden suscribirse bajo un solo código. Puesto que no es posible que un número telefónico de siete dígitos comience con 0 o 1, hay $8 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 8,000,000$ de números telefónicos diferentes posibles.

Antes de los teléfonos celulares, equipos de fax e Internet, todos los números gratuitos tenían un prefijo de 800. Pasaron 29 años antes de que todos los números 800 fueran asignados. Se introdujo el prefijo 888 con el fin de contribuir a satisfacer la demanda de números gratuitos, pero se estimó que pasarían solo 2.5 años para que los números 888 se agotaran. Lo que sigue a futuro: números gratuitos con el prefijo 877. Las técnicas de conteo de esta sección se usan para determinar la cantidad de números gratuitos distintos posibles con un prefijo dado, ya que es necesario poder satisfacer las demandas del futuro.

cinco números que resultan cuando se lleva a cabo el sorteo, la probabilidad de ganar el premio mayor es 1 dividido entre el número de distintas formas posibles de seleccionar cinco números de 39. En esta sección se presentan métodos para calcular números de resultados, como el número de las distintas formas posibles de seleccionar cinco números entre 1 y 39.

Esta sección presenta diferentes métodos para el cálculo de números de distintos resultados posibles, sin hacer listas directamente y contar las posibilidades. Comencemos por la *regla fundamental de conteo*.

Regla fundamental de conteo

Para una secuencia de dos sucesos en la que el primero puede ocurrir de m formas y el segundo puede ocurrir de n formas, los sucesos juntos pueden ocurrir un total de $m \cdot n$ formas.

La regla fundamental de conteo se extiende fácilmente a situaciones que implican más de dos sucesos, como se ilustra en los siguientes ejemplos.

EJEMPLO Robo de identidad Es aconsejable no revelar los números del seguro social, ya que a menudo los criminales los utilizan para robar la identidad y disponer del dinero de otras personas. Suponga que descubre que un criminal está utilizando su número de seguro social, quien asegura que generó los números de manera aleatoria. ¿Cuál es la probabilidad de obtener su número de seguro social al generar aleatoriamente nueve dígitos? ¿Es probable que lo que afirma el criminal sea verdad?

SOLUCIÓN Cada uno de los 9 dígitos tiene 10 resultados posibles: 0, 1, 2, ..., 9. Si aplicamos la regla fundamental de conteo, obtenemos

$$10 \cdot 10 = 1,000,000,000$$

Sólo una de las 1,000,000,000 posibilidades corresponden a su número de seguro social, de manera que la probabilidad de generar al azar un número de seguridad social y obtener el suyo es de $1/1,000,000,000$. Es extremadamente improbable que un criminal genere su número de seguro social al azar, suponiendo que sólo genera uno. (Incluso si el criminal pudiera generar miles de números del seguro social y tratara de utilizarlos, es muy poco probable que generara el suyo). Si se descubre que alguien está utilizando su número de seguro social, lo más probable es que lo haya obtenido por algún otro medio, como espiando transacciones en Internet o buscando en su correo o en la basura.

EJEMPLO Cotinina en fumadores El conjunto de datos 4 del apéndice B lista niveles de cotinina medidos en una muestra de personas en tres grupos: fumadores (denotados aquí por F), no fumadores que están expuestos al humo del tabaco (denotados por E) y no fumadores que no están expuestos al humo del tabaco (denotados por N). Cuando el cuerpo absorbe la nicotina, se produce cotinina. Si calculamos la media del nivel de cotinina

de cada uno de los tres grupos y luego acomodamos esas medias en orden creciente, obtenemos la secuencia de sucesos NEF. Un cabildero en contra del tabaquismo afirma que esto es evidencia de que consumir tabaco daña la salud, porque la presencia de cotinina aumenta conforme la exposición al tabaco y su consumo se incrementan. ¿De cuántas formas pueden acomodarse los tres grupos que se denotan con N, E y F? Si se selecciona al azar un arreglo, ¿cuál es la probabilidad de obtener la secuencia NEF? La probabilidad es lo suficientemente baja como para concluir que la secuencia NEF indica que la presencia de cotinina aumenta a medida que la exposición al tabaco y su consumo también se incrementan?

SOLUCIÓN Al hacer arreglos de secuencias de los grupos N, E y F, hay tres posibles opciones para el primer grupo, dos opciones para el segundo grupo y sólo una opción para el tercer grupo. El número total de arreglos posibles es entonces

$$3 \cdot 2 \cdot 1 = 6$$

Existen seis maneras diferentes de acomodar los grupos N, E y F (que pueden listarse como NEF, NFE, EFN, ENF, FNE y FEN). Si seleccionamos al azar una de las seis secuencias posibles, la probabilidad de obtener la secuencia NEF es de 1/6. Puesto que la probabilidad de 1/6 (o 0.167) es relativamente alta, sabemos que la secuencia NEF puede ocurrir con facilidad por azar. La probabilidad no es suficientemente baja como para concluir que la secuencia NEF indique que la presencia de cotinina aumenta a medida que se incrementan la exposición al tabaco y su consumo. Necesitaríamos tener una probabilidad más baja; por ejemplo, de 0.01.

En el ejemplo anterior, encontramos que tres grupos pueden acomodarse en $3 \cdot 2 \cdot 1 = 6$ formas diferentes. Esta solución particular puede generalizarse utilizando la siguiente notación para el símbolo ! y la siguiente *regla factorial*.

Notación

El **símbolo factorial** ! denota el producto de números enteros positivos decrecientes. Por ejemplo, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$. Por definición especial, $0! = 1$.

Regla factorial

Una colección de n elementos distintos se puede acomodar de $n!$ diferentes maneras. (Esta **regla factorial** refleja el hecho de que el primer elemento se puede seleccionar de n maneras distintas, el segundo se puede seleccionar de $n - 1$ maneras, y así sucesivamente).

Los problemas de ruta con frecuencia implican la aplicación de la regla factorial. Verizon quiere hacer llamadas telefónicas a través de las redes más cortas. Federal Express quiere encontrar las rutas más cortas para sus entregas. American Airlines quiere encontrar la ruta más corta para regresar a los miembros de la tripulación a sus casas. Vea el siguiente ejemplo.



Ganar centavos de la lotería

Muchas personas gastan grandes cantidades de dinero comprando billetes de lotería al no tener un sentido realista de sus oportunidades de ganar. El hermano Donald Kelly del Colegio Marista propone esta analogía: ¡ganar la lotería es equivalente a recoger atinadamente el centavo “ganador” de una columna de centavos que tiene una altura de 21 millas! Los aviones comerciales por lo regular vuelan a una altitud de seis millas, así que trate de imaginar una columna de centavos de una altura de más del triple de la alcanzada por esos aviones a reacción e imagínese escogiendo el centavo de esa columna que representa un billete de lotería ganador. Usando los métodos de esta sección, calcule la probabilidad de ganar la lotería de su estado y luego determine la altura de la columna de centavos correspondiente.



Elección de códigos de seguridad

Todos utilizamos códigos de seguridad personales para tener acceso a cajeros automáticos, cuentas de Internet y sistemas de seguridad para casas. La seguridad de estos códigos depende del gran número de posibilidades diferentes, pero ahora los piratas informáticos cuentan con complejas herramientas que pueden superar este obstáculo con creces. Los investigadores encontraron que usando variaciones del nombre y apellidos del usuario, además de otros 1800 nombres, podrían identificar del 10 al 20% de las contraseñas de sistemas de cómputo típicos. Cuando escoja una contraseña, *no use* variaciones de ningún nombre, ni una palabra del diccionario, ni una secuencia con menos de siete caracteres, ni números telefónicos, ni números del seguro social. Incluya caracteres no alfabéticos, como números o símbolos de puntuación.

EJEMPLO Rutas de atracciones Usted está planeando un viaje a Disney World y desea disfrutar de las siguientes cinco atracciones el primer día: Space Mountain, Tower of Terror, Rock 'n' Roller Coaster, Mission Space y Dinosaur. A veces las atracciones requieren largos períodos de espera que varían en el transcurso del día, de manera que la planeación de una ruta eficiente permite aumentar al máximo la diversión. ¿Cuántas rutas diferentes posibles existen?

SOLUCIÓN Si aplicamos la regla factorial, sabemos que 5 atracciones diferentes se pueden ordenar de $5!$ maneras distintas. El número de rutas diferentes es $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.

El ejemplo anterior es una variación del problema clásico que se conoce como *problema del vendedor viajero*. Puesto que los problemas de rutas son tan importantes para muchas empresas, y como el número de las diversas rutas a menudo es considerable, existe un esfuerzo continuo por simplificar el método para encontrar las rutas más eficientes.

De acuerdo con la regla factorial, n diferentes elementos pueden acomodarse de $n!$ diferentes maneras. Algunas veces tenemos n elementos diferentes, pero necesitamos seleccionar sólo algunos de ellos en vez de todos. Por ejemplo, si debemos realizar encuestas en capitales estatales de Estados Unidos, pero sólo tenemos tiempo de visitar cuatro capitales, el número de posibles rutas diferentes es $50 \cdot 49 \cdot 48 \cdot 47 = 5,527,200$. Otra forma de obtener este mismo resultado es evaluar

$$\frac{50!}{46!} = 50 \cdot 49 \cdot 48 \cdot 47 = 5,527,200$$

Observe que en este cálculo, al dividir el número factorial del numerador entre el número factorial del denominador, sólo permanecen los factores de 50, 49, 48 y 47. Podemos generalizar este resultado observando que si tenemos n elementos disponibles diferentes y queremos seleccionar un número r de ellos, el número de combinaciones posibles es $n!/(n-r)!$ como en $50!/46!$ Esta generalización se conoce como *regla de las permutaciones*.

Regla de las permutaciones (cuando todos los elementos son diferentes)

Requisitos

1. Existen n elementos *diferentes* disponibles. (Esta regla no se aplica si algunos de los elementos son idénticos a otros).
2. Seleccionamos r de los n elementos (sin reemplazo).
3. Consideramos que los reordenamientos de los mismos elementos son secuencias diferentes. (La permutación de *ABC* difiere de la de *CBA* y se cuentan de forma separada).

Si se satisfacen los requisitos anteriores, el número de **permutaciones** (o secuencias) de r elementos seleccionados entre n elementos disponibles (sin reemplazo) es

$${}_nP_r = \frac{n!}{(n-r)!}$$

Cuando utilizamos los términos *permutaciones*, *acomodos* o *secuencias*, implicamos que *se toma en cuenta el orden*, en el sentido de que diferentes ordenamientos de los mismos elementos se cuentan por separado. Las letras *ABC* se pueden acomodar de seis formas distintas: *ABC, ACB, BAC, BCA, CAB, CBA*. (Más adelante nos referiremos a las *combinaciones*, en las que tales acomodos no se cuentan por separado). En el siguiente ejemplo se nos pide calcular el número total de secuencias distintas posibles. Eso sugiere el uso de la regla de las permutaciones.

EJEMPLO Prueba clínica de un nuevo fármaco Cuando se prueba un nuevo fármaco, la fase I incluye sólo a 8 voluntarios; el objetivo consiste en evaluar la seguridad del fármaco. Para ser muy cuidadoso, usted planea tratar a los 8 sujetos en secuencia, de manera que cualquier efecto dañino específico permita detener los tratamientos antes de aplicarlos a otros sujetos. Si se dispone de 10 voluntarios, de los cuales se seleccionarán 8, ¿cuántas secuencias diferentes de 8 sujetos son posibles?

SOLUCIÓN Tenemos $n = 10$ diferentes sujetos disponibles y planeamos elegir a $r = 8$ de ellos sin reemplazo. El número de arreglos diferentes se calcula de la siguiente manera:

$${}_nP_r = \frac{n!}{(n-r)!} = \frac{10!}{(10-8)!} = 1,814,400$$

Hay 1,814,400 arreglos posibles de 8 sujetos elegidos de los 10 que están disponibles. El tamaño de ese resultado indica que no es práctico hacer una lista de las secuencias o considerar cada una de manera individual.

En ocasiones necesitamos calcular el número de permutaciones, pero algunos de los elementos son idénticos a otros. La siguiente variación de la regla de las permutaciones se aplica a estos casos.

Regla de las permutaciones (cuando algunos elementos son idénticos a otros)

Requisitos

1. Existen n elementos disponibles, y algunos de ellos son idénticos a otros.
2. Seleccionamos todos los n elementos (sin reemplazo).
3. Consideramos que los reordenamientos de los mismos elementos son secuencias diferentes.

Si los requisitos anteriores se satisfacen y si existen n elementos con n_1 iguales, n_2 iguales, ..., n_k iguales, el número de permutaciones de los n elementos es

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$



¿Cuántas veces hay que barajar?

Después de realizar extensas investigaciones, el matemático de Harvard, Persi Diaconis encontró que se necesita barajar siete veces un mazo de naipes para obtener un mezclado completo. La mezcla es completa en el sentido de que todos los arreglos posibles de los naipes son igualmente probables. Barajar más de siete veces no tendrá un efecto significativo, y menos de siete no será suficiente. Los repartidores de naipes en los casinos rara vez barajan los mazos siete veces o más, así que los mazos no quedan totalmente mezclados. Algunos jugadores expertos han podido aprovechar las mezclas incompletas que resultan de barajar menos de siete veces.



La secretaria aleatoria

Un clásico problema de probabilidad dice así: una secretaria prepara 50 cartas distintas y las dirige a 50 personas diferentes, pero las revuelve al azar antes de meterlas en los sobres. ¿Qué probabilidad hay de que al menos una carta quede en el sobre que le corresponde? Aunque podría parecer que la probabilidad es pequeña, en realidad es de 0.632. Incluso con un millón de cartas y un millón de sobres, la probabilidad es de 0.632. La solución está más allá del alcance de este texto.

EJEMPLO Selección del género Los ejemplos clásicos de la regla de permutaciones son los que demuestran que las letras de la palabra *Mississippi* se pueden ordenar en 34,650 formas diferentes, mientras que las letras de la palabra *statistics* se pueden ordenar en 50,400 formas. Ahora consideraremos una aplicación diferente.

Al diseñar una prueba de un método de selección del género con 10 parejas, un investigador sabe que existen 1024 secuencias diferentes de género posibles cuando nacen 10 bebés. (Utilizando la regla fundamental del conteo, el número de posibilidades es $2 \cdot 2 = 1024$). Diez parejas utilizan un método de selección del género, y los 10 nacimientos incluyen 8 niñas y 2 niños.

- ¿De cuántas maneras se pueden acomodar en secuencia 8 niñas y 2 niños?
- ¿Cuál es la probabilidad de tener 8 niñas y 2 niños en 10 nacimientos?
- ¿La probabilidad del inciso b) sirve para evaluar la eficacia del método de selección del género?

SOLUCIÓN

- Tenemos $n = 10$ nacimientos, con $n_1 = 8$ iguales (niñas) y $n_2 = 2$ (niños) que son iguales. El número de permutaciones se calcula de la siguiente manera:

$$\frac{n!}{n_1!n_2!} = \frac{10!}{8!2!} = \frac{3,628,800}{80,640} = 45$$

- Puesto que hay 45 formas diferentes de ordenar a 8 niñas y 2 niños, y existe total de 1024 arreglos diferentes posibles, la probabilidad de 8 niñas y 2 varones está dada por $P(8 \text{ niñas y } 2 \text{ niños}) = 45/1024 = 0.0439$.
- La probabilidad de 0.0439 *no* es la probabilidad que se debe usar para evaluar la eficacia del método de selección del género. En vez de la probabilidad de que haya 8 niñas en 10 nacimientos, debemos considerar la probabilidad de 8 o más niñas en 10 nacimientos, que es 0.0547. (En la sección 5-2 aclararemos la razón del uso de la probabilidad de 8 o más niñas en vez de la probabilidad de *exactamente* 8 niñas en 10 nacimientos).

El ejemplo anterior considera n elementos, cada uno perteneciente a una de dos categorías. Cuando sólo hay dos categorías, podemos estipular que x de los elementos son iguales y que los otros $n - x$ elementos también son iguales, de manera que la fórmula de las permutaciones se simplifica a

$$\frac{n!}{(n-x)!x!}$$

Este resultado en particular se usará para el análisis de probabilidades binomiales, que se explica en la sección 5-3.

Cuando deseamos seleccionar r elementos a partir de n elementos diferentes *sin tomar en cuenta el orden*, lo que nos preocupa en realidad son las combinaciones posibles más que las permutaciones. Es decir, **cuando diferentes ordenamientos de los mismos elementos se cuentan por separado, tenemos un problema de permutaciones, pero cuando los diferentes ordenamientos de los mismos elementos no se cuentan por separado, tenemos un problema de combinaciones** y se aplica la siguiente regla:

Regla de las combinaciones

Requisitos

1. Tenemos un total de n elementos *diferentes* disponibles.
2. Seleccionamos r de los n elementos (sin reemplazo).
3. Consideramos que los reordenamientos de los mismos elementos son iguales. (La combinación ABC es igual que CBA).

Si se satisfacen los requisitos anteriores, el número de **combinaciones** de r elementos seleccionados a partir de n elementos diferentes es

$${}_nC_r = \frac{n!}{(n - r)!r!}$$

Como tal vez resulte confuso escoger entre la regla de las permutaciones y la regla de las combinaciones, damos el siguiente ejemplo con la intención de enfatizar la diferencia entre ellas.

EJEMPLO Fase I de una prueba clínica Cuando se prueba un nuevo fármaco en seres humanos, generalmente se realiza una prueba clínica en tres fases. La fase I se lleva a cabo con un número relativamente pequeño de voluntarios sanos. Suponga que deseamos tratar a 8 seres humanos saludables con un nuevo fármaco y que tenemos 10 sujetos voluntarios disponibles.

- a. Si los sujetos se seleccionan y se tratan *en secuencia*, de manera que la prueba se suspenda si alguno de ellos presenta cualquier reacción adversa, ¿cuántos arreglos secuenciales diferentes son posibles si se seleccionan 8 personas de las 10 disponibles?
- b. Si se seleccionan 8 sujetos de los 10 disponibles, y los 8 sujetos elegidos se tratan al mismo tiempo, ¿cuántos grupos de tratamiento diferentes son posibles?

SOLUCIÓN Observe que en el inciso *a*) el orden es importante porque los sujetos se tratan de manera secuencial y la prueba se suspende si alguno de ellos muestra alguna reacción adversa específica. Sin embargo, en el inciso *b*), el orden de selección es irrelevante porque todos los sujetos reciben tratamiento al mismo tiempo.

- a. Puesto que el orden es importante, buscamos el número de *permutaciones* de $r = 8$ personas seleccionadas de las $n = 10$ disponibles. En un ejemplo anterior de esta sección, encontramos que el número de permutaciones es de 1,814,400.
- b. Puesto que el orden *no* es importante, buscamos el número de *combinaciones* de $r = 8$ personas seleccionadas de las $n = 10$ disponibles. Obtenemos

$${}_nC_r = \frac{n!}{(n - r)!r!} = \frac{10!}{(10 - 8)!8!} = 45$$

Cuando se toma cuenta el orden, hay 1,814,400 permutaciones, pero cuando no se toma en cuenta el orden, existen 45 combinaciones.



Muy pocos códigos de barras

En 1974 un paquete de goma de mascar fue el primer artículo en ser escaneado en un supermercado. El escaneo requirió que la goma de mascar estuviera identificada con un código de barras. Los códigos de barras o códigos universales de productos se utilizan para identificar artículos individuales que serán adquiridos. Los códigos de barras usaban 12 dígitos que permitían a los escáneres listar y registrar automáticamente el precio de cada artículo comprado. El uso de 12 dígitos se volvió insuficiente al aumentar el número de productos diferentes, por lo que recientemente se modificaron para incluir 13 dígitos.

Problemas similares surgen cuando los códigos telefónicos de área se dividen porque existen demasiados teléfonos diferentes para un código de área en una región. Los métodos de conteo se utilizan para diseñar sistemas que permitan acomodar futuros números de unidades que deben procesarse o conectarse.

En esta sección presentamos las siguientes cinco herramientas para calcular el número total de resultados: la regla fundamental del conteo, la regla factorial, la regla de las permutaciones, la regla de las permutaciones cuando algunos elementos son idénticos y la regla de las combinaciones. No todos los problemas de conteo se pueden resolver con alguna de estas cinco reglas, pero dan un fundamento firme para muchas aplicaciones reales e importantes.

4-7 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Permutaciones y combinaciones.** ¿Cuál es la diferencia básica entre una situación que requiere la aplicación de la regla de las permutaciones y una que requiere la aplicación de la regla de las combinaciones?
- Conteo.** Al tratar de calcular la probabilidad de ganar la lotería Fantasy 5 de California, es necesario obtener el número de los distintos resultados que pueden ocurrir cuando se seleccionan 5 números entre 1 y 39. ¿Por qué no se puede calcular ese número al hacer una lista de todas las posibilidades?
- Frecuencia relativa.** Un investigador está analizando una muestra grande de texto para calcular la frecuencia relativa de la palabra “zip” entre las palabras de tres letras. Es decir, el investigador quiere estimar la probabilidad de encontrar la palabra “zip” cuando se selecciona una palabra de tres letras al azar de un texto típico en inglés. ¿Se puede calcular tal probabilidad utilizando los métodos de esta sección?
- Probabilidad.** Una persona piensa que cuando se lanza una moneda, existen tres posibles resultados: resulta una cara o resulta una cruz o la moneda cae de canto. Con tres resultados en cada lanzamiento, la regla fundamental del conteo sugiere que hay 9 posibilidades ($desde 3 \cdot 3 = 9$) en dos lanzamientos de una moneda. De esto se deduce que la probabilidad de obtener dos caras en dos lanzamientos es de $1/9$. ¿Es correcto este razonamiento? Si no es así, ¿cuál es el error?

Cálculo de factoriales, combinaciones y permutaciones. En los ejercicios 5 a 12, evalúe las expresiones dadas y exprese todos los resultados utilizando el formato acostumbrado para la escritura de números (en vez de la notación científica).

$$\begin{array}{llll} 5. 5! & 6. 8! & 7. {}_{24}C_4 & 8. {}_{24}P_4 \\ 9. {}_{52}P_2 & 10. {}_{52}C_2 & 11. {}_{30}C_3 & 12. {}_{10}P_3 \end{array}$$

Probabilidad de ganar la lotería. Puesto que la lotería Fantasy 5 de California se gana al seleccionar los cinco números correctos (en cualquier orden) entre 1 y 39, existen 575,757 combinaciones diferentes de cinco números que podrían elegirse, y la probabilidad de ganar esta lotería es de $1/575,757$. En los ejercicios 13 a 16, calcule la probabilidad de ganar la lotería indicada.

13. Massachusetts Mass Cash Lottery: Seleccione los cinco números ganadores entre 1, 2, ..., 35.
14. New York Lotto: Seleccione los seis números ganadores entre 1, 2, ..., 59.
15. Pennsylvania Lucky for Life Lotto: Seleccione los seis números ganadores entre 1, 2, ..., 38.
16. Texas Cash Five: Seleccione los cinco números ganadores entre 1, 2, ..., 37.
17. California Fantasy 5: Esta lotería se gana al seleccionar los cinco números correctos entre 1, 2, ..., 39. La probabilidad de ganar el juego es de $1/575,757$. ¿Cuál es la

probabilidad de ganar si se modifican las reglas y además de seleccionar los cinco números correctos, deben seleccionarse en el mismo orden en que salen?

18. **Nucleótidos de ADN.** El ADN (ácido desoxirribonucleico) está hecho de nucleótidos, y cada uno puede contener cualquiera de las siguientes bases de nitrógeno: A (adenina), G (guanina), C (citosina), T (timina). Si tenemos que elegir una de las cuatro bases (A, G, C, T) tres veces para formar un terceto lineal, ¿cuántos triples diferentes son posibles? Observe que se pueden seleccionar las cuatro bases para cada uno de los tres componentes del terceto.
19. **Discriminación por edad.** La empresa Cytertronics Communications Company redujo su personal de gerencia de 15 a 10 gerentes. La compañía afirmó que seleccionó a cinco gerentes al azar para despacharlos. Sin embargo, los cinco gerentes elegidos son los cinco gerentes de mayor edad entre los 15 contratados. Calcule la probabilidad de que, cuando se seleccionan cinco gerentes al azar de un grupo de 15, se seleccione a los cinco de mayor edad. ¿La probabilidad es lo suficientemente baja como para acusar a la empresa de que, en vez de usar una selección aleatoria, en realidad sólo despachó a los gerentes de mayor edad?
20. **Diseño de computadoras.** En el diseño de una computadora, si un *byte* se define como una secuencia de 8 bits y cada bit debe ser 0 o 1, ¿cuántos bytes diferentes son posibles? (Con frecuencia se usa un byte para representar un carácter individual, como una letra, un dígito o un símbolo de puntuación. Por ejemplo, cierto sistema de codificación representa la letra A como 01000001). ¿Existen suficientes bytes diferentes para los caracteres que usamos comúnmente, incluyendo letras minúsculas, letras mayúsculas, dígitos, símbolos de puntuación, signo de pesos y algunos otros?
21. **Experimento de crecimiento de árboles.** Al diseñar un experimento para estudiar el crecimiento de los árboles, se utilizaron los siguientes cuatro tratamientos: ninguno, sólo riego, sólo fertilización, riego y fertilización. Una fila de 10 árboles se extiende desde una zona húmeda hasta una área de tierra seca. Si se asigna uno de los tratamientos al azar a cada uno de los 10 árboles, ¿cuántos arreglos de tratamientos diferentes son posibles?
22. **Diseñar experimentos.** Al diseñar un experimento que implica un tratamiento aplicado a 12 sujetos de prueba, los investigadores planean utilizar una muestra aleatoria simple de 12 sujetos, elegidos de un grupo de 20 individuos disponibles. (Recuerde que en un muestreo aleatorio simple todas las muestras del mismo tamaño tienen la misma posibilidad de ser elegidas). ¿Cuántas muestras aleatorias simples diferentes son posibles? ¿Cuál es la probabilidad de cada muestra aleatoria simple en este caso?
23. **Probabilidad de píldoras defectuosas.** Un lote de píldoras consta de 7 aceptables y 3 defectuosas (porque contienen la cantidad incorrecta del fármaco).
 - a. ¿Cuántas permutaciones diferentes son posibles cuando se seleccionan al azar las 10 píldoras (sin reemplazo)?
 - b. Si se eligen al azar 3 píldoras sin reemplazo, calcule la probabilidad de seleccionar las tres píldoras defectuosas.
24. **Rutas aéreas.** Usted acaba de inaugurar su propia línea aérea llamada Air Me (Su lema: "Para nosotros, usted no es sólo otra estadística"). Hasta ahora, usted cuenta con un avión para una ruta que conecta Austin, Boise y Chicago. Una ruta es Austin-Boise-Chicago y una segunda ruta es Chicago-Boise-Austin. ¿Cuántas rutas diferentes son posibles si el servicio se expande para incluir un total de ocho ciudades?
25. **Prueba de una afirmación.** Mike afirma que ha desarrollado la habilidad de obtener un 6 casi siempre que tira un dado. Usted prueba su afirmación haciendo que Mike tire un dado cinco veces, y él obtiene el 6 cada vez. Si Mike no tiene posibilidad de afectar los resultados, calcule la probabilidad de que lance el dado cinco veces consecutivas y obtenga 6 en todas. ¿La probabilidad es lo bastante baja como para descartar el azar como explicación de los resultados de Mike?

- 26. Selección del género.** En una prueba de un método de selección del género, nacen 14 bebés y 10 de ellos son niñas.
- Calcule el número de diferentes secuencias de género posibles cuando nacen 14 bebés.
 - ¿En cuántas formas se pueden ordenar en secuencia 10 niñas y 4 niños?
 - Si se seleccionan al azar 14 bebés, ¿cuál es la probabilidad de que sean 10 niñas y 4 niños?
 - ¿Al parecer el método de selección del género produce un resultado significativamente diferente del que se esperaría debido al azar?
- 27. Elección en un consejo de directores.** En el Consejo de directores del Newport General Hospital hay 12 miembros.
- Si deben elegir a un presidente, un primer vicepresidente, un segundo vicepresidente y un secretario, ¿cuántas listas de candidatos diferentes son posibles?
 - Si deben conformar un subcomité de ética de cuatro miembros, ¿cuántos subcomités diferentes son posibles?
- 28. Sopa de letras.** Muchos periódicos incluyen una “sopa de letras”, un crucigrama en el que el lector debe descifrar letras para formar palabras. Por ejemplo, las letras TAISER se incluyeron en un periódico del día en que se escribió este ejercicio. ¿De cuántas formas se pueden acomodar las letras TAISER? Identifique la palabra codificada y luego determine la probabilidad de obtener este resultado seleccionando al azar un arreglo de las letras dadas.
- 29. Cálculo del número de melodías posibles.** En el *Directorio de melodías y temas musicales* de Dennys Parsons, se listan melodías de más de 14,000 canciones de acuerdo con el siguiente esquema: la primera nota de cada canción se representa por un asterisco (*) y las notas sucesivas se representan por R (para repetir la nota previa), S (para una nota que sube), o B (para una nota que baja). La quinta sinfonía de Beethoven comienza como *RRB. Se representan melodías clásicas a través de las primeras 16 notas. Con este esquema, ¿cuántas melodías clásicas diferentes son posibles?
- 30. Candados de combinación.** Un candado “de combinación” común se abre con la secuencia correcta de tres números entre 0 y 49, inclusive. (Un número puede usarse más de una vez). ¿Cuál es la probabilidad de adivinar estos tres números y abrir el candado en el primer intento?
- 31. Cálculo del número de códigos de área.** El reportero Paul Wiseman del diario *USA Today* describió las antiguas reglas para los códigos de área telefónicos con tres dígitos al escribir acerca de “códigos de área posibles con 1 o 0 en el segundo dígito. (Se excluyen los códigos terminados en 00 y 11, para llamadas gratuitas, servicios de emergencia y otros usos especiales)”. Los códigos que empiezan con 0 o 1 también deben excluirse. ¿Cuántos códigos de área distintos eran posibles bajo estas antiguas reglas?
- 32. Huevos rotos.** Una caja contiene 12 huevos, 3 de los cuales están rotos. Si seleccionamos al azar 5 de los huevos para cocerlos, ¿cuál es la probabilidad de los siguientes sucesos?
- Todos los huevos seleccionados están rotos.
 - Ninguno de los huevos seleccionados está roto.
 - Dos de los huevos seleccionados están rotos.
- 33. Torneo de básquetbol de la NCAA.** Cada año, 64 equipos universitarios de básquetbol compiten en el torneo de la NCAA. Recientemente, Sandbox.com ofreció un premio de \$10 millones a cualquiera que pudiera escoger al ganador en todos y cada uno de los juegos del torneo. El presidente de esa compañía también prometió que, además del premio en efectivo, él se comería una cubeta de gusanos. ¡Qué asco!
- ¿Cuántos juegos se requieren para obtener un equipo campeón en un campo de 64 equipos?

- b. Si alguien hace conjeturas al azar para cada juego del torneo, calcule la probabilidad de escoger al ganador de cada juego.
- c. En un artículo acerca del premio de \$10 millones, el *New York Times* publicó que “aún un experto en básquetbol colegial que pudiera escoger juegos con acierto en una porción del 70% tiene una probabilidad de 1 en _____ de elegir todos los juegos acertadamente”. Llene el espacio.
- 34. Cajero automático.** Usted quiere obtener dinero en efectivo en un cajero automático, pero está oscuro y no puede ver su tarjeta cuando la inserta. La tarjeta debe insertarse con la parte frontal hacia arriba y de manera que el inicio de su nombre ingrese primero.
- ¿Cuál es la probabilidad de seleccionar una posición aleatoria e insertar la tarjeta correctamente?
 - ¿Cuál es la probabilidad de seleccionar al azar la posición de la tarjeta y descubrir que la insertó incorrectamente en el primer intento, pero que la insertó correctamente en el segundo intento?
 - ¿Cuántas selecciones aleatorias se requieren para estar completamente seguro de que la tarjeta funciona bien porque se insertó correctamente?
- 35. Lotería de California.** En el juego de lotería Super Lotto Plus de California, ganar el premio mayor requiere que usted seleccione los cinco números correctos del 1 al 47, inclusive, y que, por separado, también seleccione un solo número correcto entre 1 y 27, inclusive. Calcule la probabilidad de ganar el premio mayor.
- 36. Power Ball Lottery.** La Power Ball Lottery se juega en 27 entidades de Estados Unidos. Para ganar el premio mayor de esta lotería es necesario seleccionar los cinco números correctos entre el 1 y el 53, inclusive, y que, por separado, también seleccione un solo número correcto entre 1 y 42, inclusive. Calcule la probabilidad de ganar el premio mayor.

4-7 MÁS ALLÁ DE LO BÁSICO

- 37. Cálculo del número de nombres de variables de cómputo.** Una regla común de programación de computadoras es que los nombres de las variables deben tener una longitud de 1 a 8 caracteres. El primer carácter puede ser cualquiera de las 26 letras, mientras que los caracteres sucesivos pueden ser cualquiera de las 26 letras o cualquiera de los 10 dígitos. Por ejemplo, A, BBB y M3477K son nombres permitidos de variables. ¿Cuántos nombres diferentes son posibles?
- 38. Saludos y mesas redondas**
- Cinco gerentes se reúnen para una junta. Si cada gerente saluda estrechando la mano a cada uno de los otros gerentes exactamente una vez, ¿cuál es el número total de saludos?
 - Si n gerentes saludan a cada uno de los otros exactamente una vez, ¿cuál es el número total de saludos?
 - De cuántas maneras diferentes se pueden sentar cinco gerentes en torno a una mesa redonda? (Suponga que si cada uno se mueve a la derecha el acomodo es el mismo).
 - De cuántas formas diferentes se pueden sentar n gerentes en torno a una mesa redonda?
- 39. Evaluación de factoriales grandes.** Muchas calculadoras o computadoras no pueden calcular directamente el número $70!$ o un número factorial mayor. Cuando n es grande, $n!$ puede aproximarse a $n = 10^k$, donde $K = (n + 0.5) \log n + 0.39908993 - 0.43429448n$.
- Usted ha sido contratado para visitar la capital de cada una de las 50 entidades de Estados Unidos. ¿Cuántas rutas diferentes son posibles? Evalúe la respuesta usando la tecla factorial de una calculadora y también usando la aproximación dada aquí.
 - El Bureau of Fisheries una vez pidió ayuda a los Laboratorios Bell con el fin de encontrar la ruta más corta para obtener muestras en 300 emplazamientos del Golfo de México. Si usted calcula el número de posibles rutas diferentes, ¿cuántos dígitos se necesitan para escribir el número?

- 40. Inteligencia artificial.** ¿Las computadoras pueden “pensar”? De acuerdo con la *prueba Turing*, se considera que una computadora piensa si, cuando una persona se comunica con ella, cree que se está comunicando con otra persona y no con una computadora. En un experimento en el museo de computadoras de Boston, cada uno de 10 jueces se comunicó con cuatro computadoras y cuatro personas; luego se les pidió que distinguieran entre unas y otras.
- Suponga que el primer juez no puede distinguir entre las cuatro computadoras y las cuatro personas. Si este juez hace conjeturas al azar, ¿cuál es la probabilidad de que identifique correctamente las cuatro computadoras y las cuatro personas?
 - Suponga que ninguno de los 10 jueces puede distinguir entre las computadoras y las personas, por lo que hacen conjeturas al azar. Con base en el resultado del inciso a), ¿cuál es la probabilidad de que los 10 jueces acierten en todas sus conjeturas? (Este suceso nos permitiría concluir que las computadoras no pueden “pensar” cuando, de acuerdo con la prueba Turing, sí pueden).
- 41. Cambio de un dólar.** ¿De cuántas maneras diferentes se puede obtener cambio de un dólar?

4-8 Teorema de Bayes (en CD-ROM)

El CD-ROM que complementa este libro incluye otra sección sobre la probabilidad condicional. Esta sección adicional estudia aplicaciones del *teorema de Bayes* (o *regla de Bayes*), que se utiliza para revisar un valor de probabilidad con base en información adicional que se obtiene posteriormente. Revise el análisis, los ejemplos y los ejercicios que describen las aplicaciones del teorema de Bayes en el CD-ROM.

Repaso

Iniciamos este capítulo con el concepto básico de probabilidad, el cual es de suma importancia para los métodos de estadística inferencial que se estudiarán más adelante en este libro. El concepto más importante que aprendimos en este capítulo es la regla del suceso infrecuente para la estadística inferencial: si, bajo una suposición dada, la probabilidad de un suceso en particular es muy pequeña, concluimos que probablemente la suposición no sea correcta. Como ejemplo del método básico utilizado, considere la prueba de un método de selección de género. Si realizamos una prueba de una técnica de selección del género y obtenemos 20 niñas en 20 nacimientos, podemos hacer una de dos inferencias a partir de estos resultados muestrales:

- La técnica de selección del género no es efectiva, y la serie de 20 niñas consecutivas es un suceso que puede ocurrir fácilmente debido al azar.
- La técnica de selección del género es efectiva (o existe alguna otra explicación de por qué los niños y las niñas no nacen con la misma frecuencia).

Los especialistas en estadística usan la regla del suceso infrecuente cuando deciden cuál inferencia es correcta: en este caso, la probabilidad de obtener 20 niñas consecutivas es tan pequeña ($1/1,048,576$) que la inferencia de una técnica efectiva de selección del género es la mejor opción. Aquí podemos ver el importante papel de la probabilidad en los métodos estándar de inferencia estadística.

En la sección 4-2 presentamos definiciones y notaciones básicas, incluyendo la representación de sucesos por letras como A. Debemos saber que un valor de probabilidad, que se expresa como un número entre 0 y 1, refleja la posibilidad de algún suceso. Definimos

las probabilidades de un suceso simple como

$$P(A) = \frac{\text{número de veces que ocurre A}}{\text{número de veces que se repite el experimento}} \quad (\text{frecuencia relativa})$$

$$P(A) = \frac{\text{número de formas en que puede ocurrir A}}{\text{número de eventos simples diferentes}} = \frac{s}{n} \quad (\text{para resultados igualmente probables})$$

Señalamos que la probabilidad de cualquier suceso imposible es 0, la probabilidad de cualquier suceso inevitable es 1, y que para cualquier suceso A , $0 \leq P(A) \leq 1$. También explicamos que \bar{A} denota el complemento del suceso A ; es decir, \bar{A} indica que el suceso A *no* ocurre.

En las secciones 4-3, 4-4 y 4-5 consideramos sucesos compuestos, los cuales son sucesos que combinan dos o más sucesos simples. Asociamos el uso de la palabra “o” con la suma y asociamos el uso de la palabra “y” con la multiplicación. Siempre tome en cuenta las siguientes consideraciones fundamentales:

- Cuando se realiza un ensayo, ¿queremos conocer la probabilidad del suceso A o B ? Si así es, use la regla de la suma, pero sea cuidadoso para no contar un resultado más de una sola vez.
- Cuando se busca la probabilidad de que el suceso A ocurra en un ensayo y el suceso B ocurra en un segundo ensayo, use la regla de la multiplicación. Multiplique la probabilidad del suceso A por la probabilidad del suceso B . *Precaución:* Cuando calcule la probabilidad del suceso B , asegúrese de tomar en cuenta el hecho de que el suceso A ya ocurrió.

En la sección 4-6 se describieron técnicas de simulación que a menudo son útiles para determinar valores de probabilidad, especialmente en situaciones donde las fórmulas o los cálculos teóricos son sumamente difíciles.

En algunos problemas de probabilidad, el mayor obstáculo es encontrar el número total de resultados posibles. La sección 4-7 se dedicó a las siguientes técnicas de conteo:

- Regla fundamental de conteo.
- Regla factorial.
- Regla de las permutaciones (cuando todos los elementos son diferentes).
- Regla de las permutaciones (cuando algunos elementos son iguales a otros).
- Regla de las combinaciones

Conocimientos estadísticos y pensamiento crítico

1. **Valor de probabilidad.** Un estudiante de estadística reporta que, al lanzar una moneda que no está cargada, la probabilidad de que salga una cara es de 50-50. ¿Qué es incorrecto en esta afirmación? ¿Cuál sería la afirmación correcta?
2. **Interpretación de un valor de probabilidad.** Investigadores médicos realizan la prueba clínica de un nuevo fármaco diseñado para bajar los niveles de colesterol. Ellos determinan que existe una probabilidad de 0.27 de que sus resultados ocurran por azar. Con base en ese valor de probabilidad, ¿se puede descartar al azar como una explicación razonable? ¿Por qué?
3. **Probabilidad de vida en Alfa Romeo.** Los astrónomos identifican un nuevo planeta en un sistema solar muy, muy lejano. Un astrónomo considera que en este planeta hay o no hay vida. Puesto que hay dos resultados (existe vida o no existe vida), concluye que la probabilidad de que haya vida en este planeta es de 1/2 o 0.5. ¿Es correcto su razonamiento? ¿Por qué?
4. **Sucesos disjuntos y sucesos independientes.** ¿A qué nos referimos cuando decimos que dos sucesos son disjuntos? ¿A qué nos referimos cuando decimos que dos sucesos son independientes?

Ejercicios de repaso

Prueba clínica de Lipitor. En los ejercicios 1 a 8, utilice los datos de la siguiente tabla (según datos de Parke-Davis). El fármaco Lipitor para disminuir los niveles de colesterol consiste en calcio atorvastatin.

	Tratamiento	
	10 mg Atorvastatin	Placebo
Con dolor de cabeza	15	65
Sin dolor de cabeza	17	3

- Si se selecciona al azar a uno de los 100 sujetos, calcule la probabilidad de elegir a alguien que tuvo dolor de cabeza.
- Si se selecciona al azar a uno de los 100 sujetos, calcule la probabilidad de elegir a alguien que fue tratado con 10 mg de atorvastatin.
- Si se selecciona al azar a uno de los 100 sujetos, calcule la probabilidad de elegir a alguien que tuvo dolor de cabeza o que fue tratado con 10 mg de atorvastatin.
- Si se selecciona al azar a uno de los 100 sujetos, calcule la probabilidad de elegir a alguien que recibió un placebo o que no tuvo dolor de cabeza.
- Si se seleccionan al azar dos sujetos diferentes, calcule la probabilidad de que ambos recibieran placebos.
- Si se seleccionan al azar dos sujetos diferentes, calcule la probabilidad de que ambos hayan tenido dolor de cabeza.
- Si se selecciona al azar a un sujeto, calcule la probabilidad de que haya tenido dolor de cabeza, dado que el sujeto fue tratado con 10 mg de atorvastatin.
- Si se selecciona al azar a un sujeto, calcule la probabilidad de que haya sido tratado con 10 mg de atorvastatin, dado que el sujeto tuvo dolor de cabeza.
- Día nacional de la estadística**
 - Si se selecciona a una persona al azar, calcule la probabilidad de que su cumpleaños sea el 18 de octubre, que es el día nacional de la estadística en Japón. Ignore los años bisiestos.
 - Si se selecciona a una persona al azar, calcule la probabilidad de que su cumpleaños sea en octubre. Ignore los años bisiestos.
 - Estime la probabilidad subjetiva del suceso de seleccionar a un adulto estadounidense y que éste sepa que el 18 de octubre es el día nacional de la estadística en Japón.
 - ¿Es infrecuente seleccionar al azar a un adulto estadounidense y que éste sepa que el 18 de octubre es el día nacional de la estadística en Japón?
- Encuesta de pastel de frutas.** En una encuesta realizada por Bruskin-Goldring Research, se pidió a los participantes que indicaran las formas de utilizar un pastel de frutas. De todos los sujetos, 132 indicaron que debe usarse como tope para una puerta, mientras que 880 expresaron otros usos, como comida para aves, relleno sanitario y un regalo. Si se elige al azar a uno de estos participantes, ¿cuál es la probabilidad de seleccionar a alguien que usaría el pastel de frutas como tope para una puerta?
- Prueba de una afirmación.** La empresa Biogene Research Company afirma que ha desarrollado una técnica para asegurar que un bebé sea una niña. En una prueba de la técnica, 12 parejas engendran niñas. Calcule la probabilidad de que nazcan 12 niñas al azar, suponiendo que los niños y las niñas son igualmente probables, y que el género

- de cualquiera de los bebés es independiente del de los otros. ¿El resultado sustenta la afirmación de la compañía?
12. **Seguro de vida.** La aseguradora New England Life Insurance Company expide pólizas de un año a 12 hombres de 27 años de edad. Con base en datos del Departamento de Salud y Servicios Humanos, cada uno de estos hombres tiene una probabilidad del 99.82% de sobrevivir todo el año. ¿Cuál es la probabilidad de que todos sobrevivan el año?
13. **Electrizante.** Al hacer una prueba de corriente eléctrica en un cable con cinco alambres codificados con colores, el autor utilizó un medidor para probar dos cables al mismo tiempo. ¿Cuál es la probabilidad de que los dos alambres vivos se localicen en la primera selección aleatoria de dos alambres?
14. **Muestreo de aceptación.** Con un método de muestreo de aceptación, se elige al azar una muestra de artículos sin reemplazo, y el lote completo se rechaza si hay al menos uno defectuoso. La Medtyme Pharmaceutical Company acaba de fabricar 2500 aspirinas en tableta, y el 2% de ellas están defectuosas porque contienen demasiada o muy poca aspirina. Si se seleccionan y prueban 4 tabletas, ¿cuál es la probabilidad de que todo el lote sea rechazado?
15. **Tasa de clamidia.** En un año reciente, se reportó que la tasa de prevalencia de clamidia era de 278.32 por cada 100,000 habitantes.
- Calcule la probabilidad de que una persona seleccionada al azar tenga clamidia.
 - Si se seleccionan dos personas al azar, calcule la probabilidad de que ambas tengan clamidia; exprese el resultado utilizando tres dígitos.
 - Si se seleccionan dos personas al azar, calcule la probabilidad de que ninguna de ellas tenga clamidia; exprese los resultados con siete decimales.
16. **Códigos de barras.** El 1 de enero de 2005 se modificaron los códigos de barras que aparecen en los productos de venta al detalle, de manera que ahora incluyen 13 dígitos en vez de 12. ¿Cuántos productos diferentes se pueden identificar ahora con los nuevos códigos de barras?

Ejercicios de repaso acumulativo

1. **Tratamiento del síndrome de fatiga crónica.** Una muestra de pacientes que padecen el síndrome de fatiga crónica fue tratada con medicamentos. Luego, el cambio en su nivel de fatiga se midió en una escala de -7 a $+7$, donde los valores positivos representaban una mejoría y el 0 representaba que no había habido ningún cambio. Los resultados se listan abajo (según datos de “The Relation Between Neurally Mediated Hypotension and the Chronic Fatigue Syndrome”, de Bou-Holaigah, Rowe, Kan y Calkins, *Journal of the American Medical Association*, vol. 274, núm. 12).

6 5 0 5 6 7 3 3 2 4 4 0 7 3 4 3 6 0 5 5 6

- Calcule la media.
- Calcule la mediana.
- Calcule la desviación estándar.
- Calcule la varianza.
- Con base en los resultados, ¿pareciera que el tratamiento fue efectivo?
- Si se selecciona aleatoriamente un valor de esta muestra, calcule la probabilidad de que sea positivo.
- Si se seleccionan aleatoriamente dos diferentes valores de esta muestra, calcule la probabilidad de que ambos sean positivos.
- Ignore los tres valores de 0 y suponga que sólo son posibles valores positivos o negativos. Suponga que el tratamiento no es efectivo y que los valores positivos y negativos son igualmente probables; calcule la probabilidad de que 18 sujetos

tengan valores positivos (como en este grupo muestral). ¿Esta probabilidad es lo bastante baja como para justificar el rechazo de la suposición de que el tratamiento no es efectivo? Al parecer, ¿el tratamiento es efectivo?

- 2. Temperaturas máximas.** Las temperaturas máximas reales (en grados Fahrenheit) durante septiembre se describen con el siguiente resumen de los cinco números: 62, 72, 76, 80, 85. (Los valores se basan en el conjunto 8 de datos del apéndice B). Utilice esos valores del resumen de los cinco números para responder lo siguiente:
- ¿Cuál es la mediana?
 - Si se calcula una temperatura máxima de algún día seleccionado al azar en el mes de septiembre, calcule la probabilidad de que oscile entre 72 y 76°F.
 - Si se tiene una temperatura máxima de algún día seleccionado al azar en el mes de septiembre, calcule la probabilidad de que sea menor que 72°F o mayor que 76°F.
 - Si se seleccionan al azar dos días diferentes de septiembre, calcule la probabilidad de que ambos días tengan temperaturas máximas entre 72 y 76°F.
 - Si se seleccionan al azar dos días *consecutivos* de septiembre, ¿los sucesos de obtener temperaturas máximas por arriba de 80°F son independientes? ¿Por qué?

Actividades de cooperación en equipo

- Actividad en clase** Forme equipos de tres o cuatro estudiantes y utilice lanzamientos de monedas para desarrollar una simulación que imite al reino que se atiene a este decreto: “Después de que una madre dé a luz a un varón, no tendrá ningún otro hijo”. Si este decreto se obedece, ¿se incrementará la proporción de mujeres?
- Actividad en clase** Forme grupos de tres o cuatro personas y use tachuelas reales para estimar la probabilidad de que cuando se dejen caer, una tachuela quede con la punta hacia arriba. ¿Cuántos intentos son necesarios para obtener un resultado que parezca ser razonablemente preciso, cuando se redondea al primer espacio decimal?
- Actividad fuera de clase** Los biólogos marinos con frecuencia usan el *método de captura-recaptura* como procedimiento para estimar el tamaño de una población, por ejemplo, el número de peces en un lago. Este método supone capturar una muestra de la población, etiquetar a cada uno de los miembros de la muestra y luego reincorporarlos a la población. Más tarde, se captura una segunda muestra y se cuentan los miembros etiquetados, junto con el tamaño total de esta segunda muestra. Los resultados se pueden utilizar para estimar el tamaño de la población.

En vez de capturar peces reales, simule el procedimiento utilizando un conjunto uniforme de artículos como botones, cuentas de colores, dulces M&M, piezas de cereal de aros de frutas o tarjetas. Comience con una colección grande de estos artículos. Obtenga una muestra de 50 y use un marcador para “etiquetar” a cada

uno. Remplace los artículos marcados, revuelva la población completa, seleccione una segunda muestra y proceda a estimar el tamaño de la población. Compare el resultado con el tamaño real de la población que se obtiene contando todos los artículos.

- Actividad en clase** Forme equipos de dos estudiantes. Remítase al ejercicio 15 en la sección 4-6 para tener una descripción del “problema Monty Hall”. Simule el concurso y registre los resultados de no cambiar y de cambiar la decisión, después determine cuál de estas dos estrategias es mejor.
- Actividad fuera de clase** Forme equipos de dos estudiantes con el propósito de hacer un experimento diseñado con el fin de mostrar un enfoque para el manejo de preguntas de encuesta delicadas, como aquellas relacionadas con el consumo de drogas, actividad sexual (o inactividad), robo o estafa. En vez de usar realmente una pregunta polémica que podría ocasionar ira contra el autor, usaremos esta inocua pregunta: “¿Nació usted en un mes que tiene la letra *r*?”. Alrededor de 2/3 de todas las respuestas deben ser “sí”, pero vamos a suponer que la pregunta toca un tema muy delicado y esos sujetos de encuesta son reticentes a contestar con honestidad. Realice la encuesta pidiendo a las personas que lancen una moneda al aire y respondan como sigue:
 - Responda “sí” si la moneda cae en cruz *o* usted nació en un mes que tiene la letra *r*.
 - Responda “no” si la moneda cae en cara *y* usted nació en un mes que no contiene la letra *r*.

continúa

Se presume que quienes contestan tienden a ser más honestos porque sienten que lanzar la moneda al aire protege su privacidad. Encueste personas y analice los resultados para determinar la proporción de personas que nacieron en un mes que contiene la letra *r*. La exactitud de los resultados se puede cotejar con las fechas

de nacimiento reales, que se pueden obtener con una segunda pregunta. El experimento puede repetirse con una pregunta que sea más delicada, pero aquí no se plantean preguntas de este tipo porque el autor ya recibe suficientes mensajes de correo electrónico.

Proyecto tecnológico

Uso de simulaciones para probabilidades y variación de manufactura

Por lo general, los estudiantes consideran que el tema de la probabilidad es el más difícil en un curso de introducción a la estadística. A veces algunos problemas de probabilidad parecen sencillos, pero sus soluciones son sumamente complejas. En este capítulo identificamos varias reglas básicas e importantes que suelen utilizarse para calcular probabilidades, pero en este proyecto utilizaremos un enfoque muy diferente que puede resolver gran parte de la dificultad que enfrentamos al aplicar las reglas formales. Este enfoque alternativo consiste en el desarrollo de una simulación, un proceso que se comporta de la misma forma que el procedimiento, de manera que produce resultados similares. (Véase la sección 4-6).

En el ejercicio 11 de la sección 4-6, nos referimos a un proceso de fabricación de teléfonos celulares. Supusimos que un lote consiste en 500 teléfonos celulares y que la tasa general de aparatos defectuosos es del 2%. Es posible realizar una simulación generando 500 números, cada uno de ellos entre 1 y 100, inclusive. Puesto que la tasa de defectos es del 2%, podemos considerar que cualquier resultado de 1 o 2 represente un teléfono celular defectuoso, en tanto que resultados de 3, 4, 5, . . . , 100, representen teléfonos celulares sin defectos. El número medio de defectos en los lotes de 500 debe ser 10. Sin embargo, algunos lotes tendrán exactamente 10 defectos, aunque otros lotes tendrán menos de 10 defectos y otros más de 10.

- a. Utilice un recurso tecnológico, como Minitab, Excel, STATDISK, SPSS, SAS o una calculadora TI-83/84

Plus para simular la fabricación de 500 teléfonos celulares. Registre el número de defectos en este lote simulado. [Sugerencia: Sería útil ordenar los resultados, de manera que los defectos (representados por los resultados de 1 o 2) puedan identificarse con facilidad].

- b. Repita el inciso a) 19 veces más, para generar un total de 20 lotes simulados. Haga una lista del número de defectos en cada uno de los 20 lotes.
- c. Utilice los resultados del inciso b) para estimar la probabilidad de que el número de defectos en un lote sea exactamente 10. ¿Cree usted que esta estimación es hasta cierto punto exacta? ¿Por qué?
- d. Utilice los resultados del inciso b) para estimar la probabilidad de que el número de defectos en un lote sea exactamente 9.
- e. Después de examinar los resultados del inciso b), ¿qué tanto varían los números de los defectos? ¿Los números de defectos en los lotes son hasta cierto punto predecibles o varían de forma considerable?
- f. Un ingeniero de control de calidad afirma que un nuevo proceso de fabricación reduce el número de defectos, y una prueba del nuevo proceso da por resultado un lote sin defectos. Con base en los resultados del inciso b), ¿la ausencia de defectos en un lote parece sugerir que el nuevo método es mejor? ¿O el azar puede ser una explicación razonable para la ausencia de teléfonos celulares defectuosos? Explique.

De los datos a la decisión

Pensamiento crítico: Si fuera médico, ¿qué debe decir a una mujer después de que se hizo una prueba de embarazo?

Para una mujer es vital saber si está embarazada para disminuir ciertas actividades, dejar de tomar medicamentos, evitar exponerse a tóxicos en el trabajo, dejar de fumar o de consumir alcohol, ya que todas estas situaciones pueden dañar al bebé. Las pruebas de embarazo, como casi todas las pruebas de salud, no producen resultados 100% precisos. En los ensayos clínicos de una prueba sanguínea

de embarazo, los resultados que se muestran en la siguiente tabla se obtuvieron por medio de la prueba sanguínea Abbot (de acuerdo con datos de “Specificity and Detection Limit of Ten pregnancy Tests”, por Tiitinen y Stenman, *Scandinavian Journal of Clinical Laboratory Investigation*, vol. 53, suplemento 216). Existen otras pruebas más confiables que las que se presentan en esta tabla.

Análisis de resultados

1. Con base en los resultados de la tabla, ¿cuál es la probabilidad de que una mujer esté embarazada si la prueba

indica un resultado negativo? Si usted fuera médico y tuviera una paciente con un resultado negativo, ¿qué consejo le daría?

2. Con base en los resultados de la tabla, ¿cuál es la probabilidad de un falso positivo? Es decir, ¿cuál es la probabilidad de obtener un resultado positivo si la mujer realmente no está embarazada? Si usted fuera médico y tuviera una paciente con un resultado positivo, ¿qué consejo le daría?

Resultados de pruebas de embarazo

	Resultado de prueba positivo (indica embarazo)	Resultado de prueba negativo (indica que no hay embarazo)
La mujer está embarazada	80	5
La mujer no está embarazada	3	11



Proyecto de Internet

Cálculo de probabilidades

Calcular las probabilidades cuando se tiran dados es fácil. Con un dado existen seis posibles resultados, y cada uno (por ejemplo, obtener un 2) tiene una probabilidad de 1/6. Un juego de naipes implica más cálculos, aunque éstos siguen siendo manejables. Pero, ¿qué pasa con un juego más complicado, como el juego de mesa Monopolio? ¿Cuál es la probabilidad de aterrizar en un lugar en particular del tablero? La probabilidad depende del lugar que su pieza ocupe en el momento, del resultado de los dados, de tomar cartas, así como también de otros factores. Ahora considere un ejemplo más representativo de la vida real, como el de la probabilidad de tener un accidente automovilístico. El número de factores implicados

es tan grande como para siquiera considerarlo; no obstante, probabilidades de este tipo suelen mencionarse en el caso de las compañías de seguros.

El proyecto de Internet para este capítulo considera métodos para calcular probabilidades en situaciones complicadas. Consulte el proyecto de Internet para este capítulo, que puede encontrar en el sitio:

<http://www.pearsoneducation.net/triola>

Se le guiará en la investigación de probabilidades para un juego de mesa. Después calcule usted mismo las probabilidades. Finalmente, usted efectuará una estimación de una probabilidad relacionada con la salud utilizando datos empíricos.

La estadística en el trabajo

“Debemos tener un conocimiento sólido de la teoría estadística, una buena comprensión de los diseños experimentales . . . y saber cómo se aplica el razonamiento estadístico a las diversas etapas del proceso del desarrollo de un fármaco”.



Christy Chuang-Stein

Directora Senior de Pfizer Inc.

Christy es consultora estadística en un grupo llamado Statistical Research and Consulting Center (SRCC). Todos los miembros de SRCC trabajan para brindar asesoría estratégica y táctica en temas relacionados con políticas y aplicaciones estadísticas dentro de Pfizer. Además, los miembros de SRCC realizan investigación independiente y en colaboración sobre problemas que cubren las necesidades de negocios de la empresa.

¿Cómo utiliza la estadística en su trabajo y qué conceptos estadísticos específicos emplea?

Al trabajar en una empresa farmacéutica utilizamos ampliamente la estadística para apoyar el descubrimiento y desarrollo de nuevos productos médicos. Esto incluye identificar compuestos prometedores, probar los compuestos, investigar la seguridad y eficacia de productos potenciales en ensayos clínicos y fabricar los productos de acuerdo con especificaciones predeterminadas. Algunos de los conceptos estadísticos que utilizamos son el muestreo, la variabilidad, la eficiencia, el control del sesgo, la reducción de fuentes de variabilidad, la estimación de parámetros y la prueba de hipótesis.

Por favor, describa un ejemplo específico en el que el uso de la estadística haya servido para mejorar un producto o servicio.

La tasa de agotamiento de los compuestos en la industria farmacéutica es extremadamente alta. Menos del 12% de los compuestos que llegan a la fase de prueba en seres humanos salen finalmente al mercado. Ante la elevada tasa de agotamiento y los costos tan altos del desarrollo de un producto farmacéutico, un importante factor de éxito consiste en tomar buenas decisiones de aceptación o rechazo con respecto a los productos potenciales, y hacerlo lo más pronto posible. Hemos utilizado con éxito diseños de grupo secuenciales para completar las pruebas en menos tiempo, cuando los ensayos clínicos tienen pocas probabilidades de cumplir sus objetivos, incluso si siguieran adelante. El hecho de detener las pruebas a tiempo nos ha permitido ahorrar recursos que después podemos asignar para el desarrollo de otros productos farmacéuticos prometedores.

¿Qué aspectos de su trabajo considera emocionantes, interesantes o gratificantes?

El aspecto más gratificante de mi trabajo es saber que al lanzar al mercado medicamentos nuevos e innovadores estamos ayudando a que millones de personas tengan una vida más larga y con mayor calidad. Durante los últimos 50 años, se ha demostrado con claridad el valor de la medicina en aplicaciones como el combate a la diabetes, las enfermedades cardíacas, la osteoporosis, el cáncer, las infecciones por VIH, la esquizofrenia y la epilepsia, así como en procesos de inmunización mediante las vacunas infantiles.

¿Considera que su empresa ve con mejores ojos a los aspirantes que cuentan con estudios de estadística?

Esto depende del empleo al que aspire el individuo. Sin embargo, puesto que los principios estadísticos se pueden aplicar a un gran número de funciones no estadísticas en una empresa farmacéutica (como la evaluación de un portafolios, la administración y mejora de proyectos, el estudio y manejo de datos, la búsqueda de sistemas de medición), las personas que tienen cierta capacitación en estadística a menudo destacan en trabajos que requieren habilidades cuantitativas y razonamiento deductivo. Como consecuencia, creo que los aspirantes que comprenden aspectos básicos de estadística tendrán una imagen más favorable en muchas áreas dentro de la empresa.

¿Usted recomendaría a los estudiantes universitarios de hoy que estudien estadística? ¿Por qué?

Definitivamente, recomiendo el estudio de la estadística a los estudiantes que desean trabajar en un ambiente que incluya actividades de investigación y desarrollo. Es sorprendente el gran uso que tienen conceptos básicos como población, muestra, variabilidad, sesgo y estimación incluso en un entorno laboral promedio.



Distribuciones de probabilidad discreta

5



- 5-1 Panorama general**
- 5-2 Variables aleatorias**
- 5-3 Distribuciones de probabilidad binomial**
- 5-4 Media, varianza y desviación estándar para la distribución binomial**
- 5-5 Distribuciones de probabilidad de Poisson**

¿Los métodos estadísticos pueden demostrar que el proceso de selección de un jurado es discriminatorio?

Después de que un acusado ha sido condenado por algún crimen, en ocasiones se interpone una apelación con el argumento de que el acusado fue condenado por un jurado de personas diferentes a él. Uno de los criterios es que el proceso de selección del jurado debe garantizar que los miembros representen a la población de la región. En un caso célebre, el doctor Benjamin Spock, escritor del libro *Baby and Child Care*, fue condenado por conspiración al fomentar la resistencia al reclutamiento durante la guerra de Vietnam. Su defensor argumentó que el doctor Spock estaba en desventaja, pues los 12 miembros del jurado eran hombres. Las mujeres se habrían mostrado más comprensivas, ya que, en general, se oponían más a la guerra; además, el doctor Spock era muy reconocido entre el público femenino como médico infantil. Un especialista en estadística testificó que el jurado tenía una proporción consistentemente menor de mujeres que los otros seis jurados del mismo distrito. La condena del doctor Spock fue anulada por otras razones. En la actualidad los integrantes de los jurados de las cortes federales se deben elegir de manera aleatoria.

En 1972, Rodrigo Partida, mexicano-estadounidense, fue condenado por robo con intento de violación. Su condena fue dictada en el condado de Hidalgo, que se localiza en Texas, en la frontera con México. En el condado de Hidalgo había 181,535 personas que podían formar parte del jurado, y el 80% de ellas eran mexicano-estadounidenses. (Como el autor acaba de renovar su licencia poética, en este capítulo usará la cifra 80% en vez del valor más exacto del 79.1%). De las 870 personas llamadas a servir como jueces, el 39% (339) eran mexicano-estadounidenses. Tiempo después, se apeló la condena de Partida (*Castaneda contra Partida*) con base en la gran discrepancia entre el 80% de mexicano-estadounidenses disponi-

bles para fungir como jueces y el hecho de que sólo fuera seleccionado el 39% de este grupo.

En este capítulo analizaremos el problema de *Castaneda contra Partida*, en especial a partir las siguientes preguntas fundamentales::

1. Puesto que los mexicano-estadounidenses constituyen el 80% de la población y dado que Partida fue sentenciado por un jurado de 12 personas, de las que sólo el 58% (7 jueces) eran mexicano-estadounidenses, ¿podemos concluir que este jurado fue elegido en un proceso que discrimina a los mexicano-estadounidenses?
2. Dado que los mexicano-estadounidenses constituyen el 80% de la población total de 181,535 habitantes y que durante un periodo de más de 11 años sólo el 39% de los individuos llamados a servir como jueces eran mexicano-estadounidenses, ¿podemos concluir que el proceso de selección del jurado discriminó a este grupo? (Sabemos que, debido al azar, las muestras varían naturalmente hasta cierto punto de lo que se esperaría a nivel teórico. Sin embargo, ¿la discrepancia entre la tasa del 80% de mexicano-estadounidenses en la población y la tasa del 39% de mexicano-estadounidenses llamados a servir como jueces es lo suficientemente grande para explicarse por el azar?)

Este ejemplo ilustra la importancia de una comprensión básica de los métodos estadísticos en el terreno legal. Es probable que los abogados que carecen de conocimientos estadísticos no puedan ofrecer un buen servicio a sus clientes. Una ocasión el autor testificó en la Suprema Corte del estado de Nueva York y, al analizar la situación, se dio cuenta de que la falta de comprensión de conceptos estadísticos básicos puede ser muy perjudicial para el cliente de un abogado.

5-1 Panorama general

En este capítulo combinamos los métodos de *estadística descriptiva* presentados en los capítulos 2 y 3 con los de *probabilidad* que estudiamos en capítulo 4. La figura 5-1 presenta un resumen visual de los objetivos de este capítulo. Como se observa en la figura, utilizando los métodos de los capítulos 2 y 3, podríamos lanzar en repetidas ocasiones un dado para reunir datos muestrales y luego describirlos con gráficas (como un histograma o una gráfica de cuadro), medidas de tendencia central (como la media) y medidas de variación (como la desviación estándar). Si empleamos los métodos del capítulo 4 podríamos calcular la probabilidad de cada resultado. En este capítulo combinaremos esos conceptos mientras creamos distribuciones de probabilidad que describan lo que *probablemente* sucederá, en vez de lo que en realidad *sucedió*. En el capítulo 2 elaboramos tablas de frecuencias e histogramas utilizando valores muestrales *observados* que se reunieron en realidad; en este capítulo construiremos distribuciones de probabilidad presentando los resultados posibles junto con las frecuencias relativas que *esperamos*. En este capítulo estudiaremos las distribuciones de probabilidad *discretas* y en el capítulo 6 las distribuciones de probabilidad *continuas*.

La tabla que se encuentra en el extremo derecho de la figura 5-1 representa una distribución de probabilidad que sirve como modelo para una distribución de frecuencias poblacional teóricamente perfecta. En esencia, podemos describir la tabla de frecuencias relativas para un dado que se lanzó un número infinito de veces. Con este conocimiento de los resultados de la población, somos capaces de calcular sus características importantes, como la media y la desviación estándar. El resto de este libro y la esencia de la estadística inferencial se basan en el conocimiento de las distribuciones de probabilidad. Comenzamos examinando el concepto de una variable aleatoria y después estudiaremos distribuciones importantes que tienen muchas aplicaciones reales.

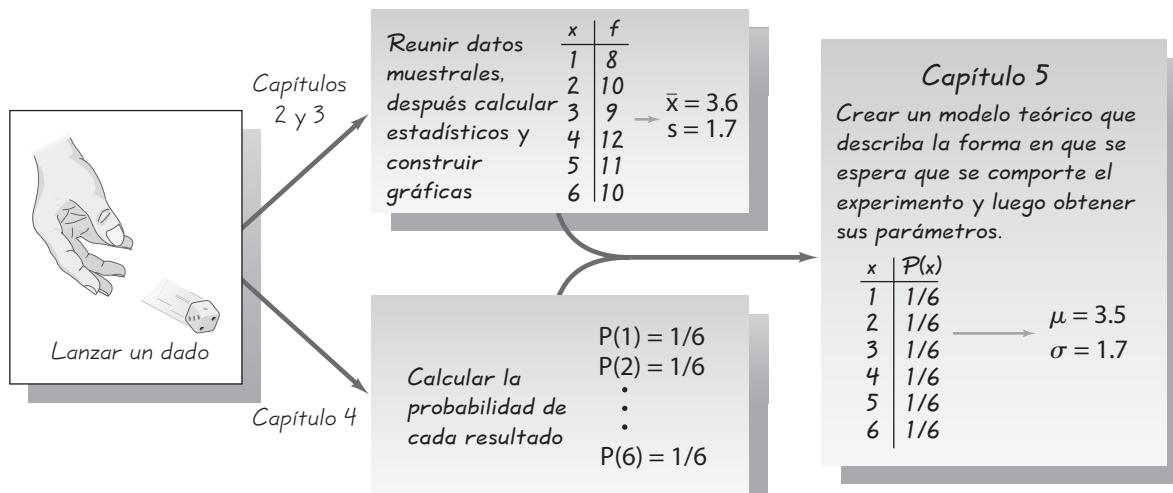


Figura 5-1 Combinación de métodos descriptivos y probabilidades para formar un modelo teórico de comportamiento



5-2 Variables aleatorias

Concepto clave En esta sección se presenta el importante concepto de una distribución de probabilidad, que indica la probabilidad de cada valor de una variable que está determinada por el azar. Por otro lado, incluye procedimientos para el cálculo de la media y de la desviación estándar para una distribución de probabilidad. Además del concepto de una distribución de probabilidad, se debe poner especial atención a los métodos que se utilizan para distinguir entre resultados que pueden ocurrir por azar y los resultados que son “poco comunes” en el sentido de que no tienen probabilidad de ocurrir por azar.

Iniciamos con los conceptos relacionados *variable aleatoria* y *distribución de probabilidad*.

Definiciones

Una **variable aleatoria** es aquella (casi siempre representada por x) que tiene un solo valor numérico determinado por el azar, para cada resultado de un procedimiento.

Una **distribución de probabilidad** es una distribución que indica la probabilidad de cada valor de la variable aleatoria. A menudo se expresa como gráfica, tabla o fórmula.



EJEMPLO Selección de miembros del jurado Se elegirá al azar a 12 integrantes del jurado de una población en la que el 80% de los habitantes son mexico-estadounidenses. Si suponemos que los miembros del jurado se seleccionan al azar sin sesgo y si permitimos que x = número de mexico-estadounidenses en un total de 12 miembros del jurado entonces x es una variable aleatoria porque su valor depende del azar. Los valores posibles de x son 0, 1, 2, ..., 12. La tabla 5-1 incluye los valores de x , junto con las probabilidades correspondientes. Los valores de probabilidad que son muy pequeños, como 0.000000123, están representados por 0+. (En la sección 5-3 aprenderemos a calcular los valores de probabilidad, como los que se listan en la tabla 5-1). Puesto que la tabla 5-1 incluye la probabilidad para cada valor de la variable aleatoria x , esa tabla describe una distribución de probabilidad.

En la sección 1-2 hicimos una distinción entre los datos discretos y continuos. Las variables aleatorias también pueden ser discretas o continuas, y las siguientes dos definiciones son consistentes con las que se presentan en la sección 1-2.

Definiciones

Una **variable aleatoria discreta** tiene un número finito de valores o un número de valores contable, donde “contable” se refiere al hecho de que podría haber un número infinito de valores, pero que pueden asociarse con un proceso de conteo.

Una **variable aleatoria continua** tiene un número infinito de valores, y esos valores pueden asociarse con mediciones en una escala continua, de manera que no existan huecos o interrupciones.

Tabla 5-1

Distribución de probabilidad: probabilidades de números de mexico-estadounidenses en un jurado de 12 miembros, suponiendo que los miembros se seleccionan al azar de una población en la que el 80% de los habitantes son mexico-estadounidenses

x (mexico-estadounidenses)	$P(x)$
0	0+
1	0+
2	0+
3	0+
4	0.001
5	0.003
6	0.016
7	0.053
8	0.133
9	0.236
10	0.283
11	0.206
12	0.069



Elección de números de lotería

En una lotería estatal tradicional, usted selecciona seis números diferentes. Después de una selección aleatoria, los boletos con la combinación correcta comparten el premio. Como los números ganadores se seleccionan al azar, cualquier elección de seis números tendrá la misma probabilidad que otra, pero algunas combinaciones son mejores que otras. La combinación de 1, 2, 3, 4, 5, 6 es una mala elección, ya que muchas personas tienden a seleccionarla. En una lotería de Florida, con un premio de \$105 millones, 52,000 boletos incluían 1, 2, 3, 4, 5, 6; si esta combinación hubiera ganado, el premio hubiera sido de tan sólo \$1000. Es más sensato elegir combinaciones que no seleccionen muchas otras personas. Evite combinaciones que formen un patrón en el boleto.



a) Variable aleatoria discreta: contador del número de asistentes al cine.



b) Variable aleatoria continua: voltaje medido de una batería de un detector de humo.

Figura 5-2 Aparatos que se utilizan para contar y medir variables aleatorias discretas y continuas

Este capítulo se refiere exclusivamente a variables aleatorias discretas, pero en los siguientes capítulos se estudiarán las variables aleatorias continuas.

EJEMPLO Los siguientes son ejemplos de variables aleatorias discretas y continuas:

1. Sea $x =$ número de huevos que una gallina pone en un día. Ésta es una variable aleatoria *discreta* porque sus únicos valores posibles son 0 o 1 o 2, etcétera. Ninguna gallina puede poner 2.343115 huevos, lo que sería posible si los datos provinieran de una escala continua.
2. El conteo del número de estudiantes de estadística que asisten a una clase es un número entero y , por lo tanto, una variable aleatoria discreta. El aparato de conteo que se muestra en la figura 5-2a) es capaz de indicar únicamente un número finito de valores, por lo que se utiliza para obtener valores de una variable aleatoria *discreta*.
3. Sea $x =$ cantidad de leche que produce una vaca en un día. Ésta es una variable aleatoria *continua*, ya que puede tomar cualquier valor en un tramo continuo. En un solo día, una vaca produce una cantidad de leche cuyo valor puede ser cualquiera entre 0 galones y 5 galones. Es posible obtener 4.123456 galones, ya que la vaca no está restringida a las cantidades discretas de 0, 1, 2, 3, 4 o 5 galones.

4. La medida del voltaje de una batería de un detector de humo puede ser cualquier valor entre 0 y 9 volts. Por lo tanto, se trata de una variable aleatoria continua. El voltímetro que se ilustra en la figura 5-2b) indica valores en una escala continua, de manera que permite obtener valores de una variable aleatoria *continua*.

Gráficas

Existen varias formas para graficar una distribución de probabilidad, pero aquí consideraremos solamente al **histograma de probabilidad**. La figura 5-3 es un histograma de probabilidad muy similar al histograma de frecuencias relativas estudiado en el capítulo 2, pero la escala vertical indica *probabilidades* en vez de frecuencias relativas basadas en resultados muestrales reales.

En la figura 5-3, observe que a lo largo del eje horizontal, los valores de 0, 1, 2, . . . , 12 se localizan en el centro de los rectángulos. Esto implica que cada uno de los rectángulos mide una unidad, de manera que las áreas de los rectángulos son 0+, 0+, 0+, 0+, 0.001, 0.003, . . . , 0.069. Las áreas de estos rectángulos son iguales a las *probabilidades* en la tabla 5-1. En el capítulo 6 y en capítulos posteriores veremos que esta correspondencia entre el área y la probabilidad es muy útil en estadística.

Toda distribución de probabilidad debe satisfacer cada uno de los dos siguientes requisitos.

Requisitos de una distribución de probabilidad

1. $\sum P(x) = 1$ donde x asume todos los valores posibles. (Es decir, la suma de todas las probabilidades debe ser 1).
2. $0 \leq P(x) \leq 1$ para cada valor individual de x . (Es decir, cada valor de probabilidad debe ubicarse entre 0 y 1, inclusive).

El primer requisito surge del simple hecho de que la variable aleatoria x representa todos los sucesos posibles en el espacio muestral completo, de manera que tenemos la certeza (con probabilidad 1) de que uno de los sucesos ocurrirá. (En la

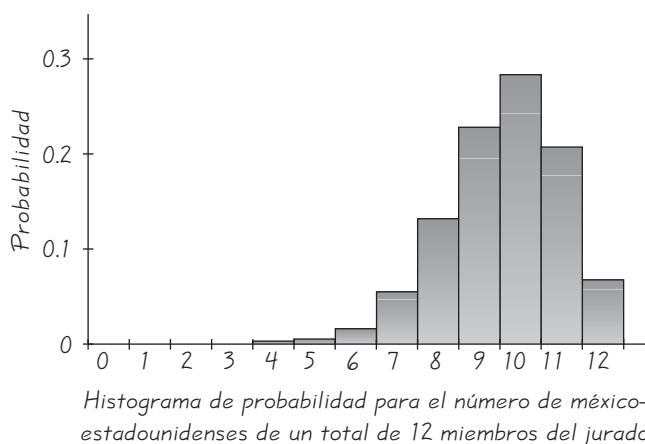


Figura 5-3 Histograma de probabilidad del número de mexicano-estadounidenses de un total de 12 miembros del jurado

tabla 5-1, la suma de todas las probabilidades es 1, pero en otros casos, valores tales como 0.999 o 1.001 son aceptables porque son el resultado de errores de redondeo). Asimismo, la regla de probabilidad que establece que $0 \leq P(x) \leq 1$ para cualquier suceso A , implica que $P(x)$ debe estar entre 0 y 1 para cualquier valor de x . Puesto que la tabla 5-1 satisface ambos requisitos, es un ejemplo de una distribución de probabilidad. Una distribución de probabilidad puede describirse como una tabla (por ejemplo, la tabla 5-1), una gráfica (como la figura 5-3) o una fórmula.

Tabla 5-2 Probabilidades de una variable aleatoria	
x	$P(x)$
0	0.2
1	0.5
2	0.4
3	0.3

EJEMPLO ¿La tabla 5-2 describe una distribución de probabilidad?

SOLUCIÓN Para ser una distribución de probabilidad, $P(x)$ debe satisfacer los dos requisitos anteriores. Pero

$$\begin{aligned}\sum P(x) &= P(0) + P(1) + P(2) + P(3) \\ &= 0.2 + 0.5 + 0.4 + 0.3 \\ &= 1.4 \quad [\text{lo que demuestra que } \sum P(x) \neq 1]\end{aligned}$$

Como no se satisface el primer requisito, concluimos que la tabla 5-2 no describe una distribución de probabilidad.

EJEMPLO ¿Determina $P(x) = x/3$ (donde x puede ser 0, 1 o 2) una distribución de probabilidad?

SOLUCIÓN Para la función dada, encontramos que $P(0) = 0/3$, $P(1) = 1/3$ y $P(2) = 2/3$, de manera que

1. $\sum P(x) = \frac{0}{3} + \frac{1}{3} + \frac{2}{3} = \frac{3}{3} = 1$

2. Cada uno de los valores $P(x)$ se encuentra entre 0 y 1.

Puesto que ambos requisitos se satisfacen, la función $P(x)$ de este ejemplo es una distribución de probabilidad.

Media, varianza y desviación estándar

En el capítulo 2 describimos las siguientes características importantes de los datos (que pueden recordarse por medio de las siglas CVDVT “Cuidado con los Virus que Destruyen el Valioso Trabajo”): **1.** centro, **2.** variación, **3.** distribución, **4.** valores extremos y **5.** tiempo (características de los datos que cambian con el tiempo). El histograma de probabilidad puede darnos información acerca de la naturaleza o forma de la distribución. Además, a menudo podemos calcular la media, la varianza y la desviación estándar de los datos, los cuales brindan información acerca de otras características. La media, la varianza y la desviación estándar de una distribución de probabilidad se calcula al aplicar las fórmulas 5-1, 5-2, 5-3 y 5-4.

Fórmula 5-1 $\mu = \sum [x \cdot P(x)]$

media de una distribución de probabilidad

Fórmula 5-2 $\sigma^2 = \sum [(x - \mu)^2 \cdot P(x)]$

varianza de una distribución de probabilidad

Fórmula 5-3 $\sigma^2 = \sum [x^2 \cdot P(x)] - \mu^2$

varianza de una distribución de probabilidad

Fórmula 5-4 $\sigma = \sqrt{\sum [x^2 \cdot P(x)] - \mu^2}$

desviación estándar de una distribución de probabilidad

Precaución: Evalúe $\Sigma[x^2 \cdot P(x)]$ elevando al cuadrado primero cada valor de x , multiplicando después cada cuadrado por la probabilidad $P(x)$ correspondiente y después sumando.

Fundamentos de las fórmulas 5-1 a la 5-4

En vez de aceptar y aplicar fórmulas a ciegas, es mucho mejor comprender por qué funcionan. La fórmula 5-1 logra lo mismo que la fórmula de la media para una tabla de frecuencias. (Recuerde que f representa la frecuencia de clase y N representa el tamaño de la población). Al reescribir la fórmula de la media de una tabla de frecuencias, de manera que se aplique a una población, y luego cambiando su forma, obtenemos

$$\mu = \frac{\Sigma(f \cdot x)}{N} = \Sigma\left[\frac{f \cdot x}{N}\right] = \Sigma\left[x \cdot \frac{f}{N}\right] = \Sigma[x \cdot P(x)]$$

En la fracción f/N , el valor de f es la frecuencia con que ocurre el valor x y N es el tamaño de la población, así que f/N es la probabilidad del valor de x .

Un razonamiento similar nos permite tomar la fórmula de la varianza del capítulo 3 y aplicarla a una variable aleatoria para una distribución de probabilidad; el resultado es la fórmula 5-2. La fórmula 5-3 es una versión abreviada que siempre producirá el mismo resultado que la fórmula 5-2. Aun cuando la fórmula 5-3 suele ser más fácil de usar, la fórmula 5-2 es más fácil de comprender directamente. Con base en la fórmula 5-2, podemos expresar la desviación estándar como

$$\sigma = \sqrt{\Sigma[(x - \mu)^2 \cdot P(x)]}$$

o como la forma equivalente dada en la fórmula 5-4.

Cuando utilice las fórmulas 5-1 a la 5-4, aplique esta regla para redondear los resultados.

Regla de redondeo para μ , σ , y σ^2

Redondee los resultados llevando una posición decimal más que el número de posiciones decimales utilizadas para la variable aleatoria x . Si los valores de x son enteros, redondee μ , σ y σ^2 a una posición decimal.

En ocasiones es necesario usar una regla diferente de redondeo ante circunstancias especiales, tales como resultados que requieren más decimales para ser significativos. Por ejemplo, para aviones de propulsión a chorro de cuatro motores, el número medio de motores que funcionan adecuadamente durante un vuelo es de 3.999714286, que se convierte en 4.0 cuando se redondea a una posición decimal más que los datos originales. Aquí, el 4.0 sería confuso, ya que sugiere que todos los motores del avión de propulsión a chorro siempre funcionan bien. Necesitamos más precisión para reflejar correctamente la media verdadera, como la precisión en el número 3.999714.

Identificación de resultados poco comunes con la regla práctica del intervalo

La regla práctica del intervalo (que se estudió en la sección 3-3) también resulta útil para interpretar los valores de una desviación estándar. Según la regla práctica del intervalo, la mayoría de los valores deben caer dentro de 2 desviaciones estándar de la

media; no es común que un valor difiera de la media en más de dos desviaciones estándar. (El uso de dos desviaciones estándar no es un valor absolutamente rígido, y en su lugar se pueden emplear otros valores como 3). De esta manera, podemos identificar valores “poco comunes” si se determina que caen fuera de los siguientes límites:

Regla práctica del intervalo

$$\text{valor máximo común} = \mu + 2\sigma$$

$$\text{valor mínimo común} = \mu - 2\sigma$$

EJEMPLO La tabla 5-1 describe la distribución de probabilidad del número de mexicano-estadounidenses que integran un jurado de 12 miembros en el condado de Hidalgo, Texas. Suponiendo que repetimos el proceso de selección aleatoria de los 12 miembros del jurado y contamos el número de mexicano-estadounidenses cada vez, calcule la media del número de mexicano-estadounidenses (de un total de 12), la varianza y la desviación estándar. Utilice esos resultados y la regla práctica del intervalo para calcular el valor máximo común y el valor mínimo común. Con base en los resultados, determine si un jurado que consta de 7 mexicano-estadounidenses de un total de 12 miembros es común o poco común.

SOLUCIÓN En la tabla 5-3, las dos columnas a la izquierda describen la distribución de probabilidad que se presentó antes en la tabla 5-1, y creamos las tres columnas a la derecha para realizar los cálculos requeridos.

Al utilizar las fórmulas 5-1 y 5-3, así como los resultados de la tabla, obtenemos

$$\mu = \Sigma[x \cdot P(x)] = 9.598 = 9.6 \quad (\text{redondeado})$$

$$\begin{aligned} \sigma^2 &= \Sigma[x^2 \cdot P(x)] - \mu^2 \\ &= 94.054 - 9.598^2 = 1.932396 = 1.9 \quad (\text{redondeado}) \end{aligned}$$

La desviación estándar es la raíz cuadrada de la varianza, de manera que

$$\sigma = \sqrt{1.932396} = 1.4 \quad (\text{redondeado})$$

Sabemos que cuando elegimos al azar a 12 miembros del jurado, la media del número de mexicano-estadounidenses es 9.6, la varianza es 1.9 “mexicano-estadounidenses al cuadrado” y la desviación estándar es 1.4 mexicano-estadounidenses. Si utilizamos la regla práctica del intervalo, ahora podemos calcular el valor máximo común y el valor mínimo común de la siguiente manera:

$$\text{valor máximo común: } \mu + 2\sigma = 9.6 + 2(1.4) = 12.4$$

$$\text{valor mínimo común: } \mu - 2\sigma = 9.6 - 2(1.4) = 6.8$$

INTERPRETACIÓN Con base en estos resultados, concluimos que, para grupos de 12 miembros del jurado elegidos al azar en el condado de Hidalgo, el número de mexicano-estadounidenses debe caer entre 6.8 y 12.4. Si un jurado consta de 7 mexicano-estadounidenses, no se trata de un suceso poco común y no serviría como base para afirmar que el jurado se eligió de una manera que discrimina a los mexicano-estadounidenses. (El jurado que sentenció a Roger Partida incluyó a 7 mexicano-estadounidenses, pero la acusación de un proceso de selección injusto se basó en el proceso utilizado para seleccionar al Gran Jurado y no al jurado específico que lo sentenció).

Tabla 5-3 Cálculo de μ , σ y σ^2 para una distribución de probabilidad				
x	$P(x)$	$x \cdot P(x)$	x^2	$x^2 \cdot P(x)$
0	0+	0.000	0	0.000
1	0+	0.000	1	0.000
2	0+	0.000	4	0.000
3	0+	0.000	9	0.000
4	0.001	0.004	16	0.016
5	0.003	0.015	25	0.075
6	0.016	0.096	36	0.576
7	0.053	0.371	49	2.597
8	0.133	1.064	64	8.512
9	0.236	2.124	81	19.116
10	0.283	2.830	100	28.300
11	0.206	2.266	121	24.926
12	0.069	0.828	144	9.936
Total		9.598 ↑ $\Sigma[x \cdot P(x)]$	94.054 ↑ $\Sigma[x^2 \cdot P(x)]$	

Identificación de resultados *infrecuentes* con probabilidades

Recomendación importante: Tome su tiempo para leer con cuidado y comprender la regla del suceso infrecuente y el párrafo que le sigue. Esta breve discusión presenta un método sumamente importante que se utiliza a menudo en estadística.

Regla del suceso infrecuente

Si, bajo un supuesto dado (como el supuesto de que una moneda está balanceada), la probabilidad de un suceso particular observado (como 992 caras en 1000 lanzamientos de una moneda) es extremadamente pequeña, concluimos que el supuesto probablemente no sea correcto.

Las probabilidades se pueden utilizar para aplicar la regla del suceso infrecuente de la siguiente manera:

Uso de las probabilidades para determinar resultados infrecuentes

- **Número de éxitos inusualmente alto:** x éxitos en n ensayos es un número *inusualmente alto* de éxitos si $P(x \text{ o más}) \leq 0.05$.*
- **Número de éxitos inusualmente bajo:** x éxitos en n ensayos es un número *inusualmente bajo* de éxitos si $P(x \text{ o menos}) \leq 0.05$.

Suponga que lanza una moneda para determinar si se ven favorecidas las caras y suponga que 1000 lanzamientos dan como resultado 501 caras. Esto no es evidencia de que la moneda favorezca las caras, ya que es muy fácil obtener un resultado de

*El valor de 0.05 se utiliza de forma regular, pero no es absolutamente rígido. Se podrían usar otros valores, como 0.01, para distinguir entre sucesos que pueden ocurrir con facilidad por azar y sucesos que tienen muy pocas probabilidades de ocurrir por azar.

501 caras en 1000 lanzamientos por el azar. Sin embargo, la probabilidad de obtener *exactamente* 501 caras en 1000 lanzamientos es bastante baja: 0.0252. Esta baja probabilidad refleja el hecho de que, con 1000 lanzamientos, cada número *específico* de caras tendrá una probabilidad sumamente baja. Sin embargo, no consideramos que 501 caras en 1000 lanzamientos sea un suceso *infrecuente*, ya que la probabilidad de obtener *al menos* 501 caras es alta: 0.487.



EJEMPLO Selección de miembros del jurado Si el 80% de las personas que pueden fungir como miembros del jurado en el condado de Hidalgo son mexicano-estadounidenses, entonces un jurado de 12 individuos seleccionados al azar debe incluir 9 o 10 mexicano-estadounidenses. (La media del número de mexicano-estadounidenses en los jurados es de 9.6). ¿Siete jueces mexicano-estadounidenses, de un total de 12, es un número excepcionalmente bajo? ¿La selección de sólo 7 mexicano-estadounidenses en un total de 12 miembros del jurado sugiere que existe discriminación en el proceso de selección?

SOLUCIÓN Usaremos el criterio de que 7 mexicano-estadounidenses en un total de 12 miembros del jurado es excepcionalmente bajo si $P(7 \text{ o menos mexicano-estadounidenses}) \leq 0.05$. Si nos remitimos a la tabla 5-1, obtenemos el siguiente resultado:

$$\begin{aligned} P(7 \text{ o menos mexicano-estadounidenses en un total de 12 miembros del jurado}) \\ &= P(7 \text{ o } 6 \text{ o } 5 \text{ o } 4 \text{ o } 3 \text{ o } 2 \text{ o } 1 \text{ o } 0) \\ &= P(7) + P(6) + P(5) + P(4) + P(3) + P(2) + P(1) + P(0) \\ &= 0.053 + 0.016 + 0.003 + 0.001 + 0 + 0 + 0 + 0 \\ &= 0.073 \end{aligned}$$

INTERPRETACIÓN Puesto que la probabilidad de 0.073 es mayor que 0.05, concluimos que el resultado de 7 mexicano-estadounidenses *no es poco común*. Existe una alta probabilidad (0.073) de seleccionar a 7 mexicano-estadounidenses por azar. (Sólo una probabilidad de 0.05 o menor indicaría que el suceso es poco común). Ningún tribunal de justicia declararía que, en esas circunstancias, la selección de sólo 7 mexicano-estadounidenses es discriminatoria.

Valor esperado

La media de una variable aleatoria discreta es el resultado medio teórico de un número infinito de ensayos. Podemos considerar esa media como el *valor esperado* en el sentido de que constituye el valor promedio que esperaríamos obtener si los ensayos pudieran continuar de manera indefinida. Los usos del valor esperado (también llamado *esperanza* o *esperanza matemática*) son extensos y variados, y desempeñan un papel muy importante en una área de aplicación denominada *téoría de la decisión*.



Definición

El **valor esperado** de una variable aleatoria discreta se denota por E y representa el valor promedio de los resultados. Se obtiene calculando el valor de $\Sigma[x \cdot P(x)]$.

$$E = \Sigma[x \cdot P(x)]$$

De la fórmula 5-1 vemos que $E = \mu$. Es decir, la media de una variable aleatoria discreta es la misma que su valor esperado. Vea la tabla 5-3 y observe que al seleccionar a 12 miembros del jurado de una población en la que el 80% de los habitantes son mexico-estadounidenses, la *media* del número de mexico-estadounidenses es 9.6, por lo que se deduce que el *valor esperado* del número de mexico-estadounidenses también es 9.6.

EJEMPLO Lotería Kentucky Pick 4 Si usted apuesta \$1 en el juego de lotería Kentucky Pick 4, pierde \$1 o gana \$4999. (El premio ganador es de \$5,000, pero no le devuelven su apuesta de \$1, por lo que la ganancia neta es de \$4999). El juego consiste en seleccionar un número de cuatro dígitos entre 0000 y 9999. Si usted apuesta \$1 al 1234, ¿cuál es el valor esperado de ganar o perder?

SOLUCIÓN Para esta apuesta existen dos resultados: usted pierde \$1 o gana \$4999. Como existen 10,000 posibilidades de números de cuatro dígitos y sólo una de ellas representa el número ganador, la probabilidad de perder es de 9,999/10,000 y la probabilidad de ganar es de 1/10,000. La tabla 5-4 resume la distribución de probabilidad, y podemos ver que el valor esperado es $E = -50$ centavos.

Tabla 5-4 Lotería Pick 4 de Kentucky			
Suceso	x	$P(x)$	$x \cdot P(x)$
Pérdida	-\$1	0.9999	-\$0.9999
Ganancia (neta)	\$4999	0.0001	\$0.4999
Total			-\$0.50 (o 50 centavos)

INTERPRETACIÓN En cualquier juego individual, usted pierde \$1 u obtiene una ganancia neta de \$4999, pero el valor esperado indica que, a largo plazo, usted espera perder un promedio de 50 centavos por cada apuesta de \$1. Tal vez esta lotería tenga cierto valor de entretenimiento limitado, pero, sin duda, se trata de una inversión económica extremadamente inadecuada.

En esta sección aprendimos que una variable aleatoria tiene un valor numérico asociado a cada resultado de algún procedimiento aleatorio, y que una distribución de probabilidad tiene una probabilidad asociada a cada valor de una variable aleatoria. Examinamos métodos para calcular la media, la varianza y la desviación estándar de una distribución de probabilidad. Vimos que el valor esperado de una variable aleatoria es, en realidad, igual a la media. Por último, un concepto sumamente importante de esta sección es el uso de probabilidades para determinar cuándo los resultados son *poco comunes*.

5-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Distribución de probabilidad.** Considere el ensayo de lanzar un dado, con los resultados 1, 2, 3, 4, 5 y 6. Construya la tabla que represente la distribución de probabilidad.

- 2. Distribución de probabilidad.** Uno de los requisitos de una distribución de probabilidad es que la suma de las probabilidades debe ser 1 (se permite una pequeña cantidad de variación por errores de redondeo). ¿Cuál es la justificación de este requisito?
- 3. Distribución de probabilidad.** Un jugador profesional afirma que cargó un dado para que los resultados de 1, 2, 3, 4, 5, 6 tengan probabilidades correspondientes de 0.1, 0.2, 0.3, 0.4, 0.5 y 0.6. ¿Realmente será cierto lo que dice? Una distribución de probabilidad se describe haciendo una lista de los resultados junto con sus probabilidades correspondientes?
- 4. Valor esperado.** Un investigador calcula el valor esperado del número de niñas en cinco nacimientos y obtiene un resultado de 2.5. Luego, redondea los resultados a 3, al afirmar que no es posible que nazcan 2.5 niñas en cinco nacimientos. ¿Es correcto este razonamiento?

Identificación de variables aleatorias discretas y continuas. En los ejercicios 5 y 6, identifique si la variable aleatoria dada es discreta o continua.

- 5. a.** La estatura de una jirafa que vive en Kenia, elegida al azar
b. El número de águilas calvas que habitan en el estado de Nueva York
c. El tiempo exacto que se requiere para sumar $27 + 72$
d. El número de autores de libros de texto que ahora están sentados ante una computadora
e. El número de estudiantes de estadística que ahora están leyendo un libro
- 6. a.** El costo de realizar un experimento de genética
b. El número de supermodelos que ayer comieron pizza
c. El tiempo de vida exacto de un gato
d. El número de profesores de estadística que leen un periódico cada día
e. El peso de una pluma

Identificación de distribuciones de probabilidad. En los ejercicios 7 a 12, determine si se trata de una distribución de probabilidad. En los casos en que no se describe una distribución de probabilidad, identifique los requisitos que no se satisfacen. Para los casos en los que se describe una distribución de probabilidad, calcule su media y desviación estándar.

- 7. Trastorno genético.** Cada uno de tres hombres que tienen un trastorno genético relacionado con el cromosoma X tiene un hijo. La variable aleatoria x es el número de hijos de los tres hombres que heredan el trastorno genético relacionado con el cromosoma X.

x	$P(x)$
0	0.4219
1	0.4219
2	0.1406
3	0.0156

- 8. Números de niñas.** Un investigador reporta que, cuando se seleccionan al azar grupos de cuatro niños de una población de parejas que cumplen ciertos criterios, la distribución de probabilidad del número de niñas es como la que se presenta en la siguiente tabla.

x	$P(x)$
0	0.502
1	0.365
2	0.098
3	0.011
4	0.001

- 9. Experimento de genética.** Un experimento de genética incluye vástagos de guisantes en grupos de cuatro. Un investigador reporta que, para un grupo, el número de plantas de guisantes con flores blancas tiene una distribución de probabilidad como la que se presenta en la siguiente tabla.

x	$P(x)$
0	0.04
1	0.16
2	0.80
3	0.16
4	0.04

- 10. Estudio de mortalidad.** Para un grupo de cuatro hombres, la distribución de probabilidad del número x que sobreviven al año siguiente es como la que se presenta la siguiente tabla.

x	$P(x)$
0	0.0000
1	0.0001
2	0.0006
3	0.0387
4	0.9606

- 11. Número de juegos en una Serie Mundial de béisbol.** Con base en resultados pasados encontrados en el *Information Please Almanac*, existe una probabilidad del 0.1818 de que la Serie Mundial de béisbol dure cuatro juegos, una probabilidad del 0.2121 de que dure cinco juegos, una probabilidad de 0.2323 de que dure seis juegos y una probabilidad del 0.3737 de que dure siete juegos. ¿Será infrecuente que un equipo “arrase” al ganar cuatro juegos?

- 12. Reconocimiento de marca.** En un estudio de reconocimiento de la marca Sony se entrevistaron grupos de cuatro consumidores. Si x es el número de personas en el grupo que reconocen la marca Sony, entonces x puede ser 0, 1, 2, 3 o 4, y las probabilidades correspondientes son 0.0016, 0.0250, 0.1432, 0.3892 y 0.4096. ¿Será infrecuente seleccionar al azar a cuatro consumidores y descubrir que ninguno de ellos reconoce la marca Sony?

- 13. Determinar si un proceso de selección de miembros de un jurado es discriminatorio.** Suponga que se seleccionan 12 jueces al azar de una población en la que el 80% de los habitantes son mexicano-estadounidenses. Remítase la tabla 5-1 y calcule las probabilidades indicadas.

- a. Calcule la probabilidad de que haya exactamente 5 mexicano-estadounidenses en un total de 12 miembros del jurado.
- b. Calcule la probabilidad de que haya 5 o menos mexicano-estadounidenses en un total de 12 miembros del jurado.
- c. ¿Qué probabilidad es relevante para determinar si 5 jueces de un total de 12 son excepcionalmente pocos: el resultado del inciso a) o el del inciso b)?
- d. ¿Cinco mexicano-estadounidenses de un total de 12 miembros del jurado sugieren que el proceso de selección discrimina a los mexicano-estadounidenses? ¿Por qué?

- 14. Determinar si un proceso de selección de miembros de un jurado es discriminatorio.** Suponga que se seleccionan 12 jueces al azar de una población en la que el 80% de los habitantes son mexicano-estadounidenses. Remítase a la tabla 5-1 y calcule las probabilidades indicadas.

- a. Calcule la probabilidad de que haya exactamente 6 mexicano-estadounidenses en un total de 12 miembros del jurado.
- b. Calcule la probabilidad de que haya 6 o menos mexicano-estadounidense en un total de 12 miembros del jurado.
- c. ¿Qué probabilidad es relevante para determinar si 6 jueces de un total de 12 son excepcionalmente pocos: el resultado del inciso a) o el del inciso b)?
- d. ¿Seis mexicano-estadounidenses de un total de 12 miembros del jurado sugieren que el proceso de selección discrimina a los mexicano-estadounidenses? ¿Por qué?

- 15. Determinar si un proceso de selección de miembros de un jurado es discriminatorio.** Suponga que se seleccionan 12 jueces al azar de una población en la que el 80% de los habitantes son mexicano-estadounidenses. Remítase la tabla 5-1 y calcule las probabilidades indicadas.

- a. Utilice los valores de probabilidad de la tabla 5-1 para calcular el valor de probabilidad que se debe emplear para determinar si el resultado de 8 mexicano-estadounidenses en un total de 12 miembros del jurado es inusualmente bajo.
- b. ¿El resultado de 8 mexicano-estadounidenses sugiere que el proceso de selección discrimina a los mexicano-estadounidenses? ¿Por qué?

- 16. Determinar si un proceso de selección de miembros de un jurado es discriminatorio.** Suponga que se seleccionan 12 jueces al azar de una población en la que el 80% de los habitantes son mexicano-estadounidenses. Remítase a la tabla 5-1 y calcule las probabilidades indicadas.

- a. Utilice los valores de probabilidad de la tabla 5-1 para calcular el valor de probabilidad que se debe emplear para determinar si el resultado de 11 mexicano-estadounidenses en un total de 12 miembros del jurado es inusualmente alto.
- b. ¿El resultado de 11 mexicano-estadounidenses sugiere que el proceso de selección favorece a los mexicano-estadounidenses? ¿Por qué?
17. **Cálculo del valor esperado en la ruleta.** Cuando usted apuesta \$5 al número 7 en la ruleta en el casino Venetian de Las Vegas, tiene una probabilidad de 37/38 de perder \$5, y una probabilidad de 1/38 de obtener una ganancia neta de \$175. (El premio es de \$180, incluyendo su apuesta de \$5, de manera que la ganancia neta es de \$175). Si apuesta \$5 a que el resultado es un número impar, la probabilidad de perder \$5 es de 20/38 y la probabilidad de obtener una ganancia neta de \$5 es de 18/38. (Si usted apuesta \$5 a un número impar y gana, recibe \$10 incluyendo su apuesta, de manera que la ganancia neta es de \$5).
- a. Si apuesta \$5 al número 7, ¿cuál es su valor esperado?
- b. Si apuesta \$5 a que el resultado es un número impar, ¿cuál es su valor esperado?
- c. ¿Cuál de estas opciones es mejor: apostar al siete, apostar a número impar o no apostar? ¿Por qué?
18. **Cálculo del valor esperado en los dados en un casino.** Cuando usted apuesta \$5 en un casino en la “línea de pase” en el juego de dados, existe una probabilidad de 251/495 de que pierda \$5 y una probabilidad de 244/495 de que obtenga una ganancia neta de \$5. (Si usted gana, el casino le da \$5 y conserva su apuesta de \$5, de manera que la ganancia neta es de \$5). ¿Cuál es su valor esperado? A la larga, ¿cuánto pierde por cada dólar que apueste?
19. **Cálculo del valor esperado para una póliza de seguro de vida.** La compañía de seguros CNA le cobra a un hombre de 21 años \$250 por un año de una póliza de seguro de vida de \$100,000. Un hombre de 21 años tiene una probabilidad del 0.9985 de sobrevivir durante un año (según datos del U.S. National Center for Health Statistics).
- a. Desde la perspectiva del hombre de 21 años (o de su estado), ¿cuáles son los valores de los dos resultados diferentes?
- b. ¿Cuál es el valor esperado para un hombre de 21 años que compra el seguro?
- c. ¿Cuál sería el costo de la póliza del seguro si la compañía sale a mano (a la larga eso sucede con muchas pólizas), en vez de obtener una ganancia?
- d. Dado que el valor esperado es negativo (de manera que la compañía obtiene una ganancia), ¿por qué debería un hombre de 21 años o cualquier otra persona adquirir seguros de vida?
20. **Cálculo del valor esperado de la rifa organizada por una revista.** La revista *Reader's Digest* realizó una rifa en la que los premios se listaron junto con las probabilidades de ganar: \$1,000,000 (1 posibilidad en 90,000,000), \$100,000 (1 posibilidad en 110,000,000), \$25,000 (1 posibilidad en 110,000,000), \$5000 (1 posibilidad en 36,667,000) y \$2500 (1 posibilidad en 27,500,000).
- a. Suponiendo que no hay un costo por participar en la rifa, calcule el valor esperado de la cantidad a ganar con un boleto.
- b. Calcule el valor esperado si el costo de un boleto en esta rifa equivale al de una estampilla postal. ¿Vale la pena participar en esta rifa?
21. **Cálculo de la media y la desviación estándar.** Sea x la variable aleatoria que represente el número de niñas en una familia de tres hijos. Construya una tabla que describa la distribución de probabilidad, después calcule la media y la desviación estándar. (*Sugerencia:* Liste los distintos resultados posibles). ¿Es poco común que una familia de tres hijos incluya tres niñas?
22. **Cálculo de la media y la desviación estándar.** Sea x la variable aleatoria que represente el número de niñas en una familia de cuatro hijos. Construya una tabla que describa la distribución de probabilidad, después calcule la media y la desviación estándar. (*Sugerencia:* Liste los distintos resultados posibles). ¿Es poco común que una familia de cuatro hijos incluya cuatro niñas?

- 23. Encuestas telefónicas.** Con frecuencia se utilizan computadoras para generar dígitos aleatorios de números telefónicos y realizar encuestas. Cada dígito tiene la misma probabilidad de resultar seleccionado. Construya una tabla que represente la distribución de probabilidad de los dígitos seleccionados, calcule la media y la desviación estándar, y describa la forma del histograma de probabilidad.
- 24. Ventas de casas.** Remítase al número de recámaras en casas vendidas, tal como aparecen en el conjunto de datos 18 del apéndice B. Utilice la distribución de frecuencias para construir una tabla que represente la distribución de probabilidad; luego, calcule la media y la desviación estándar. Además, describa la forma del histograma de probabilidad.

5-2 MÁS ALLÁ DE LO BÁSICO

- 25. Distribución de frecuencias y distribución de probabilidad.** ¿Cuál es la principal diferencia entre una distribución de frecuencias (tal como se definió en la sección 2-2) y una distribución de probabilidad (tal como se definió en esta sección)?
- 26. Bonos especulativos.** Kim Hunter tiene \$1000 para invertir, y su analista financiero le recomienda dos tipos de bonos especulativos. Los bonos A tienen un rendimiento anual del 6%, con una tasa de incumplimiento del 1%. Los bonos B tienen un rendimiento anual del 8%, con una tasa de incumplimiento del 5%. (Si el bono incumple, se pierden los \$1000). ¿Cuál de los dos bonos es mejor? ¿Por qué? ¿Kim debería elegir cualquiera de los dos bonos? ¿Por qué?
- 27. Partes defectuosas: cálculo de la media y la desviación estándar.** Sky Ranch es un proveedor de partes para aeronaves. Sus existencias incluyen 8 altímetros que están correctamente calibrados y 2 que no lo están. Se seleccionan 3 altímetros al azar y sin reemplazo. Sea x la variable aleatoria que represente el número de aparatos que no están calibrados correctamente. Calcule la media y la desviación estándar de la variable aleatoria x .
- 28. Colocar marcas en dados para obtener una distribución uniforme.** Suponga que tiene dos dados en blanco, de manera que puede marcar las 12 caras con los números que deseé. Describa de qué manera marcaría los dados para que, cuando tire ambos, los totales de los dos dados se distribuyan de manera uniforme y cada uno de los resultados de 1, 2, 3, . . . , 12 tenga una probabilidad de 1/12. (Véase “Can One Load a Set of Dice so that the Sum is Uniformly Distributed?” de Chen, Rao y Shreve, *Mathematics Magazine*, vol. 70, núm. 3).



5-3 Distribuciones de probabilidad binomial

Concepto clave En la sección 4-2 estudiamos las distribuciones discretas de probabilidad en general, pero en esta sección nos enfocaremos en un tipo específico: la distribución de probabilidad binomial. Puesto que las distribuciones de probabilidad binomial implican proporciones que se utilizan con métodos de estadística inferencial, que se analizan más adelante en este libro, es importante entender las propiedades fundamentales de esta clase particular de distribuciones de probabilidad. Esta sección presenta una definición básica de una distribución de probabilidad binomial y la notación que se utiliza; también presenta métodos para calcular valores de probabilidad.

Las distribuciones de probabilidad binomial nos permiten enfrentar circunstancias en las que los resultados pertenecen a *dos* categorías relevantes, tales como aceptable/defectuoso o sobrevivió/murió. En la siguiente definición se plantean otros requisitos.

Definición

Una **distribución de probabilidad binomial** resulta de un procedimiento que cumple con todos los siguientes requisitos:

1. El procedimiento tiene un número *fijo de ensayos*.
2. Los ensayos deben ser *independientes*. (El resultado de cualquier ensayo individual no afecta las probabilidades de los demás ensayos).
3. Todos los resultados de cada ensayo deben estar clasificados en dos *categorías* (generalmente llamadas *éxito* y *fracaso*).
4. La probabilidad de un éxito permanece igual en todos los ensayos.

Si un procedimiento satisface estos cuatro requisitos, la distribución de la variable aleatoria x (número de éxitos) se denomina *distribución de probabilidad binomial* (o *distribución binomial*), en la que suele usarse la siguiente notación:

Notación para distribuciones de probabilidad binomial

E y F (éxito y fracaso) denotan las dos categorías posibles de todos los resultados; p y q denotan las probabilidades de E y F, respectivamente, de manera que

$P(S) = p$	(p = probabilidad de un éxito)
$P(F) = 1 - p = q$	(q = probabilidad de un fracaso)
n	denota el número fijo de ensayos.
x	denota un número específico de éxitos en n ensayos, de manera que x puede ser cualquier número entero entre 0 y n inclusive.
p	denota la probabilidad de éxito en <i>uno</i> de n ensayos.
q	denota la probabilidad de fracaso en <i>uno</i> de n ensayos.
$P(x)$	denota la probabilidad de lograr exactamente x éxitos en los n ensayos.

El término *éxito* que se utiliza aquí es arbitrario y no necesariamente tiene una connotación positiva. Cualquiera de las dos categorías posibles puede denominarse el éxito E, siempre y cuando su probabilidad se identifique como p . Una vez que se designa una categoría como éxito E, asegúrese de que p es la probabilidad de un éxito y que x es el número de éxitos. Es decir, *asegúrese de que los valores p y x se refieren a la misma categoría designada como un éxito*. (El valor de q se puede calcular siempre al restar p de 1; si $p = 0.95$, entonces $q = 1 - 0.95 = 0.05$). He aquí una recomendación importante para trabajar con problemas de probabilidad binomial:

Asegúrese de que x y p se refieren a la misma categoría denominada como un éxito.

Cuando seleccionamos una muestra (como una encuesta) para algún análisis estadístico, por lo general realizamos el muestreo sin reemplazo, y el muestreo sin reemplazo implica sucesos dependientes, lo cual viola el segundo requisito de la definición anterior. Sin embargo, a menudo se utiliza la siguiente regla práctica (porque los errores son insignificantes): cuando se hace un muestreo sin reemplazo, los sucesos pueden tratarse como si fueran independientes, si el tamaño de la muestra no es mayor que el 5% del tamaño de la población.

Cuando realice un muestreo sin reemplazo, considere los sucesos como independientes si $n \leq 0.05N$.

**PROBLEMA
DEL
CAPÍTULO**
5

EJEMPLO Selección de jueces En el caso de *Castaneda contra Partida* se señaló que, aunque el 80% de la población de un condado en Texas es mexico-estadounidense, sólo el 39% de quienes fueron llamados para integrar el jurado pertenecían a este grupo. Supongamos que necesitamos seleccionar a 12 jueces de una población integrada en un 80% por mexico-estadounidenses, y que deseamos calcular la probabilidad de que, de 12 jueces elegidos al azar, exactamente 7 sean mexico-estadounidenses.

- Este proceso dará por resultado una distribución binomial?
- Si este proceso da por resultado una distribución binomial, identifique los valores de n , x , p y q .

SOLUCIÓN

- Este procedimiento sí satisface los requisitos de una distribución binomial, como se indica a continuación.
 - El número de ensayos (12) es fijo.
 - Los 12 ensayos son independientes. (Técnicamente, los 12 ensayos implican una selección sin reemplazo y no son independientes, pero podemos suponer independencia porque estamos seleccionando al azar sólo a 12 miembros de una población muy grande.)
 - Cada uno de los 12 ensayos tiene dos categorías de resultados posibles: el miembro del jurado elegido es mexico-estadounidense o no lo es.
 - Para cada miembro del jurado elegido, la probabilidad de que sea mexico-estadounidense es de 0.8 (porque el 80% de la población es mexico-estadounidense). Esa probabilidad de 0.8 es la misma para cada uno de los miembros del jurado.
- Una vez que concluimos que el procedimiento dado sí da por resultado una distribución binomial, ahora procedemos a identificar los valores de n , x , p y q .
 - Con 12 jueces elegidos, tenemos que $n = 12$.
 - Buscamos la probabilidad de exactamente 7 mexico-estadounidenses, entonces $x = 7$.
 - La probabilidad de éxito (elegir a un mexico-estadounidense) en una selección es 0.8, por lo tanto, $p = 0.8$.
 - La probabilidad de fracaso (no elegir a un mexico-estadounidense) es 0.2, por lo tanto, $q = 0.2$.



¿Hay alguien en casa?

Los encuestadores no pueden ignorar simplemente a quienes no estaban en casa cuando acudieron por primera vez. Una solución implica regresar varias veces hasta localizar a la persona. Alfred Politz y Willard Simmons describen una forma para compensar los resultados faltantes, sin tener que regresar varias veces. Ellos sugieren ponderar los resultados con base en la frecuencia en que las personas no se encuentran en su casa. Por ejemplo, una persona que está en su casa sólo dos de seis días a la semana, tendrá una probabilidad de $2/6$ o $1/3$ de estar allí en la primera visita. Cuando se localiza a esa persona por primera vez, sus resultados se ponderan de manera que se cuenten tres veces, respecto a un individuo que siempre está en su casa. Esta ponderación compensa a los demás individuos similares que permanecen en casa dos de seis días a la semana y que no respondieron cuando se les buscó por primera vez. Esta inteligente solución se presentó inicialmente en 1949.

continúa

Una vez más, es muy importante asegurarse de que tanto x como p se refieran al mismo concepto de “éxito”. En este ejemplo, usamos x para contar el número de mexico-estadounidenses, de manera que p debe ser la probabilidad de un mexico-estadounidense. Por consiguiente, x y p sí usan aquí el mismo concepto de éxito (mexico-estadounidense).

Ahora presentaremos tres métodos para calcular las probabilidades correspondientes a la variable aleatoria x en una distribución binomial. El primer método implica realizar cálculos mediante la *fórmula de probabilidad binomial* y es la base de los otros dos métodos. El segundo método implica el uso de la tabla A-1, y el tercero el uso de un programa de cómputo o de una calculadora. Si está utilizando alguna de estas dos herramientas que produzcan de forma automática probabilidades binomiales, le recomendamos que resuelva uno o dos ejercicios por medio del método 1, para asegurarse de que comprende los fundamentos de estos cálculos. La comprensión es siempre mucho mejor que la aplicación a ciegas de las fórmulas.

Método 1: Uso de la fórmula de probabilidad binomial En una distribución binomial de probabilidad, las probabilidades pueden calcularse mediante la fórmula de probabilidad binomial.

$$\text{Fórmula 5-5} \quad P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x} \quad \text{para } x = 0, 1, 2, \dots, n$$

donde

n = número de ensayos

x = número de éxitos en n ensayos

p = probabilidad de éxito en cualquier ensayo

q = probabilidad de fracaso en cualquier ensayo ($q = 1 - p$)

El símbolo de factorial $!$, que se presentó en la sección 4-7, denota el producto de factores decrecientes. Dos ejemplos de factoriales son $3! = 3 \cdot 2 \cdot 1 = 6$ y $0! = 1$ (por definición).



EJEMPLO Selección de miembros de jurado Utilice la fórmula de probabilidad binomial para calcular la probabilidad de seleccionar exactamente a 7 mexico-estadounidenses cuando se eligen al azar 12 miembros del jurado de una población en la que el 80% de los habitantes son mexico-estadounidenses. Es decir, calcule $P(7)$ dado que $n = 12$, $x = 7$, $p = 0.8$ y $q = 0.2$.

SOLUCIÓN Al emplear los valores dados de n , x , p y q en la fórmula de probabilidad binomial (fórmula 5-5), obtenemos

$$\begin{aligned} P(7) &= \frac{12!}{(12-7)!7!} \cdot 0.8^7 \cdot 0.2^{12-7} \\ &= \frac{12!}{5!7!} \cdot 0.2097152 \cdot 0.00032 \\ &= (792)(0.2097152)(0.00032) = 0.0531502203 \end{aligned}$$

La probabilidad de elegir exactamente a 7 mexico-estadounidenses de un total de 12 miembros del jurado seleccionados al azar es 0.0532 (redondeado a los tres dígitos significativos).

Sugerencia para el cálculo: Cuando se calcula una probabilidad con la fórmula de probabilidad binomial, es útil obtener un solo número para $n!/(n-x)!x!$, un solo número para p^x y un solo número para q^{n-x} , y luego simplemente multiplicar los tres factores, como se hizo al final de los cálculos del ejemplo anterior. No redondee demasiado al calcular esos tres factores; redondee únicamente al final.

Método 2: Uso de la tabla A-1 del apéndice A En algunos casos podemos calcular fácilmente las probabilidades binomiales con sólo remitirnos a la tabla A-1 del apéndice A (parte de la tabla A-1 se presenta al margen). Primeiro localice n y el valor de x deseado correspondiente. En este paso se debe aislar un renglón de números. Ahora alinee ese renglón con la probabilidad correspondiente de p , usando la columna que cruza. El número aislado representa la probabilidad deseada. Una probabilidad tan pequeña como 0.000064 se indica como 0+.

EJEMPLO Use la parte de la tabla A-1 (para $n = 12$ y $p = 0.8$) que está al margen, para calcular lo siguiente:

- La probabilidad de exactamente 7 éxitos.
- La probabilidad de 7 o menos éxitos.

SOLUCIÓN

- En la tabla A-1 que aparece al margen se observa que cuando $n = 12$ y $p = 0.8$, la probabilidad de $x = 7$ está dada por $P(7) = 0.053$, que es el mismo valor (excepto por el redondeo) calculado con la fórmula de probabilidad binomial en el ejemplo anterior.
- “7 o menos éxitos” significa que el número de éxitos es 7 o 6 o 5 o 4 o 3 o 2 o 1 o 0.

$$\begin{aligned} P(7 \text{ o menos}) &= P(7 \text{ o } 6 \text{ o } 5 \text{ o } 4 \text{ o } 3 \text{ o } 2 \text{ o } 1 \text{ o } 0) \\ &= P(7) + P(6) + P(5) + P(4) + P(3) + P(2) + P(1) + P(0) \\ &= 0.053 + 0.016 + 0.003 + 0.001 + 0 + 0 + 0 + 0 \\ &= 0.073 \end{aligned}$$

Puesto que la probabilidad de 0.073 no es pequeña (no es 0.05 o menor), sugiere que si se eligen al azar 12 miembros del jurado, el resultado de 7 mexicano-estadounidenses no es excepcionalmente bajo y es fácil que ocurra por azar.

En el inciso b) de la solución anterior, si quisieramos calcular $P(7 \text{ o menos})$ por medio de la fórmula de probabilidad binomial, habríamos necesitado aplicar la fórmula ocho veces para calcular ocho probabilidades diferentes, que después deberían sumarse. Al poder elegir entre la fórmula y la tabla, es más lógico emplear esta última. Por desgracia, la tabla A-1 incluye sólo un número limitado de valores de n y de p , por lo que no siempre sirve; en tal caso debemos calcular las probabilidades mediante la fórmula de probabilidad binomial, un programa de cómputo o una calculadora, como en el siguiente método.

De la Tabla A-1:

n	x	p 0.80
12	0	0+
	1	0+
	2	0+
	3	0+
	4	0.001
	5	0.003
	6	0.016
	7	0.053
	8	0.133
	9	0.236
	10	0.283
	11	0.206
	12	0.069



Distribución de probabilidad binomial para $n = 12$ y $p = 0.8$

x	p
0	0+
1	0+
2	0+
3	0+
4	0.001
5	0.003
6	0.016
7	0.053
8	0.133
9	0.236
10	0.283
11	0.206
12	0.069

Método 3: Uso de herramientas tecnológicas STATDISK, Minitab, Excel, SPSS, SAS y la calculadora TI-83/84 Plus pueden usarse para calcular probabilidades binomiales. (SPSS y SAS son más difíciles de utilizar, ya que, en vez de dar directamente probabilidades para valores individuales de x , dan probabilidades *acumulativas* de x o menos éxitos). He aquí pantallas comunes que listan las probabilidades binomiales para $n = 12$ y $p = 0.8$.

STATDISK

Binomial Probability			
Num Trials, n:	12	Evaluate	
Success Prob, p:	0.8		
Mean:	9.6000		
St Dev:	1.3856		
Variance:	1.9200		
x	$P(x)$	$P(x \text{ or fewer})$	$P(x \text{ or greater})$
0	0.0000000	0.0000000	1.0000000
1	0.0000002	0.0000002	1.0000000
2	0.0000043	0.0000045	0.9999998
3	0.0000517	0.0000092	0.9999955
4	0.0005190	0.0005192	0.9999378
5	0.0033219	0.0039031	0.9994168
6	0.0155021	0.0194053	0.9980968
7	0.0531502	0.0725555	0.9905947
8	0.1328756	0.2054311	0.9744445
9	0.2343222	0.4416543	0.7949689
10	0.2834688	0.7251221	0.5593457
11	0.2061584	0.9312895	0.2748779
12	0.068719	1.0000000	0.0000000

Excel

	A	B
1	0	4.096E-09
2	1	1.966E-07
3	2	4.325E-06
4	3	5.767E-05
5	4	0.000519
6	5	0.0033219
7	6	0.0155021
8	7	0.0531502
9	8	0.1328756
10	9	0.2362232
11	10	0.2834678
12	11	0.2061584
13	12	0.0687195

Minitab

Binomial with $n = 12$ and $p = 0.8$	
x	$P(X = x)$
0	0.000000
1	0.000000
2	0.000004
3	0.000058
4	0.000519
5	0.003322
6	0.015502
7	0.053150
8	0.132876
9	0.236223
10	0.283468
11	0.206158
12	0.068719

TI-83/84 Plus

L1	L2	L3	Z
0	4.1E-9		
1	2E-7		
2	4.3E-6		
3	5.8E-5		
4	5.2E-4		
5	.00332		
6	.0155		
L3(1) =			

L1	L2	L3	Z
7	.05315		
8	.13288		
9	.23622		
10	.28347		
11	.20616		
12	.06872		
L2(14) =			

Como ahora conocemos tres métodos diferentes para calcular probabilidades binomiales, he aquí una estrategia efectiva y eficiente:

1. Utilice un programa de cómputo o una calculadora TI-83/84 Plus, si están disponibles.
2. Si no dispone de un programa de cómputo ni de la calculadora TI-83/84 Plus, utilice la tabla A-1.
3. Si no dispone de un programa de cómputo ni de calculadora y no puede encontrar las probabilidades en la tabla A-1, entonces utilice la fórmula de probabilidad binomial.

Fundamentos de la fórmula de probabilidad binomial

La fórmula de probabilidad binomial es la base de los tres métodos presentados en esta sección. En vez de aceptar y usar la fórmula a ciegas, veamos cómo funciona.

Antes en esta sección utilizamos la fórmula de probabilidad binomial para calcular la probabilidad de obtener exactamente 7 mexico-estadounidenses como miembros del jurado, cuando se eligen 12 sujetos al azar de una población en la que el 80% de los habitantes son mexico-estadounidenses. Para cada selección, la probabilidad de elegir a un mexico-estadounidense es de 0.8. Si empleamos la regla de la multiplicación de la sección 4-4, obtenemos el siguiente resultado:

$$\begin{aligned}
 P(\text{seleccionar 7 mexico-estadounidenses, seguidos de 5 personas} \\
 \text{que no son mexico-estadounidenses}) \\
 &= 0.8 \cdot 0.2 \cdot 0.2 \cdot 0.2 \cdot 0.2 \cdot 0.2 \\
 &= 0.8^7 \cdot 0.2^5 \\
 &= 0.0000671
 \end{aligned}$$

Este resultado es incorrecto porque supone que los *primeros* 7 jueces son mexico-estadounidenses y que los *últimos* 5 no lo son, en tanto que existen otros acomodos posibles para siete mexico-estadounidenses y cinco personas que no lo sean.

En la sección 4-7 vimos que con siete elementos idénticos (como mexico-estadounidenses) y otros cinco sujetos idénticos entre sí (como sujetos que no son mexico-estadounidenses), el número total de acomodos (permutaciones) es $12! / [(7 - 5)!7!]$ o 792. Cada uno de estos 792 acomodos diferentes tiene una probabilidad de $0.8^7 \cdot 0.2^5$, de manera que la probabilidad total es la siguiente:

$$P(7 \text{ jueces mexico-estadounidenses de un total de } 12) = \frac{12!}{(12 - 7)!7!} \cdot 0.8^7 \cdot 0.2^5$$

Generalice los resultados como sigue: reemplace n por 12, reemplace x por 7, reemplace p por 0.8, reemplace q por 0.2 y exprese el exponente de 5 como $12 - 7$, que puede ser reemplazado por $n - x$. El resultado es la fórmula de probabilidad binomial. Es decir, esta fórmula es una combinación de la regla de la multiplicación de probabilidad y la regla de conteo para el número de acomodos de n elementos,

cuando x de ellos son idénticos entre sí, y los otros $n - x$ son idénticos entre sí. (Véase los ejercicios 13 y 14).

El número de resultados con exactamente x éxitos en n ensayos

La probabilidad de x éxitos en n ensayos, para cualquier orden

$$P(x) = \frac{\overbrace{n!}^n}{\overbrace{(n-x)!x!}^{(n-x)}} \cdot p^x \cdot q^{n-x}$$

Uso de la tecnología

El método 3 de esta sección incluyó el uso del STATDISK, Minitab, Excel o de una calculadora TI-83/84 Plus. Las pantallas que aparecen en el método 3 ilustran resultados típicos que se obtienen al aplicar los siguientes procedimientos para el cálculo de probabilidades binomiales.

STATDISK Seleccione **Analysis** del menú principal, después seleccione la opción **Binomial Probabilities**. Introduzca los valores requeridos de n y p , y aparecerá la distribución de probabilidad completa. Las otras columnas representan las probabilidades acumulativas que se obtienen al sumar los

valores de $P(x)$, conforme sube o baja a lo largo de la columna.

MINITAB Primero introduzca la columna C1 de los valores x de los que desea las probabilidades (tales como 0, 1, 2, 3, 4), después seleccione **Calc** del menú principal y proceda a seleccionar los elementos **Probability Distributions** y **Binomial**. Seleccione **Probabilities**, introduzca el número de ensayos, la probabilidad de éxito y C1 en la columna de entrada; después haga clic en **OK**.

EXCEL Liste los valores de x en la columna A (tales como 0, 1, 2, 3, 4). Haga clic en la celda B1, luego en f_x de la barra de herramientas y seleccione la categoría de función **Statistical** y luego **BINOMDIST**. En el cuadro de diálogo introduzca A1 para el número de éxitos, introduzca el número de ensayos, la probabilidad y 0 para la distribución binomial (en vez de 1 para la distribución binomial acumulativa). Debe aparecer un valor en la celda B1. Haga clic y arrastre la esquina derecha inferior de la celda B1 hacia abajo de la columna para emparejarla con los datos de la columna A, después suelte el botón del ratón.

TI-83/84 PLUS Presione **2nd VARS** (para obtener **DISTR**, que denota “distribuciones”), después seleccione la opción identificada por **binompdf**. Complete **binompdf** (**n,p,x**) con los valores específicos de n , p y x , después presione **ENTER** y el resultado será la probabilidad de obtener x éxitos en n ensayos.

También podría elegir **binompdf** (**n,p,x**) para obtener una lista de *todas* las probabilidades correspondientes a $x = 1, 2, \dots, n$. Puede almacenar esta lista en L2 si presiona **STO → L2**. Después podría introducir los valores de 0, 1, 2, ..., n en la lista L1, lo cual le permitiría calcular estadísticos (con **STAT**, **CALC** y luego **L1, L2**) o ver la distribución en formato de tabla (presionando **STAT** y luego **EDIT**).

El comando **binomcdf** da probabilidades *acumulativas* a partir de una distribución binomial. El comando **binomcdf(n, p, x)** da la suma de todas las probabilidades, desde $x = 0$ hasta el valor específico indicado para x .



5-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Notación.** Si utilizamos la distribución de probabilidad binomial para analizar conjeturas en un examen de opción múltiple, ¿qué error cometeríamos si usamos p para denotar la probabilidad de obtener una respuesta correcta y x para contar el número de respuestas incorrectas?
- Independencia.** Suponga que deseamos utilizar la distribución de probabilidad binomial para analizar el género cuando se seleccionan 12 sujetos al azar de una población grande de jueces potenciales. Si la selección se realiza sin reemplazo, ¿las selecciones son independientes? ¿Podemos tratar las selecciones como si fueran independientes para utilizar la distribución de probabilidad binomial?
- Tabla A-1.** Puesto que las probabilidades binomiales de la tabla A-1 son muy fáciles de calcular, ¿por qué no utilizamos esa tabla cada vez que necesitamos calcular una probabilidad binomial?

- 4. Probabilidades binomiales.** Al tratar de calcular la probabilidad de obtener exactamente dos 6 al lanzar un dado cinco veces, ¿por qué no es posible obtener la respuesta de la siguiente manera?: Usar la regla de la multiplicación para calcular la probabilidad de obtener dos 6, seguidos por los tres resultados que no son 6, que es $(1/6)(1/6)(5/6)(5/6)(5/6)$?

Identificación de distribuciones binomiales. En los ejercicios 5 a 12, determine si el procedimiento indicado produce una distribución binomial. En los casos en que las distribuciones no sean binomiales, identifique al menos un requisito que no se cumpla.

5. Seleccionar al azar a 12 jueces y registrar su nacionalidad
6. Encuestar a 12 miembros del jurado y registrar si responden de manera negativa cuando se les pregunta si han sido condenados por un delito
7. Tratar a 50 fumadores con Nicorette y preguntarles cómo sienten su boca y garganta
8. Tratar a 50 fumadores con Nicorette y registrar si responden de manera afirmativa cuando se les pregunta si sienten algún malestar en la boca o en la garganta
9. Registrar el género de 250 bebés recién nacidos
10. Registrar el número de hijos en 250 familias
11. Encuestar a 250 parejas casadas y registrar si responden de manera afirmativa cuando se les pregunta si tienen hijos
12. Determinar si 500 desfibriladores son aceptables o están defectuosos
13. **Cálculo de probabilidades con respuestas de adivinación.** Cada pregunta de opción múltiple tiene cinco posibles respuestas(a, b, c, d y e), una de las cuales es la correcta. Suponga que adivina las respuestas de tres de estas preguntas.
 - a. Utilice la regla de la multiplicación para calcular la probabilidad de que las dos primeras conjeturas sean incorrectas y que la tercera sea correcta. Es decir, calcule $P(IIC)$, donde C denota una respuesta correcta e I una incorrecta.
 - b. Inicie con IIC y haga una lista completa de los distintos acomodos posibles de dos respuestas incorrectas y una correcta; después calcule la probabilidad de cada dato en la lista.
 - c. Con base en los resultados anteriores, ¿cuál es la probabilidad de tener exactamente una respuesta correcta cuando se hacen tres conjeturas?
14. **Cálculo de probabilidades con respuestas de adivinación.** Un examen consiste en preguntas de opción múltiple con 4 respuestas posibles (a, b, c y d), una de las cuales es la correcta. Suponga que adivina las respuestas a seis de estas preguntas.
 - a. Utilice la regla de la multiplicación para calcular la probabilidad de que las dos primeras conjeturas sean incorrectas y que las cuatro últimas sean correctas. Es decir, calcule $P(IICCCC)$, donde C denota una respuesta correcta e I una incorrecta.
 - b. Inicie con IICCCC y haga una lista completa de los distintos acomodos posibles de dos respuestas incorrectas y cuatro correctas; después calcule la probabilidad de cada dato en la lista.
 - c. Con base en los resultados anteriores, ¿cuál es la probabilidad de tener exactamente cuatro respuestas correctas cuando se hacen seis adivinaciones?

Uso de la tabla A-1. En los ejercicios 15 a 20, suponga que un procedimiento produce una distribución binomial con un ensayo repetido n veces. Utilice la tabla A-1 para calcular la probabilidad de x éxitos, dada la probabilidad p de éxito en un ensayo dado.

15. $n = 3, x = 0, p = 0.05$

16. $n = 4, x = 3, p = 0.30$

- 17.** $n = 8, x = 4, p = 0.05$ **18.** $n = 8, x = 7, p = 0.20$
19. $n = 14, x = 2, p = 0.30$ **20.** $n = 15, x = 12, p = 0.90$

Uso de la fórmula de probabilidad binomial. En los ejercicios 21 a 24, suponga que un procedimiento produce una distribución binomial con un ensayo repetido n veces. Utilice la fórmula de probabilidad binomial para calcular la probabilidad de x éxitos, dada la probabilidad p de éxito en un solo ensayo.

- 21.** $n = 5, x = 2, p = 0.25$ **22.** $n = 6, x = 4, p = 0.75$
23. $n = 9, x = 3, p = 1/4$ **24.** $n = 10, x = 2, p = 2/3$

Uso de resultados de computadora. En los ejercicios 25 a 28, remítase a la pantalla de Minitab que aparece abajo. Las probabilidades se obtuvieron al introducir los valores de $n = 6$ y $p = 0.167$. En una prueba del fármaco Lipitor, el 16.7% de los sujetos tratados con 10 mg de atorvastatin tuvieron dolor de cabeza (según datos de Parke-Davis). En cada caso, suponga que se selecciona a 6 sujetos al azar, los cuales fueron tratados con 10 mg de atorvastatin, y calcule la probabilidad indicada.

MINITAB

Binomial with $n = 6$ and
 $p = 0.167000$

x	$P(X = x)$
0.00	0.3341
1.00	0.4019
2.00	0.2014
3.00	0.0538
4.00	0.0081
5.00	0.0006
6.00	0.0000

- 25.** Calcule la probabilidad de que al menos cinco de los sujetos tengan dolor de cabeza. ¿Es infrecuente que al menos cinco de los seis sujetos tengan dolor de cabeza?
- 26.** Calcule la probabilidad de que, a lo sumo, dos sujetos tengan dolor de cabeza. ¿Es infrecuente que a lo sumo dos de seis sujetos tengan dolor de cabeza?
- 27.** Calcule la probabilidad de que más de un sujeto tenga dolor de cabeza. ¿Es infrecuente que más de uno de seis sujetos tengan dolor de cabeza?
- 28.** Calcule la probabilidad de que al menos un sujeto tenga dolor de cabeza. ¿Es infrecuente que al menos uno de seis sujetos tenga dolor de cabeza?
- 29. Encuestas a televidentes.** El programa de televisión *60 minutos*, de la CBS, ha sido exitoso por muchos años. Recientemente tuvo un índice de audiencia de 20, lo que significa que de todos los televisores encendidos, el 20% estaban sintonizados en *60 minutos* (según datos de Nielsen Media Research). Suponga que un anunciante desea verificar ese valor del 20% realizando su propia encuesta, y que inicia una encuesta piloto con 10 hogares que tienen el televisor encendido en el momento en que se transmite el programa *60 minutos*.
- a. Calcule la probabilidad de que ninguno de los hogares esté sintonizando *60 minutos*.
b. Calcule la probabilidad de que al menos uno de los hogares esté sintonizando *60 minutos*.

- c. Calcule la probabilidad de que a lo sumo uno de los hogares esté sintonizando *60 minutos*.
 - d. Si a lo sumo un hogar está sintonizando *60 minutos*, ¿será incorrecto el valor de un índice de audiencia del 20%? ¿Por qué?
30. **Auditorías de la IRS.** La Hemingway Financial Company prepara devoluciones de impuestos para individuos. (Su lema: “También escribimos extraordinarias novelas de ficción”). Según el Internal Revenue Service, los individuos que ganan entre \$25,000 y \$50,000 se auditán en una proporción del 1%. La Hemingway Company prepara cinco devoluciones de impuestos para individuos que están en esa categoría de impuestos, y se audita a tres de ellos.
- a. Calcule la probabilidad de que, cuando se seleccione al azar a cinco personas que ganan entre \$25,000 y \$50,000, se audite exactamente a tres de ellas.
 - b. Calcule la probabilidad de que se audite al menos a tres.
 - c. Con base en los resultados anteriores, ¿qué se puede concluir acerca de los clientes de Hemingway? ¿Sólo son desafortunados o están siendo blanco de las auditorías?
31. **Muestreo de aceptación.** La compañía Medassist Pharmaceutical Company recibe grandes embarques de tabletas de aspirina y usa el siguiente plan de muestreo de aceptación: seleccionar al azar y probar 24 tabletas, después aceptar el grupo completo sólo si hay una o cero tabletas que no cumplan con las especificaciones requeridas. Si un embarque particular de miles de tabletas de aspirina tiene en realidad una tasa de defectos del 4%, ¿cuál es la probabilidad de que el embarque completo sea aceptado?
32. **Programas de acción afirmativa.** Se realizó un estudio para determinar si existían diferencias significativas entre estudiantes de medicina aceptados a través de programas especiales (como el de acción afirmativa) y estudiantes de medicina aceptados a través de los criterios regulares de admisión. Se encontró que el 94% de los estudiantes de medicina aceptados a través de programas especiales se graduaron (según datos del *Journal of the American Medical Association*).
- a. Si se seleccionan al azar 10 de los estudiantes de los programas especiales, calcule la probabilidad de que al menos nueve se gradúen.
 - b. ¿Sería infrecuente que de 10 estudiantes de los programas especiales, seleccionados al azar, solamente se graduaran siete? ¿Por qué?
33. **Vuelos sobresaturados.** Air America tiene la política de registrar tantas como 15 personas en un avión en el que sólo hay cupo para 14. (Estudios anteriores han revelado que sólo el 85% de los pasajeros registrados llegan para tomar el vuelo). Calcule la probabilidad de que, si Air America registra a 15 personas, no haya suficientes asientos disponibles. ¿Será la probabilidad lo suficientemente baja para que la sobreventa no sea un problema real para los pasajeros?
34. **Máquina tragamonedas del autor.** El autor compró una máquina tragamonedas configurada de tal forma que existe una probabilidad de 1/2000 de ganarse el premio mayor en cualquier ensayo individual. Aun cuando nadie consideraría seriamente hacer trampa al autor, suponga que un invitado afirma haber jugado con la máquina cinco veces y haber ganado en dos ocasiones.
- a. Calcule la probabilidad de exactamente dos premios en cinco ensayos.
 - b. Calcule la probabilidad de al menos dos premios en cinco ensayos.
 - c. ¿Parece ser válida la afirmación del invitado de dos triunfos en cinco juegos? Explique.
35. **Identificación de la discriminación por género.** Después de ser rechazada para un empleo, Kim Kelly se entera que la Bellevue Advertising Company sólo contrató a dos mujeres entre los últimos 20 empleados nuevos. También se entera de que el grupo de solicitantes es muy grande y que incluye un número aproximadamente igual de hombres y mujeres calificados. Ayúdale a presentar cargos por discriminación por género, calculando la probabilidad de que dos o menos mujeres sean incluidas en una contratación de 20 personas, suponiendo que no existe discriminación basada en el género. ¿Apoya la probabilidad resultante esos cargos?

36. Mejoría de la calidad. La empresa Write Right Company fabrica bolígrafos y ha estado registrando una tasa del 5% de bolígrafos defectuosos. Se hacen modificaciones al proceso de manufactura para mejorar la calidad, y el gerente afirma que el procedimiento modificado es mejor, ya que una prueba de 50 bolígrafos indica que sólo uno está defectuoso:

- Suponiendo que la tasa de defectos del 5% no ha cambiado, calcule la probabilidad de que, en 50 bolígrafos, exactamente uno esté defectuoso.
- Suponiendo que la tasa de defectos del 5% no ha cambiado, calcule la probabilidad de que, en 50 bolígrafos, ninguno esté defectuoso.
- ¿Qué valor de probabilidad se debe usar para determinar si el proceso modificado produce una tasa de defectos menor al 5%?
- ¿Qué concluye usted acerca de la efectividad del proceso de producción modificado?

5-3 MÁS ALLÁ DE LO BÁSICO

37. Distribución geométrica. Si un procedimiento cumple con todas las condiciones de una distribución binomial, excepto que el número de ensayos no es fijo, entonces se puede utilizar una **distribución geométrica**. La probabilidad de obtener el primer éxito en el ensayo x -ésimo está dada por $P(x) = p(1 - p)^{x-1}$, donde p es la probabilidad de éxito en cualquier ensayo. Suponga que la probabilidad de un componente de computadora defectuoso es de 0.2. Calcule la probabilidad de que el primer defecto se descubra en el séptimo componente probado.

38. Distribución hipergeométrica. Si realizamos un muestreo sin reemplazo de una población finita pequeña, no debe usarse la distribución binomial porque los eventos no son independientes. Si el muestreo se hace sin reemplazo y los resultados pertenecen a uno de dos tipos, podemos usar la **distribución hipergeométrica**. Si una población tiene A objetos de un tipo, mientras que los objetos B restantes son de otro tipo, y si se muestrean sin reemplazo n objetos, entonces la probabilidad de obtener x objetos del tipo A y $n - x$ objetos del tipo B es

$$P(x) = \frac{A!}{(A-x)!x!} \cdot \frac{B!}{(B-n+x)!(n-x)!} \div \frac{(A+B)!}{(A+B-n)!n!}$$

En la Lotería 54, un participante selecciona seis números del 1 al 54 (sin repetición); después se selecciona al azar una combinación de seis números ganadores. Calcule la probabilidad de obtener

- los seis números ganadores.
- exactamente cinco de los números ganadores.
- exactamente tres de los números ganadores.
- ningún número ganador.

39. Distribución multinomial. La distribución binomial se aplica únicamente a casos que implican dos tipos de resultados, mientras que la **distribución multinomial** supone más de dos categorías. Suponga que tenemos tres tipos de resultados mutuamente excluyentes, denotados por A, B y C. Sean $P(A) = p_1$, $P(B) = p_2$ y $P(C) = p_3$. En n ensayos independientes, la probabilidad de x_1 resultados tipo A, x_2 resultados tipo B y x_3 resultados tipo C está dada por

$$\frac{n!}{(x_1!)(x_2!)(x_3!)} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3}$$

Un experimento en genética incluye seis genotipos mutuamente excluyentes identificados como A, B, C, D, E y F, todos igualmente probables. Si se prueba a 20 descendientes, calcule la probabilidad de obtener exactamente cinco A, cuatro B, tres C, dos D, tres E y tres F, al expandir la expresión anterior de manera que se aplique a seis tipos de resultados y no sólo a tres.

Media, varianza y desviación estándar para la distribución binomial

Concepto clave En la sección 5-3 se estudió la distribución de probabilidad binomial; en la presente sección estudiaremos características importantes de una distribución binomial, incluyendo medidas de tendencia central, medidas de variación y la naturaleza de la distribución. Es decir, dada una distribución de probabilidad binomial en particular, presentaremos métodos para calcular su media, varianza y desviación estándar. Al igual que en secciones anteriores, el objetivo no consiste simplemente en calcular esos valores, sino también en *interpretarlos y entenderlos*.

En la sección 5-2 incluimos métodos para calcular la media, la varianza y la desviación estándar de un distribución de probabilidad discreta. Puesto que una distribución binomial es un tipo especial de distribución de probabilidad, podríamos utilizar las fórmulas 5-1, 5-3 y 5-4 (de la sección 5-2) para calcular la media, la varianza y la desviación estándar, pero esas fórmulas pueden simplificarse mucho para las distribuciones binomiales, como se observará a continuación.

Para cualquier distribución de probabilidad discreta

$$\text{Fórmula 5-1} \quad \mu = \sum [x \cdot P(x)]$$

$$\text{Fórmula 5-3} \quad \sigma^2 = \sum [x^2 \cdot P(x)] - \mu^2$$

$$\text{Fórmula 5-4} \quad \sigma = \sqrt{\sum [x^2 \cdot P(x)] - \mu^2}$$

Para distribuciones binomiales

$$\text{Fórmula 5-6} \quad \mu = np$$

$$\text{Fórmula 5-7} \quad \sigma^2 = npq$$

$$\text{Fórmula 5-8} \quad \sigma = \sqrt{npq}$$

Al igual que en las secciones anteriores, es adecuado calcular los valores para μ y σ , pero es especialmente importante *interpretar y entender* tales valores, de manera que la regla práctica del intervalo resulta muy útil. Por medio de la regla práctica del intervalo podemos considerar que los valores son infrecuentes si caen fuera de los límites que se obtienen de la siguiente manera:

$$\begin{aligned} &\text{valor máximo común } \mu + 2\sigma \\ &\text{valor mínimo común } \mu - 2\sigma \end{aligned}$$

EJEMPLO Selección de miembros del jurado En la sección 5-2 incluimos un ejemplo para ilustrar los cálculos de μ y σ . Utilizamos el ejemplo de la variable aleatoria x para representar el número de mexicano-estadounidenses en un jurado de 12 miembros. (Estamos suponiendo que los miembros del jurado se seleccionan de una población en la que el 80% de los habitantes son mexicano-estadounidenses. Véase a tabla 5-3 en la página 207, donde se incluyen los cálculos para las fórmulas 5-1 y 5-4). Utilice las fórmulas 5-6 y 5-8 para calcular la media y la desviación estándar del número de miembros del jurado elegidos de esta población donde el 80% de los habitantes son mexicano-estadounidenses.

SOLUCIÓN Utilizando los valores $n = 12$, $p = 0.8$ y $q = 0.2$, las fórmulas 5-6 y 5-8 se aplican de la siguiente manera

$$\mu = np = (12)(0.8) = 9.6$$

$$\sigma = \sqrt{npq} = \sqrt{(12)(0.8)(0.2)} = 1.4 \quad (\text{redondeado})$$

Si compara estos cálculos con los de la tabla 5-3, es evidente que las fórmulas 5-6 y 5-8 son mucho más fáciles de usar.

La fórmula 5-6 para la media tiene una lógica intuitiva. Si el 80% de una población es mexicano-estadounidense, y se selecciona al azar a 12 personas, esperamos obtener alrededor de $12 \cdot 0.8 = 9.6$ mexicano-estadounidenses. Este resultado se puede generalizar con facilidad como $\mu = np$. La varianza y la desviación estándar no se justifican fácilmente, y omitiremos las complicadas manipulaciones algebraicas que conducen a las fórmulas 5-7 y 5-8. En vez de ello, remítase de nuevo al ejemplo anterior y a la tabla 5-3 para verificar que, para una distribución binomial, las fórmulas 5-6, 5-7 y 5-8 producen los mismos resultados que las fórmulas 5-1, 5-3 y 5-4.



EJEMPLO Selección de un Gran Jurado El problema del capítulo señala que en el condado de Hidalgo, Texas, el 80% de las personas que podían fungir como miembros del jurado eran mexicano-estadounidenses. También se señaló que durante un periodo de 11 años, 870 personas fueron elegidas para integrar un Gran Jurado.

- Suponiendo que se seleccionan grupos de 870 miembros del jurado, calcule la media y la desviación estándar del número de mexicano-estadounidenses.
- Utilice la regla práctica del intervalo para calcular el valor mínimo común y el valor máximo común de mexicano-estadounidenses. Con base en esos números, ¿podemos concluir que el resultado real de 339 mexicano-estadounidenses es *poco común*? ¿Sugiere esto que el proceso de selección discrimina a los mexicano-estadounidenses?

SOLUCIÓN

- Suponiendo que los miembros del jurado se eligieron al azar, tenemos $n = 870$ personas elegidas con $p = 0.80$ y $q = 0.20$. Podemos calcular la media y la desviación estándar del número de mexicano-estadounidenses por medio de las fórmulas 5-6 y 5-8 de la siguiente manera:

$$\begin{aligned}\mu &= np = (870)(0.80) = 696.0 \\ \sigma &= \sqrt{npq} = \sqrt{(870)(0.80)(0.20)} = 11.8\end{aligned}$$

Para grupos de 870 miembros del jurado elegidos al azar, el número medio de mexicano-estadounidenses es 696.0, y la desviación estándar es 11.8.

- Ahora debemos interpretar los resultados para determinar si 339 mexicano-estadounidenses es un resultado que puede ocurrir fácilmente por azar, o si ese resultado es tan poco probable que el proceso de selección parece ser discriminatorio. Usaremos la regla práctica del intervalo de la siguiente manera:

$$\text{valor máximo común: } \mu + 2\sigma = 696.0 + 2(11.8) = 719.6$$

$$\text{valor mínimo común: } \mu - 2\sigma = 696.0 - 2(11.8) = 672.4$$

INTERPRETACIÓN Según la regla práctica del intervalo, los valores se consideran poco comunes si están entre 672.4 y 719.6, de manera que 339 mexicano-estadounidenses es un resultado poco común porque no se encuentra entre esos dos valores. Es muy poco probable que se obtengan 339 mexicano-estadounidenses únicamente por azar. De hecho, la Suprema Corte estableció que el resultado de únicamente 339 mexicano-estadounidenses era evidencia suficiente para deter-

minar que el proceso de selección de un jurado fue sesgado. La decisión del caso *Castaneda contra Partida* se convirtió en un caso judicial importante, y esto se basó en la aplicación de la distribución de probabilidad binomial.

Recuerde que es importante calcular los valores de la media μ y de la desviación estándar σ , pero es especialmente importante ser capaz de *interpretar* esos valores utilizando herramientas como la regla práctica del intervalo para identificar un rango de valores comunes.

5-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Identificación de valores poco comunes.** Si consideramos un experimento en el que hay 100 nacimientos y se registra el género de los bebés, el número medio de niñas es de 50 y la desviación estándar es de 5 niñas. ¿Sería *poco común* que hubiera 70 niñas en 100 nacimientos? ¿Por qué?
- Identificación de valores poco comunes.** Un proceso de manufactura tiene una tasa de defectos del 10%, lo que significa que el 10% de los artículos producidos están defectuosos. Si se producen lotes de 80 artículos, el número medio de defectos por lote es 8.0 y la desviación estándar es 2.7. ¿Sería *poco común* que hubiera sólo cinco defectos en un lote? ¿Por qué?
- Varianza.** Un investigador planea un diseño experimental en el que, cuando se seleccionan al azar grupos de tratamiento de personas, el número medio de mujeres es 3.0 y la desviación estándar es de 1.2 mujeres. ¿Cuál es la varianza? (Expresé la respuesta en las unidades apropiadas).
- Media y desviación estándar.** Un investigador realiza un estudio observacional y luego utiliza los métodos de esta sección para calcular que la media es 5.0 y la desviación estándar es -2.0. ¿Qué es incorrecto en estos resultados?

Cálculo de μ , σ y valores poco comunes. En los ejercicios 5 a 8, suponga que un procedimiento produce una distribución binomial con n ensayos, y que la probabilidad de éxito de un ensayo es p . Utilice los valores de n y p dados para calcular la media μ y la desviación estándar σ . Además, use la regla práctica del intervalo para calcular el valor mínimo común $\mu - 2\sigma$ y el valor máximo común $\mu + 2\sigma$.

5. $n = 200, p = 0.4$

6. $n = 60, p = 0.25$

7. $n = 1492, p = 1/4$

8. $n = 1068, p = 2/3$

- Respuestas de adivinación.** Varios estudiantes de psicología no están preparados para un examen sorpresa de verdadero/falso de 16 preguntas, por lo que todas sus respuestas son conjeturas.
 - Calcule la media y la desviación estándar del número de respuestas correctas de esos estudiantes.
 - ¿Sería poco común que un estudiante aprobara el examen adivinando y que obtenga al menos 10 respuestas correctas? ¿Por qué?
- Respuestas de adivinación.** Varios estudiantes de economía no están preparados para un examen de opción múltiple de 25 preguntas, por lo que todas sus respuestas son conjeturas. Cada pregunta tiene cinco respuestas posibles y sólo una de ellas es correcta.

continúa

- a. ¿Sería poco común que un estudiante aprobara el examen adivinando y que obtenga al menos 10 respuestas correctas? ¿Por qué?
- b. ¿Sería poco común que un estudiante apruebe el examen adivinando y que obtenga al menos 15 respuestas correctas? ¿Por qué?
- 11. ¿El 20% de los dulces M&M son anaranjados?** Mars, Inc., afirma que el 20% de sus dulces M&M son anaranjados; se selecciona al azar una muestra de 100 de esos dulces.
- Calcule la media y la desviación estándar del número de dulces anaranjados en grupos de 100 como éste.
 - El conjunto de datos 13 del apéndice B consiste en una muestra aleatoria de 100 M&M, de los cuales 25 son anaranjados. ¿Es poco común este resultado? ¿Será incorrecta la tasa del 20%?
- 12. ¿El 14% de los dulces M&M son amarillos?** Mars, Inc., afirma que el 14% de sus dulces M&M son amarillos; se selecciona al azar una muestra de 100 de estos dulces.
- Calcule la media y la desviación estándar del número de dulces amarillos en grupos de 100 como éste.
 - El conjunto de datos 13 del apéndice B consiste en una muestra aleatoria de 100 M&M, de los cuales 8 son amarillos. ¿Es poco común este resultado? ¿Será incorrecta la tasa del 14%?
- 13. Selección del género.** En una prueba del método MicroSort de selección del género, un grupo de parejas que desean tener niñas dio a luz a 325 bebés, de los cuales 295 fueron niñas (según datos del Genetics & IVF Institute).
- Si el método de selección del género no tiene efecto, y los niños y las niñas tienen las mismas probabilidades de nacer, calcule la media y la desviación estándar del número de niñas nacidas en grupos de 325 bebés.
 - ¿El resultado de 295 niñas es poco frecuente? ¿Sugiere que el método de selección del género parece ser efectivo?
- 14. Selección del género.** En una prueba del método MicroSort de selección del género, un grupo de parejas que desean tener niños dio a luz a 51 bebés, de los cuales 39 son varones (según datos del Genetics & IVF Institute).
- Si el método de selección del género no tiene efecto, y los niños y las niñas tienen las mismas probabilidades de nacer, calcule la media y la desviación estándar del número de varones nacidos en grupos de 51 bebés.
 - ¿El resultado de 39 niños es poco frecuente? ¿Sugiere que el método de selección del género parece ser efectivo?
- 15. Mensajes descifrados.** La Central Intelligence Agency (CIA) tiene especialistas que analizan la frecuencia de letras del alfabeto, en un intento por descifrar mensajes interceptados. En un texto estándar en inglés, la letra *r* se utiliza en una proporción del 7.7%.
- Calcule la media y la desviación estándar del número de veces que la letra *r* aparecerá en una página común de 2600 caracteres.
 - En un mensaje interceptado que iba hacia Irak, se encontró que en una página con 2600 caracteres la letra *r* aparecía 175 veces. ¿Es esto infrecuente?
- 16. Genética mendeliana.** Cuando Mendel realizó sus famosos experimentos de genética con guisantes, una muestra consistió en 580 vástagos; Mendel teorizó que el 25% de ellos darían guisantes amarillos.
- Si la teoría de Mendel es correcta, calcule la media y la desviación estándar del número de plantas de guisantes amarillos en grupos de 580 vástagos de guisantes.
 - Los resultados reales consistieron en 152 plantas de guisantes amarillos. ¿Es poco común este resultado? ¿Qué sugiere este resultado acerca de la teoría de Mendel?

- 17. Votación.** En una elección presidencial pasada, la tasa de votos emitidos fue del 61%. En una encuesta, se preguntó a 1002 sujetos si habían votado en la elección presidencial.
- Calcule la media y la desviación estándar del número de votantes reales en grupos de 1002 individuos.
 - En la encuesta de 1002 personas, 701 *dijeron* que habían votado en la última elección presidencial (según datos de ICT Research Group). ¿Es consistente este resultado con la tasa real de votos emitidos, o es improbable que ocurra en una tasa real del 61%? ¿Por qué?
 - Con base en esos resultados, ¿parece que los resultados exactos de votación se pueden obtener preguntando a los votantes cómo actuaron?
- 18. Teléfonos celulares y cáncer cerebral.** En un estudio de 420,095 usuarios de teléfono celular en Dinamarca, se encontró que 135 desarrollaron cáncer cerebral o del sistema nervioso. Si suponemos que este tipo de cáncer no se ve afectado por los teléfonos celulares, la probabilidad de que una persona adquiera esta enfermedad es de 0.000340.
- Suponiendo que los teléfonos celulares no están relacionados con el cáncer, calcule la media y la desviación estándar del número de personas, en grupos de 420,095, que pueden esperar tener cáncer cerebral o del sistema nervioso.
 - Con base en los resultados del inciso a), ¿será poco común que, entre 420,095 personas, existan 135 casos de cáncer cerebral o del sistema nervioso? ¿Por qué?
 - ¿Qué sugieren estos resultados sobre la preocupación pública de que los teléfonos celulares son dañinos para la salud porque incrementan el riesgo de cáncer cerebral o del sistema nervioso?
- 19. Fármaco que reduce el colesterol.** En una prueba clínica del Lipitor (atorvastatin), un fármaco común utilizado para disminuir el colesterol, 863 pacientes recibieron un tratamiento de 10 mg de tabletas de atorvastatin. Este grupo incluyó a 19 pacientes que experimentaron síntomas de gripe (según datos de Pfizer, Inc.). La probabilidad de que una persona que no recibe tratamiento alguno presente síntomas de gripe es de 0.019.
- Suponiendo que el Lipitor no tiene efectos sobre los síntomas de la gripe, calcule la media y la desviación estándar del número de personas en grupos de 863 individuos que se esperaría presentaran estos síntomas.
 - Con base en resultados del inciso a), ¿será poco común encontrar que, de 863 personas, 19 experimenten síntomas de gripe? ¿Por qué?
 - Con base en los resultados anteriores, ¿parece que los síntomas de gripe son una reacción adversa que debe preocupar a los usuarios de Lipitor?
- 20. Prueba de la terapia de contacto.** Emily Rosa, de 9 años de edad, realizó la siguiente prueba: un terapeuta profesional de contacto colocó sus manos a través de un separador de cartón y Emily lanzó una moneda para elegir al azar una de las manos. La niña colocaba su mano arriba de la mano del terapeuta, quien debía identificar la mano que Emily había elegido. Los terapeutas de contacto creen que son capaces de sentir el campo de energía y, en consecuencia, pensaban que podrían identificar la mano que Emily había seleccionado. La prueba se repitió 280 veces. (Según datos de “A Close Look at Therapeutic Touch” de Rosa *et al.*, *Journal of the American Medical Association*, vol. 279, núm. 13).
- Suponiendo que el terapeuta no posee poderes especiales y que hace conjjeturas, calcule la media y la desviación estándar del número de respuestas correctas en grupos de 280 pruebas.
 - Los terapeutas profesionales de contacto identificaron la mano correcta 123 ocasiones en las 280 pruebas. ¿Es poco común este resultado? ¿Qué sugiere el resultado sobre la habilidad de los terapeutas de contacto para elegir la mano correcta al sentir el campo de energía?

5-4 MÁS ALLÁ DE LO BÁSICO

- 21. Uso de la regla empírica.** Se diseña un experimento para probar la eficacia del método de selección de género MicroSort, y 100 parejas intentan concebir niñas por medio de este método. Suponga que los niños y las niñas son igualmente probables, y también suponga que el método de selección de género no tiene efecto alguno.
- Con el uso de los métodos de esta sección, ¿cuáles son los valores máximos y mínimos comunes de niñas en grupos de 100 bebés elegidos al azar?
 - La regla empírica (véase la sección 3-3) se aplica a distribuciones normales. ¿La distribución de probabilidad binomial de este experimento es (aproximadamente) normal? ¿Cómo lo sabe?
 - Suponiendo que la distribución es normal, ¿qué tan probable es que el número de niñas caiga entre 40 y 60 (según la regla empírica)?
- 22. Productos aceptables/defectuosos.** Mario's Pizza Parlor acaba de inaugurar. Debido a la falta de entrenamiento de los empleados, existe sólo un 0.8 de probabilidad de que una pizza sea comestible. Se acaban de ordenar cinco pizzas. ¿Cuál es el número mínimo de pizzas que deben prepararse para estar al menos 99% seguros de que habrá cinco comestibles?

Distribuciones de probabilidad

5-5 de Poisson

Concepto clave Esta sección presenta la *distribución de Poisson*, que es una distribución de probabilidad discreta importante, ya que a menudo se utiliza para describir comportamientos que ocurren en raras ocasiones (con probabilidades pequeñas). Debemos conocer los requisitos para el uso de la distribución de Poisson y también debemos saber cómo calcular probabilidades utilizando la fórmula 5-9. También es importante saber que, cuando se aplica la distribución de Poisson a una variable con una media μ , la desviación estándar es $\sqrt{\mu}$.

La distribución de Poisson se utiliza para describir comportamientos tales como el decaimiento radiactivo, la llegada de pasajeros en una línea, la reproducción de águilas en una región, los pacientes que llegan a la sala de emergencias y los usuarios de Internet que visitan un sitio Web. Por ejemplo, suponga que en el hospital local, la media de los pacientes que llegan a la sala de emergencias los viernes entre las 10:00 P.M. y las 11:00 P.M. es de 2.3. Podemos calcular la probabilidad de que un viernes elegido al azar, entre las 10:00 P.M. y las 11:00 P.M. lleguen exactamente cuatro pacientes. Para ello utilizamos la distribución de Poisson, que se define de la siguiente manera.

Definición

La **distribución de Poisson** es una distribución de probabilidad discreta que se aplica a las ocurrencias de algún suceso *durante un intervalo específico*. La variable aleatoria x es el número de veces que ocurre un suceso en un intervalo. El intervalo puede ser tiempo, distancia, área, volumen o alguna unidad similar. La probabilidad de que el suceso ocurra x veces durante un intervalo está dada por la fórmula 5-9.

Fórmula 5-9 $P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!}$ donde $e \approx 2.71828$

Requisitos de la distribución de Poisson

- La variable aleatoria x es el número de veces que ocurre un suceso *durante un intervalo*.
- Las ocurrencias deben ser *aleatorias*.
- Las ocurrencias deben ser *independientes* entre sí.
- Las ocurrencias deben estar *uniformemente distribuidas* dentro del intervalo empleado.

Parámetros de la distribución de Poisson

- La media es μ .
- La desviación estándar es $\sigma = \sqrt{\mu}$.

La distribución de Poisson difiere de una distribución binomial en estas formas fundamentales:

1. La distribución binomial es afectada por el tamaño de la muestra n y la probabilidad p , mientras que la distribución de Poisson sólo se ve afectada por la media μ .
2. En una distribución binomial, los valores posibles de la variable aleatoria x son $0, 1, \dots, n$, pero los valores posibles x de una distribución de Poisson son $0, 1, 2, \dots$, sin límite superior.

EJEMPLO Bombas de la Segunda Guerra Mundial Al analizar los impactos de las bombas V-1 en la Segunda Guerra Mundial, el sur de Londres se subdividió en 576 regiones, cada una con área de 0.25 km^2 . En total, 535 bombas impactaron el área combinada de 576 regiones.

- a. Si se selecciona al azar una región, calcule la probabilidad de que haya sido impactada exactamente en dos ocasiones.
- b. Con base en la probabilidad calculada en el inciso a), ¿cuántas de las 576 regiones se esperaría que fueran impactadas exactamente dos veces?

SOLUCIÓN

- a. Aplicamos la distribución de Poisson, ya que estamos tratando con las ocurrencias de un suceso (impactos de bombas) dentro de un intervalo (una región con una área de 0.25 km^2). El número medio de impactos por región es

$$\mu = \frac{\text{número de impactos de bomba}}{\text{número de regiones}} = \frac{535}{576} = 0.929$$

Puesto que buscamos la probabilidad de exactamente dos impactos en una región, $x = 2$, y utilizamos la fórmula 5-9 de la siguiente manera:

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{0.929^2 \cdot 2.71828^{-0.929}}{2!} = \frac{0.863 \cdot 0.395}{2} = 0.170$$

La probabilidad de que una región particular sea impactada exactamente dos veces es $P(2) = 0.170$.



Filas

La teoría de las filas es una rama de las matemáticas que usa probabilidad y estadística. El estudio de las filas o líneas de espera es importante para negocios como supermercados, bancos, restaurantes de comida rápida, líneas aéreas y parques de diversiones. Los supermercados Grand Union tratan de mantener filas en las cajas de no más de tres compradores. Wendy's introdujo el sistema "Express Pak" para agilizar el servicio a los numerosos clientes que atiende en sus automóviles. Disney realiza extensos estudios de filas en sus parques de diversiones para poder mantener contentos a sus visitantes y planear su expansión. Los laboratorios Bell aplican la teoría de las filas para optimizar el uso de las redes telefónicas, en tanto que las fábricas la emplean para diseñar líneas de producción eficientes.

continúa

- b. Puesto que existe una probabilidad de 0.170 de que una región sea impactada exactamente dos veces, esperamos que entre las 576 regiones, el número de regiones impactadas exactamente dos veces sea $576 \cdot 0.170 = 97.9$.

En el ejemplo anterior también podemos calcular las probabilidades y los valores esperados para 0, 1, 2, 3, 4 y 5 impactos. (Nos detenemos en $x = 5$ porque ninguna región fue impactada más de cinco ocasiones, y las probabilidades de $x > 5$ son 0.000, cuando se redondea a tres decimales). Esas probabilidades y valores esperados se mencionan en la tabla 5-5. La cuarta columna de la tabla 5-5 describe los resultados reales de la Segunda Guerra Mundial. Hubo 229 regiones sin impactos, 211 regiones que fueron impactadas una vez, y así sucesivamente. Ahora podemos comparar las frecuencias *predichas* por medio de la distribución de Poisson (tercera columna) con las frecuencias *reales* (cuarta columna), para concluir que existe una coincidencia muy alta. En este caso, la distribución de Poisson sirve para predecir los resultados que ocurrieron en realidad. (La sección 11-2 describe un procedimiento estadístico para determinar si tales frecuencias esperadas constituyen un buen “ajuste” de las frecuencias reales. Ese procedimiento sugiere que, en este caso, existe un buen ajuste).

Distribución de Poisson como aproximación de la distribución binomial

En ocasiones, la distribución de Poisson se utiliza para aproximar la distribución binomial, cuando n es grande y p es pequeña. Una regla práctica consiste en utilizar una aproximación como éstas cuando se satisfacen las siguientes dos condiciones.

Requisitos para utilizar la distribución de Poisson como una aproximación a la distribución binomial

1. $n \geq 100$
2. $np \leq 10$

Si se cumplen estas condiciones y deseamos utilizar la distribución de Poisson como aproximación de la distribución binomial, necesitamos un valor de μ , y ese valor puede calcularse utilizando la fórmula 5-6 (presentada originalmente en la sección 5-4):

Fórmula 5-6

$$\mu = np$$

Tabla 5-5 Impactos de bombas V-1 en 576 regiones del sur de Londres

Número de impactos de bomba	Probabilidad	Número esperado de regiones	Número real de regiones
0	0.395	227.5	229
1	0.367	211.4	211
2	0.170	97.9	93
3	0.053	30.5	35
4	0.012	6.9	7
5	0.002	1.2	1

EJEMPLO Juego Pick 4 de Kentucky En el juego Pick 4 de Kentucky, usted paga \$1 para seleccionar una secuencia de cuatro dígitos, como 2283. Si participa en este juego una vez al día, calcule la probabilidad de ganar exactamente una vez en 365 días.

SOLUCIÓN El intervalo de tiempo es de 365 días, así que $n = 365$. Puesto que existe un conjunto ganador de números entre los 10,000 posibles (del 0000 al 9999), $p = 1/10,000$. Se satisfacen las condiciones $n \geq 100$ y $np \leq 10$, de manera que podemos utilizar la distribución de Poisson como aproximación de la distribución binomial. Primero necesitamos el valor de μ , que se calcula de la siguiente manera:

$$\mu = np = 365 \cdot \frac{1}{10,000} = 0.0365$$

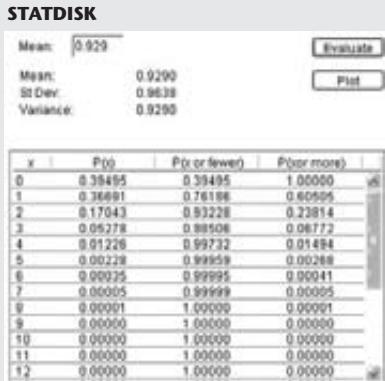
Luego de calcular el valor de μ , podemos calcular $P(1)$:

$$P(1) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{0.0365^1 \cdot 2.71828^{-0.0365}}{1!} = \frac{0.0352}{1} = 0.0352$$

Si aplicamos la distribución de Poisson como aproximación de la distribución binomial, encontramos que existe una probabilidad de 0.0352 de ganar exactamente una vez en 365 días. Si utilizamos la distribución binomial, nuevamente obtenemos 0.0352, de manera que observamos que la aproximación de Poisson es bastante buena aquí. (Si usamos más decimales, vemos que la aproximación de Poisson produce 0.03519177, y el resultado binomial más exacto es 0.03519523).

Uso de la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, después seleccione **Poisson Probabilities** y proceda a introducir el valor de la media μ . Haga clic en el botón **Evaluate** y desplace la pantalla donde aparecen los valores que no se ajustan a la ventana inicial. Vea la siguiente pantalla de Statdisk que utiliza la media de 0.929 del primer ejemplo de esta sección.



MINITAB Primero introduzca el valor deseado de x en la columna C1. Ahora seleccione **Calc** de la barra del menú principal, luego **Probability Distributions** y finalmente **Poisson**. Introduzca el valor de la media μ y C1 en la columna de entrada.

EXCEL Haga clic en **fx** de la barra del menú principal, después seleccione la categoría **Statistical**, luego seleccione **POISSON** y haga clic en **OK**. En el cuadro del diálogo introduzca los valores de x y la media, luego introduzca 0 para “Cumulative”. (Introducir 1 en “Cumulative” da como resultado la probabilidad de los valores hasta el valor introducido de x inclusive).

TI-83/84 PLUS Presione **2nd VARS** (para obtener **DISTR**), después seleccione la opción **B: poissonpdf**. Ahora presione **ENTER** y después proceda a introducir μ , x (incluida la coma). Para μ , introduzca el valor de la media; para x introduzca el número deseado de ocurrencias.

5-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Distribución de Poisson.** ¿Cuáles son las condiciones necesarias para usar la distribución de Poisson?
- Distribución de Poisson.** La variable aleatoria x representa el número de llamadas telefónicas recibidas en una hora, y tiene una distribución de Poisson con una media de 9. ¿Cuál es la desviación estándar? ¿Cuál es su varianza?
- Parámetros.** Al tratar de aplicar la distribución de Poisson, ¿cuáles de los siguientes elementos debemos conocer: la media, la desviación estándar, la varianza, la forma de la distribución?
- Poisson/binomial.** Un experimento implica lanzar un dado 6 veces y contar el número de veces que resulta un 2. Si calculamos la probabilidad de $x = 0$ ocurrencias de 2 por medio de la distribución de Poisson, obtenemos 0.368; sin embargo, con la distribución binomial obtenemos 0.335. ¿Cuál es la probabilidad correcta de no obtener ningún 2 cuando lanzamos un dado 6 veces? ¿Por qué la otra probabilidad es incorrecta?

Uso de una distribución de Poisson para calcular la probabilidad. En los ejercicios 5 a 8, suponga que se puede aplicar la distribución de Poisson y proceda a emplear la media dada para calcular la probabilidad indicada.

- | | |
|--|---|
| <ol style="list-style-type: none"> Si $\mu = 5$, calcule $P(4)$. Si $\mu = 0.5$, calcule $P(3)$. | <ol style="list-style-type: none"> Si $\mu = 3/4$, calcule $P(2)$. Si $\mu = 3.25$, calcule $P(5)$. |
|--|---|

En los ejercicios 9 a 16, utilice la distribución de Poisson para calcular las probabilidades indicadas.

- Dientes de león.** Los dientes de león se estudian para conocer sus efectos sobre los cultivos y el crecimiento del césped. En una región se descubrió que el número medio de dientes de león por metro cuadrado es de 7.0 (según datos de Manitoba Agriculture and Food).
 - Calcule la probabilidad de que no haya dientes de león en una área de 1 m^2 .
 - Calcule la probabilidad de al menos un diente de león en una área de 1 m^2 .
 - Calcule la probabilidad de dos dientes de león, cuando mucho, en una área de 1 m^2 .
- Llamadas telefónicas.** El autor descubrió que en un mes (30 días), hizo 47 llamadas por teléfono celular, las cuales se distribuyeron de la siguiente manera: durante 17 días no hubo llamadas, 7 días hizo una llamada cada día, 2 días hizo 3 llamadas cada día, 2 días hizo 4 llamadas cada día, un día hizo 12 llamadas y otro día hizo 14 llamadas.
 - Calcule la media de llamadas por día.
 - Utilice la distribución de Poisson para calcular la probabilidad de un día sin llamadas y compare el resultado con la frecuencia relativa real del número de días sin llamadas.
 - Utilice la distribución de Poisson para calcular la probabilidad de hacer una llamada en un día y compare el resultado con la frecuencia relativa real del número de días con una llamada.
 - Con base en los resultados anteriores, ¿las llamadas que el autor hizo por teléfono celular en un día parecen ajustarse razonablemente bien a la distribución de Poisson? ¿Por qué?
- Decaimiento radiactivo.** Los átomos radiactivos son inestables porque tienen demasiada energía. Cuando liberan su energía sobrante, se dice que decaen. Al estudiar el cesio 137, se descubre que durante el curso del decaimiento en 365 días, 1,000,000 de átomos radiactivos se reducen a 997,287 átomos radiactivos.
 - Calcule el número medio de átomos radiactivos perdidos durante el decaimiento en un día.
 - Calcule la probabilidad de que en un día dado decaigan 50 átomos radiactivos.

12. Muertes por coches de caballos. Un ejemplo clásico de la distribución de Poisson implica el número de muertes de hombres del ejército prusiano causadas por coches de caballo entre 1875 y 1894. Se combinaron datos de 14 cadáveres durante el periodo de 20 años, y los 280 años-cadáveres incluyeron un total de 196 muertes. Después de calcular el número medio de muertes por año-cadáver, calcule la probabilidad de que un año-cadáver, seleccionado al azar, tenga el siguiente número de muertes:

- a. 0 b. 1 c. 2 d. 3 e. 4

Los resultados reales incluyen las siguientes frecuencias: 0 muertes (en 144 años-cadáveres); 1 muerte (en 91 años-cadáveres); 2 muertes (en 32 años-cadáveres); 3 muertes (en 11 años-cadáveres); 4 muertes (en 2 años-cadáveres). Compare los resultados reales con los esperados de las probabilidades de Poisson. ¿Sirve la distribución de Poisson como una buena herramienta para predecir los resultados reales?

13. Muertes por homicidio. En un año hubo 116 muertes por homicidio en Richmond, Virginia (según “A Classroom Note on the Poisson Distribution: A Model for Homicidal deaths in Richmond, VA for 1991”, de Winston A. Richards en *Mathematics and Computer Education*). Para un día seleccionado al azar, calcule la probabilidad de que el número de muertes por homicidio sea

- a. 0 b. 1 c. 2 d. 3 e. 4

Compare las probabilidades calculadas con los siguientes resultados reales: 268 días (ningún homicidio); 79 días (1 homicidio); 17 días (2 homicidios); 1 día (3 homicidios); no hubo días con más de 3 homicidios.

14. Terremotos. Durante un periodo reciente de 100 años hubo 93 terremotos importantes en el mundo (de al menos 6.0 grados en la escala de Richter) (según datos del *World Almanac and Book of Facts*). Suponiendo que la distribución de Poisson es un modelo adecuado, calcule el número medio de terremotos importantes por año, después calcule la probabilidad de que el número de terremotos en un año seleccionado al azar sea

- a. 0 b. 1 c. 2 d. 3 e. 4
f. 5 g. 6 h. 7

Los resultados reales son: 47 años (ningún terremoto importante); 31 años (1 terremoto importante); 13 años (2 terremotos importantes); 5 años (3 terremotos importantes); 2 años (4 terremotos importantes); 0 años (5 terremotos importantes); 1 año (6 terremotos importantes); 1 año (7 terremotos importantes). Después de comparar las probabilidades calculadas con los resultados reales, ¿es un buen modelo la distribución de Poisson?

5-5 MÁS ALLÁ DE LO BÁSICO

15. Aproximación de Poisson a una distribución binomial. La distribución de Poisson pueden emplearse para aproximar una distribución binomial si $n \geq 100$ y $np \leq 10$. Suponga una distribución binomial con $n = 100$ y $p = 0.1$. Es imposible obtener 101 éxitos con una distribución como ésta, pero podemos calcular la probabilidad de $x = 101$ con la aproximación de Poisson. Calcule ese valor. ¿Qué tanto coincide el resultado con la imposibilidad de que $x = 101$ en una distribución binomial?

16. Aproximación de Poisson a una distribución binomial. Para una distribución binomial con $n = 10$ y $p = 0.5$, no debemos usar la aproximación de Poisson, ya que las condiciones $n \geq 100$ y $np \leq 10$ no se satisfacen. Suponga que de cualquier manera empleamos la aproximación de Poisson. ¿Son aproximaciones inaceptables las probabilidades resultantes? ¿Por qué?

Repaso

El concepto de distribución de probabilidad es un elemento fundamental de la estadística. Una distribución de probabilidad describe la probabilidad de cada valor de una variable aleatoria. Este capítulo incluyó sólo distribuciones de probabilidad discreta, pero los siguientes capítulos abarcarán distribuciones de probabilidad continua. Se estudiaron los siguientes conceptos básicos:

- Una *variable aleatoria* posee valores que están determinados por el azar.
- Una *distribución de probabilidad* consiste en todos los valores de una variable aleatoria, junto con sus probabilidades correspondientes. Una distribución de probabilidad debe cumplir dos requisitos: $\sum P(x) = 1$, para cada valor de x , $0 \leq P(x) \leq 1$.
- Se pueden explorar características importantes de una *distribución de probabilidad* construyendo un histograma de probabilidad y calculando su media y desviación estándar por medio de las siguientes fórmulas:

$$\mu = \sum [x \cdot P(x)]$$

$$\sigma = \sqrt{\sum [x^2 \cdot P(x)] - \mu^2}$$

- En una *distribución binomial*, existen dos categorías de resultados y un número fijo de ensayos independientes con una probabilidad constante. La probabilidad de x éxitos en n ensayos puede calcularse empleando la fórmula de probabilidad binomial, la tabla A-1, un programa de cómputo (como STATIDISK, Minitab o Excel) o una calculadora TI-83/84 Plus.
- En una distribución binomial, la media y la desviación estándar pueden obtenerse fácilmente calculando los valores de $\mu = np$ y $\sigma = \sqrt{npq}$.
- Una *distribución de probabilidad de Poisson* se aplica a ocurrencias de algún suceso durante un intervalo específico; sus probabilidades se pueden calcular con la fórmula 5-9.
- *Resultados infrecuentes*: Este capítulo enfatizó la importancia de interpretar los resultados a través de la distinción entre los resultados que son comunes y aquellos que son poco comunes. Utilizamos dos criterios diferentes: la regla práctica del intervalo y el uso de probabilidades.

Al utilizar la regla práctica del intervalo para identificar valores poco comunes, tenemos:

$$\text{valor máximo común} = \mu + 2\sigma$$

$$\text{valor mínimo común} = \mu - 2\sigma$$

Al utilizar probabilidades para identificar valores poco comunes, tenemos que:

- **Número de éxitos inusualmente alto:** x éxitos en n ensayos es un número *inusualmente alto* de éxitos si $P(x \text{ o más}) \leq 0.05$.*
- **Número de éxitos inusualmente bajo:** x éxitos en n ensayos es un número *inusualmente bajo* de éxitos si $P(x \text{ o menos}) \leq 0.05$.

*El valor de 0.05 se utiliza de forma regular, pero no es absolutamente rígido. Es posible usar otros valores, como 0.01, para distinguir entre sucesos que pueden ocurrir fácilmente por azar y sucesos que tienen muy pocas probabilidades de ocurrir por azar.

Conocimientos estadísticos y pensamiento crítico

- 1. Distribución de probabilidad.** ¿Qué es una distribución de probabilidad?
- 2. Distribución de probabilidad.** ¿Cuáles son los requisitos de una distribución de probabilidad?
- 3. Distribución de probabilidad discreta.** Las distribuciones de probabilidad descritas en este capítulo son *discretas*. ¿Qué las hace *discretas*? ¿Qué otro tipo de distribuciones de probabilidad existe?
- 4. Distribuciones de probabilidad.** En este capítulo se describe el concepto de una distribución de probabilidad discreta, y después se describen las distribuciones de probabilidad binomial y de Poisson. ¿Todas las distribuciones de probabilidad discretas son binomiales o de Poisson? ¿Por qué?

Ejercicios de repaso

- 1. Examen de opción múltiple.** Las preguntas de opción múltiple son fáciles de corregir, así que es común que se utilicen en exámenes para las clases, exámenes SAT, exámenes MCAT para las escuelas de medicina y en muchas otras circunstancias. La tabla al margen describe la distribución de probabilidad del número de respuestas correctas cuando alguien trata de adivinar las respuestas a 10 preguntas de opción múltiple del examen SAT. Cada pregunta tiene 5 respuestas posibles (a, b, c, d y e), una de las cuales es la correcta. Suponga que se hacen conjeturas para responder las 10 preguntas.

x	$P(x)$
0	0.107
1	0.268
2	0.302
3	0.201
4	0.088
5	0.026
6	0.006
7	0.001
8	0+
9	0+
10	0+

- Verifique que la tabla satisfaga los requisitos necesarios para una distribución de probabilidad.
 - Calcule el número medio de respuestas correctas.
 - Calcule la desviación estándar del número de respuestas correctas cuando muchos sujetos diferentes hacen conjeturas para responder las 10 preguntas.
 - Cuál es la probabilidad de que alguien tenga al menos la mitad de las preguntas correctas?
 - Cuando alguien hace conjeturas para las 10 respuestas, ¿cuál es el número esperado de respuestas correctas?
 - ¿Cuál es la probabilidad de tener al menos una respuesta correcta?
 - Si alguien tiene al menos una respuesta correcta, ¿esto significa que la persona sabe algo acerca de la materia del examen?

- 2. Audiencia de televidentes.** El programa de televisión *Cold Case* tiene un índice de audiencia de 15; es decir, mientras se está transmitiendo, el 15% de los televisores encendidos está sintonizando ese programa (según datos de Nielsen Media Research). Un grupo de enfoque consiste en 12 hogares seleccionados al azar (cada uno de ellos con el televisor encendido durante la transmisión del programa *Cold Case*).

- ¿Cuál es el número esperado de televisores sintonizados en el programa *Cold Case*?
- En tales grupos de 12 hogares, ¿cuál es el número medio de televisores que están sintonizando el programa *Cold Case*?
- En tales grupos de 12, ¿cuál es la desviación estándar del número de televisores que están sintonizando el programa *Cold Case*?
- Para tal grupo de 12, calcule la probabilidad de que exactamente tres televisores estén sintonizando el programa *Cold Case*.
- Para tal grupo de 12, ¿sería poco común descubrir que ningún televisor está sintonizando el programa *Cold Case*? ¿Por qué?

- 3. Razones de despido.** “La incapacidad para llevarse bien con otras personas” es la razón que se cita en el 17% de despidos de empleados (según datos de Robert Half International, Inc.). Preocupado por las condiciones de trabajo de su compañía, el gerente de personal de la compañía Boston Finance planea investigar los cinco despidos que ocurrieron durante el año anterior.
- Suponiendo que se aplica la tasa del 17%, calcule la probabilidad de que, de esos cinco empleados, el número de despidos por la incapacidad de llevarse bien con otras personas sea de al menos cuatro.
 - Si el gerente de personal realmente descubre que al menos cuatro de los despidos se deben a la incapacidad de llevarse bien con otras personas, ¿será esta compañía muy diferente de otras compañías típicas? ¿Por qué?
- 4. Muertes.** En la actualidad, un promedio de siete residentes del pueblo de Westport (población 760) mueren cada año (según datos del U.S. National Center for Health Statistics).
- Calcule el número medio de muertes por día.
 - Calcule la probabilidad de que en un día dado no haya muertes.
 - Calcule la probabilidad de que en un día dado haya una muerte.
 - Calcule la probabilidad de que en un día dado haya más de una muerte.
 - Con base en los resultados anteriores, ¿debería Westport tener un plan de contingencia para enfrentar más de una muerte al día? ¿Por qué?

Ejercicio de repaso acumulativo

x	f
0	7
1	14
2	5
3	11
4	8
5	4
6	5
7	6
8	12
9	8

- 1. Pesos: análisis de los últimos dígitos.** La tabla al margen lista los últimos dígitos de los pesos de los sujetos incluidos en el conjunto de datos 1 del apéndice B. Los últimos dígitos de un conjunto de datos en ocasiones se pueden utilizar para determinar si éstos fueron medidos o simplemente reportados. La presencia desproporcionada de los números 0 y 5 suele ser un buen indicador de que los datos fueron reportados y no medidos.
- Calcule la media la desviación estándar de los últimos dígitos.
 - Construya la tabla de frecuencias relativas que corresponda a la tabla de frecuencias dada.
 - Construya una tabla para la distribución de probabilidad de dígitos seleccionados al azar que sean igualmente probables. Liste los valores de la variable aleatoria x ($0, 1, 2, \dots, 9$) junto con sus probabilidades correspondientes ($0.1, 0.1, 0.1, \dots, 0.1$), luego calcule la media y la desviación estándar de esta distribución de probabilidad.
 - Reconociendo que los datos muestrales se desvían naturalmente de los resultados que esperamos en teoría, ¿parece que los últimos dígitos dados coinciden aproximadamente con la distribución que esperaríamos con una selección aleatoria? ¿O al parecer hay algo en los datos muestrales (como una cantidad desproporcionada de números 0 y 5) que sugiere que los últimos dígitos dados no son aleatorios? (En el capítulo 11 presentaremos un método para responder este tipo de preguntas de manera mucho más objetiva).
- 2. Determinación de la eficacia de un programa de prevención del VIH.** El Departamento de Salud del estado de Nueva York reporta una tasa del 10% del virus VIH para la población “en riesgo”. En una región se utiliza un programa educativo intensivo en un intento por disminuir esa tasa del 10%. Después de aplicar programa, se realiza un estudio de seguimiento de 150 individuos en riesgo.

- a. Suponiendo que el programa no tiene efecto, calcule la media en la desviación estándar del número de casos de VIH en grupos de 150 personas en riesgo.
- b. De las 150 personas que participaron en el estudio de seguimiento, el 8% (o 12 personas) resultaron positivas para el virus del VIH. Si el programa no tiene efecto, ¿es extraordinariamente baja esta tasa? ¿Este resultado sugiere que el programa es efectivo?
- 3. Uso de la tarjeta de crédito.** Un alumno del autor realizó una encuesta sobre el uso de la tarjeta de crédito con 25 amigos suyos. Se preguntó a cada sujeto cuántas veces había utilizado una tarjeta de crédito durante los siete días anteriores, y los resultados se muestran como frecuencias relativas en la tabla al margen.
- a. ¿La tabla es una distribución de probabilidad? ¿Por qué?
- b. Suponiendo que la tabla describe una distribución de probabilidad, ¿a qué población representa? ¿Se trata de la población de todos los usuarios de tarjeta de crédito de Estados Unidos?
- c. ¿El tipo de muestreo limita la utilidad de los datos?
- d. Calcule la media de los sujetos que usaron una tarjeta de crédito durante los últimos siete días.
- e. Calcule la desviación estándar, suponiendo que se trata de una tabla de frecuencias relativas que resume los resultados de una muestra de 25 sujetos.

<i>x</i>	<i>Frecuencia relativa</i>
0	0.16
1	0.24
2	0.40
3	0.16
4	0.04

Actividades de cooperación en equipo

1. **Actividad en clase** ¡Gane \$1,000,000! La James Randi Educational Foundation ofrece un premio de \$1,000,000 a quien pueda demostrar “en condiciones de observación apropiadas, evidencias de cualquier poder o suceso paranormal, supernatural u oculto”. Formen grupos de tres estudiantes y seleccionen a uno a quien se le hará una prueba de percepción extrasensorial (PES); la prueba consiste en tratar de identificar correctamente un dígito seleccionado al azar por otro miembro del grupo. El tercer integrante del grupo debe registrar el dígito seleccionado al azar, el dígito adivinado por el sujeto, y si la adivinación fue correcta o incorrecta. Construya la tabla de la distribución de probabilidad para dígitos generados al azar, la tabla de frecuencias relativas para dígitos aleatorios seleccionados realmente y la tabla de frecuencias relativas para las adivinaciones. Después de comparar las tres tablas, ¿qué concluye? ¿Qué proporción de las adivinaciones fue correcta? ¿Parecería que el sujeto tiene la habilidad de seleccionar el dígito correcto significativamente con mayor frecuencia de lo que se esperaría por el azar?
2. **Actividad en clase** Vea la actividad anterior y *diseñe un experimento* que serviría para poner a prueba la aseveración de una persona que afirma tener la habilidad de identificar el color de una carta seleccionada de una baraja estándar. Describa el experimento con mucho detalle. Como está en juego el premio de \$1,000,000, queremos ser cuidadosos para evitar el grave error de concluir que la persona tiene poderes paranormales

cuento en realidad no los tiene. Existe la probabilidad de que el sujeto conjeture y acierte cada vez, por lo que es necesario que identifique una probabilidad que sea razonable para el suceso de que el sujeto pase la prueba adivinando. Asegúrese de diseñar la prueba de manera que esta probabilidad sea igual o menor que el valor de probabilidad elegido.

3. **Actividad en clase** Suponga que deseamos identificar la distribución de probabilidad del número de hijos de parejas elegidas al azar. Para cada estudiante de la clase, registre el número de hermanos y hermanas, así como el número total de hijos (incluyendo al alumno) en cada familia. Construya la tabla de frecuencias relativas con el resultado obtenido. (Los valores de la variable aleatoria x serán 1, 2, 3, . . .) ¿Por qué sería incorrecto utilizar esta tabla de frecuencias relativas como un estimado de la distribución de probabilidad del número de hijos de parejas elegidas al azar?
4. **Actividad fuera de clase** Vea el ejercicio 1 de la sección de ejercicios de repaso acumulativos, el cual sugiere que un análisis de los últimos dígitos de datos en ocasiones revela si éstos fueron realmente medidos o si fueron reportados por los sujetos. Remítase a un almanaque o a Internet y busque un conjunto de datos (como las longitudes de ríos del mundo); luego, analice la distribución de los últimos dígitos para determinar si los valores se obtuvieron por medio de mediciones reales.

Proyecto tecnológico

El vuelo 179 de American Airlines de Nueva York a San Francisco utiliza un Boeing 767-300 con 213 asientos. Como algunas personas con reservación no se presentan, American Airlines acostumbra vender más lugares para el vuelo y aceptar más de 213 reservaciones. Sin esta práctica, la aerolínea perdería utilidades debido a los asientos vacíos, pero si vende demasiados asientos y tiene que rechazar a algunos pasajeros, la aerolínea pierde dinero por la compensación que debe pagarles. Suponga que existe una probabilidad de 0.0995 de que un pasajero con reservación no llegue al vuelo (según datos del trabajo de investigación de IBM “Passenger-Base Predictive Modeling on Airline No-Show Rates”, de Lawrence, Hong y Cherrier). Suponga también que la aerolínea acepta 236 reservaciones para los 213 asientos disponibles.

Calcule la probabilidad de que, cuando el vuelo 179 acepte 236 reservaciones, se presente un número de pasajeros mayor

al de asientos disponibles. Es decir, calcule la probabilidad de que lleguen más de 213 personas con reservación, suponiendo que se aceptaron 236 reservaciones. La tabla A-1 no se puede utilizar debido a los valores implicados, y los cálculos por medio de la fórmula de la probabilidad binomial serían demasiado largos y tediosos. El mejor método consiste en utilizar un programa de cómputo o una calculadora TI-83/84 Plus. Consulte la sección 5-3, donde encontrará las instrucciones para el uso de STATDISK, Minitab, Excel o una calculadora TI-83/84 Plus. ¿La probabilidad de vender más de los lugares disponibles en el vuelo es lo suficientemente pequeña para que no ocurra con mucha frecuencia, o es tan alta que se deben hacer cambios para disminuirla? Ahora utilice el ensayo y error para calcular el número máximo de reservaciones que deben aceptarse para que la probabilidad de tener más pasajeros que asientos sea de 0.05 o menos.

De los datos a la decisión

Pensamiento crítico: Determinación de criterios para concluir que un método de selección del género es efectivo

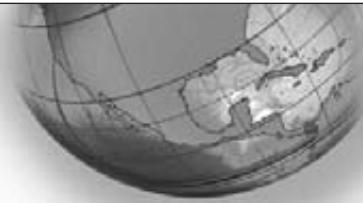
Usted es el responsable de analizar los resultados de una prueba clínica sobre la efectividad de un nuevo método de selección del género. Suponga que ya se estableció un tamaño muestral de $n = 50$ parejas, y que cada pareja tiene un hijo. Además, suponga que cada una de las parejas será sometida a

un tratamiento que, al parecer, incrementa la probabilidad de que su bebé sea niña.

Es muy peligroso obtener primero los resultados y después obtener conclusiones acerca de ellos. Si los resultados se aproximan a mostrar la eficacia de un tratamiento, puede ser tentador concluir que hay un efecto cuando en realidad no existe. Es mejor establecer un criterio *antes* de obtener los resultados. Al emplear los métodos de este capítulo, identifique los criterios que se deben utilizar para concluir que el tratamiento es efectivo para incrementar la probabilidad de que nazca una niña. De los 50

nacimientos, ¿cuántas niñas debería haber para concluir que el procedimiento de selección del género es efectivo? Explique cómo llegó a este resultado.





Proyecto de Internet

Distribuciones de probabilidad y simulaciones

Las distribuciones de probabilidad se utilizan para predecir el resultado de los sucesos que modelan. Por ejemplo, si lanzamos una moneda balanceada, la distribución del resultado es una probabilidad de 0.5 para las caras y 0.5 para las cruces. Si lanzamos la moneda 10 veces consecutivas, esperamos 5 caras y cinco cruces. Tal vez no obtengamos este resultado exacto, pero a la larga, después de cientos o miles de lanzamientos, esperamos que la proporción de caras y cruces sea muy cercana a “50-50”. Visite el sitio de Internet de este libro de texto:

<http://www.pearsoneducacion.net/triola>

Localice el proyecto de Internet del capítulo 5, donde encontrará dos exploraciones. En la primera se le pide crear una distribución de probabilidad para un experimento sencillo y utilizar esa distribución para predecir el resultado de ensayos repetidos del experimento. En la segunda exploración analizaremos una situación más complicada: las rutas de canicas que ruedan, mientras se mueven de forma similar al *pinball* o billar romano, a través de un grupo de obstáculos. En cada caso, una simulación visual dinámica le permitirá comparar los resultados predichos con un conjunto de resultados experimentales.

La estadística en el trabajo

“Nuestro programa en realidad es un programa de educación, pero es ampliamente reconocido porque los resultados se han hecho públicos”.



Barbara Carvalho

Directora de Marist College Poll

Lee Miringoff

Director del Marist College Institute for Public Opinion

Barbara Carvalho y Lee Miringoff reportan los resultados de sus encuestas en muchas entrevistas para medios impresos y electrónicos, incluyendo programas de noticias de NBC, CBS, ABC, FOX y de la televisión pública. Lee Miringoff aparece regularmente en el programa *Today* de la NBC.

¿A qué se dedican?

Realizamos encuestas públicas. Hacemos encuestas sobre asuntos públicos, estimaciones de aprobación de funcionarios públicos en la ciudad de Nueva York, en el estado de Nueva York y a lo largo de toda la nación. No somos partidarios de realizar encuestas para partidos políticos, candidatos políticos o grupos de cabildeo. Recibimos fondos de manera independiente del Marist College y no tenemos ingresos externos que pudieran sugerir que hacemos investigación para algún grupo particular o sobre un tema específico.

¿Cómo seleccionan a los individuos que encuestan?

En una encuesta estatal, seleccionamos a los sujetos en proporción a los registros de votantes de los condados. Los distintos condados tienen diferentes tasas de rechazo y si seleccionáramos al azar a personas en todo el territorio del estado, obtendríamos un modelo desigual de este último. Hacemos estratos por condado y usamos marcación aleatoria de dígitos, de manera que obtenemos números que se incluyen y no se incluyen en el directorio telefónico.

Acaban de mencionar las tasas de rechazo, ¿constituyen éstas un verdadero problema?

Uno de los aspectos que tenemos que enfrentar constantemente es el hecho de que la gente no responde a las encuestas. Este fenómeno se incrementa con el paso del tiempo y recibe mucha atención por parte de la comunidad de investigación por en-

cuesta. Como centro de investigación, nos va bastante bien en comparación con otros. Pero cuando se hacen entrevistas cara a cara y se tienen tasas de rechazo del 25 al 50%, existe una verdadera preocupación por descubrir quién se rehusa, por qué no quiere responder y el efecto que tiene en la representatividad de los estudios que realizamos.

¿Recomendarían a los estudiantes tomar un curso de estadística?

Totalmente. Los números no se crean todos de la misma forma. Sin importar su campo de estudio o sus intereses profesionales, es una gran ventaja poseer la habilidad para evaluar de forma crítica la información de investigaciones que se les presente, utilizar datos para mejorar servicios o interpretar resultados para diseñar estrategias. Las encuestas, en particular, están por todas partes. Es vital que como empleados, gerentes y ciudadanos sea mos capaces de evaluar su precisión y valor. La estadística cubre todas las disciplinas. Los estudiantes se encontrarán con ella inevitablemente en sus carreras en algún momento.

¿Tienen alguna otra recomendación para los estudiantes?

Es importante que los estudiantes aprovechen cualquier oportunidad para desarrollar sus habilidades de comunicación y presentación. No es suficiente mejorar sus habilidades para hablar y escribir, sino que también deben incrementar su nivel de familiaridad con las nuevas tecnologías.



Distribuciones de probabilidad normal

6



- 6-1 Panorama general**
- 6-2 La distribución normal estándar**
- 6-3 Aplicaciones de las distribuciones normales**
- 6-4 Distribuciones muestrales y estimadores**
- 6-5 El teorema del límite central**
- 6-6 La distribución normal como aproximación de la distribución binomial**
- 6-7 Determinación de la normalidad**

¿Cómo identificamos límites de seguridad para los pasajeros?

“Tenemos una emergencia en el vuelo 54-80 de Midwest Air”, dijo la piloto Katie Leslie un momento antes de que su avión chocara en Charlotte, Carolina del Norte. El accidente del avión Beech 1900 cobró la vida de las 21 personas que iban a bordo. Posteriores investigaciones despertaron la sospecha de que el peso de los pasajeros había contribuido al accidente. Esto provocó que la Federal Aviation Administration pidiera a las aerolíneas que reunieran información referente al peso en vuelos elegidos al azar, con el fin de actualizar los antiguos supuestos sobre los pesos de los pasajeros.

Recientemente se hundió un taxi acuático en el Inner Harbor de Baltimore. De las 25 personas a bordo, 5 murieron y 16 resultaron lesionadas. Una investigación reveló que la carga segura de pasajeros del taxi acuático era de 3500 libras. Suponiendo un peso medio de 140 libras por pasajero, el barco tenía permitido llevar 25 pasajeros, pero la media de 140 libras fue determinada hace 44 años, cuando la gente no pesaba tanto como ahora. (Se descubrió que el peso medio de los 25 pasajeros que viajaban en el barco que se hundió era de 168 libras). El National Transportation and Safety Board sugirió que la antigua media estimada de 140 libras se actualizara a 174 libras, de manera que la carga segura de 3500 libras ahora sólo permitiría 20 pasajeros en vez de 25. En este capítulo investigaremos pesos de pasajeros y el papel que desempeñan esos pesos en el establecimiento de límites de carga seguros para los medios de transporte.

Los ejemplos de los accidentes del avión y el taxi acuático ilustran aspectos sumamente importantes que nos afectan a todos. Uno de ellos es el cambio en el peso de la gente a través del tiempo. En el capítulo 2 señalamos que, además de las características centrales, de variación, de distribución y los valores extremos de una población, otro aspecto relevante son los cambios que pueden ocurrir con el paso del tiempo. Resultados de la National Health and Nutrition Examination Survey revelan que los estadounidenses adultos pesan alrededor de 25 libras más que en 1960. Por esa razón, el uso continuado de los pesos calculados hace muchos años puede dar por resultado cálculos incorrectos y circunstancias de inseguridad.

Los problemas que surgen al determinar cargas seguras en aviones y barcos son ejemplos del tipo de problemas que se estudian en una disciplina relativamente nueva llamada *ergonomía*, que es el estudio del ajuste de las personas en su entorno. Un buen diseño ergonómico da como resultado un entorno seguro, funcional, eficiente y cómodo. La ergonomía tiene una gran cantidad de aplicaciones, incluyendo el diseño de tableros de automóvil, ataúdes, kayacs, cascos para ciclismo, tapas para botellas, manijas para puertas, tapas para alcantarillas, teclados, centros de control de tráfico aéreo y líneas de ensamblado de computadoras. El trabajo con el tema de cargas seguras de pasajeros ilustrará una experiencia real en relación con el papel que desempeña la estadística en la ergonomía.

6-1 Panorama general

En el capítulo 2 consideramos la distribución de datos y en el capítulo 3 estudiamos algunas medidas importantes de conjuntos de datos, incluyendo las medidas de tendencia central y de variación. En el capítulo 4 analizamos principios básicos de probabilidad, y en el capítulo 5 presentamos los siguientes conceptos:

- Una *variable aleatoria* es una variable que tiene un valor numérico único, determinado por el azar, para cada resultado de algún procedimiento.
- Una *distribución de probabilidad* describe la probabilidad para cada valor de la variable aleatoria.
- Una variable aleatoria *discreta* tiene un número finito de valores o un número contable de valores. Es decir, el *número* de valores posibles que x puede asumir es 0, o 1, o 2, etcétera.
- Una variable aleatoria *continua* tiene un número infinito de valores, los cuales suelen estar asociados con mediciones en una escala continua, sin huecos ni interrupciones.

En el capítulo 5 consideramos únicamente distribuciones de probabilidad *discretas*, pero en este capítulo presentamos las distribuciones de probabilidad *continuas*. Aun cuando iniciamos con una distribución uniforme, la mayor parte del capítulo se enfoca en las *distribuciones normales*. Las distribuciones normales son sumamente importantes puesto que ocurren con gran frecuencia en las aplicaciones reales y porque desempeñan un papel fundamental en los métodos de estadística inferencial. Las distribuciones normales se utilizarán a menudo a lo largo del resto de este libro.

Definición

Si una variable aleatoria continua tiene una distribución con una gráfica simétrica y en forma de campana, como la de la figura 6-1, y puede expresarse por medio de la fórmula 6-1, decimos que tiene una **distribución normal**.

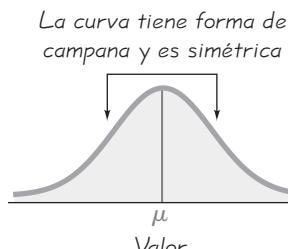
Fórmula 6-1

$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

La fórmula 6-1 quizá resulte intimidante y compleja, pero tenemos buenas noticias: en realidad no es necesario utilizarla. Mostramos la fórmula 6-1 para indicar que cualquier distribución normal en particular está determinada por dos parámetros: la media μ y la desviación estándar σ . Una vez que se han seleccionado valores específicos para μ y σ , podemos graficar la fórmula 6-1 como graficaríamos cualquier ecuación que relaciona x con y ; el resultado es una distribución de probabilidad continua con forma de campana, como la que se ilustra en la figura 6-1.

Figura 6-1

La distribución normal



6-2 La distribución normal estándar

Concepto clave En esta sección estudiaremos la distribución normal estándar, la cual tiene las siguientes tres propiedades: **1.** presenta forma de campana (como se observa en la figura 6-1); **2.** posee una media igual a 0; **3.** tiene una desviación estándar igual a 1. Es muy importante que en esta sección el lector desarrolle la habilidad para calcular áreas (o probabilidades o frecuencias relativas) correspondientes a diversas regiones debajo de la gráfica de la distribución normal estándar. También es importante calcular valores de la variable z que corresponden a áreas debajo de la curva. Podría parecer que, dada la naturaleza de la fórmula 6-1, se trata de una tarea inalcanzable; sin embargo, será relativamente *fácil* lograrlo.

Este capítulo se enfoca en el concepto de una distribución de probabilidad normal, pero comenzamos con una *distribución uniforme*. La distribución uniforme nos facilita ver estas dos propiedades muy importantes: **1.** el área bajo la curva de una distribución de probabilidad es igual a 1; **2.** existe una correspondencia entre el área y la probabilidad (o frecuencia relativa), de manera que algunas probabilidades se pueden calcular al identificar las áreas correspondientes.

Distribuciones uniformes

Definición

Una variable aleatoria continua tiene una **distribución uniforme** si sus valores se dispersan *uniformemente* a través del rango de posibilidades. La gráfica de una distribución uniforme tiene forma rectangular.

EJEMPLO Duración de la clase Un profesor de estadística planea sus clases con tanto cuidado que sus duraciones están distribuidas uniformemente entre 50.0 y 52.0 min. (Como las clases de estadística son tan interesantes, generalmente dan la impresión de ser más cortas). Esto es, cualquier tiempo entre 50.0 y 52.0 min es posible, y todos los valores posibles tienen la misma probabilidad. Si seleccionamos aleatoriamente una de las clases y permitimos que x sea la variable aleatoria que representa la duración de esa clase, entonces x tiene una distribución que puede graficarse como en la figura 6-2.

Cuando estudiamos las distribuciones de probabilidad *discretas* en la sección 5-2, identificamos dos requisitos: **1.** $\sum P(x) = 1$ y **(2)** $0 \leq P(x) \leq 1$

continúa

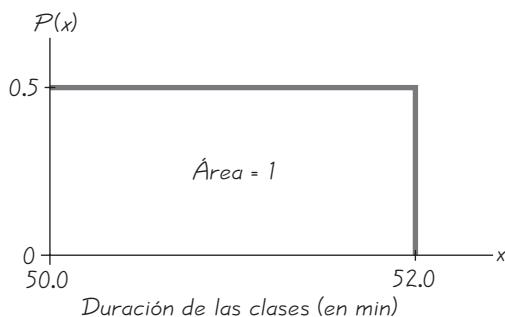


Figura 6-2 Distribución uniforme de duración de las clases

LA ESTADÍSTICA EN LAS NOTICIAS

Buen consejo para los periodistas

El columnista Max Frankel escribió en el *New York Times* que “la mayoría de las escuelas de periodismo dan poca importancia a la estadística y algunos permiten que los estudiantes se gradúen sin ningún tipo de formación referente al manejo de números. ¿Cómo pueden estos reporteros escribir con fundamento sobre temas como el comercio, la asistencia social y el crimen, o sobre tarifas aéreas, la atención a la salud y la nutrición? El uso sentimental que hacen los medios de comunicación de los números acerca de la incidencia de accidentes o muertes atemoriza a las personas y las deja vulnerables a las exageraciones periodísticas, la demagogia política y el fraude comercial”. El analista cita varios casos, incluyendo el ejemplo de un artículo de página completa acerca del déficit de la ciudad de Nueva York con la promesa del alcalde de esa ciudad de reducir el déficit presupuestal de \$2,700 millones; en el artículo nunca se menciona el monto *total* del presupuesto, de manera que la cifra de \$2,700 millones carece de contexto.

para todos los valores de x . También en la sección 5-2 establecimos que la gráfica de una distribución de probabilidad discreta se denomina *histograma de probabilidad*. La gráfica de una distribución de probabilidad continua, como la que se incluye en la figura 6-2, se llama *curva de densidad*, y debe satisfacer dos propiedades similares a los requisitos de las distribuciones de probabilidad discretas, tal como se plantea en la siguiente definición.

Definición

Una **curva de densidad** es una gráfica de una distribución de probabilidad continua. Debe satisfacer las siguientes propiedades:

1. El área total bajo la curva debe ser igual a 1.
2. Cada punto de la curva debe tener una altura vertical igual o mayor que 0.
(Es decir, la curva no puede estar por debajo del eje x).

Si establecemos que la altura del rectángulo de la figura 6-2 es 0.5, obligamos a que el área circunscrita sea $2 \times 0.5 = 1$, como se requiere. (En general, el área del rectángulo se convierte en 1 cuando igualamos su altura al valor de 1/rango). Esta propiedad (área = 1) facilita mucho la solución de problemas de probabilidad, de manera que la siguiente afirmación es importante:

Puesto que el área total bajo de la curva de densidad es igual a 1, existe una correspondencia entre área y probabilidad.

EJEMPLO Duración de la clase Kim, quien tiene el hábito de vivir siempre de prisa, se comprometió a acudir a una entrevista de trabajo inmediatamente después de su clase de estadística. Si la clase dura más de 51.5 minutos, llegará tarde a la entrevista de trabajo. Dada la distribución uniforme de la figura 6-2, calcule la probabilidad de que una clase seleccionada al azar dure más de 51.5 minutos.

Solución Observe la figura 6-3, donde la región sombreada representa duraciones mayores de 51.5 minutos. Puesto que el área total bajo la curva de densidad es igual a 1, existe una correspondencia entre área y probabilidad.

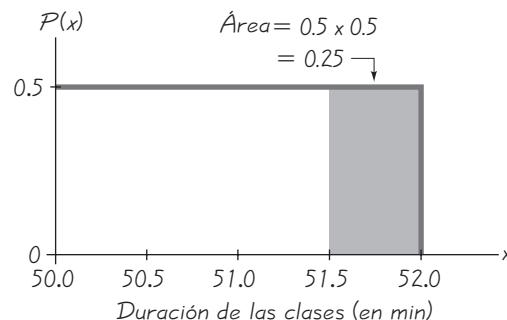


Figura 6-3 Uso del área para el cálculo de probabilidad

Por lo tanto, podemos calcular la probabilidad deseada utilizando áreas de la siguiente manera:

$P(\text{clase con duración mayor de } 51.5 \text{ minutos}) = \text{área de región sombreada de la figura 6-3}$

$$\begin{aligned} &= 0.5 \times 0.5 \\ &= 0.25 \end{aligned}$$

INTERPRETACIÓN La probabilidad de seleccionar al azar una clase que dure más de 51.5 minutos es de 0.25. Como esa probabilidad es demasiado alta, Kim debe considerar elaborar un plan de contingencia que le permita llegar a su entrevista de trabajo a tiempo. Nadie debe llegar tarde a una entrevista de trabajo.

Distribución normal estándar

La curva de densidad de una distribución uniforme es una línea horizontal, de manera que es fácil calcular el área de cualquier región rectangular: multiplicando anchura por altura. La curva de densidad de una distribución normal tiene una forma de campana más complicada, como se observa en la figura 6-1, por lo que es más difícil calcular áreas, pero el principio básico es el mismo: *existe una correspondencia entre área y probabilidad*.

Así como existen muchas distribuciones uniformes diferentes (con distintos rangos de valores), también existen muchas distribuciones normales diferentes, las cuales dependen de dos parámetros: la media poblacional μ y la desviación estándar poblacional σ . (Recuerde que en el capítulo 1 vimos que un *parámetro* es una medida numérica que describe alguna característica de una *población*). La figura 6-4 ilustra curvas de densidad de estaturas de hombres y mujeres adultos. Debido a que los hombres tienen una estatura media mayor, la cima de la curva de densidad de los hombres está ubicada hacia la derecha. Puesto que las estaturas de los hombres tienen una desviación estándar ligeramente mayor, su curva de densidad es un poco más ancha. La figura 6-4 presenta dos posibles distribuciones normales diferentes. Existe una infinidad de posibilidades, pero una es de especial interés.

Definición

La **distribución normal estándar** es una distribución normal de probabilidad con $\mu = 0$ y $\sigma = 1$, y el área total debajo de su curva de densidad es igual a 1. (Véase la figura 6-5).

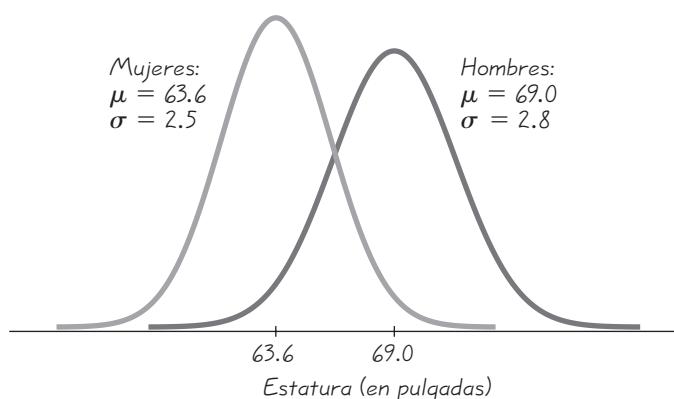
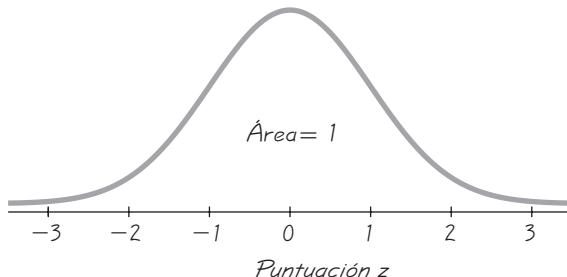


Figura 6-4
Estaturas de hombres
y mujeres adultos

Figura 6-5

Distribución normal estándar:
 $\mu = 0$ y $\sigma = 1$



Suponga que fuimos contratados para realizar cálculos empleando la fórmula 6-1. Rápidamente veríamos que los valores más fáciles para μ y para σ son $\mu = 0$ y $\sigma = 1$. Al permitir que $\mu = 0$ y $\sigma = 1$, los matemáticos han calculado muchas áreas diferentes bajo la curva. Como se observa en la figura 6-5, el área bajo la curva es 1, y esto nos permite establecer la correspondencia entre área y probabilidad, tal como hicimos en el ejemplo anterior con la distribución uniforme.

Cálculo de probabilidades con puntuaciones z dadas

Si empleamos la tabla A-2 (en el apéndice A y en la tarjeta con *fórmulas y tablas*), podemos calcular áreas (o probabilidades) para muchas regiones diferentes. Tales áreas pueden calcularse utilizando la tabla A-2, una calculadora TI-83/84 Plus o programas de cómputo como STATDISK, Minitab o Excel. Las características más importantes de los distintos métodos se resumen en la tabla 6-1. No es necesario conocer los cinco métodos; usted sólo necesita aprender el método que utilizará para la clase y los exámenes. Puesto que las calculadoras o los programas de cómputo ofrecen resultados más exactos que la tabla A-2, se recomienda el uso de la tecnología. (Cuando haya discrepancias, las respuestas en el apéndice E generalmente incluirán tanto resultados basados en la tabla A-2 como en recursos tecnológicos).

Tabla 6-1 Métodos para el cálculo de las áreas de la distribución normalTabla A-2, STATDISK, Minitab, Excel

Da el área acumulativa de la izquierda hasta una línea vertical por encima de un valor específico de z .

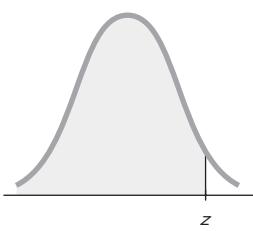


Tabla A-2: El procedimiento para usar la tabla A-2 se describe en el texto.

STATDISK: Seleccione **Analysis, Probability Distributions, Normal Distribution**. Ingrese el valor z y luego haga clic en **Evaluate**.

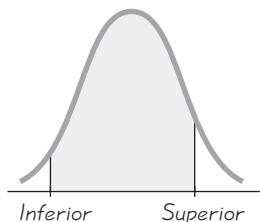
Minitab: Seleccione **Calc, Probability Distributions, Normal**. En el cuadro de diálogo seleccione **Cumulative Probability, Input Constant**.

Excel: Seleccione **fx, Statistical, NORMDIST**. En el cuadro de diálogo registre el valor y la media, la desviación estándar y "true".

TI-83/84: Presiones **2nd|VARS|2: normal cdf()**, después ingrese las dos puntuaciones z separadas por una coma, como en (puntuación z izquierda, puntuación z derecha).

Calculadora TI-83/84 Plus

Da el área limitada a la izquierda y a la derecha por líneas verticales por arriba de cualesquier valores específicos.



Si utiliza la tabla A-2, es esencial que comprenda los siguientes puntos:

1. La tabla A-2 está diseñada únicamente para la distribución normal *estándar*, que tiene una media de 0 y una desviación estándar de 1.
2. La tabla A-2 abarca dos páginas, una para las puntuaciones z *negativas* y la otra para las puntuaciones z *positivas*.
3. Cada valor en la tabla es una *área acumulativa desde la izquierda* hasta un límite vertical por arriba de una puntuación z específica.
4. Cuando construya una gráfica, evite la confusión entre puntuaciones z y las áreas.

Puntuación z : *Distancia a lo largo de la escala horizontal de la distribución normal estándar; remítase a la columna de la extrema izquierda y al renglón superior de la tabla A-2.*

Área: *Región bajo la curva; remítase a los valores de la tabla A-2.*

5. La parte de la puntuación z que denota centésimas se encuentra en el renglón superior de la tabla A-2.

El siguiente ejemplo requiere que calculemos la probabilidad asociada con un valor menor que 1.58. Comience con la puntuación z de 1.58, localizando 1.5 en la columna izquierda; después encuentre el valor en el renglón adjunto de probabilidad que está directamente debajo de 0.08, como se muestra en el extracto de la tabla A-2 que aparece al margen.

El valor del área (o probabilidad) de 0.9429 indica que existe una probabilidad de 0.9429 de seleccionar aleatoriamente una puntuación z menor que 1.58. (En las siguientes secciones consideraremos casos en los que la media no es 0 o la desviación estándar no es 1).

z 0.08
.	.
.	.
.	.
1.5 0.9429

EJEMPLO Termómetros científicos La Precision Scientific Instrument Company fabrica termómetros que se supone deben dar lecturas de 0°C al punto de congelación del agua. Las pruebas de una muestra grande de estos instrumentos revelaron que en el punto de congelación del agua, algunos termómetros daban lecturas por debajo de 0° (denotadas con números negativos), y otros daban lecturas por encima de 0° (denotadas con números positivos). Suponga que la lectura media es 0°C y que la desviación estándar de las lecturas es 1.00°C . También suponga que las lecturas se distribuyen de manera normal. Si se elige al azar un termómetro, calcule la probabilidad de que, al punto de congelación del agua, la lectura sea menor que 1.58° .

SOLUCIÓN La distribución de probabilidad de las lecturas es una distribución normal estándar, ya que las lecturas se distribuyen de forma normal, con $\mu = 0$ y $\sigma = 1$. Necesitamos encontrar el área que está debajo de $z = 1.58$,

continúa

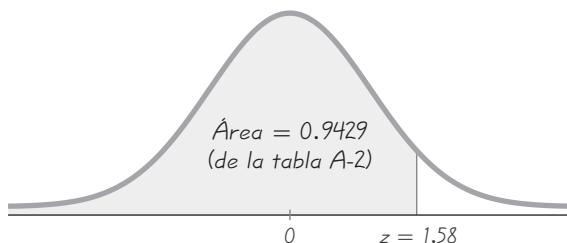


Figura 6-6 Cálculo del área por debajo de $z = 1.58$

en la figura 6-6. El área por debajo de $z = 1.58$ es igual a la *probabilidad* de seleccionar al azar un termómetro con una lectura menor que 1.58° . En la tabla A-2 encontramos que esta área es 0.9429.

INTERPRETACIÓN La *probabilidad* de seleccionar al azar un termómetro con una lectura menor que 1.58° (en el punto de congelación del agua) es igual al *área* de 0.9429, que aparece como la región sombreada en la figura 6-6. Otra forma de interpretar este resultado es concluir que el 94.29% de los termómetros tendrán lecturas por debajo de 1.58° .

EJEMPLO Termómetros científicos Utilice los termómetros del ejemplo anterior y calcule la probabilidad de seleccionar al azar un termómetro con una lectura (en el punto de congelación del agua) por arriba de -1.23° .

SOLUCIÓN Nuevamente, calculamos la *probabilidad* deseada encontrando el *área* correspondiente. Buscamos el área de la región sombreada en la figura 6-7, pero la tabla A-2 está diseñada para aplicarse únicamente en áreas acumulativas desde la *izquierda*. Si nos remitimos a la tabla A-2, en la página con puntuaciones z *negativas*, encontramos que el área acumulativa de la izquierda hasta $z = -1.23$ es 0.1093, tal como se observa. Sabiendo que el área total bajo la curva es 1, podemos calcular el área sombreada si restamos 0.1093 de 1. El resultado es 0.8907. Aun cuando la tabla A-2 está diseñada únicamente para áreas acumulativas a partir de la *izquierda*, podemos utilizarla para calcular áreas acumulativas desde la *derecha*, tal como se muestra la figura 6-7.

INTERPRETACIÓN Debido a la correspondencia entre probabilidad y área, podemos concluir que la *probabilidad* de seleccionar aleatoriamente un termómetro con una lectura por arriba de -1.23° , en el punto de congelación del agua, es de 0.8907 (que corresponde al *área* a la derecha de $z = -1.23$). En otras palabras, el 89.07% de los termómetros tienen lecturas por encima de -1.23° .

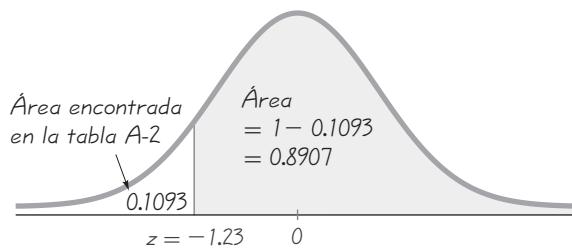


Figura 6-7 Cálculo del área por arriba de $z = -1.23$

El ejemplo anterior ilustra una de las formas en que podemos utilizar la tabla A-2 para calcular de manera indirecta una área acumulativa a partir de la derecha. El siguiente ejemplo ilustra otra manera para calcular alguna área utilizando la tabla A-2.

EJEMPLO Termómetros científicos Una vez más, haga una selección aleatoria de la misma muestra de termómetros y calcule la probabilidad de que el termómetro elegido tenga lecturas (en el punto de congelación del agua) entre -2.00° y 1.50° .

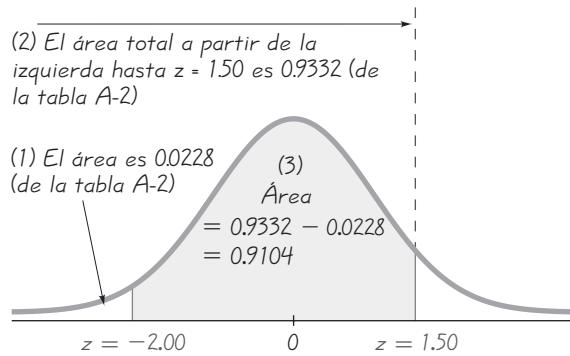
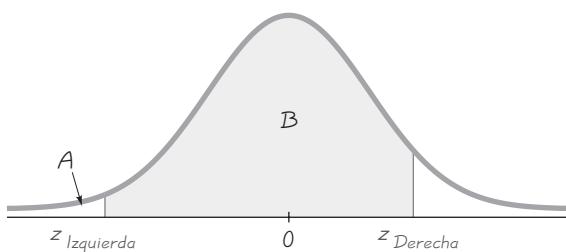


Figura 6-8 Cálculo del área entre dos valores

SOLUCIÓN Nuevamente tratamos con valores distribuidos de manera normal, con una media de 0° y una desviación estándar de 1° . La probabilidad de seleccionar un termómetro con una lectura comprendida entre -2.00° y 1.50° corresponde al área sombreada de la figura 6-8. La tabla A-2 no puede utilizarse para calcular el área de forma directa, pero podemos emplearla para encontrar que $z = -2.00$ corresponde al área de 0.0228, y que $z = 1.50$ corresponde al área de 0.9332, como se observa en la figura. Remítase a la figura 6-8 y note que el área sombreada corresponde a la diferencia entre 0.9332 y 0.0228. El área sombreada es, por lo tanto, $0.9332 - 0.0228 = 0.9104$.

INTERPRETACIÓN Considerando la correspondencia entre probabilidad y área, concluimos que existe una probabilidad de 0.9104 de seleccionar al azar uno de los termómetros con una lectura entre -2.00° y 1.50° , en el punto de congelación del agua. Otra forma de interpretar este resultado es afirmar que si se seleccionan muchos termómetros para probarlos en el punto de congelación del agua, entonces 0.9104 (o el 91.04%) de ellos tendrán lecturas entre -2.00° y 1.50° .

El ejemplo anterior puede generalizarse como una regla que establece que el área correspondiente a la región localizada entre dos puntuaciones z específicas puede obtenerse al calcular la diferencia entre las dos áreas localizadas en la tabla A-2. Observe la figura 6-9, la cual muestra que la región sombreada B puede obtenerse calculando la *diferencia* entre dos áreas de la tabla A-2: las áreas A y B combinadas (que en la tabla A-2 aparecen como las áreas correspondientes a z_{Derecha}) y el área

Figura 6-9
Cálculo del área entre dos puntuaciones z

$$\text{Área sombreada } B = (\text{áreas } A \text{ y } B \text{ combinadas}) - (\text{área } A) = (\text{área de la tabla A-2, usando } z_{\text{Derecha}}) - (\text{área de la tabla A-2, usando } z_{\text{Izquierda}})$$

A (que en la tabla A-2 aparece como el área correspondiente a $z_{\text{izquierda}}$). *Recomendación:* No trate de memorizar una regla o una fórmula para este caso, ya que es infinitamente mejor *comprender* el procedimiento. Trate de comprender cómo funciona la tabla A-2, después dibuje una gráfica, sombree el área deseada y piense en una forma para calcular el área, ya que la tabla A-2 proporciona sólo áreas acumulativas desde la izquierda.

El ejemplo anterior concluyó con la afirmación de que la probabilidad de una lectura comprendida entre -2.00° y 1.50° es de 0.9104. Probabilidades como ésta también pueden expresarse con la siguiente notación.

Notación

- | | |
|----------------|--|
| $P(a < z < b)$ | denota la probabilidad de que la puntuación z esté entre a y b . |
| $P(z > a)$ | denota la probabilidad de que la puntuación z sea mayor que a . |
| $P(z < a)$ | denota la probabilidad de que la puntuación z sea menor que a . |

Con esta notación podemos expresar el resultado del último ejemplo de la siguiente manera: $P(-2.00 < z < 1.50) = 0.9104$, lo cual establece en símbolos que la probabilidad de que una puntuación z caiga entre -2.00 y 1.50 es de 0.9104. Con una distribución de probabilidad continua, tal como la distribución normal, la probabilidad de obtener cualquier valor *exacto* es 0. Es decir, $P(z = a) = 0$. Por ejemplo, existe una probabilidad 0 de seleccionar al azar a una persona cuya estatura sea exactamente de 68.12345678 pulgadas. En la distribución normal, cualquier punto único sobre la escala horizontal está representado, no por una región bajo la curva, sino por una línea vertical por arriba del punto. Para $P(z = 1.50)$ tenemos una línea vertical que está por arriba de $z = 1.50$, pero esta línea vertical, por sí misma, no contiene una área, de manera que $P(z = 1.50) = 0$. Para cualquier variable aleatoria continua, la probabilidad de un valor exacto es 0, y se infiere que $P(a \leq z \leq b) = P(a < z < b)$. También se deduce que la probabilidad de obtener una puntuación z de *a lo sumo* b es igual a la probabilidad de obtener una puntuación z *menor que* b . Es importante interpretar correctamente frases clave como *a lo sumo, al menos, mayor que, no mayor que*, etcétera.

Cálculo de puntuaciones z de áreas conocidas

Hasta ahora, todos los ejemplos de esta sección que implican la distribución normal estándar han seguido el mismo formato: dadas puntuaciones z , calculamos áreas bajo la curva; estas áreas corresponden a probabilidades. En muchos otros casos, buscamos el proceso contrario: conocemos el área (o probabilidad), pero necesitamos calcular la puntuación z correspondiente. En tales casos, es muy importante evitar una confusión entre las puntuaciones z y las áreas. Recuerde, las puntuaciones z son *distancias* a lo largo de la escala horizontal y están representadas por los números de la tabla A-2 que se encuentran en la columna de la extrema izquierda y en el cruce del renglón superior. Las áreas (o probabilidades) son regiones bajo la curva y están representadas por los valores en el cuerpo de la tabla A-2. Además, las puntuaciones z ubicadas en la mitad izquierda de la curva siempre son negativas. Si ya conocemos una probabilidad y deseamos determinar la puntuación z correspondiente, la calculamos de la siguiente forma.

Procedimiento para el cálculo de una puntuación z a partir de una área conocida

- Dibuje una curva en forma de campana e identifique la región bajo la curva que corresponde a la probabilidad dada. Si no se trata de una región acumulativa a partir de la izquierda, trabaje con una región acumulativa conocida de la izquierda.
- Usando el área acumulativa de la izquierda, localice la probabilidad más cercana en el *cuerpo* de la tabla A-2 e identifique la puntuación z correspondiente.

EJEMPLO Termómetros científicos Use los mismos termómetros anteriores, con lecturas de temperatura al punto de congelación del agua distribuidas normalmente, con una media de 0°C y una desviación estándar de 1°C . Calcule la temperatura correspondiente a P_{95} , el percentil 95. Es decir, calcule la temperatura que separa el 95% inferior del 5% superior. Observe la figura 6-10.

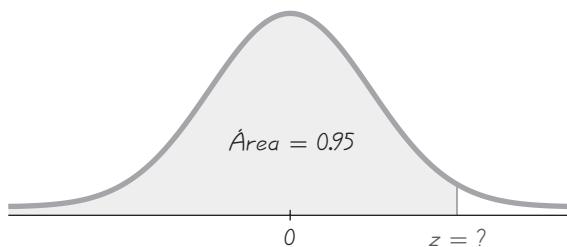


Figura 6-10 Cálculo del percentil 95

SOLUCIÓN La figura 6-10 incluye la puntuación z que corresponde al percentil 95, con el 95% del área (o 0.95) por debajo de ella. Importante: Cuando se remita a la tabla A-2, recuerde que el cuerpo de la tabla incluye las *áreas acumulativas a partir de la izquierda*. Al remitirnos a la tabla A-2 buscamos el área de 0.95 en el *cuerpo* de la tabla y después buscamos la puntuación z correspondiente. En la tabla encontramos las áreas de 0.9495 y 0.9505, pero hay un asterisco con una nota especial que indica que 0.9500 corresponde a una puntuación z de 1.645. Ahora podemos concluir que la puntuación z en la figura 6-10 es 1.645, por lo que el percentil 95 es la lectura de la temperatura de 1.645°C .

INTERPRETACIÓN Al probar los termómetros a la temperatura de congelación, el 95% de las lecturas serán menores o iguales que 1.645°C , y el 5% de ellas serán mayores o iguales que 1.645°C .

Note que en la solución anterior, la tabla A-2 indicó una puntuación z de 1.645, que está a la mitad de 1.64 y 1.65. Con la tabla A-2, generalmente podemos evitar la interpolación si tan sólo seleccionamos el valor más cercano. Existen casos especiales, listados en la tabla adjunta, que son importantes ya que se utilizan

Puntuaciones z	Área acumulativa de la izquierda
1.645	0.9500
-1.645	0.0500
2.575	0.9950
-2.575	0.0050

con frecuencia en una amplia variedad de aplicaciones. (El valor de $z = 2.576$ da una área ligeramente más cercana a la de 0.9950, pero $z = 2.575$ tiene la ventaja de ser el valor intermedio entre $z = 2.57$ y $z = 2.58$). Con la excepción de estos casos especiales, podemos seleccionar el valor más cercano en la tabla. (Si un valor deseado se encuentra entre dos valores de la tabla, seleccione el valor más grande). Además, para las puntuaciones z por arriba de 3.49, podemos utilizar 0.9999 como aproximación del área acumulativa de la izquierda; para puntuaciones z por debajo de -3.49, podemos utilizar 0.0001 como aproximación del área acumulativa a partir de la izquierda.

EJEMPLO Termómetros científicos Utilice los mismos termómetros y calcule las temperaturas que separan el 2.5% inferior y el 2.5% superior.

SOLUCIÓN Remítase a la figura 6-11, que presenta las puntuaciones z requeridas. Para encontrar la puntuación z localizada a la izquierda, remítase a la tabla A-2 y busque una área de 0.025 en el *cuerpo de la tabla*. El resultado es $z = -1.96$. Para encontrar la puntuación z localizada a la derecha, remítase al *cuerpo de la tabla* y busque una área de 0.975. (Recuerde que la tabla A-2 siempre da áreas acumulativas a partir de la *izquierda*). El resultado es $z = 1.96$. Los valores de $z = -1.96$ y $z = 1.96$ separan el 2.5% inferior y el 2.5% superior, como se observa en la figura 6-11.

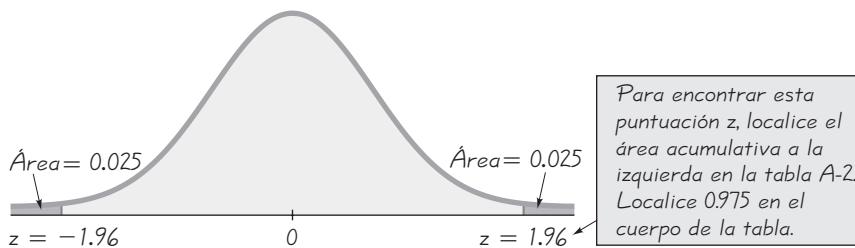


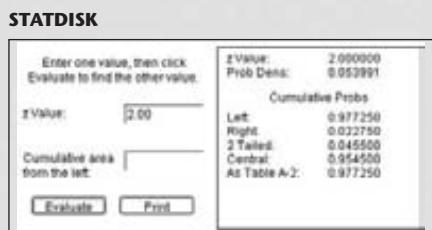
Figura 6-11 Cálculo de puntuaciones z

INTERPRETACIÓN Al probar los termómetros a la temperatura de congelación, el 2.5% de las lecturas serán iguales o menores que -1.96° , y el 2.5% de las lecturas serán iguales o mayores que 1.96° . Otra interpretación es que, al punto de congelación del agua, el 95% de todas las lecturas de los termómetros estarán entre -1.96° y 1.96° .

Los ejemplos de esta sección se elaboraron de forma que la media de 0 y la desviación estándar de 1 coincidieran exactamente con los parámetros de la distribución normal estándar. En realidad, es raro encontrar parámetros tan convenientes, ya que las distribuciones normales típicas incluyen medias distintas de 0 y desviaciones estándar distintas de 1. En la siguiente sección presentamos métodos para trabajar con este tipo de distribuciones normales, que son más realistas.

Uso de la tecnología

STATDISK Seleccione **Analysis, Probability Distributions, Normal Distribution**. Registre la puntuación z para calcular áreas correspondientes o registre el área acumulativa de la izquierda para calcular la puntuación z . Despues de introducir un valor, haga clic en el botón **Evaluate**. Vea la siguiente representación visual de STATDISK para un valor $z = 2.00$.



MINITAB

- Para encontrar el área acumulativa que está a la izquierda de una puntuación z (como en la tabla A-2), seleccione **Calc, Probability Distributions, Normal, Cumulative probabilities**, registre la media de 0 y la desviación estándar de 1, después haga clic en el botón de **Input Constant** e ingrese la puntuación z .
- Para encontrar la puntuación z correspondiente a una probabilidad conocida, seleccione **Calc, Probability Distributions, Normal**, después seleccione **Inverse cumulative probabilities** y la opción **Input constant**. Para la constante de entrada, indique el área total que se encuentra a la izquierda del valor dado.

EXCEL

- Para encontrar el área acumulativa a la izquierda de una puntuación z (como en la tabla A-2), haga clic en **fx**, después selec-

cione **Statistical, NORMDIST** e indique la puntuación z .

- Para encontrar la puntuación z correspondiente a una probabilidad conocida, seleccione **fx, Statistical, NORMSINV** e ingrese el área total que se encuentra a la izquierda del valor dado.

TI-83/84 PLUS

- Para calcular el área entre dos puntuaciones z , presione **2nd VARS, 2** (para **normalcdf**), después proceda a registrar las dos puntuaciones z , separadas por una coma, como en (puntuación z izquierda, puntuación z derecha).
- Para encontrar una puntuación z correspondiente a una probabilidad conocida, presione **2nd VARS, 3** (para **invNorm**), y proceda a indicar el área total a la izquierda del valor, la media y la desviación estándar con el formato (área total izquierda, media, desviación estándar) incluyendo las comas.

6-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Distribución normal.** Cuando nos referimos a una distribución “normal”, ¿el término “normal” tiene el mismo significado que en el lenguaje cotidiano o tiene un significado especial en estadística? ¿Qué es exactamente una distribución normal?
- Distribución normal.** Una distribución normal se describe de manera informal y aproximada como una distribución de probabilidad con “forma de campana” cuando se grafica. ¿Qué es la “forma de campana”?
- Distribución normal estándar.** ¿Qué requisitos son necesarios para que una distribución de probabilidad normal sea una distribución de probabilidad normal *estándar*?
- Áreas.** Si usted determina que para la gráfica de una distribución normal estándar el área acumulativa a la izquierda de una puntuación z es 0.4, ¿cuál es el área acumulativa que se localiza a la derecha de esa puntuación z ?

Distribución uniforme continua. En los ejercicios 5 a 8, remítase a la distribución uniforme continua descrita en la figura 6-2; suponga que se selecciona al azar una clase con una duración entre 50.0 y 52.0 min, y calcule la probabilidad de seleccionar el tiempo dado.

- Menor que 51.5 min
- Entre 50.5 min y 51.5 min
- Mayor que 50.5 min y 51.5 min
- Entre 51.5 min y 51.6 min

Distribución normal estándar. En los ejercicios 9 a 28, suponga que las lecturas de termómetros se distribuyen normalmente, con una media de 0° y una desviación estándar de 1.00°C . Se selecciona aleatoriamente un termómetro y se prueba. En cada caso, dibuje un bosquejo y calcule la probabilidad de cada lectura. (Los valores están en grados Celsius).

- Menor que -1.00
- Menor que -2.50

- | | |
|---|---|
| <p>11. Menor que 1.00
 13. Mayor que 1.25
 15. Mayor que -1.75
 17. Entre 1.00 y 2.00
 19. Entre -2.45 y -2.00
 21. Entre -2.11 y 1.55
 23. Entre -1.00 y 4.00
 25. Mayor que 3.52
 27. Mayor que 0</p> | <p>12. Menor que 2.50
 14. Mayor que 1.96
 16. Mayor que -1.96
 18. Entre 0.50 y 1.50
 20. Entre 1.05 y 2.05
 22. Entre -1.80 y 2.08
 24. Entre -3.90 y 1.50
 26. Menor que -3.75
 28. Menor que 0</p> |
|---|---|

Bases de la regla empírica. En los ejercicios 29 a 32, calcule el área bajo la curva indicada de la distribución normal estándar, después conviértala en porcentaje y complete el espacio en blanco. Los resultados conforman la base de la regla empírica explicada en la sección 3-3.

- 29.** Aproximadamente ____% del área está entre $z = -1$ y $z = 1$ (o dentro de una desviación estándar a partir de la media).
- 30.** Aproximadamente ____% del área está entre $z = -2$ y $z = 2$ (o dentro de 2 desviaciones estándar a partir de la media).
- 31.** Aproximadamente ____% del área está entre $z = -3$ y $z = 3$ (o dentro de 3 desviaciones estándar a partir de la media).
- 32.** Aproximadamente ____% del área está entre $z = -3.5$ y $z = 3.5$ (o dentro de 3.5 desviaciones estándar a partir de la media).

Cálculo de probabilidad. En los ejercicios 33 a 36, suponga que las lecturas de termómetros se distribuyen normalmente, con una media de 0° y una desviación estándar de 1.00° . Calcule la probabilidad indicada, donde z es la lectura en grados.

- 33.** $P(-1.96 < z < 1.96)$ **34.** $P(z > 1.645)$
35. $P(z < -2.575 \text{ o } z > 2.575)$ **36.** $P(z < -1.96 \text{ o } z > 1.96)$

Cálculo de valores de temperatura. En los ejercicios 37 a 40, suponga que las lecturas de termómetros se distribuyen normalmente, con una media de 0° y una desviación estándar de 1.00°C . Se selecciona al azar un termómetro y se prueba. En cada caso, dibuje un bosquejo y calcule la lectura de la temperatura correspondiente a la información dada.

- 37.** Calcule P_{90} , el percentil 90. Ésta es la lectura de temperatura que separa el 90% inferior del 10% superior.
- 38.** Calcule P_{25} , el percentil 25.
- 39.** Si se rechaza el 2.5% de los termómetros porque tienen lecturas demasiado altas y se rechaza otro 2.5% porque tienen lecturas demasiado bajas, calcule las dos lecturas que son valores de corte que separan a los termómetros rechazados de los otros.
- 40.** Si se rechaza el 1.0% de los termómetros porque tienen lecturas demasiado altas, y otro 1.0% se rechaza porque tiene lecturas demasiado bajas, calcule las dos lecturas que son los valores de corte que separan a los termómetros rechazados de los otros.

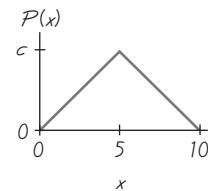
6-2 MÁS ALLÁ DE LO BÁSICO

41. Para una distribución normal estándar, calcule el porcentaje de datos que están
- dentro de 1 desviación estándar a partir de la media.
 - dentro de 1.96 desviaciones estándar a partir de la media.
 - entre $\mu - 3\sigma$ y $\mu + 3\sigma$.
 - entre 1 desviación estándar por debajo de la media y 2 desviaciones estándar por encima de la media.
 - a más de 2 desviaciones estándar a partir de la media.
42. Si una distribución uniforme continua tiene parámetros de $\mu = 0$ y $\sigma = 1$, entonces el mínimo es $-\sqrt{3}$ y el máximo es $\sqrt{3}$.
- Para esta distribución calcule $P(-1 < x < 1)$.
 - Calcule $P(-1 < x < 1)$ si usted supone de manera incorrecta que la distribución es normal en vez de uniforme.
 - Compare los resultados de los incisos a) y b). ¿Afecta mucho la distribución a los resultados?
43. Suponga que puntuaciones z se distribuyen normalmente, con una media de 0 y una desviación estándar de 1.
- Si $P(0 < z < a) = 0.3907$, calcule a .
 - Si $P(-b < z < b) = 0.8664$, calcule b .
 - Si $P(z > c) = 0.0643$, calcule c .
 - Si $P(z > d) = 0.9922$, calcule d .
 - Si $P(z < e) = 0.4500$, calcule e .
44. En una distribución uniforme continua,

$$\mu = \frac{\text{mínimo} + \text{máximo}}{2} \quad y \quad \sigma = \frac{\text{rango}}{\sqrt{12}}$$

Calcule la media y la desviación estándar de la distribución uniforme representada en la figura 6-2.

45. Haga el bosquejo de una gráfica que represente una distribución acumulativa de a) una distribución uniforme y b) una distribución normal.
46. Remítase a la gráfica de la distribución de probabilidad triangular de la variable continua aleatoria x . (Véase la gráfica al margen).
- Calcule el valor de la constante c .
 - Calcule la probabilidad de que x esté entre 0 y 3.
 - Calcule la probabilidad de que x esté entre 2 y 9.



Aplicaciones de las distribuciones

6-3 normales

Concepto clave En la sección anterior se estudiaron las aplicaciones de la distribución normal, pero todos los ejemplos y ejercicios estuvieron basados en la distribución normal *estándar*, con $\mu = 0$ y $\sigma = 1$. En esta sección se presentan métodos para trabajar con distribuciones normales que no son estándar, es decir, cuya media no es 0 o cuya desviación estándar no es 1, o bien, ambas condiciones. El concepto clave consiste en que podemos utilizar una conversión simple (fórmula 6-2)

LA ESTADÍSTICA EN LAS NOTICIAS

Ganadores múltiples de la lotería

Evelyn Marie Adams ganó la lotería de Nueva Jersey dos veces en cuatro meses. Este feliz suceso fue reportado en los medios de comunicación como una increíble coincidencia con tan sólo una posibilidad en 17 billones. Sin embargo, los matemáticos Persi Diaconis y Frederick Mosteller de Harvard señalan que existe una posibilidad en 17 billones de que una persona en particular, que posea un boleto para cada una de las dos loterías de Nueva Jersey, gane dos veces. Sin embargo, existe aproximadamente una posibilidad en 30 de que alguien en Estados Unidos gane la lotería dos veces durante un periodo de cuatro meses. Diaconis y Mosteller analizaron las coincidencias y concluyeron que “con una muestra lo suficientemente grande, cualquier cosa sorprendente puede suceder”. Según el *Detroit News*, Joe y Dolly Hornick ganaron la lotería de Pensylvania cuatro veces en 12 años, con premios de \$2.5 millones, \$68,000, \$206,217 y \$71,037.

que nos permite estandarizar cualquier distribución normal para poder utilizar los mismos métodos de la sección anterior. Por lo tanto, este proceso de estandarización nos permite hacer aplicaciones mucho más realistas y significativas.

Para trabajar con una distribución normal que no es estándar, simplemente estandarice los valores para poder continuar utilizando los mismos procedimientos de la sección 6-2.

Si convertimos valores en puntuaciones estándar, empleando la fórmula 6-2, entonces los procedimientos para trabajar con todas las distribuciones normales son los mismos que los de la distribución normal estándar.

Fórmula 6-2
$$z = \frac{x - \mu}{\sigma}$$
 (redondear las puntuaciones z hasta dos decimales)

El uso continuo de la tabla A-2 requiere la comprensión y la aplicación del principio anterior. (Si usted utiliza ciertas calculadoras o programas de cómputo, no es necesaria la transformación a puntuaciones z , ya que las probabilidades pueden calcularse de manera directa). Sin importar el método que utilice, debe comprender con claridad el principio básico anterior, ya que constituye un fundamento importante de los conceptos que se presentarán en los siguientes capítulos.

La figura 6-2 ilustra la transformación de una distribución no estándar a una distribución normal estándar. El área de cualquier distribución normal, limitada por alguna puntuación x (como en la figura 6-12a), es igual que el área limitada por la puntuación z equivalente en la distribución normal estándar (como en la figura 6-12b). Esto significa que cuando se trabaja con una distribución normal no estándar, se puede utilizar la tabla A-2 de la misma forma que se empleó en la sección 6-2, siempre y cuando se conviertan primero los valores a puntuaciones z . Cuando calcule áreas en una distribución normal no estándar, utilice este procedimiento:

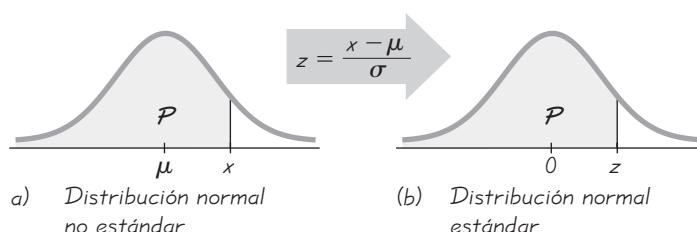
1. Dibuje una curva normal, indique la media y los valores específicos de x , después *sombree* la región que representa la probabilidad deseada.
2. Para cada valor relevante de x que sea un límite de la región sombreada, utilice la fórmula 6-2 para convertir el valor a la puntuación z equivalente.
3. Remítase a la tabla A-2 para encontrar el área de la región sombreada, que constituye la probabilidad deseada.

El siguiente ejemplo aplica estos tres pasos e ilustra la relación entre una distribución no normal típica y la distribución normal estándar.

EJEMPLO Pesos de pasajeros de taxis acuáticos En el problema del capítulo señalamos que la carga segura para un taxi acuático se calculó en 3500 libras. También indicamos que se suponía que el peso medio de un pasajero era de 140 libras. Supongamos “el peor de los casos”, en el que todos los

Figura 6-12

Transformación de una distribución normal no estándar a una distribución normal estándar



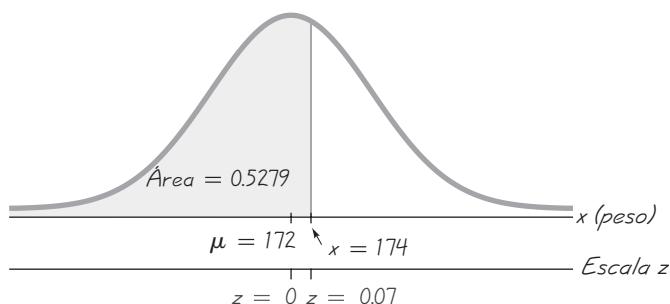


Figura 6-13 Pesos de hombres (en libras)

pasajeros son hombres adultos. (Esto podría ocurrir fácilmente en una ciudad donde se realizan convenciones en las que personas del mismo género suelen viajar en grupos). En concordancia con los datos de la National Health and Nutrition Examination Survey, suponga que los pesos de hombres se distribuyen normalmente, con una media de 172 lb y una desviación estándar de 29 lb. Si seleccionamos al azar a un hombre, calcule la probabilidad de que pese menos de 174 lb (el valor sugerido por el National Transportation and Safety Board).

SOLUCIÓN

- Paso 1: Observe la figura 6-13, que incluye la siguiente información: los hombres tienen pesos que se distribuyen normalmente, con una media de 172 libras y una desviación estándar de 29 libras, y la región sombreada representa a los hombres con pesos menores de 174 libras.
- Paso 2: Para usar la tabla A-2 primero debemos aplicar la fórmula 6-2 para transformar la distribución normal no estándar a una distribución normal estándar. El peso de 174 libras se convierte a una puntuación z de la siguiente manera:

$$z = \frac{x - \mu}{\sigma} = \frac{174 - 172}{29} = 0.07$$

- Paso 3: Si nos remitimos a la tabla A-2 y utilizamos $z = 0.07$, encontramos que el área acumulativa a la izquierda de $z = 0.07$ es una área de 0.5279.

INTERPRETACIÓN Existe una probabilidad de 0.5279 de elegir al azar a un hombre que pese menos de 174 lb. Se infiere que el 52.79% de los hombres pesan menos de 174 lb. También se infiere que el 47.21% de los hombres pesan más de 174 lb.

EJEMPLO Puntuaciones de CI Un psicólogo está diseñando un experimento para probar la eficacia de un nuevo programa de capacitación para vigilantes de seguridad aeroportuaria. Desea comenzar con un grupo homogéneo de sujetos con puntuaciones de CI comprendidas entre 85 y 125. Dado que las puntuaciones de CI se distribuyen normalmente, con una media de 100 y una desviación estándar de 15, ¿qué porcentaje de la gente tiene una puntuación de CI entre 85 y 125?



Atajo para ensayo clínico

¿Qué haría usted si estuviera probando un tratamiento y, antes de que su estudio termine, se diera cuenta de que es claramente eficaz? Debe acortar el estudio e informar a todos los participantes acerca de la eficacia del tratamiento. Esto fue lo que sucedió cuando se probó la hidroxiurea como tratamiento para la anemia falciforme. El estudio estaba programado para durar cerca de 40 meses, pero la eficacia del tratamiento se hizo evidente y el estudio se detuvo después de 36 meses (véase "Trial Halted as Sickle Cell Treatment Proves Itself", de Charles Marwick, *Journal of the American Medical Association*, vol. 273, núm. 8).

continúa



Encuestas sobre temas delicados

En ocasiones, las personas que contestan las encuestas se rehúsan a responder con honestidad preguntas sobre un tema delicado, como el robo al empleador o la conducta sexual. Stanley Warner (York University, Ontario) diseñó un sistema que produce resultados más exactos en casos como éstos. Como ejemplo, pregunte a algunos empleados si cometieron un robo durante el año anterior, y pídale que lancen una moneda. Instrúyalos para que respondan que no en caso de que no hayan robado y la moneda caiga en cara. De otra forma, deben responder afirmativamente. Los empleados suelen ser más honestos debido a que el lanzamiento de la moneda los ayuda a proteger su privacidad. Después, se puede utilizar la teoría de la probabilidad para analizar las respuestas, de manera que se obtengan resultados más exactos.

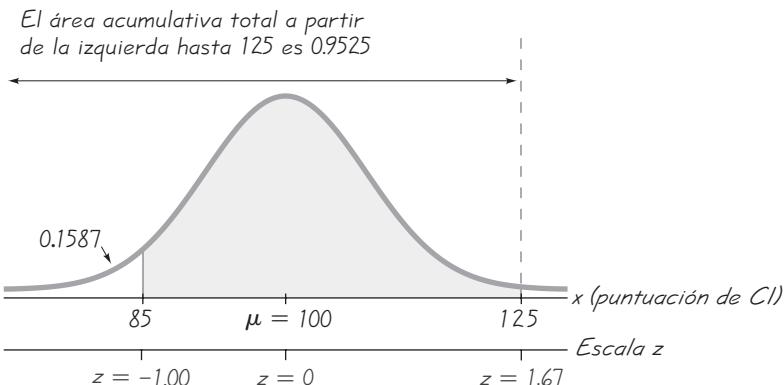


Figura 6-14 Puntuaciones de CI

SOLUCIÓN Observe la figura 6-14, que muestra la región sombreada de las puntuaciones de CI entre 85 y 125. No podemos encontrar esa área sombreada directamente en la tabla A-2, pero podemos obtenerla de manera indirecta utilizando los procedimientos básicos presentados en la sección 6-2. He aquí la estrategia para calcular el área sombreada: primero calcule el área acumulativa desde la izquierda hasta 85 y luego el área acumulativa desde la izquierda hasta 125; después obtenga la diferencia entre ambas áreas.

Obtención del área acumulativa hasta 85:

$$z = \frac{x - \mu}{\sigma} = \frac{85 - 100}{15} = -1.00$$

Si usamos la tabla A-2, encontramos que $z = -1.00$ corresponde a una área de 0.1587, como se ve en la figura 6-14.

Obtención del área acumulativa hasta 125:

$$z = \frac{x - \mu}{\sigma} = \frac{125 - 100}{15} = 1.67$$

Si usamos la tabla A-2, encontramos que $z = 1.67$ corresponde a una área de 0.9525, como se ve en la figura 6-14.

Obtención del área entre 85 y 125:

$$\text{Área sombreada} = 0.9525 - 0.1587 = 0.7938$$

INTERPRETACIÓN Si expresamos el resultado en porcentaje, concluimos que el 79.38% de las personas tienen una puntuación de CI entre 85 y 125.

Cálculo de valores de áreas conocidas

A continuación presentamos útiles sugerencias para los casos en que conocemos el área (o probabilidad o porcentaje) y debemos calcular el valor (o valores) relevante(s).

1. *No confunda las puntuaciones z y las áreas.* Recuerde que las puntuaciones z son *distancias* a lo largo de la escala horizontal, en tanto que las áreas son *regiones* bajo la curva normal. La tabla A-2 lista puntuaciones z en las columnas de la izquierda y a lo largo del renglón superior, pero las áreas se localizan en el cuerpo de la tabla.
2. *Elija el lado correcto de la gráfica (derecho/izquierdo).* Un valor que separa el 10% superior del resto se encontrará localizado en el lado derecho de la

gráfica, pero un valor que separa el 10% inferior se ubicará en el lado izquierdo de la gráfica.

3. Una puntuación z debe ser *negativa* siempre que esté localizada en la mitad *izquierda* de la distribución normal.
4. Las áreas (o probabilidades) son positivas o tienen valores de 0, pero nunca son negativas.

Las gráficas son sumamente útiles para visualizar, comprender y trabajar con éxito con las distribuciones de probabilidad normal, por lo tanto, deben emplearse siempre que sea posible.

Procedimiento para calcular valores con la tabla A-2 y la fórmula 6-2

1. Dibuje una curva de distribución normal, anote la probabilidad o porcentaje dados en la región apropiada de la gráfica e identifique el valor (o valores) x que se busca(n).
2. Utilice la tabla A-2 para encontrar la puntuación z correspondiente al área izquierda acumulativa, limitada por x . Remítase al *cuerpo* de la tabla A-2 para localizar el área más cercana, después identifique la puntuación z correspondiente.
3. Para emplear la fórmula 6-2, sustituya los valores de μ , σ y la puntuación z obtenida en el paso 2, después despeje x . Con base en el formato de la fórmula 6-2, podemos despejar x de la siguiente manera:

$$x = \mu + (z \cdot \sigma) \quad (\text{otra forma de la fórmula 6-2})$$

↑

(Si z está localizada a la izquierda de la media, asegúrese de que sea un número negativo).

4. Remítase al dibujo de la curva para verificar que la solución sea lógica en el contexto de la gráfica y en el contexto del problema.

El siguiente ejemplo utiliza el procedimiento que se acaba de describir.



EJEMPLO Pesos de pasajeros de taxis acuáticos En un ejemplo anterior de esta sección se demostró que el 52.79% de los hombres pesan menos del valor de 174 libras establecidas por el National Transportation and Safety Board. ¿Qué peso separa al 99.5% de los hombres menos pesados del 0.5% de los hombres más pesados? Nuevamente, suponga que los pesos de los hombres se distribuyen de manera normal, con una media de 172 libras y una desviación estándar de 29 libras.

SOLUCIÓN

- Paso 1: Comenzamos con la gráfica de la figura 6-15. Ya incluimos la media de 172 libras, sombreado el área que representa al 99.5% de los hombres menos pesados e identificamos el valor deseado como x .
- Paso 2: En el *cuerpo* de la tabla A-2 buscamos una área de 0.9950. (El área de 0.9950 que aparece en la figura 6-15 es una área acumulativa de la izquierda, y ése es exactamente el tipo de área que se incluye en la tabla A-2). El área de 0.9950 se localiza entre las áreas de 0.9949 y 0.9951 de la tabla A-2, pero hay un asterisco y una nota al pie de página que indican que el área de 0.9950 corresponde a una puntuación z de 2.575.

continúa

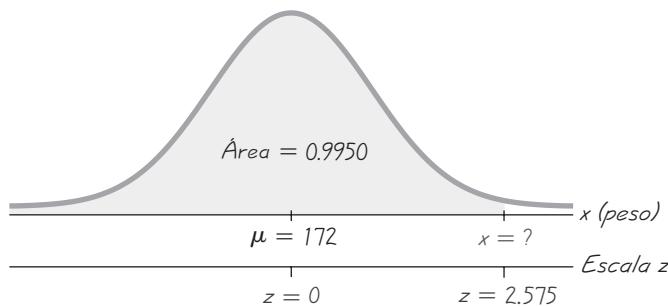


Figura 6-15 Cálculos del peso

Paso 3: Con $z = 2.575$, $\mu = 172$ y $\sigma = 29$, podemos despejar x por medio de la fórmula 6-2:

$$z = \frac{x - \mu}{\sigma} \quad \text{se convierte en} \quad 2.575 = \frac{x - 172}{29}$$

El resultado de $x = 246.675$ lb se puede calcular directamente o utilizando la siguiente versión de la fórmula 6-2:

$$x = \mu + (z \cdot \sigma) = 172 + (2.575 \cdot 29) = 246.675$$

Paso 4: La solución de $x = 247$ lb (redondeado) de la figura 6-15 es razonable porque es mayor que la media de 172 libras.

INTERPRETACIÓN El peso de 247 libras separa al 99.5% de los hombres menos pesados del 0.5% de los hombres más pesados.

EJEMPLO Diseño de tableros de automóviles Al diseñar la ubicación de un reproductor de CD en un nuevo modelo de automóviles, los ingenieros deben considerar el alcance frontal del conductor. Si el reproductor de CD se coloca más allá del alcance, el conductor tendría que mover su cuerpo, lo cual lo distraería y resultaría peligroso (no queremos que alguien se lastime al tratar de escuchar lo mejor de Barry Manilow). Los diseñadores deciden que el reproductor debe ubicarse de manera que esté dentro del alcance del 95% de las mujeres. Las mujeres tienen alcances frontales distribuidos normalmente, con una media de 27.0 pulgadas y una desviación estándar de 1.3 pulgadas (según datos de una encuesta antropométrica de Gordon, Churchill, *et al.*). Calcule el alcance frontal de las mujeres que separa al 95% superior del resto.

SOLUCIÓN

Paso 1: Iniciamos con la gráfica de la figura 6-16. Ya incluimos la media de 27.0 pulgadas e identificamos el área que representa el 95% superior de los alcances frontales. Aun cuando el problema se refiere al 95% superior, la tabla A-2 requiere que trabajemos con una área *izquierda* acumulativa, por lo que restamos 0.95 de 1 para obtener 0.05, que se indica como la región sombreada.

Paso 2: En el *cuerpo* de la tabla A-2 buscamos una área de 0.05. Las áreas más cercanas a 0.05 son 0.0505 y 0.0495, pero existe un asterisco que indica que una área de 0.05 corresponde a una puntuación z de -1.645 .

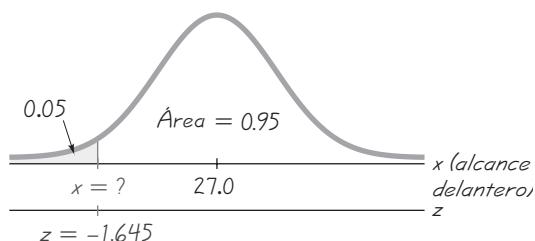


Figura 6-16 Cálculo del valor que separa al 95% superior

Paso 3: Con $z = -1.645$, $\mu = 27.0$ y $\sigma = 1.3$, calculamos x empleando la fórmula 6-2 de manera directa o utilizando la siguiente versión de la fórmula:

$$x = \mu + (z \cdot \sigma) = 27.0 + (-1.645 \cdot 1.3) = 24.8615$$

Paso 4: Si permitimos que $x = 24.8615$ en la figura 6-16, vemos que esta solución es razonable, ya que el alcance frontal que separa al 95% superior del 5% inferior debe ser menor que la media de 27.0 pulgadas.

INTERPRETACIÓN El alcance frontal de 24.9 pulgadas (redondeado) separa al 95% superior del resto, ya que el 95% de las mujeres tienen alcances frontales mayores que 24.9 pulgadas, y el 5% de las mujeres tienen alcances frontales menores que 24.9 pulgadas.

Uso de la tecnología

STATDISK Seleccione **Analysis, Probability Distributions, Normal Distribution**. Indique la puntuación z para calcular las áreas correspondientes, o introduzca el área acumulativa desde la izquierda para calcular la puntuación z . Después de ingresar un valor, haga clic en el botón **Evaluate**.

MINITAB

- Para encontrar el área acumulativa que está a la izquierda de una puntuación z (como en la tabla A-2), seleccione **Calc, Probability Distributions, Normal, Cumulative probabilities**, registre la media y la desviación estándar, después haga clic en el botón de **Input Constant** e ingrese el valor.

- Para encontrar un valor correspondiente a una área conocida, seleccione **Calc, Probability Distributions, Normal**, después seleccione **Inverse cumulative probabilities** e indique la media y la desviación estándar. Seleccione la opción **Input constant** y registre el área total que se encuentra a la izquierda del valor dado.

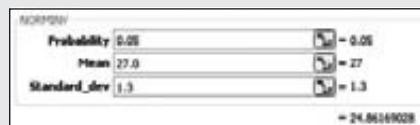
EXCEL

- Para encontrar el área acumulativa a la izquierda de un valor (como en la tabla A-2), haga clic en **fx**, después seleccione **Statistical, NORMDIST**. En el cuadro de diálogo, anote el valor de x , la media y la desviación estándar, y finalmente 1 en el espacio “cumulative”.
- Para encontrar el valor correspondiente a una área conocida, seleccione **fx, Statistical, NORMINV** y proceda a introducir la información en el cuadro de diálogo. Cuando anote el valor de probabilidad, registre el área total a la izquierda del valor dado. Observe la pantalla de Excel que aparece a continuación referente al último ejemplo de esta sección.

TI-83/84 PLUS

- Para calcular el área entre dos valores, presione **2nd VARS, 2** (para **normalecdf**), después proceda a registrar los dos valores, la media y la desviación estándar, todos separados por comas, como en (valor izquierdo, valor derecho, media, desviación estándar).
- Para encontrar un valor correspondiente a una área conocida, presione **2nd, VARS, 3** (para **invNorm**) y proceda a anotar el área total a la izquierda del valor, la media y la desviación estándar con el formato (área total a la izquierda, media, desviación estándar) incluyendo las comas.

Excel



6-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Distribuciones normales.** ¿Cuál es la diferencia entre una distribución normal estándar y una distribución normal no estándar?
2. **Distribuciones normales.** La distribución de puntuaciones de CI es una distribución normal no estándar con una media de 100 y una desviación estándar de 15. ¿Cuáles son los valores de la media y de la desviación estándar después de estandarizar todas las puntuaciones de CI por medio de $z = (x - \mu)/\sigma$?
3. **Dígitos aleatorios.** A menudo las computadoras se utilizan para generar dígitos aleatorios de números telefónicos al realizar una encuesta. ¿Se pueden usar los métodos de esta sección para calcular la probabilidad de que, cuando se genere aleatoriamente un dígito, éste sea menor que 5? ¿Por qué? ¿Cuál es la probabilidad de obtener un dígito menor que 5?
4. **Puntuaciones z y áreas.** ¿Cuál es la diferencia entre una puntuación z y el área bajo la curva de una distribución de probabilidad normal? ¿Una puntuación z puede ser negativa? ¿Una área puede ser negativa?

Puntuaciones de CI. En los ejercicios 5 a 12, suponga que sujetos adultos tienen puntuaciones de CI distribuidas normalmente, con una media de 100 y una desviación estándar de 15 (como en la prueba Wechsler). (Sugerencia: Dibuje una gráfica en cada caso).

5. Calcule la probabilidad de que un adulto seleccionado al azar tenga un CI menor que 130.
6. Calcule la probabilidad de que un adulto seleccionado al azar tenga un CI mayor que 131.5 (el requisito para ser miembro de la organización Mensa, que agrupa a personas con un elevado cociente intelectual).
7. Calcule la probabilidad de que un adulto seleccionado al azar tenga un CI entre 90 y 110 (denominado rango *normal*).
8. Calcule la probabilidad de que un adulto seleccionado al azar tenga un CI entre 110 y 120 (denominado *normal brillante*).
9. Calcule P_{10} , que es la puntuación de CI que separa al 10% inferior del 90% superior.
10. Calcule P_{60} , que es la puntuación de CI que separa al 60% inferior del 40% superior.
11. Calcule la puntuación de CI que separa al 35% superior del resto.
12. Calcule la puntuación de CI que separa al 85% del resto.

En los ejercicios 13 a 16, use la siguiente información (según datos de la National Health Survey).

- La estatura de los **hombres** se distribuye normalmente, con una media de 69.0 pulgadas y una desviación estándar de 2.8 pulgadas.
 - La estatura de las **mujeres** se distribuye normalmente, con una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas.
13. **Estaturas requerida por el Club Beanstalk.** El Club Beanstalk, una organización social para personas altas, requiere que las mujeres midan al menos 70 pulgadas (o 5 pies 10 pulgadas). ¿Qué porcentaje de las mujeres cumple con este requisito?
 14. **Estaturas requeridas para mujeres soldados.** El ejército de Estados Unidos requiere que las mujeres midan entre 58 y 80 pulgadas. Calcule el porcentaje de mujeres que cumplen con este requisito. ¿Se negará a muchas mujeres la oportunidad de unirse al ejército por ser demasiado bajas o demasiado altas?

- 15. Diseño de entradas.** La altura estándar de la entrada de una puerta es de 80 pulgadas.
- ¿Qué porcentaje de hombres son demasiado altos para pasar por la entrada de una puerta estándar sin tener que agacharse, y qué porcentaje de mujeres son demasiado altas para pasar por la entrada de una puerta estándar sin tener que agacharse? Con base en esos resultados, ¿parece que el diseño actual de las entradas de las puertas es adecuado?
 - Si un especialista en estadística diseña una casa de tal manera que la entrada de las puertas tenga una altura suficiente para todos los hombres, con excepción del 5% de los más altos, ¿cuál sería la altura de la entrada?
- 16. Diseño de ataúdes.** El ataúd estándar tiene una longitud interna de 78 pulgadas.
- ¿Qué porcentaje de hombres son demasiado altos para caber en un ataúd estándar, y qué porcentaje de mujeres son demasiado altas para caber en un ataúd estándar? Con base en esos resultados, ¿parece que el tamaño de un ataúd estándar es adecuado?
 - Un fabricante de ataúdes desea reducir los costos de producción fabricando ataúdes más pequeños. ¿En qué longitud interna cabrían todos los hombres con excepción del 1% de los más altos?
- 17. Pesos al nacer.** En Estados Unidos, los pesos al nacer se distribuyen normalmente, con una media de 3420 g y una desviación estándar de 495 g. Si un hospital planea establecer condiciones especiales de observación para el 2% de los bebés menos pesados, ¿qué peso se utilizaría para establecer un punto de corte que separe al 2% de los bebés menos pesados de los demás?
- 18. Pesos al nacer.** En Noruega los pesos al nacer se distribuyen normalmente, con una media de 3570 g y una desviación estándar de 500 g. Repita el ejercicio 17 con los bebés que nacen en Noruega. ¿Difiere mucho el resultado del que se obtuvo en el ejercicio 17?
- 19. Contacto visual.** En un estudio de conducta facial, se toma el tiempo de contacto visual entre los participantes en un grupo de control, durante un periodo de 5 minutos. Los tiempos se distribuyen normalmente, con una media de 184.0 s y una desviación estándar de 55.0 s (según datos de “Ethological Study of Facial Behavior in Nonparanoid and Paranoid Schizophrenic Patients”, de Pittman, Olk, Orr y Singh, *Psychiatry*, vol. 144, núm. 1). Para una persona elegida al azar del grupo de control, calcule la probabilidad de que el tiempo de contacto visual sea mayor que 230.0 s, que es la media de los esquizofrénicos paranoides.
- 20. Temperaturas corporales.** Con base en los resultados muestrales del conjunto de datos 2 del apéndice B, suponga que las temperaturas corporales humanas se distribuyen normalmente, con una media de 98.20°F y una desviación estándar de 0.62°F.
- El hospital Bellevue en la ciudad de Nueva York establece que la temperatura más baja considerada como fiebre es de 100.6°F. ¿Qué porcentaje de personas normales y saludables se consideraría que tienen fiebre? ¿Sugiere este porcentaje que un punto de corte de 100.6°F es apropiado?
 - Los médicos desean seleccionar una temperatura mínima como requisito para solicitar más exámenes médicos. ¿Cuál debe ser esa temperatura, si deseamos que sólo el 5.0% de las personas saludables la excedan? (Un resultado como éste es un *falso positivo*, lo que significa que el resultado de la prueba es positivo, pero el sujeto no está realmente enfermo).
- 21. Duración de embarazos.** La duración de los embarazos se distribuye normalmente, con una media de 268 días y una desviación estándar de 15 días.
- Un uso clásico de la distribución normal está inspirado por una carta dirigida a “Dear Abby”, en la que una mujer afirmaba haber dado a luz 308 días después de una breve visita de su esposo, quien trabajaba en la Marina. Dada esta información, calcule la probabilidad de que un embarazo dure 308 días o más. ¿Qué sugiere el resultado?
 - Si estipulamos que un bebé es *prematuro* cuando la duración del embarazo se encuentra en el 4% inferior, calcule la duración que separa a los bebés prematuros de aquellos que no lo son. Los bebés prematuros suelen requerir cuidados especiales y este resultado podría ser útil para que los administradores de hospitales planeen esos cuidados.

- 22. Ancho de cadera y asientos de aeronaves.** Un grupo de ingenieros desean diseñar asientos para aviones comerciales, de tal manera que sean lo suficientemente amplios para que quepa el 98% de los hombres. (Para abarcar al 100% de los hombres se requerirían asientos muy amplios que serían demasiado costosos). Las anchuras de cadera de los hombres se distribuyen normalmente, con una media de 14.4 pulgadas y una desviación estándar de 1.0 pulgadas (según datos de una encuesta antropomórfica de Gordon, Clauser, *et al.*). Calcule P_{98} , es decir, calcule la anchura de cadera de los hombres que separa al 98% de los individuos con caderas más pequeñas del 2% de individuos con caderas más grandes.
- 23. Diseño de cascos.** Los ingenieros deben tomar en cuenta la anchura de las cabezas de los hombres cuando diseñan cascos para motocicletas. La anchura de las cabezas de los hombres se distribuye normalmente, con una media de 6.0 in y una desviación estándar de 1.0 in (según datos de una encuesta antropométrica de Gordon, Churchill, *et al.*). Debido a limitaciones económicas, los cascos serán diseñados para que se ajusten a todos los hombres, excepto al 2.5% con anchuras más pequeñas y al 2.5% con anchuras más grandes. Calcule las anchuras de cabeza mínima y máxima que se ajustarán a los cascos.
- 24. Distancia entre asientos.** Un requisito común de diseño es que un artículo (como el asiento de una aeronave o de un teatro) debe ajustarse al rango de individuos que caen entre el quinto percentil para las mujeres y el percentil 95 para los hombres. Si se adopta este requisito, ¿cuál es la mínima distancia al estar sentado y cuál es la máxima distancia al estar sentado? Al considerar la distancia al estar sentado, considere la longitud desde el glúteo hasta la rodilla. Los hombres tienen longitudes del glúteo a la rodilla que se distribuyen normalmente, con una media de 23.5 in y una desviación estándar de 1.1 in. Las mujeres tienen longitudes del glúteo a la rodilla que se distribuyen normalmente, con una media de 22.7 pulgadas y una desviación estándar de 1.0 pulgadas.
- 25. Conjunto de datos del apéndice B: presión sanguínea sistólica.** Remítase al conjunto de datos 1 del apéndice B y utilice los niveles de presión sanguínea sistólica de los hombres.
- Utilice los niveles de presión sanguínea sistólica de los hombres para calcular la media y la desviación estándar y verifique que los datos tengan una distribución aproximadamente normal.
 - Suponiendo que los niveles de presión sanguínea sistólica de los hombres se distribuyen normalmente, calcule el quinto percentil y el percentil 95. [Considere los estadísticos del inciso a) como si fueran parámetros de la población]. Este tipo de percentiles resultan útiles cuando los médicos tratan de determinar si los niveles de presión sanguínea son demasiado bajos o demasiado altos.
- 26. Conjunto de datos del apéndice B: presión sanguínea sistólica.** Repita el ejercicio 25 para las mujeres.

6-3 MÁS ALLÁ DE LO BÁSICO

- 27. Unidades de medición.** Las estaturas de las mujeres se distribuyen normalmente.
- Si las estaturas de mujeres individuales se expresan en pulgadas, ¿cuáles serían las unidades utilizadas para las puntuaciones z correspondientes a las estaturas individuales?
 - Si las estaturas de todas las mujeres se convierten a puntuaciones z , ¿cuál es la media, la desviación estándar y la distribución de esas puntuaciones z ?
- 28. Uso de la corrección por continuidad.** Existen muchas situaciones en las que una distribución normal puede utilizarse como una buena aproximación de una variable aleatoria que tiene sólo valores *discretos*. En tales casos podemos emplear esta *corrección por continuidad*: represente cada número entero con el intervalo que se extiende desde 0.5 por debajo del número hasta 0.5 por arriba de él. Suponga que las puntuaciones de CI son números enteros con una distribución aproximadamente normal, con una media de 100 y una desviación estándar de 15.
- Sin utilizar la corrección por continuidad, calcule la probabilidad de seleccionar al azar a alguien con una puntuación de CI mayor que 105.

- b. Utilice la corrección por continuidad y calcule la probabilidad de seleccionar al azar a alguien con una puntuación de CI mayor que 105.
- c. Compare los resultados de los incisos a) y b).
29. **Normalización de calificaciones de un examen.** Una profesora de estadística aplica un examen y descubre que las calificaciones se distribuyen normalmente, con una media de 25 y una desviación estándar de 5. Ella planea normalizar las calificaciones.
- a. Si las normaliza sumando 50 a cada calificación, ¿cuál es la nueva media? ¿Cuál es la nueva desviación estándar?
- b. ¿Será justo normalizarlas sumando 50 a cada calificación? ¿Por qué?
- c. Si las calificaciones se normalizan según el siguiente esquema (en vez de sumar 50), calcule los límites numéricos de cada calificación.
- A: 10% superior
B: calificaciones por arriba del 70% inferior y por debajo del 10% superior
C: calificaciones por arriba del 30% inferior y por debajo del 30% superior
D: calificaciones por arriba del 10% inferior y por debajo del 70% superior
F: 10% inferior
- d. ¿Cuál método de normalización de las calificaciones es más justo: sumar 50 a cada calificación o emplear el esquema dado en el inciso c)? Explique.
30. **Pruebas SAT y ACT.** Las calificaciones de mujeres en la prueba SAT-I se distribuyen de manera normal, con una media de 998 y una desviación estándar de 202. Las calificaciones de mujeres en la prueba ACT se distribuyen de manera normal, con una media de 20.9 y una desviación estándar de 4.6. Suponga que las dos pruebas emplean escalas distintas para medir la misma habilidad.
- a. Si una mujer obtiene una calificación en la prueba SAT que corresponde al percentil 67, calcule su calificación real en la prueba SAT y su calificación equivalente en la prueba ACT.
- b. Si una mujer obtiene una calificación de 1220 en la prueba SAT, calcule su calificación equivalente en la prueba ACT.



Distribuciones y estimadores muestrales

Concepto clave El principal objetivo de esta sección es comprender el concepto de una *distribución muestral de un estadístico*, que es la distribución de todos los valores de ese estadístico cuando se obtienen todas las muestras posibles del mismo tamaño de la misma población. En específico, analizaremos la distribución muestral de la proporción y la distribución muestral de la media. También veremos que algunos estadísticos (como la proporción y la media) son útiles para estimar valores de parámetros poblacionales, mientras que otros estadísticos (como la mediana) no son buenos estimadores de los parámetros poblacionales.

En una encuesta realizada por el Education and Resources Institute, se eligieron al azar 750 estudiantes universitarios, y el 64% (o 0.64) de ellos dijeron que tenían al menos una tarjeta de crédito. Esta encuesta incluye únicamente a 750 participantes de una población de aproximadamente 15 millones de estudiantes universitarios. Sabemos que los estadísticos muestrales varían de manera natural de una muestra a otra, de manera que otra muestra de 750 estudiantes universitarios tal vez genere una proporción muestral diferente del 64% encontrado en la primera encuesta. En este contexto, debemos considerar a la proporción muestral como una variable aleatoria.

Dado que la muestra de 750 estudiantes universitarios representa un porcentaje tan pequeño de la población (0.005%), ¿realmente podemos esperar que la proporción muestral sea un estimado razonable de la proporción real de todos los estudiantes universitarios que tienen tarjetas de crédito? ¡Sí! Los profesionales de la estadística, al ser



¿Prevalecen los niños o las niñas en las familias?

El autor de este libro, sus hermanos y sus sobrinos suman un total de 11 hombres y sólo una mujer. ¿Será este ejemplo un fenómeno en el que un género particular prevalece en una familia? Este tema fue estudiado examinando una muestra aleatoria de 8770 hogares de Estados Unidos. Los resultados se reportaron en la revista *Chance*, en el artículo “Does Having Boys or Girls Run in the Family?”, escrito por Joseph Rodgers y Debby Doughty. Parte de su análisis implica el uso de la distribución de probabilidad binomial. Ellos concluyeron que no encontraron “evidencias contundentes de que un sexo prevalezca más en la familia”.

individuos tan inteligentes, han diseñado métodos para utilizar resultados muestrales y estimar parámetros poblacionales con bastante exactitud. ¿Cómo lo hacen? Su enfoque se basa en la comprensión del comportamiento de la estadística. Al entender la *distribución* de una proporción muestral, los especialistas en estadística pueden determinar qué tan exacta puede ser una proporción muestral individual. Además, ellos comprenden la distribución de medias muestrales. En general, conocen claramente el concepto de la *distribución muestral de un estadístico*.

Definición

La **distribución muestral de un estadístico** (como una proporción muestral o una media muestral) es la distribución de todos los valores del estadístico cuando se obtienen todas las muestras posibles del mismo tamaño n de la misma población. (La distribución muestral de un estadístico generalmente se representa como la distribución de probabilidad en el formato de tabla, histograma de probabilidad o fórmula).

Distribución muestral de proporciones

La definición anterior se puede aplicar al estadístico específico de una proporción muestral.

Definición

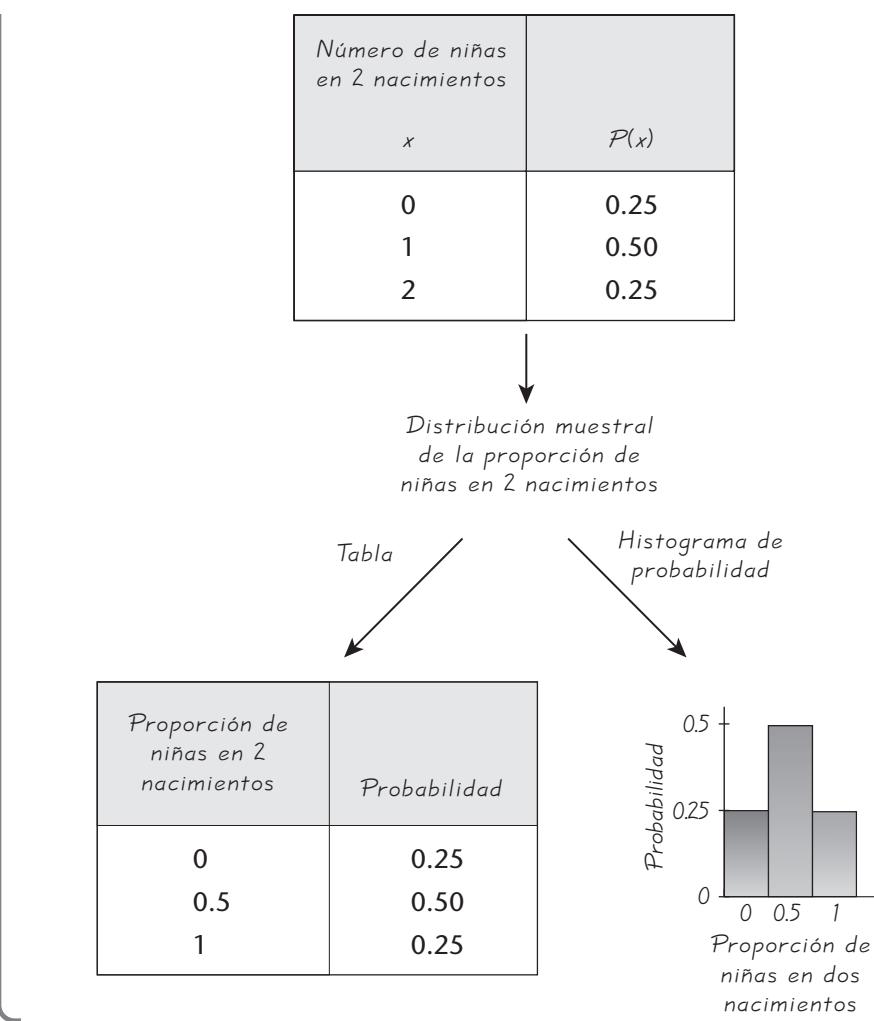
La **distribución muestral de la proporción** es la distribución de probabilidad de proporciones muestrales, donde todas las muestras tienen el mismo tamaño muestral n y provienen de la misma población.

Comprenderemos mejor el importante concepto de una distribución muestral de la proporción si consideramos algunos ejemplos específicos.

EJEMPLO Distribución muestral de la proporción de niñas en dos nacimientos Cuando se eligen dos nacimientos al azar, el espacio muestral es hh, hm, mh, mm. Esos cuatro resultados igualmente probables sugieren que la probabilidad de 0 niñas es de 0.25, la probabilidad de una niña es de 0.50 y la probabilidad de 2 niñas es de 0.25. La imagen que se muestra más adelante indica la distribución de probabilidad del *número* de niñas, seguida por dos formatos diferentes (tabla y gráfica) que describen la distribución muestral de la *proporción* de niñas.

Una distribución muestral también se puede expresar como una fórmula (véase el ejercicio 15), además de los formatos de tabla y gráfica, o se puede describir de otra manera, como la siguiente: “La distribución muestral de la media de la muestra es una distribución normal con $\mu = 100$ y $\sigma = 15$ ”. En esta sección, generalmente describimos una distribución muestral utilizando una tabla que lista los valores del estadístico de muestra, junto con sus probabilidades correspondientes. En capítulos posteriores usaremos algunas de las otras descripciones.

Aunque por lo general las encuestas implican tamaños muestrales de alrededor de 1000 a 2000, y tamaños poblacionales de hasta millones, el siguiente ejemplo implica una población de sólo tres valores, de manera que es fácil listar cada muestra posible.



EJEMPLO Distribución muestral de proporciones Un mariscal de campo lanzó 1 intercepción en su primer juego, 2 intercepciones en su segundo juego, 5 intercepciones en su tercer juego y después se retiró. Considere la *población* consistente en los valores 1, 2, 5. Observe que dos de los valores (1 y 5) son impares, de manera que la proporción de números impares en la población es $2/3$.

- Liste todas las muestras diferentes posibles de tamaño $n = 2$ seleccionadas con reemplazo. (Posteriormente explicaremos por qué el muestreo *con reemplazo* es tan importante). Para cada muestra, calcule la proporción de números *impares*. Utilice una tabla para representar la distribución muestral de la proporción de números impares.
- Calcule la media de la distribución muestral para la proporción de números impares.
- Para la población de 1, 2, 5, la proporción de números impares es $2/3$. ¿La media de la distribución muestral de la proporción de números impares también es igual a $2/3$? ¿Las proporciones muestrales coinciden con el valor de la proporción poblacional? Es decir, ¿las proporciones de la muestra tienen una media igual a la proporción poblacional?



Profetas de las ganancias

Muchos libros y programas de computadoras aseguran ser útiles para predecir números ganadores de la lotería. Algunos utilizan la idea de que ciertos números están “rezagados” (y deberían seleccionarse), ya que no han salido con frecuencia; otros creen que algunos números están “fríos” (y deben evitarse), ya que no han salido con frecuencia; y aún existen otros que utilizan la astrología, la numerología o los sueños. Como la selección de las combinaciones de números ganadores de la lotería son sucesos independientes, tales suposiciones son inútiles. Un método válido es el de elegir números que sean “raros”, en el sentido de que no son seleccionados por otras personas, de manera que si usted gana, no se verá obligado a compartir sus ganancias con otros. Por esta razón, la combinación de 1, 2, 3, 4, 5 y 6 es inadecuada, ya que muchos individuos la utilizan, mientras que la combinación 12, 17, 18, 33, 40, 46 es mucho mejor, al menos hasta la publicación de este libro.

continúa

Tabla 6-2 Distribución muestral de proporciones de números impares		
Muestra	Proporción de números impares	Probabilidad
1, 1	1	1/9
1, 2	0.5	1/9
1, 5	1	1/9
2, 1	0.5	1/9
2, 2	0	1/9
2, 5	0.5	1/9
5, 1	1	1/9
5, 2	0.5	1/9
5, 5	1	1/9

SOLUCIÓN

- a. En la tabla 6-2 se listan las nueve muestras diferentes posibles de tamaño $n = 2$, obtenidas con reemplazo de la población de 1, 2, 5. Esta tabla también contiene el número de valores muestrales que son números impares, incluyendo sus probabilidades. (Como existen 9 muestras igualmente probables, cada muestra tiene una probabilidad de 1/9). La tabla 6-3, que simplemente es una versión condensada de la tabla 6-2, representa de forma concisa la distribución muestral de la proporción de números impares.
 - b. La tabla 6-3 es una distribución de probabilidad, de manera que podemos calcular su media utilizando la fórmula 5-1 de la sección 5-2. La media de 2/3 se obtiene de la siguiente manera:
- $$\mu = \sum[x \cdot P(x)] = (0 \cdot 1/9) + (0.5 \cdot 4/9) + (1 \cdot 4/9) = 6/9 = 2/3$$
- c. La media de la distribución muestral de proporciones es 2/3, y 2/3 de los números en la población son impares. Esto no es una coincidencia. En general, la distribución muestral de proporciones tendrá una media que es igual a la proporción poblacional. A esto nos referimos cuando decimos que las proporciones de la muestra “coinciden” con la proporción poblacional.

INTERPRETACIÓN En el caso de seleccionar dos valores (con reemplazo) de la población de 1, 2, 5, ya identificamos la distribución muestral (tabla 6-3). También encontramos que la media de la distribución muestral es 2/3, que es igual a la proporción de números impares en la población. Por lo tanto, las proporciones muestrales tienden a coincidir con la proporción poblacional, en vez de tender sistemáticamente a subestimar o sobreestimar ese valor.

Tabla 6-3 Versión condensada de la tabla 6-2	
Proporción de números impares	Probabilidad
0	1/9
0.5	4/9
1	4/9

El ejemplo anterior incluye una proporción bastante pequeña, de manera que ahora consideraremos los géneros de los senadores en el CVII (centésimo séptimo) Congreso. Como sólo existen 100 miembros [13 mujeres (M) y 87 hombres (H)], podemos listar la población completa:

H M H H M H H H H H H H M H H H H H H H
 H H H H H H H H H H H H H M M H H H H H H H
 H H H M H H H H H M H H H H H H H H H H H H H
 M H H H H H H H H H H H H H H H H H M M M
 H H H M H M H H H H H H H H H H H H H H H H H

La proporción poblacional de senadoras es $p = 13/100 = 0.13$. Por lo general, no conocemos a todos los miembros de la población, por lo que debemos estimarla a partir de una muestra. Con el propósito de estudiar el comportamiento de las proporciones muestrales, listamos unas pocas muestras de tamaño $n = 10$ e indicamos la proporción correspondiente de mujeres.

Muestra 1: H M H H M H H H H H → la proporción muestral es 0.2

Muestra 2: H M H H H H H H H H → la proporción muestral es 0.1

Muestra 3: H H H H H H M H H H → la proporción muestral es 0.1

Muestra 4: H H H H H H H H H H → la proporción muestral es 0

Muestra 5: H H H H H H H H M H → la proporción muestral es 0.1

Preferimos no listar las 100,000,000,000,000,000 muestras diferentes posibles. En vez de ello, el autor seleccionó al azar 95 muestras adicionales, antes de parar de rodar las llantas de su automóvil. Si combinamos las 95 muestras adicionales con las cinco listadas antes, obtenemos las 100 muestras que se resumen en la tabla 6-4.

En la tabla 6-4 podemos ver que la media de las 100 proporciones muestrales es 0.119, pero si deseamos incluir todas las demás muestras posibles de tamaño 10, la media de las proporciones muestrales sería igual a 0.13, que es el valor de la proporción poblacional. La figura 6-17 presenta la distribución de las 100 proporciones de muestras resumidas en la tabla 6-14. La forma de esa distribución es bastante cercana a la que se habría obtenido con todas las muestras posibles de tamaño 10. Podemos observar que la distribución que se ilustra en la figura 6-17 tiene cierto sesgo hacia la derecha, pero con un poco de alargamiento podría aproximarse a una distribución normal. En la figura 6-18 mostramos los resultados obtenidos de 10,000 muestras, de tamaño 50, seleccionadas al azar y con reemplazo, de la lista anterior de 100 géneros. La figura 6-18 sugiere enfáticamente que la distribución se aproxima a la forma de campana que caracteriza a una distribución normal. Por consiguiente, los resultados de la tabla 6-4 y de la figura 6-18 sugieren lo siguiente.

Tabla 6-4
Resultados de 100 muestras

Proporción de senadoras	Frecuencia
0.0	26
0.1	41
0.2	24
0.3	7
0.4	1
0.5	1
Media:	0.119
Desviación estandar:	0.100

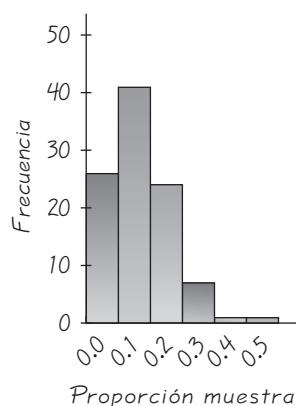
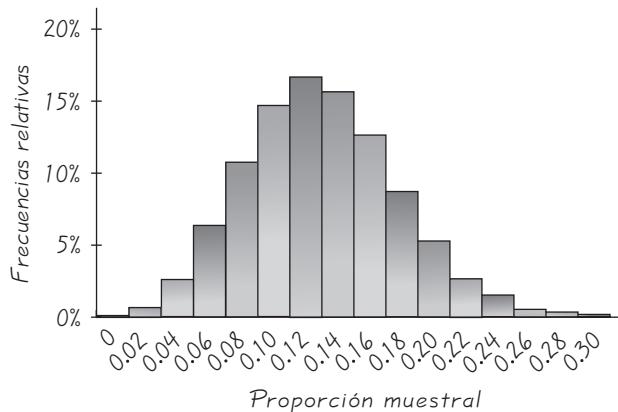


Figura 6-17

100 proporciones muestrales con $n = 10$ en cada muestra

Figura 6-18

10,000 proporciones muestrales con $n = 50$ en cada muestra



Propiedades de la distribución de proporciones muestrales

- Las proporciones muestrales tienden a coincidir con el valor de la proporción poblacional. (Es decir, todas las proporciones muestrales posibles tienen una media igual a la proporción poblacional).
- En ciertas condiciones, la distribución de la proporción muestral puede aproximarse por medio de una distribución normal.

Distribución muestral de la media

Ahora consideraremos la distribución muestral de la media.



Definición

La **distribución muestral de la media** es la distribución de medias muestrales, donde todas las medias tienen el mismo tamaño muestral n y se obtienen de la misma población. (La distribución muestral de la media generalmente se representa como una distribución de probabilidad en formato de tabla, histograma de probabilidad o fórmula).

Nuevamente, en vez de usar conceptos demasiado abstractos, usaremos una población pequeña para ilustrar las propiedades importantes de esta distribución.

EJEMPLO Distribución muestral de la media Una *población* consiste en los valores 1, 2 y 5. Observe que la media de esta población es $\mu = 8/3$.

- Liste todas las muestras posibles (con reemplazo) de tamaño $n = 2$, así como las medias muestrales y sus probabilidades individuales.
- Calcule la media de la distribución muestral.
- La media poblacional es $8/3$. ¿Las medias muestrales coinciden con el valor de la media poblacional?

SOLUCIÓN

- En la tabla 6-5 se listan las nueve muestras diferentes posibles de tamaño $n = 2$, obtenidas con reemplazo de la población 1, 2 y 5. La tabla 6-5 también presenta las medias muestrales e incluye sus probabilidades. (Como hay 9 muestras igualmente probables, cada muestra tiene una

probabilidad de $1/9$). La tabla 6-6, que es una versión condensada de la tabla 6-5, representa de forma concisa la distribución muestral de las medias muestrales.

- b. La tabla 6-6 es una distribución de probabilidad, de manera que podemos calcular su media con la fórmula 5-1 de la sección 5-2. La media de $8/3$ se obtiene de la siguiente manera:

$$\begin{aligned}\mu &= \sum[x \cdot P(x)] = (1.0 \cdot 1/9) + (1.5 \cdot 2/9) + \dots + (5.0 \cdot 1/9) \\ &= 24/9 = 8/3\end{aligned}$$

- c. La media de la distribución muestral de proporciones es $8/3$, y la media de la población también es $8/3$. Nuevamente, no se trata de una coincidencia. En general, la distribución de medias muestrales tiene una media igual a la media poblacional. Por lo tanto, las medias muestrales tienden a coincidir con la media poblacional, en vez de ser sistemáticamente demasiado bajas o demasiado altas.

INTERPRETACIÓN En el caso de seleccionar dos valores (con reemplazo) de la población 1, 2 y 5, ya hemos identificado la distribución muestral (tabla 6-6). También observamos que la media de la distribución muestral es $8/3$, que es igual a la media poblacional. Por consiguiente, las medias muestrales tienden a coincidir con la media poblacional.

Del ejemplo anterior observamos que la media de todas las medias muestrales posibles es igual a la media de la población original, $\mu = 8/3$. Podemos generalizar esto como una propiedad de las medias muestrales: para un tamaño muestral fijo, la media de todas las medias muestrales posibles es igual a la media de la población. Revisaremos esta importante propiedad en la siguiente sección.

Ahora hagamos una observación obvia, pero importante: *las medias muestrales varían*. Vea la tabla 6-5 y observe que las medias muestrales son diferentes. La primera media de muestra es 1.0, la segunda media de muestra es 1.5, etcétera. Esto nos conduce a la siguiente definición.

Definición

El valor de un estadístico, como la media muestral \bar{x} , depende de los valores particulares incluidos en la muestra y generalmente varía de una muestra a otra. Esta variabilidad de un estadístico se denomina **variabilidad de muestreo**.

En el capítulo 2 estudiamos las características importantes de un conjunto de datos: centro, variación, distribución, valores extremos y patrón temporal. Al examinar las muestras en la tabla 6-5, ya identificamos una propiedad que describe el comportamiento de las medias muestrales: la media de las medias muestrales es igual a la media de la población. Esta propiedad enfatiza la característica central; investigaremos otras características en la siguiente sección. Veremos que al incrementarse el tamaño de la muestra, la distribución muestral de medias muestrales tiende a convertirse en una *distribución normal*. En consecuencia, la distribución normal asume una importancia que va más allá de las aplicaciones ilustradas en la sección 6-3. La distribución normal se utilizará en muchos casos en los que deseamos emplear una media muestral \bar{x} con el propósito de hacer alguna inferencia acerca de una media poblacional μ .

Tabla 6-5 Distribución muestral de \bar{x}		
Muestra	Media \bar{x}	Probabilidad
1, 1	1.0	1/9
1, 2	1.5	1/9
1, 5	3.0	1/9
2, 1	1.5	1/9
2, 2	2.0	1/9
2, 5	3.5	1/9
5, 1	3.0	1/9
5, 2	3.5	1/9
5, 5	5.0	1/9



Tabla 6-6 Versión condensada de la tabla 6-5	
\bar{x}	Probabilidad
1.0	1/9
1.5	2/9
2.0	1/9
3.0	2/9
3.5	2/9
5.0	1/9

¿Cuáles estadísticos son buenos estimadores de parámetros?

En el capítulo 7 estudiaremos métodos formales para el uso de estadísticos muestrales con el fin de hacer estimaciones de los valores de parámetros de población. Algunos estadísticos funcionan mucho mejor que otros, y podemos juzgar su valor examinando sus distribuciones muestrales, como en el siguiente ejemplo.

EJEMPLO Distribuciones muestrales Una *población* consiste en los valores 1, 2 y 5. Si seleccionamos al azar muestras de tamaño 2 con reemplazo, existen nueve distintas muestras posibles, que se listan en la tabla 6-7. Como las nueve muestras distintas son igualmente posibles, cada muestra tiene una probabilidad de 1/9.

- Para cada muestra calcule la media, la mediana, el rango, la varianza, la desviación estándar y la proporción de valores muestrales impares. (Para cada estadístico esto generará nueve valores que, cuando se asocien con nueve probabilidades de 1/9 cada una, se combinarán para formar una *distribución muestral del estadístico*).
- Para cada estadístico, calcule la media de los resultados del inciso a).

Tabla 6-7 Distribuciones muestrales de estadísticos (para muestras de tamaño 2, obtenidas con reemplazo de la población 1, 2, 5)

Muestra	Media \bar{x}	Mediana	Rango	Varianza s^2	Desviación estándar s	Proporción de números impares	Proba- bilidad
1, 1	1.0	1.0	0	0.0	0.000	1	1/9
1, 2	1.5	1.5	1	0.5	0.707	0.5	1/9
1, 5	3.0	3.0	4	8.0	2.828	1	1/9
2, 1	1.5	1.5	1	0.5	0.707	0.5	1/9
2, 2	2.0	2.0	0	0.0	0.000	0	1/9
2, 5	3.5	3.5	3	4.5	2.121	0.5	1/9
5, 1	3.0	3.0	4	8.0	2.828	1	1/9
5, 2	3.5	3.5	3	4.5	2.121	0.5	1/9
5, 5	5.0	5.0	0	0.0	0.000	1	1/9
Media de valores de los estadísticos	8/3	8/3	16/9	26/9	1.3	2/3	
Parámetro poblacional	8/3	2	4	26/9	1.7	2/3	
¿Coincide el estadístico muestral con el parámetro poblacional?	Sí	No	No	Sí	No	Sí	

- c. Compare las medias del inciso b) con los parámetros de población correspondientes, después determine si cada estadístico coincide con el valor del parámetro poblacional. Por ejemplo, las medias muestrales tienden a centrarse alrededor del valor de la media poblacional, que es $8/3$, de manera que las medias muestrales coinciden con el valor de la media poblacional.

SOLUCIÓN

- Vea la tabla 6-7, que incluye los estadísticos individuales para cada muestra.
- Las medias de los estadísticos muestrales aparecen casi al final de la tabla 6-7. La media de las medias muestrales es $8/3$, la media de las medianas muestrales es $8/3$, y así sucesivamente.
- El renglón inferior de la tabla 6-7 está basado en una comparación de los parámetros poblacionales y los resultados de los estadísticos muestrales. Por ejemplo, la media poblacional de 1, 2, 5 es $\mu = 8/3$, y las medias muestrales “coinciden” con el valor de $8/3$, ya que la media de las medias muestrales también es $8/3$.

INTERPRETACIÓN Con base en los resultados de la tabla 6-7, podemos observar que cuando se utiliza un estadístico muestral para estimar un parámetro de población, algunos estadísticos son buenos en el sentido de que coinciden con el parámetro poblacional y, por lo tanto, tienden a producir buenos resultados. Estadísticos como éstos se denominan *estimadores sin sesgo*. Otros estadísticos no son tan buenos (ya que son *estimadores sesgados*). He aquí un resumen.

- **Estadísticos que coinciden con los parámetros poblacionales:** media, varianza, proporción
- **Estadísticos que no coinciden con los parámetros poblacionales:** mediana, rango, desviación estándar

Aun cuando la desviación estándar muestral no coincide con la desviación estándar poblacional, el sesgo es relativamente pequeño en muestras grandes, de manera que con frecuencia se utiliza s para estimar σ . En consecuencia, las medias, proporciones, varianzas y desviaciones estándar serán consideradas temas importantes en los siguientes capítulos, en tanto que la mediana y el rango se utilizarán en pocas ocasiones.

¿Por qué se hace el muestreo *con reemplazo*? Para muestras pequeñas como las que hemos considerado hasta ahora en esta sección, el muestreo *sin reemplazo* tiene la ventaja práctica de evitar una duplicación inútil, siempre que se selecciona el mismo elemento más de una vez. Sin embargo, estamos particularmente interesados en el muestreo *con reemplazo* por las siguientes razones: **1.** Cuando se selecciona una muestra relativamente pequeña de una población grande, no hay gran diferencia si realizamos el muestreo con reemplazo o sin él. **2.** El muestreo con reemplazo da como resultado sucesos independientes que no se ven afectados por resultados previos, y los sucesos independientes son más fáciles de analizar y derivan en fórmulas más simples. Por lo tanto, nos enfocamos en el comportamiento de muestras seleccionadas aleatoriamente *con reemplazo*. Muchos de los procedimientos estadísticos que se analizan en los siguientes capítulos se basan en el supuesto de que el muestreo se hizo con reemplazo.

El aspecto más importante de esta sección es analizar el concepto de distribución muestral de un estadístico. Considere el objetivo de tratar de calcular la temperatura corporal media de todos los adultos. Puesto que la población es demasiado grande, no

es práctico medir la temperatura de cada adulto. En vez de ello, obtenemos una muestra de temperaturas corporales y la utilizamos para estimar la media poblacional. El conjunto de datos 2 del apéndice B incluye una muestra de 106 temperaturas corporales, y la media de esta muestra es $\bar{x} = 98.20^{\circ}\text{F}$. Las conclusiones que hacemos acerca de la temperatura media poblacional de todos los adultos requieren que comprendamos el comportamiento de la distribución muestral de todas las medias muestrales de este tipo. Aunque no es práctico obtener cada muestra posible y nos conformemos con una sola muestra, podemos extraer algunas conclusiones muy importantes y significativas acerca de la población de todas las temperaturas corporales. Uno de los objetivos principales de las secciones y los capítulos siguientes es aprender el uso eficaz de una muestra para obtener conclusiones acerca de una población. En la sección 6-5 tomamos en cuenta más detalles acerca de la distribución muestral de medias muestrales, y en la sección 6-6 estudiamos más detalles acerca de la distribución muestral de las proporciones de muestra.

6-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Distribución muestral.** Responda esta pregunta con sus propias palabras: “¿Qué es una distribución muestral?”
- 2. Estimador sin sesgo.** ¿A qué nos referimos cuando decimos que las medias muestrales “coinciden” con la media poblacional o que la media muestral es un *estimador sin sesgo* de la media poblacional?
- 3. Estimadores sin sesgo.** ¿Cuáles de los siguientes estadísticos son estimadores sin sesgo?
 - a. La media muestral utilizada para estimar una media poblacional
 - b. La mediana muestral utilizada para estimar una mediana poblacional
 - c. La proporción muestral utilizada para estimar una proporción poblacional
 - d. La varianza muestral utilizada para estimar una varianza poblacional
 - e. La desviación estándar muestral utilizada para estimar una desviación estándar poblacional
 - f. El rango muestral utilizado para estimar un rango poblacional
- 4. Muestreo con reemplazo.** Dé al menos una razón por la que los métodos estadísticos tienden a estar basados en la suposición de que el muestreo se realiza *con reemplazo* y no *sin reemplazo*.
- 5. Encuesta de votantes.** Con base en una muestra aleatoria de $n = 400$ votantes, la división de noticias de la NBC predice que el candidato demócrata a la presidencia obtendrá el 49% de los votos, aunque en realidad obtiene el 51%. ¿Debemos concluir que la encuesta se realizó incorrectamente? ¿Por qué?
- 6. Distribución muestral de temperaturas corporales.** El conjunto de datos 2 del apéndice B incluye una muestra de 106 temperaturas corporales de adultos. Si construimos un histograma para describir la forma de la distribución de esa muestra, ¿el histograma tendría la forma de una *distribución muestral* de medias muestrales? ¿Por qué?

En los ejercicios 7 a 14, represente distribuciones muestrales en un formato de tabla que liste los diferentes valores del estadístico muestral junto con sus probabilidades correspondientes.

- 7. Centro telefónico.** La Nome Ice Company abrió únicamente durante tres días (adivine por qué). He aquí el número de llamadas recibidas durante cada uno de esos días: 10, 6, 5. Suponga que se seleccionan al azar muestras de tamaño 2, *con reemplazo*, de esta población de tres valores.
 - a. Liste las 9 muestras diferentes posibles y calcule la media de cada una.
 - b. Identifique la probabilidad de cada muestra y describa la distribución muestral de las medias muestrales (*Sugerencia:* Véase la tabla 6-6).

- c. Calcule la media de la distribución muestral.
- d. ¿Será igual la media de la distribución muestral [del inciso c)] a la media de la población de los tres valores listados? ¿Serán estas medias *siempre* iguales?
- 8. Telemercadeo.** A continuación se presenta el número de ventas por día de Kim Ryan, un cortés vendedor que trabajó durante cuatro días antes de ser despedido: 1, 11, 9, 3. Suponga que se seleccionan al azar muestras de tamaño 2, *con reemplazo*, de esta población de cuatro valores.
- Liste las 16 diferentes muestras posibles y calcule la media de cada una.
 - Identifique la probabilidad de cada muestra y después describa la distribución muestral de las medias muestrales (*Sugerencia:* Véase la tabla 6-6).
 - Calcule la media de la distribución muestral.
 - ¿Es la media de la distribución muestral [del inciso c)] igual a la media de la población de los cuatro valores listados? ¿Serán estas medias *siempre* iguales?
- 9. La gente más adinerada.** Los bienes (en miles de millones de dólares) de las cinco personas más adineradas de Estados Unidos son 47 (Bill Gates), 43 (Warren Buffet), 21 (Paul Allen), 20 (Alice Walton) y 20 (Helen Walton). Suponga que se seleccionan al azar muestras de tamaño 2 *con reemplazo* de esta población de cinco valores.
- Después de listar las 25 muestras diferentes posibles y de calcular la media de cada muestra, utilice una tabla para describir la distribución muestral de las medias muestrales. (*Sugerencia:* Véase la tabla 6-6).
 - Calcule la media de la distribución muestral.
 - ¿Es la media de la distribución muestral [del inciso b)] igual a la media de la población de los cinco valores listados? ¿Serán estas medias *siempre* iguales?
- 10. Presidentes militares.** A continuación se lista la población de los cinco presidentes de Estados Unidos que tuvieron profesiones militares, junto con sus edades en el momento de tomar posesión: Eisenhower (62), Grant (46), Harrison (68), Taylor (64) y Washington (57). Suponga que se seleccionan al azar muestras de tamaño 2, *con reemplazo*, de la población de las cinco edades.
- Después de identificar las 25 distintas muestras posibles y calcular la media de cada una, use una tabla para describir la distribución muestral de las medias muestrales. (*Sugerencia:* Véase la tabla 6-6).
 - Calcule la media de la distribución muestral.
 - ¿Es la media de la distribución muestral [del inciso b)] igual a la media de la población de los cinco valores listados? ¿Serán estas medias *siempre* iguales?
- 11. Genética.** Un experimento en genética incluye una población de moscas de la fruta consistente en un macho llamado Mike y tres hembras llamadas Anna, Barbara y Chris. Suponga que se seleccionan dos moscas de la fruta al azar, *con reemplazo*.
- Después de identificar las 16 diferentes muestras posibles, calcule la proporción de hembras en cada muestra, luego use una tabla para describir la distribución muestral de las proporciones de hembras. (*Sugerencia:* Véase la tabla 6-3).
 - Calcule la media de la distribución muestral.
 - ¿Es la media de la distribución muestral [del inciso b)] igual a la proporción poblacional de hembras? ¿Coincide *siempre* la media de la distribución muestral de proporciones con la proporción poblacional?
- 12. Control de calidad.** Después de construir una máquina nueva, se producen 5 faros prototipo para automóvil y se descubre que 2 son defectuosos (D) y tres son aceptables (A). Suponga que se seleccionan dos faros al azar, *con reemplazo*, de esta población.
- Después de identificar las 25 distintas muestras posibles, calcule la proporción de defectos en cada una y luego use una tabla para describir la distribución muestral de proporciones de defectos (*Sugerencia:* Véase la tabla 6-3).

continúa

- b. Calcule la media de la distribución muestral.
 - c. ¿Es la media de la distribución muestral [del inciso b)] igual a la proporción poblacional de defectos? ¿Coincide *siempre* la media de la distribución muestral de proporciones con la proporción poblacional?
- 13. Clasificación de competidores olímpicos de triatlón.** Mujeres estadounidenses compitieron en el triatlón de los Juegos Olímpicos realizados en Atenas, y al final ocuparon los lugares 3, 9 y 23. Suponga que se seleccionan al azar muestras de tamaño 2 con reemplazo.
- a. Utilice una tabla para describir la distribución muestral de las medias muestrales.
 - b. Puesto que los datos se refieren a lugares de clasificación, ¿realmente tiene sentido identificar la distribución muestral de las medias muestrales?
- 14. Mediana y lunas de Júpiter.** Júpiter tiene 4 lunas grandes y 12 lunas pequeñas. Los tiempos de órbita de las 4 lunas grandes son los siguientes (en días): 1.8 (Io), 3.6 (Europa), 7.2 (Ganímedes) y 16.7 (Calixto). Suponga que se seleccionan al azar dos de estos valores con reemplazo.
- a. Despues de identificar las 16 diferentes muestras posibles, calcule la mediana de cada una y luego utilice una tabla para describir la distribución muestral de las medianas.
 - b. Calcule la media de la distribución muestral.
 - c. ¿Es la media de la distribución muestral [del inciso b)] igual a la mediana de la población? ¿Es la mediana un estimador sin sesgo de la mediana poblacional?

6-4 MÁS ALLÁ DE LO BÁSICO

- 15. Uso de una fórmula para describir una distribución muestral.** Vea el primer ejemplo de esta sección, el cual incluye una tabla y una gráfica que describen la distribución muestral de las proporciones de niñas en dos nacimientos. Considere la fórmula que aparece abajo y evalúela utilizando proporciones muestrales x de 0, 0.5 y 1. Con base en los resultados, ¿la fórmula describe la distribución muestral? ¿Por qué?

$$P(x) = \frac{1}{2(2 - 2x)!(2x)!} \quad \text{donde } x = 0, 0.5, 1$$

- 16. Desviación media absoluta.** La población de 1, 2 y 5 se utilizó para elaborar la tabla 6-7. Identifique la distribución muestral de la desviación media absoluta (definida en la sección 3-3), después determine si la desviación media absoluta de una muestra es un buen estadístico para estimar la desviación media absoluta de la población.



6-5 El teorema del límite central

Concepto clave En la sección 6-4 se incluyó cierto análisis de la distribución muestral de \bar{x} , y en esta sección se describen procedimientos para utilizarla en situaciones muy reales y prácticas. Los procedimientos de esta sección constituyen la base para la estimación de parámetros poblacionales y la prueba de hipótesis, temas que se analizarán con profundidad en los siguientes capítulos. Cuando se selecciona una muestra aleatoria simple de una población con media μ y desviación estándar σ , es esencial conocer los siguientes principios:

1. Si $n > 30$, entonces las medias muestrales tienen una distribución que se puede aproximar por medio de una distribución normal, con una media μ y una desviación estándar σ/\sqrt{n} . (Éste es el lineamiento que suele utilizarse, independientemente de la distribución de la población original).

2. Si $n \leq 30$ y la población original tiene una distribución normal, entonces las medias muestrales tienen una distribución normal con una media μ y una desviación estándar σ/\sqrt{n} .
3. Si $n \leq 30$, pero la población original no tiene una distribución normal, entonces no se aplican los métodos de esta sección.

Trate de conservar la siguiente idea en mente: cuando tomamos muestras de una población, deseamos conocer el comportamiento de las medias muestrales. El *teorema del límite central* nos dice que si el tamaño de una muestra es lo suficientemente grande, la distribución de las medias muestrales se puede aproximar por medio de una *distribución normal*, aun cuando la población original no esté distribuida de forma normal. Aunque hablamos de un “teorema”, no incluimos pruebas rigurosas, sino que nos enfocamos en los *conceptos* y en su aplicación. He aquí los puntos clave que conforman una base importante para los siguientes capítulos.

El teorema del límite central y la distribución muestral de \bar{x}

Dado que:

1. La variable aleatoria x tiene una distribución (que puede o no ser normal) con media μ y desviación estándar σ .
2. Todas las muestras aleatorias del mismo tamaño n se seleccionan de la población. (Las muestras se seleccionan de manera que todas las muestras posibles de tamaño n tengan la misma probabilidad de ser seleccionadas).

Conclusiones:

1. Conforme el tamaño de la muestra aumenta, la distribución de las medias muestrales \bar{x} se aproximarán a una distribución *normal*.
2. La media de todas las medias muestrales es la media poblacional μ . (Es decir, la distribución normal de la conclusión 1 tiene una media μ).
3. La desviación estándar de todas las medias muestrales es σ/\sqrt{n} . (Es decir, la distribución normal de la conclusión 1 tiene una desviación estándar σ/\sqrt{n} .)

Reglas prácticas de uso común

1. Si la población original no está distribuida normalmente, la siguiente es una directriz común: para muestras de tamaño n mayores que 30, la distribución de las medias muestrales puede aproximarse razonablemente bien por medio de una distribución normal. (Existen excepciones, como las poblaciones con distribuciones muy diferentes a la normal, que requieren tamaños de muestra mucho más grandes que 30, aunque tales excepciones son relativamente raras). La aproximación mejora conforme el tamaño muestral n se incrementa.
2. Si la población original se distribuye normalmente, entonces las medias muestrales estarán distribuidas normalmente para *cualquier* tamaño de muestra n (no sólo los valores de n mayores que 30).

El teorema del límite central implica dos distribuciones diferentes: la distribución de la población original y la distribución de las medias muestrales. Igual que en capítulos anteriores, utilizamos los símbolos μ y σ para denotar la media y la desviación estándar de la población original, pero ahora necesitamos nuevas notaciones para la media y la desviación estándar de la distribución de las medias muestrales.

Notación para la distribución muestral de \bar{x}

Si se seleccionan todas las muestras aleatorias posibles de tamaño n de una población con media μ y desviación estándar σ , la media de las medias muestrales se denota con $\mu_{\bar{x}}$, de manera que

$$\mu_{\bar{x}} = \mu$$

Asimismo, la desviación estándar de las medias muestrales se denota con $\sigma_{\bar{x}}$, de manera que

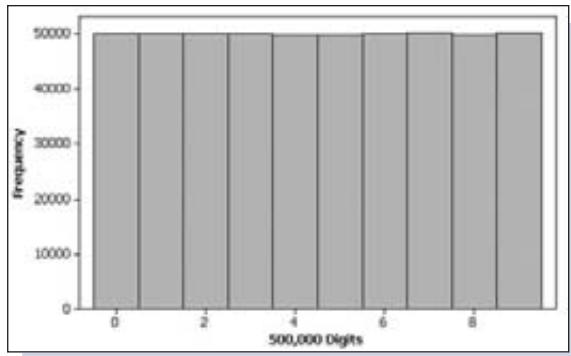
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$\sigma_{\bar{x}}$, suele denominarse el **error estándar de la media**.

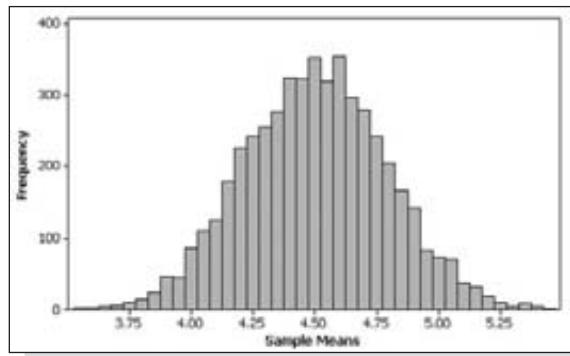
EJEMPLO Simulación con dígitos aleatorios Con frecuencia las computadoras se utilizan para generar dígitos aleatorios de números telefónicos para realizar encuestas. (Por ejemplo, el Pew Research Center genera aleatoriamente los últimos dos dígitos de números telefónicos para evitar un “sesgo de lista”). Los dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 se generan de forma que todos son igualmente probables. La imagen de Minitab que aparece a continuación incluye el histograma de 500,000 dígitos generados. Observe que la distribución tiene una apariencia uniforme, como esperábamos.

Ahora, agrupamos los 500,000 dígitos en 5000 muestras, cada una con $n = 100$ valores. Calculamos la media de cada muestra y mostramos el histograma de las 5000 medias muestrales. Observe el siguiente efecto sorprendente: ¡Aun cuando los 500,000 dígitos originales tienen una distribución *uniforme*, la distribución de las 500 medias muestrales es aproximadamente *normal*! Un fenómeno verdaderamente fascinante e intrigante de la estadística es el hecho de que al obtener muestras de cualquier distribución podamos crear una distribución de medias muestrales que es normal o al menos aproximadamente normal.

Minitab



Minitab



Aplicación del teorema del límite central

Muchos problemas prácticos importantes se resuelven mediante el teorema del límite central. Cuando trabaje con este tipo de problemas, recuerde que si el tamaño de la muestra es mayor que 30, o si la población original se distribuye normalmente,

debe tratar la distribución de medias muestrales como si fuera una distribución normal con media μ y desviación estándar σ/\sqrt{n} .

En el siguiente ejemplo, el inciso *a*) incluye un valor *individual*, pero el inciso *b*) incluye la media de una *muestra* de 20 hombres, por lo que debemos usar el teorema del límite central al trabajar con la variable aleatoria \bar{x} . Estudie este ejemplo con atención para comprender la diferencia significativa entre los procedimientos utilizados en los incisos *a*) y *b*). Observe cómo este ejemplo ilustra el siguiente procedimiento de trabajo:

- **Cuando trabaje con un valor *individual* de una población distribuida normalmente, utilice los métodos de la sección 6-3. Use $z = \frac{x - \mu}{\sigma}$.**
- **Cuando trabaje con una media de alguna *muestra* (o grupo), asegúrese de utilizar el valor de σ/\sqrt{n} para la desviación estándar de las medias muestrales. Use $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$.**



EJEMPLO Seguridad de taxis acuáticos

En el problema del capítulo señalamos que algunos pasajeros murieron cuando un taxi acuático se hundió en el Inner Harbor de Baltimore. Los hombres suelen ser más pesados que las mujeres y los niños; por lo tanto, supongamos que al cargar un taxi acuático la situación extrema es aquella en la que todos los pasajeros son hombres. En concordancia con los datos de la National Health and Nutrition Examination Survey, suponga que los pesos de los hombres se distribuyen normalmente, con una media de 172 libras y una desviación estándar de 29 libras.

- Calcule la probabilidad de que, si se selecciona un hombre al azar, su peso sea mayor que 175 libras.
- Calcule la probabilidad de que 20 hombres elegidos al azar tengan una media mayor que 175 libras (de manera que su peso total exceda la capacidad segura de 3500 libras).

SOLUCIÓN

- Enfoque:** Utilice los métodos presentados en la sección 6-3 (porque estamos trabajando con un valor individual de una población distribuida normalmente). Buscamos el área de la región sombreada en la figura 6-19a). Si utilizamos la tabla A-2, convertimos el peso de 175 a su puntuación z correspondiente:

$$z = \frac{x - \mu}{\sigma} = \frac{175 - 172}{29} = 0.10$$

Nos remitimos a la tabla A-2 y utilizamos $z = 0.10$ para encontrar que el área acumulativa a la izquierda de 175 libras es 0.5398. Por lo tanto, la región sombreada es $1 - 0.5398 = 0.4602$. La probabilidad de que un hombre elegido al azar pese más de 175 libras es de 0.4602. (Si se utiliza una calculadora o un programa de cómputo en vez de la tabla A-2, el resultado más exacto es 0.4588 en lugar de 0.4602).

continúa

b. *Enfoque:* Utilice el teorema del límite central (porque estamos trabajando con la media de una muestra de 20 hombres y no con un solo hombre). Aun cuando el tamaño de la muestra no es mayor que 30, utilizamos una distribución normal por la siguiente razón: la población original de hombres tiene una distribución normal, de manera que las muestras de *cualquier* tamaño producirán medias distribuidas normalmente. Puesto que estamos trabajando con una distribución de medias muestrales, debemos utilizar los parámetros $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$, que se evalúan de la siguiente manera:

$$\begin{aligned}\mu_{\bar{x}} &= \mu = 172 \\ \sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} = \frac{29}{\sqrt{20}} = 6.4845971\end{aligned}$$

He aquí un punto realmente importante: debemos utilizar la desviación estándar calculada de 6.4845971 y no la desviación estándar original de 29 (porque estamos trabajando con la distribución de medias muestrales cuya desviación estándar es 6.4845971, y no con la distribución de pesos individuales cuya desviación estándar es 29). Necesitamos calcular el área sombreada que se destaca en la figura 6-1b). Si usamos la tabla A-2, obtenemos la puntuación z relevante, que se calcula de la siguiente manera:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{175 - 172}{\frac{29}{\sqrt{20}}} = \frac{3}{6.4845971} = 0.46$$

Si nos remitimos a la tabla A-2, encontramos que $z = 0.46$ corresponde a una área izquierda acumulativa de 0.6772, de manera que la región sombreada es $1 - 0.6772 = 0.3228$. La probabilidad de que 20 hombres tengan un peso medio mayor que 175 libras es de 0.3228. (Si se utiliza una calculadora o un programa de cómputo, el resultado es 0.3218 en vez de 0.3228).

INTERPRETACIÓN Existe una probabilidad de 0.4602 de que un hombre pese más de 175 libras, y una probabilidad de 0.3228 de que 20 hombres tengan un peso medio mayor de 175 libras. Dado que la capacidad segura del taxi

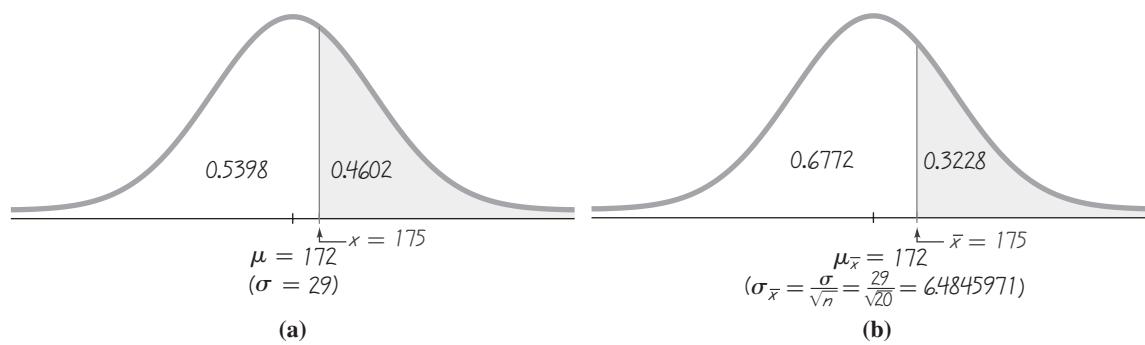


Figura 6-19 Pesos de hombres

a) Distribución de pesos individuales de hombres; b) Distribución de medias muestrales

acuático es de 3500 libras, es muy probable (con una probabilidad de 0.3228) que se sobrecargue si se transporta a 20 hombres elegidos al azar. Dado que ya han muerto 21 personas, y dada la alta probabilidad de sobrecarga, lo más pertinente sería limitar el número de pasajeros a menos de 20. La capacidad de 20 pasajeros no es suficiente.

Los cálculos que se utilizaron aquí son exactamente iguales a los cálculos que utilizan los ingenieros al diseñar teleféricos, elevadores, escaleras eléctricas, aviones y otros aparatos que transportan personas.



Interpretación de resultados

El siguiente ejemplo ilustra otra aplicación del teorema del límite central; examine con atención la conclusión a la que se llega. Este ejemplo ilustra el tipo de pensamiento que es fundamental para el importante procedimiento de prueba de hipótesis (que se estudia en el capítulo 8). Este ejemplo ilustra la regla del suceso infrecuente de la estadística inferencial, presentado inicialmente en la sección 4-1.

Regla del suceso infrecuente

Si, bajo cierto supuesto, la probabilidad de un suceso particular observado es pequeña de manera excepcional, concluimos que el supuesto probablemente no es correcto.

EJEMPLO Temperaturas corporales Suponga que la población de temperaturas corporales humanas tiene una media de 98.6°F , como suele creerse. También suponga que la desviación estándar de la población es 0.62°F (según datos de investigadores de la Universidad de Maryland). Si se selecciona al azar una muestra de tamaño $n = 106$, calcule la probabilidad de obtener una media de 98.2°F o menor. (En realidad se obtuvo el valor de 98.2°F ; vea las temperaturas de media noche del día 2 en el conjunto de datos 2 del apéndice B).

SOLUCIÓN

No tenemos información sobre la distribución de la población, pero, como el tamaño de la muestra $n = 106$ excede a 30, utilizamos el teorema del límite central y concluimos que la distribución de medias muestrales es normal, con estos parámetros:

$$\mu_{\bar{x}} = \mu = 98.6 \quad (\text{por suposición})$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.62}{\sqrt{106}} = 0.0602197$$

La figura 6-20 destaca el área sombreada (observe la pequeña cola izquierda de la gráfica), correspondiente a la probabilidad que buscamos. Una vez encontrados los parámetros que se aplican a la distribución de la figura 6-20, ahora

El efecto placebo

Durante mucho tiempo se ha creído que los placebos realmente ayudan a algunos pacientes. De hecho, algunos estudios serios han demostrado que cuando se administra un placebo (un tratamiento sin valor medicinal), muchos sujetos de prueba manifiestan cierta mejoría. Los estimados de las tasas de mejoría van por lo común de una tercera a dos terceras partes de los pacientes. Sin embargo, un estudio más reciente sugiere que los placebos no tienen efecto real. Un artículo en la *New England Journal of Medicine* (vol. 334, núm. 21) se basó en la investigación de 114 ensayos médicos durante 50 años. Los autores del artículo concluyeron que los placebos parecen tener algún efecto sólo en aliviar el dolor, pero no en otras condiciones físicas. Ellos concluyen que, excepto en ensayos clínicos, el uso de placebos “no puede recomendarse”.

continúa



El teorema del límite central borroso

En *The Cartoon Guide to Statistics*, los autores Gonick y Smith describen el teorema del límite central borroso de la siguiente manera: "Los datos que se ven influidos por efectos aleatorios muy pequeños y sin relación entre sí se distribuyen aproximadamente de manera normal. Esto explica por qué la normalidad está en todos lados: en las fluctuaciones del mercado de acciones, en los pesos de estudiantes, en los promedios anuales de temperatura y en las calificaciones del SAT. Todos son el resultado de muchos efectos diferentes". La estatura de las personas, por ejemplo, es el resultado de factores hereditarios, factores ambientales, nutrición, cuidado de la salud, región geográfica y otras influencias que, cuando se combinan, producen valores distribuidos de forma normal.

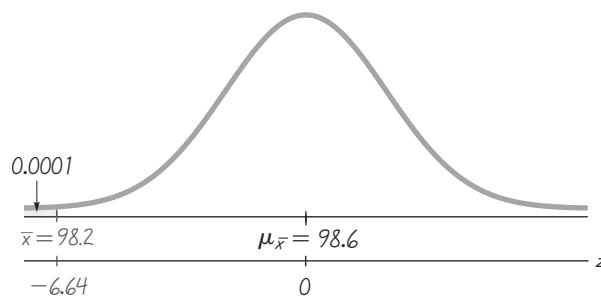


Figura 6-20

Distribución de temperaturas corporales medias, para muestras de tamaño $n = 106$

podemos encontrar el área sombreada utilizando los mismos procedimientos desarrollados en la sección 6-3. En la tabla A-2 primero encontramos la puntuación z :

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{98.20 - 98.6}{0.0602197} = -6.64$$

Si nos remitimos a la tabla A-2, encontramos que $z = -6.64$ no aparece, pero para los valores de z que están por debajo de -3.49 utilizamos una área de 0.0001 para el área izquierda acumulativa hasta $z = -3.49$. Por lo tanto, concluimos que la región sombreada de la figura 6-20 es 0.0001. (Si se utiliza una calculadora TI-83/84 Plus o un programa de cómputo, el área de la región sombreada se acerca a 0.0000000002, pero incluso estos resultados son sólo aproximaciones. Con seguridad podemos reportar que la probabilidad es muy baja, menor que 0.0001).

INTERPRETACIÓN Los resultados demuestran que si la media de nuestra temperatura corporal es en realidad 98.6°F , entonces existe una probabilidad sumamente baja de obtener una media de muestra de 98.2°F o menor cuando se seleccionan 106 sujetos al azar. Los investigadores de la Universidad de Maryland obtuvieron una media muestral como ésta, ante lo cual existen dos explicaciones posibles: o la media de la población es realmente de 98.6°F y su muestra representa un suceso aleatorio extremadamente infrecuente, o en realidad la media poblacional es menor que 98.6°F y su muestra es típica. Como la probabilidad es tan baja, parece más razonable concluir que la media poblacional es menor que 98.6°F . Éste es el tipo de razonamiento que se usa en las *pruebas de hipótesis*, que se estudiarán en el capítulo 8. Por ahora, debemos enfocarnos en el uso del teorema del límite central para calcular la probabilidad de 0.0001, pero debemos señalar que este teorema se usará posteriormente para explicar algunos conceptos muy importantes en estadística.

Corrección para una población finita

Al aplicar el teorema del límite central, el uso de $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ supone que la población tiene un número infinito de miembros. Cuando hacemos un muestreo con reemplazo (es decir, cada elemento seleccionado se reincorpora a la muestra antes de hacer la siguiente selección), la población es efectivamente infinita. Aunque muchas aplicaciones realistas implican un muestreo sin reemplazo, estas muestras sucesivas dependen de resultados previos. En la fabricación, los inspectores de

control de calidad suelen muestrear elementos de un lote finito de producción, sin reemplazarlos. Para una población finita como ésta tal vez necesitemos ajustar $\sigma_{\bar{x}}$. La siguiente es una regla práctica:

Cuando realice un muestreo sin reemplazo y el tamaño de muestra n sea mayor que el 5% de la población finita de tamaño N (es decir, $n > 0.05N$), ajuste la desviación estándar de medias muestrales $\sigma_{\bar{x}}$ multiplicándola por el factor de corrección de población finita:

$$\sqrt{\frac{N-n}{N-1}}$$

Con excepción de los ejercicios 22 y 23, los ejemplos y los ejercicios de esta sección suponen que el factor de corrección de población finita *no* se aplica, ya sea porque estamos tomando una muestra con reemplazo, porque la población es infinita, o porque el tamaño de la muestra no excede el 5% del tamaño de la población.

El teorema del límite central es muy importante porque nos permite usar los métodos básicos de la distribución normal en una amplia variedad de circunstancias. Por ejemplo, en el capítulo 7 aplicaremos el teorema cuando utilicemos datos muestrales para estimar medias de poblaciones. En el capítulo 8, por ejemplo, lo aplicaremos cuando usemos datos muestrales para probar aseveraciones acerca de medias poblacionales. Tales aplicaciones para estimar parámetros de población y probar aseveraciones constituyen usos sumamente importantes de la estadística, y el teorema del límite central los hace posibles.

6-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Error estándar de la media.** ¿Qué es el error estándar de la media?
2. **Muestra pequeña.** Si se seleccionan muestras de tamaño $n = 2$ de una población con media y desviación estándar conocidas, ¿qué requisitos se deben cumplir para suponer que la distribución de las medias muestrales es normal?
3. **Notación.** Se seleccionan al azar muestras grandes ($n > 30$) de una población con media μ y desviación estándar σ . ¿Qué notación se utiliza para la media de la población que consiste en todas las medias muestrales? ¿Qué notación se utiliza para la desviación estándar de la población que consiste en todas las medias muestrales?
4. **Muestra de conveniencia.** Un estudiante de estadística esperó hasta el último minuto para realizar un proyecto, y ahora sólo tiene tiempo suficiente para reunir las estaturas de sus amigas y parientes mujeres. Luego, calcula la media de la estatura de las mujeres de su muestra. Suponiendo que las mujeres tienen estaturas que se distribuyen normalmente, con una media de 63.6 in y una desviación estándar de 2.5 in, ¿el estudiante puede utilizar el teorema del límite central para analizar la estatura media de su muestra?

Uso del teorema del límite central. En los ejercicios 5 a 8, suponga que las estaturas de mujeres se distribuyen de manera normal, con una media dada por $\mu = 63.6$ in y una desviación estándar $\sigma = 2.5$ in (según datos de la National Health Survey).

5. a. Si se selecciona a una mujer al azar, calcule la probabilidad de que mida menos de 64 in.
- b. Si se seleccionan 36 mujeres al azar, calcule la probabilidad de que tengan una estatura media menor que 64 in.

- 6.** a. Si se selecciona a una mujer al azar, calcule la probabilidad de que su estatura sea mayor que 63 in.
b. Si se seleccionan 100 mujeres al azar, calcule la probabilidad de que tengan una estatura media mayor que 63 in.
- 7.** a. Si se selecciona a una mujer al azar, calcule la probabilidad de que mida entre 63.5 y 64.5 in.
b. Si se seleccionan 9 mujeres al azar, calcule la probabilidad de que tengan una estatura media entre 63.5 y 64.5 in.
c. ¿Por qué se puede usar el teorema del límite central en el inciso b), aunque el tamaño de la muestra no sea mayor que 30?
- 8.** a. Si se selecciona a una mujer al azar, calcule la probabilidad de que mida entre 60 y 65 in.
b. Si se seleccionan 16 mujeres al azar, calcule la probabilidad de que tengan una estatura media entre 60 y 65 in.
c. ¿Por qué se puede usar el teorema del límite central en el inciso b), aunque el tamaño de la muestra no sea mayor que 30?
- 9. Seguridad de teleférico.** Un teleférico en Vail, Colorado, lleva a los esquiadores a la cima de la montaña. En él se ve una placa que indica que su capacidad máxima es de 12 personas o 2004 libras. Esa capacidad se excedería si 12 personas tienen pesos con una media mayor que $2004/12 = 167$ lb. Puesto que los hombres suelen pesar más que las mujeres, el “peor de los casos” implicaría a 12 pasajeros hombres. Los pesos de los hombres se distribuyen normalmente, con una media de 172 lb y una desviación estándar de 29 lb (según datos de la National Health Survey).
a. Calcule la probabilidad de que, al seleccionar al azar a un hombre, su peso sea mayor que 167 lb.
b. Calcule la probabilidad de que 12 hombres seleccionados al azar tengan una media mayor que 167 lb (de manera que su peso total sea mayor que la máxima capacidad del teleférico de 2004 lb).
c. Al parecer, ¿el teleférico tiene el límite correcto de peso? ¿Por qué?
- 10. Autobuses de casino.** El nuevo casino Lucky Lady desea incrementar sus ingresos al ofrecer autobuses que puedan transportar a jugadores de otras ciudades. La investigación indica que estos jugadores tienden a ser mayores, que tienden a jugar únicamente en las máquinas tragamonedas y que registran pérdidas con una media de \$182 y una desviación estándar de \$105. Los autobuses transportan a 35 jugadores por viaje. El casino entrega a cada pasajero del autobús vales con un valor de \$50, los cuales pueden canjearse por dinero en efectivo, de manera que el casino necesita recuperar ese costo para obtener una ganancia. Calcule la probabilidad de que, si un autobús se llena con 35 pasajeros, la cantidad media perdida por un pasajero sea mayor que \$50. Con base en el resultado, ¿el casino apuesta al proporcionar este tipo de autobuses?
- 11. Cantidad de Coca-Cola.** Suponga que latas de Coca-Cola se llenan de tal manera que las cantidades reales tienen una media de 12.00 oz y una desviación estándar de 0.11 oz.
a. Calcule la probabilidad de que una muestra de 36 latas tenga una cantidad media de al menos 12.19 oz, como en el conjunto de datos 12 del apéndice B.
b. Con base en el resultado del inciso a), ¿será razonable creer que las latas en realidad contienen una media de 12.00 oz? Si la media no es de 12.00 oz, ¿se está engañando a los consumidores?
- 12. Preparación para la prueba SAT.** Las calificaciones de hombres en la parte verbal de la prueba SAT-I se distribuyen normalmente, con una media de 509 y una desviación estándar de 112 (según datos del College Board). A hombres seleccionados al

- azar se les imparte el curso Columbian Review, antes de tomar la prueba SAT. Suponga que el curso no tiene efecto alguno.
- Si se selecciona a uno de los hombres al azar, calcule la probabilidad de que su calificación sea de al menos 590.
 - Si se selecciona a 16 de los hombres al azar, calcule la probabilidad de que la calificación media sea de al menos 590.
 - En el cálculo de la probabilidad del inciso b), ¿por qué puede usarse el teorema del límite central aun cuando el tamaño muestral no excede 30?
 - Si la muestra aleatoria de 16 hombres produce una calificación media de 590, ¿habría una fuerte evidencia que apoye la afirmación de que el curso es realmente eficaz? ¿Por qué?
- 13. Diseño de luces estroboscópicas.** Se diseña una luz estroboscópica para aeronaves, de manera que los tiempos entre los destellos se distribuyen normalmente, con una media de 3.00 s y una desviación estándar de 0.40 s.
- Calcule la probabilidad de que un tiempo individual sea mayor que 4.00 s.
 - Calcule la probabilidad de que la media de 60 tiempos elegidos al azar sea mayor que 4.00 s.
 - Dado que la luz estroboscópica se diseña para que otros pilotos vean una aeronave, ¿qué resultado es más importante para evaluar la seguridad de la luz estroboscópica: el resultado del inciso a) o del inciso b)? ¿Por qué?
- 14. Diseño de cascos para motocicleta.** Los ingenieros deben tomar en cuenta la anchura de las cabezas de los hombres cuando diseñan cascos para motocicleta. Las anchuras de las cabezas de los hombres se distribuyen normalmente, con una media de 6.0 in y una desviación estándar de 1.0 in (según datos de una encuesta antropométrica de Gordon, Churchill, *et al.*).
- Si se selecciona a un hombre al azar, calcule la probabilidad de que el ancho de su cabeza sea menor que 6.2 in.
 - La compañía Safeguard Helmet planea un lote de producción inicial de 100 cascos. Calcule la probabilidad de que 100 hombres, seleccionados al azar, tengan una anchura media de cabeza menor que 6.2 in.
 - El gerente de producción observa los resultados del inciso b) y piensa que todos los cascos deben hacerse para hombres con anchuras de cabeza menores de 6.2 in, porque se ajustarían a casi todos los hombres. ¿Por qué es incorrecto este razonamiento?
- 15. Presión sanguínea.** La presión sanguínea sistólica (en mm de Hg) de mujeres entre 18 y 24 años se distribuye normalmente, con una media de 114.8 y una desviación estándar de 13.1 (según datos de la National Health Survey). La hipertensión suele definirse como una presión sistólica mayor que 140.
- Si se selecciona al azar a una mujer entre 18 y 24 años, calcule la probabilidad de que su presión sistólica sea mayor que 140.
 - Si se seleccionan al azar 4 mujeres del mismo rango de edad, calcule la probabilidad de que su presión sistólica media sea mayor que 140.
 - Dado que el inciso b) incluye un tamaño de muestra no mayor que 30, ¿por qué se puede utilizar el teorema del límite central?
 - Si un médico recibe un reporte que afirma que 4 mujeres tienen una presión sistólica media menor que 140, ¿puede concluir que ninguna de las mujeres es hipertensa (con una presión sanguínea mayor que 140)?
- 16. Sin agua caliente.** Al planear las necesidades de agua caliente, el gerente del Luxurion Hotel descubre que los huéspedes pasan una media de 11.4 minutos al día bajo la regadera (según datos de Opinion Research Corporation). Suponga que los tiempos bajo la regadera se distribuyen normalmente, con una desviación estándar de 2.6 minutos.
- Calcule el porcentaje de huéspedes que pasan más de 12 minutos bajo la regadera.

continúa

- b.** El hotel instaló un sistema que puede suministrar suficiente agua caliente, dado que la media del tiempo bajo la regadera para 84 huéspedes es menor que 12 minutos. Si en este momento el hotel cuenta con 84 huéspedes, calcule la probabilidad de que no haya agua caliente suficiente. Al parecer, ¿el sistema actual es efectivo?
- 17. Rediseño de asientos de expulsión.** Cuando se permitió que las mujeres se convirtieran en pilotos de aviones de combate, los ingenieros necesitaron rediseñar los asientos expulsores porque habían sido diseñados sólo para hombres. Los asientos ACES-II estaban diseñados para hombres que pesaran entre 140 y 211 lb. La población de mujeres tiene pesos distribuidos normalmente, con una media de 143 lb y una desviación estándar de 29 lb (según datos de la National Health Survey).
- Si se selecciona a una mujer al azar, calcule la probabilidad de que pese entre 140 y 211 lb.
 - Si se seleccionan 36 mujeres diferentes al azar, calcule la probabilidad de que su peso medio se ubique entre 140 y 211 lb.
 - Al rediseñar los asientos expulsores de aviones de combate para que se ajusten mejor a las mujeres, ¿qué probabilidad es más importante: el resultado del inciso *a*) o el resultado del inciso *b*)? ¿Por qué?
- 18. Etiquetas de paquetes de M&M.** Los dulces M&M sencillos tienen un peso medio de 0.8565 g y una desviación estándar de 0.0518 g (según el conjunto de datos 13 del apéndice B). Los dulces M&M utilizados en el conjunto de datos 13 provienen de un paquete que contenía 465 dulces y la etiqueta del paquete indicaba que su peso neto era de 396.9 g. (Si cada paquete contiene 465 dulces, el peso medio de los dulces debe exceder $396.9/465 = 0.8535$ g del contenido neto, para pesar al menos 396.9 g).
- Si se selecciona al azar un dulce M&M sencillo, calcule la probabilidad de que pese más de 0.8535 g.
 - Si se seleccionan al azar 465 dulces M&M sencillos, calcule la probabilidad de que su peso medio sea de al menos 0.8535 g.
 - Con estos resultados, ¿la compañía Mars está ofreciendo a los consumidores de M&M la cantidad indicada en la etiqueta?
- 19. Máquinas expendedoras.** En la actualidad, las monedas de 25 centavos tienen pesos que se distribuyen normalmente con una media de 5.670 g y una desviación estándar de 0.062 g. Una máquina expendedora se configura para aceptar únicamente las monedas que pesen entre 5.550 y 5.790 g.
- Si se insertan 280 monedas diferentes de 25 centavos en la máquina expendedora, ¿cuál es el número esperado de monedas rechazadas?
 - Si se insertan 280 monedas diferentes de 25 centavos en la máquina expendedora, ¿cuál es la probabilidad de que la media se ubique entre los límites de 5.550 y 5.790 g?
 - Si usted es el propietario de la máquina expendedora, ¿qué resultado le interesa más: el resultado del inciso *a*) o el resultado del inciso *b*)? ¿Por qué?
- 20. Estándares de seguridad de aeronaves.** Bajo las antiguas reglas de la Federal Aviation Administration, las aerolíneas tenían que estimar el peso de un pasajero en 185 libras. (Esa cantidad es para un adulto que viaja en invierno e incluye 20 lb de equipaje de mano). Las reglas actuales requieren un estimado de 195 lb. Los hombres tienen pesos que se distribuyen normalmente, con una media de 172 lb y una desviación estándar de 29 lb.
- Si se selecciona al azar a un hombre adulto y se supone que lleva un equipaje de mano de 20 lb, calcule la probabilidad de que su peso total sea mayor que 195 lb.
 - Si un avión Boeing 767-300 transporta a 213 pasajeros varones adultos y se supone que cada uno lleva un equipaje de mano de 20 lb, calcule la probabilidad de que el peso medio de los pasajeros (incluyendo el equipaje de mano) sea mayor que 195 lb. ¿Un piloto debe preocuparse por exceder este límite de peso?

6-5 MÁS ALLÁ DE LO BÁSICO

- 21. Diseño de asientos.** Usted necesita construir una banca en la que se sentarán 18 jugadores universitarios de fútbol americano y debe determinar primero la longitud de la banca. Los hombres tienen anchuras de cadera que se distribuyen normalmente, con una media de 14.4 in y una desviación estándar de 1.0 in.
- ¿Cuál será la longitud mínima de la banca si usted busca una probabilidad de 0.975 de que se ajuste a las anchuras de cadera combinadas de 18 hombres seleccionados al azar?
 - ¿Por qué sería incorrecto utilizar realmente el resultado del inciso *a*) como longitud de la banca?
- 22. Corrección para una población finita.** El club Boston Women necesita un elevador limitado a 8 pasajeros. El club tiene 120 miembros mujeres con pesos que se aproximan a una distribución normal, con una media de 143 lb y una desviación estándar de 29 lb. (*Sugerencia:* Véase la explicación del factor de corrección para una población finita).
- Si se seleccionan al azar 8 miembros diferentes, calcule la probabilidad de que su peso total no exceda la capacidad máxima de 1300 lb.
 - Si buscamos una probabilidad de 0.99 de que el elevador no se sobrecargue siempre que se seleccione al azar a 8 miembros como pasajeros, ¿cuál debe ser el peso máximo permitido?
- 23. Parámetros de población** Una *población* consiste en los valores: 2, 3, 6, 8, 11, 18.
- Calcule μ y σ .
 - Liste todas las muestras de tamaño $n = 2$ que pueden obtenerse con reemplazo.
 - Calcule la población de todos los valores de \bar{x} al obtener la media de cada muestra del inciso *b*).
 - Calcule la media $\mu_{\bar{x}}$ y la desviación estándar $\sigma_{\bar{x}}$ para la población de medias muestrales obtenidas en el inciso *c*).
 - Verifique que

$$\mu_{\bar{x}} = \mu \quad y \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

6-6 La distribución normal como aproximación de la distribución binomial

Concepto clave En esta sección se presenta un método para utilizar una distribución normal como aproximación de una distribución de probabilidad binomial. Si se satisfacen las condiciones $nq \geq 5$ entonces las probabilidades de una distribución de probabilidad binomial se pueden aproximar bastante bien utilizando una distribución normal con media $\mu = np$ y desviación estándar $\sigma = \sqrt{npq}$. Puesto que una distribución de probabilidad binomial generalmente usa sólo números enteros para la variable aleatoria x , mientras que la aproximación normal es continua, debemos utilizar una “corrección por continuidad”, con un número entero x representado por el intervalo de $x - 0.5$ a $x + 0.5$. *Nota importante:* En vez de utilizar una distribución normal como aproximación de una distribución de probabilidad binomial, la mayoría de las aplicaciones prácticas de la distribución binomial se pueden manejar con un programa de cómputo o una calculadora, pero esta sección expone el importante principio de que una distribución binomial se puede aproximar por medio de una distribución normal, y este principio se utilizará en capítulos posteriores.

Considere la carga de un Boeing 767-300 de American Airlines que lleva 213 pasajeros. Al analizar la carga que puede transportar con seguridad, debemos con-

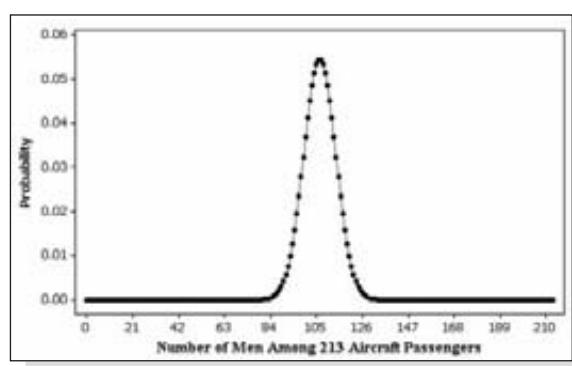


Voltaire vence a la lotería

En 1729, el filósofo Voltaire se hizo rico al diseñar un esquema para vencer a la lotería de París. El gobierno organizó una lotería para reembolsar bonos municipales que habían perdido cierto valor. La ciudad aportó grandes cantidades de dinero, con el efecto neto de que el valor total de los premios fuera mayor que el costo de todos los boletos. Voltaire organizó un grupo y compró todos los boletos de la lotería mensual y ganó durante más de un año. Por otro lado, un participante de la lotería del estado de Nueva York trató de ganar una parte de un premio excepcionalmente grande, que había crecido gracias a la falta de ganadores previos. Él quería extender un cheque por \$6,135,756, que cubriera todas las combinaciones, pero el estado se rehusó con el argumento de que esto cambiaría la naturaleza de la lotería.

siderar el peso de los pasajeros. Sabemos que un hombre típico pesa alrededor de 30 libras más que una mujer típica, de manera que el número de pasajeros varones es un tema importante. Podemos utilizar la distribución de probabilidad binomial con $n = 213$, $p = 0.5$ y $q = 0.5$ (suponiendo que los hombres y las mujeres son igualmente probables). Observe que la pantalla de Minitab que aparece abajo incluye una gráfica de la probabilidad para cada número de pasajeros varones, desde 0 hasta 213, y observe que la gráfica tiene la apariencia de una distribución normal aun cuando los puntos graficados provienen de una distribución binomial. Esta gráfica sugiere que podemos utilizar una distribución normal para aproximar la distribución binomial.

Minitab



La distribución normal como aproximación de la distribución binomial

Si una distribución de probabilidad binomial satisface los requisitos $np \geq 5$ y $nq \geq 5$, entonces la distribución de probabilidad binomial puede aproximarse con una distribución normal con media $\mu = np$ y desviación estándar $\sigma = \sqrt{npq}$, así como con el número entero discreto x ajustado con una *corrección por continuidad*, de manera que x está representada por el intervalo de $x - 0.5$ a $x + 0.5$.

Cuando se utiliza una distribución normal como aproximación de una distribución binomial se sigue el siguiente procedimiento:

Procedimiento para el uso de una distribución normal como aproximación de una distribución binomial

1. Establezca que la distribución normal es una aproximación adecuada de la distribución binomial, verificando que $np \geq 5$ y $nq \geq 5$. (Si no se satisfacen ambas condiciones, entonces debe utilizar un programa de cómputo, una calculadora, la tabla A-1 o la fórmula de probabilidad binomial).
2. Obtenga los valores de los parámetros μ y σ calculando $\mu = np$ y $\sigma = \sqrt{npq}$.
3. Identifique el valor discreto x (el número de éxitos). Cambie el valor discreto x reemplazándolo con el *intervalo* de $x - 0.5$ a $x + 0.5$ (para mayor explicación,

véase el apartado “Correcciones por continuidad”, más adelante en esta sección). Dibuje una curva normal e indique los valores de μ , σ y $x - 0.5$ o $x + 0.5$, según sea apropiado.

4. Modifique x reemplazándola por $x - 0.5$ o $x + 0.5$, según sea apropiado.
5. Utilice $x - 0.5$ o $x + 0.5$ (según sea apropiado) en vez de x , calcule el área correspondiente a la probabilidad deseada encontrando primero la puntuación z : $z = (x - \mu)/\sigma$. Ahora use esa puntuación z para encontrar el área a la izquierda del valor ajustado de x . Ahora esa área puede emplearse para identificar el área correspondiente a la probabilidad deseada.

Ilustraremos este procedimiento de aproximación normal con el siguiente ejemplo.

EJEMPLO Carga de pasajeros en un Boeing 767-300 Un avión Boeing 767-300 de American Airlines tiene 213 asientos. Cuando se llena con pasajeros, equipaje, carga y combustible, el piloto debe verificar que el peso neto no rebasa el límite máximo permitido, y el peso debe distribuirse adecuadamente para que el equilibrio de la aeronave permanezca dentro de los límites de seguridad aceptables. En vez de pesar a cada pasajero, sus pesos se estiman según las reglas de la Federal Aviation Administration. En realidad, sabemos que los hombres tienen un peso medio de 172 lb y que las mujeres tienen un peso medio de 143 lb, de manera que un número desproporcionadamente mayor de hombres podría provocar una situación insegura de sobrepeso. Suponga que, si hay por lo menos 122 hombres en una lista de 213 pasajeros, la carga debe ajustarse de alguna manera. Suponiendo que los pasajeros se registran al azar, que los hombres y las mujeres son igualmente probables, y que la aeronave está llena de adultos, calcule la probabilidad de que en un Boeing 767-300 con 213 pasajeros haya al menos 122 hombres.

SOLUCIÓN El problema implica una distribución binomial con un número fijo de ensayos ($n = 213$) que se supone son independientes, dos categorías (hombre, mujer) de resultados para cada ensayo, y la probabilidad de un hombre ($p = 0.5$) que se supone permanece constante de un ensayo a otro. Los cálculos con la fórmula de probabilidad binomial no son prácticos porque tendríamos que aplicarla 92 veces (una para cada valor de x desde 122 hasta 213, inclusive). En vez de ello, utilizamos el método de los cinco pasos para aproximar la distribución binomial con la distribución normal.

Paso 1: Verificación requerida: Primero debemos verificar que es razonable aproximar la distribución binomial con la distribución normal, porque $np \geq 5$ y $nq \geq 5$. Con $n = 213$, $p = 0.5$ y $q = 1 - p = 0.5$, verificamos las condiciones requeridas como sigue:

$$np = 213 \cdot 0.5 = 106.5 \quad (\text{Por lo tanto } np \geq 5.)$$

$$nq = 213 \cdot 0.5 = 106.5 \quad (\text{Por lo tanto } nq \geq 5.)$$

Paso 2: Ahora procedemos a calcular los valores de μ y σ , necesarios para la distribución normal. Obtenemos lo siguiente:

$$\mu = np = 213 \cdot 0.5 = 106.5$$

$$\sigma = \sqrt{npq} = \sqrt{213 \cdot 0.5 \cdot 0.5} = 7.2972598$$



¿Es seguro el paracaidismo?

De las más de 100,000 personas que realizan cerca de 2.25 millones de saltos en paracaídas, aproximadamente 30 mueren cada año. En comparación, un año típico incluye alrededor de 200 muertes en el buceo, 7000 ahogamientos, 900 muertes en bicicletas, 800 muertes por relámpagos y 1150 muertes por picaduras de abeja. Desde luego, estas cifras no significan necesariamente que el paracaidismo sea más seguro que andar en bicicleta o que la natación. En una comparación justa deben incluirse las tasas de mortalidad, no sólo el número total de fallecimientos. El autor, con gran osadía, realizó dos saltos en paracaídas, pero desistió después de no caer dentro de la amplia zona de aterrizaje, en ambas ocasiones. Él también ha volado un ala delta, un globo aerostático y en el dirigible Goodyear.

continúa

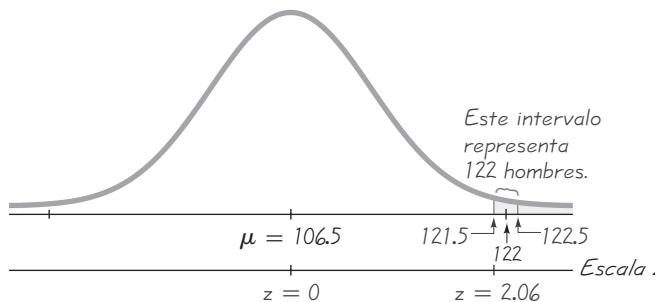


Figura 6-21 Búsqueda de la probabilidad de “al menos” 122 hombres entre 213 pasajeros

- Paso 3: Buscamos la probabilidad de “al menos 122 hombres” y el valor discreto de 122 se ajusta utilizando la corrección por continuidad de la siguiente manera: represente $x = 122$ por medio de la banda vertical limitada por 121.5 y 122.5.
- Paso 4: Puesto que buscamos la probabilidad de *al menos* 122 hombres, queremos el área que representa el número entero discreto de 122 (la región limitada por 121.5 y 122.5), así como también el área a la derecha, como se indica en la figura 6-21.
- Paso 5: Ahora podemos proceder a la búsqueda del área sombreada de la figura 6-21, utilizando los mismos métodos que se emplearon en la sección 6-3. Para usar la tabla A-2 de la distribución normal estándar, primero debemos transformar 121.5 a una puntuación z , después usar la tabla para encontrar el área a la izquierda de 121.5, que posteriormente se resta de 1. La puntuación z se obtiene como sigue:

$$z = \frac{x - \mu}{\sigma} = \frac{121.5 - 106.5}{7.2972598} = 2.06$$

Al emplear la tabla A-2, encontramos que $z = 2.06$ corresponde a una área de 0.9803, de manera que la región sombreada es $1 - 0.9803 = 0.0197$. Si se utiliza una calculadora TI-83/84 Plus o un programa de cómputo, se obtiene el resultado más exacto de 0.0199.

INTERPRETACIÓN Existe una probabilidad de 0.0197 de obtener al menos 122 hombres entre 213 pasajeros. Como esa probabilidad es demasiado baja, sabemos que una lista de 200 pasajeros pocas veces incluirá al menos 122 hombres, por lo que no será necesario ajustar la carga del avión con mucha frecuencia.

Correcciones por continuidad

El procedimiento que implica el uso de la distribución normal como aproximación de la distribución binomial incluye un ajuste en el que cambiamos un número entero discreto por un intervalo que está 0.5 por abajo y 0.5 por arriba del número discreto. Este paso en particular, denominado *corrección por continuidad*, suele ser difícil de comprender, por lo que ahora lo explicaremos con mayor detalle.

Definición

Cuando empleamos la distribución normal (que es una distribución de probabilidad *continua*) como una aproximación de la distribución binomial (que es *discreta*), se realiza una **corrección por continuidad** a un número entero discreto x en la distribución binomial, representando el valor único x en el *intervalo* de $x - 0.5$ a $x + 0.5$ (es decir, sumando y restando 0.5).

Las siguientes sugerencias prácticas le ayudarán a utilizar las correcciones por continuidad apropiadamente.

Procedimiento para correcciones por continuidad

1. Cuando use la distribución normal como aproximación de la distribución binomial, *siempre* aplique la corrección por continuidad. (Es necesario porque estamos utilizando la distribución normal *continua* para aproximar la distribución binomial *discreta*).
2. Para emplear la corrección por continuidad, primero identifique el número entero discreto x relevante al problema de probabilidad binomial. Por ejemplo, si usted está intentando calcular la probabilidad de obtener al menos 122 hombres entre 213 personas seleccionadas al azar, el número entero discreto relevante sería $x = 122$. Primero enfoque su atención en el valor x e ignore temporalmente si busca al menos x , más que x , menos que x o alguna otra condición.
3. Dibuje una distribución normal centrada alrededor de μ , después dibuje una *franja vertical* alrededor de x . Marque el lado izquierdo de la franja con el número igual a $x + 0.5$, y marque el lado derecho con el número igual a $x - 0.5$. Por ejemplo, para $x = 122$, dibuje una franja desde 121.5 hasta 122.5. *Considere el área completa de la franja para representar la probabilidad del número entero discreto x .*
4. Ahora determine si el valor de x debe incluirse en la probabilidad que busca. (Por ejemplo, “al menos x ” incluye a x , pero “más que x ” no la incluye). Despues, determine si busca la probabilidad de al menos x , a lo sumo x , más que x , menos que x o exactamente x . Sombree el área a la derecha o izquierda de la franja, según sea apropiado; también sombree el interior de la franja si y sólo si x debe incluirse. Esta región total sombreada corresponde a la probabilidad buscada.

Para ver cómo resulta este procedimiento en las correcciones por continuidad, observe los casos comunes ilustrados en la figura 6-22. Esos casos corresponden a las aseveraciones de la siguiente lista.

Aseveración	Área
Al menos 122 (incluye 122 y números mayores)	A la <i>derecha</i> de 121.5
Más que 122 (no incluye 122)	A la <i>derecha</i> de 122.5
A lo sumo 122 (incluye 122 y números menores)	A la <i>izquierda</i> de 122.5
Menos que 122 (no incluye 122)	A la <i>izquierda</i> de 121.5
Exactamente 122	Entre 121.5 y 122.5

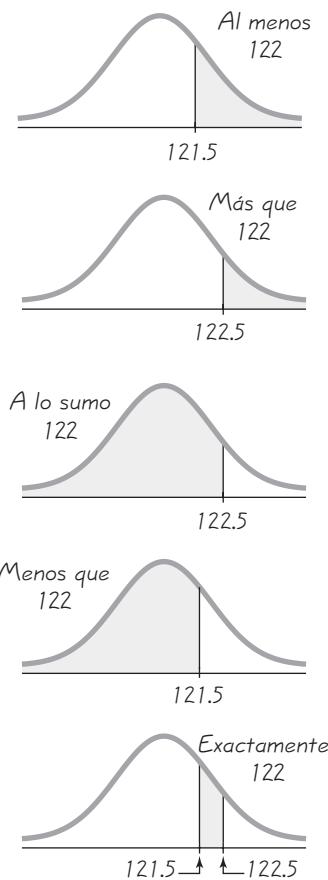


Figura 6-22
Uso de correcciones
por continuidad

EJEMPLO Uso de Internet Una encuesta reciente reveló que, de 2013 adultos elegidos al azar, 1358 (o el 67.5%) afirmaron ser usuarios de Internet (según datos del Pew Research Center). Si la proporción de todos los adultos que utilizan Internet es en realidad de $2/3$, calcule la probabilidad de que una muestra aleatoria de 2013 adultos produzca exactamente 1358 usuarios de Internet.

SOLUCIÓN Tenemos $n = 2013$ sujetos de encuesta independientes y $x = 1358$ de ellos son usuarios de Internet; suponemos que la proporción de la población es $p = 2/3$, de lo que se deduce que $q = 1/3$. Utilizaremos una distribución normal para aproximar la distribución binomial.

Paso 1: Primero verificamos los requisitos para determinar si es posible la aproximación normal:

$$\begin{aligned} np &= 2013 \cdot 2/3 = 1342 \quad (\text{Por lo tanto } np \geq 5.) \\ nq &= 2013 \cdot 1/3 = 671 \quad (\text{Por lo tanto } nq \geq 5.) \end{aligned}$$

Paso 2: Ahora procedemos a calcular los valores de μ y σ , necesarios para la distribución normal. Obtenemos lo siguiente:

$$\begin{aligned} \mu &= np = 2013 \cdot 2/3 = 1342 \\ \sigma &= \sqrt{npq} = \sqrt{2013 \cdot (2/3) \cdot (1/3)} = 21.150256 \end{aligned}$$

Paso 3: Dibujamos la curva normal de la figura 6-23. La región sombreada de la figura representa la probabilidad de obtener exactamente 1358 usuarios de Internet. Si utilizamos la corrección por continuidad, representamos $x = 1358$ por medio de la región ubicada entre 1357.5 y 1358.5.

Paso 4: He aquí el método empleado para calcular la región sombreada de la figura 6-23. Primero calcule el área total a la izquierda de 1358.5, después obtenga el área total a la izquierda de 1357.5, luego calcule la *diferencia* entre ambas áreas. Comencemos con el área total a la izquierda de 1358.5. Si deseamos utilizar la tabla A-2, primero debemos obtener la puntuación z que corresponde a 1358.5. Obtenemos

$$z = \frac{1358.5 - 1342}{21.150256} = 0.78$$

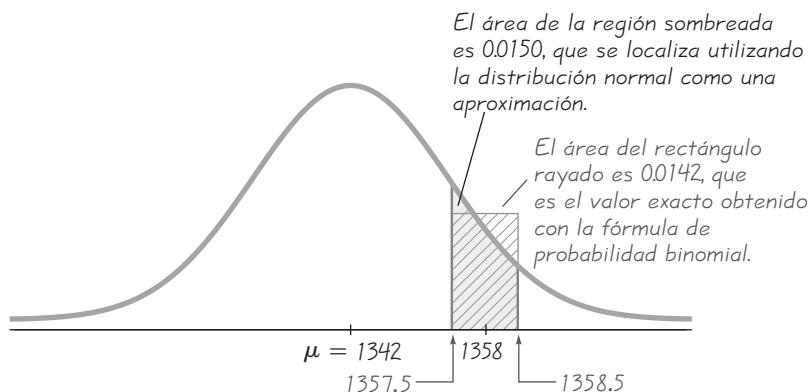


Figura 6-23 Uso de la corrección por continuidad

Usamos la tabla A-2 para encontrar que $z = 0.78$ corresponde a una probabilidad de 0.7823, que es el área total a la izquierda de 1358.5. Ahora, procedemos a obtener el área a la izquierda de 1357.5, calculando primero la puntuación z correspondiente a 1357.5:

$$z = \frac{1357.5 - 1342}{21.150256} = 0.73$$

Si utilizamos la tabla A-2, encontramos que $z = 0.73$ corresponde a una probabilidad de 0.7673, la cual es el área total a la izquierda de 1357.5. El área sombreada es $0.7823 - 0.7673 = 0.0150$. (Con un recurso tecnológico obtenemos 0.0142).

INTERPRETACIÓN Si suponemos que $2/3$ de todos los adultos utilizan Internet, la probabilidad de obtener exactamente 1358 usuarios de Internet entre 2013 personas elegidas al azar es de 0.0150. Esta probabilidad nos indica que si la proporción de usuarios de Internet en la población es de $2/3$, entonces es sumamente improbable que obtengamos exactamente 1358 usuarios de Internet al encuestar a 2013 personas. En realidad, cuando se encuesta a 2013 individuos, la probabilidad de *cualquier* número específico de usuarios de Internet es muy pequeña.

Si resolvemos el ejemplo anterior por medio de STATDISK, Minitab o una calculadora, obtenemos un resultado de 0.0142, pero el método de aproximación normal dio por resultado un valor de 0.0150. La discrepancia de 0.0008 proviene de dos factores: **1.** el uso de la distribución normal da como resultado un valor *aproximado* que corresponde al área de la región sombreada en la figura 6-23, mientras que el área correcta exacta es un rectángulo centrado por arriba de 1358 (la figura 6-23 ilustra esta discrepancia); **2.** el uso de la tabla A-2 nos obligó a calcular uno de un número limitado de valores basados en una puntuación z redondeada. El área del rectángulo es 0.0142, pero el área de la región sombreada aproximada es 0.0150.

Interpretación de los resultados

En realidad, cuando utilizamos una distribución normal como aproximación de la distribución binomial, nuestra meta no es simplemente calcular un número de probabilidad. A menudo necesitamos hacer algún *juicio* con base en el valor de probabilidad. Por ejemplo, suponga que el reportero de un periódico ve los datos muestrales del ejemplo anterior y, después de observar que el 67.5% de los adultos encuestados utilizaban Internet, escribe el encabezado “Más de $2/3$ de los adultos utilizan Internet”. Los datos muestrales no justifican ese encabezado por la siguiente razón: si la verdadera proporción poblacional es *igual a* $2/3$ (en vez de ser mayor que $2/3$), existe una elevada probabilidad (0.2327) de obtener al menos 1358 usuarios de Internet entre los 2013 adultos encuestados. (El área ubicada a la izquierda de 1357.5 es 0.7673, de manera que la probabilidad de obtener al menos 1358 usuarios de Internet es $1 - 0.7673 = 0.2327$). Es decir, con una proporción poblacional igual a $2/3$, el resultado de 1358 usuarios de Internet *no es excepcionalmente elevado*. Este tipo de conclusiones tal vez parezcan un poco confusas en este momento, pero en los siguientes capítulos se presentarán métodos sistemáticos que las harán más fáciles. Por ahora, debemos comprender que las bajas probabilidades corresponden a sucesos con pocas posibilidades, mientras que las altas probabilidades corresponden a sucesos posibles. El valor de probabilidad de 0.05 suele utilizarse

como punto de corte para distinguir entre sucesos improbables y sucesos probables. El siguiente criterio (de la sección 5-2) describe la aplicación de las probabilidades para distinguir resultados que pueden ocurrir fácilmente por azar de aquellos que son muy poco comunes.

Uso de las probabilidades para determinar cuando los resultados son poco comunes

- **Extremadamente alto:** x éxitos en n ensayos es un número *extremadamente alto* de éxitos si $P(x \text{ o más})$ es muy pequeña (como 0.05 o menos).
- **Extremadamente bajo:** x éxitos en n ensayos es un número *extremadamente bajo* de éxitos si $P(x \text{ o menos})$ es muy pequeña (como 0.05 o menos).

El papel de la aproximación normal

En realidad, casi todas las aplicaciones prácticas de la distribución de probabilidad binomial ahora se pueden trabajar bien con un programa de cómputo o una calculadora TI-83/84 Plus. En esta sección se presentan métodos para manejar casos en los que no se puede utilizar un programa de cómputo y, algo más importante, también ilustra el principio de que, en las circunstancias apropiadas, la distribución de probabilidad binomial puede aproximarse por medio de una distribución normal. Los capítulos posteriores incluyen procedimientos basados en el uso de una distribución normal como aproximación de una distribución binomial, de manera que esta sección establece las bases de esos importantes procedimientos.

6-6 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Distribución de proporciones muestrales.** Considere un estudio en el que obtenemos registros de los siguientes 50 bebés que nacen, luego calcule la proporción de niñas en esta muestra. Suponga que este estudio se repite muchas veces y que se utilizan las proporciones muestrales para construir un histograma. ¿Cuál sería la forma del histograma?
2. **Corrección por continuidad.** La prueba Wechsler se utiliza para medir puntuaciones de CI y está diseñada de tal manera que la media es 100 y la desviación estándar es 16. Se sabe que las puntuaciones de CI tienen una distribución normal. Suponga que deseamos calcular la probabilidad de que una persona elegida al azar tenga un CI igual a 107. ¿Cuál sería la corrección por continuidad y cómo se aplicaría para calcular esa probabilidad?
3. **Distribución de proporciones muestrales.** La planta Newport Bottling fabrica botellas de bebidas de cola que se empacan en grupos de seis. La probabilidad de una botella defectuosa es de 0.001. ¿Podemos aproximar la distribución de defectos en los empaques de seis como una distribución normal? ¿Por qué?
4. **Interpretación de probabilidad binomial.** En la prueba de un método de selección del género, 80 parejas reciben un tratamiento diseñado para incrementar la probabilidad de que un bebé sea niña. De 80 bebés nacidos, 47 fueron niñas. Si el método de selección del género no tiene efecto, la probabilidad de obtener exactamente 47 niñas es de 0.0264, y la probabilidad de obtener 47 niñas o más es de 0.0728. ¿Cuál probabilidad se debe usar para evaluar la eficacia del método de selección del género? ¿Parece que el método es efectivo?

Aplicación de la corrección por continuidad. En los ejercicios 5 a 12 los valores dados son discretos. Utilice la corrección por continuidad y describa la región de la distribución normal que corresponde a la probabilidad indicada. Por ejemplo, la probabilidad de “más de 20 artículos defectuosos” corresponde al área de la curva normal descrita en esta respuesta: “el área a la derecha de 20.5”.

5. La probabilidad de más de 15 personas en prisión por quitar las etiquetas de advertencia de almohadas
6. La probabilidad de al menos 12 adultos varones en un elevador en el edificio Empire State
7. La probabilidad de menos de 12 bebés llorando en un vuelo de American Airlines
8. La probabilidad de que el número de máquinas expendedoras que funcionan en Estados Unidos sea exactamente 27
9. La probabilidad de no más de 4 estudiantes ausentes en una clase de estadística
10. La probabilidad de que el número de procedimientos estadísticos incorrectos en Excel sea entre 15 y 20, inclusive
11. La probabilidad de que el número de políticos verdaderamente honestos esté entre 8 y 10, inclusive
12. La probabilidad de que exactamente 3 empleados hayan sido despedidos por no usar Internet de manera apropiada

Uso de la aproximación normal. En los ejercicios 13 a 16, haga lo siguiente. a) Calcule la probabilidad binomial indicada por medio de la tabla A-1 del apéndice A. b) Si $np \geq 5$ y $nq \geq 5$, también estime la probabilidad indicada utilizando la distribución normal como aproximación de la distribución binomial; si $np < 5$ o $nq < 5$, entonces estableza que la aproximación normal no es adecuada.

13. Con $n = 12$ y $p = 0.6$, calcule $P(7)$.
14. Con $n = 14$ y $p = 0.4$, calcule $P(6)$.
15. Con $n = 11$ y $p = 0.5$, calcule $P(\text{al menos } 4)$.
16. Con $n = 13$ y $p = 0.3$, calcule $P(\text{menor que } 5)$.
17. **Probabilidad de más de 36 niñas.** Suponiendo que los niños y las niñas son igualmente probables, estime la probabilidad de que resulten más de 36 niñas en 64 nacimientos. ¿Es infrecuente que resulten más de 36 niñas en 64 nacimientos?
18. **Probabilidad de al menos 42 niñas.** Suponiendo que los niños y las niñas son igualmente probables, estime la probabilidad de que resulten al menos 42 niñas en 64 nacimientos. ¿Es infrecuente que resulten al menos 42 niñas en 64 nacimientos?
19. **¿Los votantes mienten?** En una encuesta de 1002 individuos, 701 dijeron que habían votado en una elección presidencial reciente (según datos del ICR Research Group). Los registros de votos indican que el 61% de los votantes potenciales realmente votaron. Dado que el 61% de los votantes potenciales realmente votaron, calcule la probabilidad de que entre 1002 votantes potenciales seleccionados al azar, al menos 701 hayan votado realmente. ¿Qué sugiere el resultado?
20. **Publicidad televisiva.** El cobro por los anuncios en un programa de televisión se basa en el número de televidentes, el cual se mide por medio del índice de audiencia. El índice de audiencia es un porcentaje de la población de 110 millones de hogares con televisor. El programa de televisión *60 Minutes* de la CBS recientemente tuvo una audiencia del 7.8%, lo que indica que el 7.8% de los hogares estaban sintonizando ese programa. Un anunciante realiza una encuesta independiente de 100 hogares y descubre que sólo 4 de ellos estaban sintonizando *60 Minutes*. Suponiendo que la audiencia de 7.8 es correcta, calcule la probabilidad de encuestar de manera aleatoria 100 hogares y encontrar que 4 o menos están sintonizando *60 Minutes*. ¿El resultado sugiere que la audiencia de

7.8 es demasiado alta? ¿El anuncio tiene bases para reclamar un reembolso argumentando que el tamaño de la audiencia se exageró?

21. **Experimento de hibridación de Mendel.** Cuando Mendel realizó sus famosos experimentos de hibridación, utilizó plantas de guisantes con vainas verdes y vainas amarillas. Uno de los experimentos implicó una cruce de guisantes, de manera que se esperaba que el 25% (o 145) de los 580 vástagos de guisantes tuvieran vainas amarillas. En vez de obtener 145 plantas de guisantes con vainas amarillas, obtuvo 152. Suponiendo que el porcentaje del 25% de Mendel es correcto, estime la probabilidad de obtener al menos 152 plantas de guisantes con vainas amarillas entre los 580 vástagos de guisantes. ¿Existe una fuerte evidencia que sugiera que la probabilidad del 25% de Mendel es incorrecta?
22. **Fármaco que reduce el colesterol.** La probabilidad de que una persona que no recibe algún tratamiento tenga síntomas de gripe es de 0.019. En un ensayo clínico de Lipitor, un fármaco común utilizado para disminuir el colesterol, 863 pacientes recibieron un tratamiento con tabletas de atorvastatin de 10 mg, y 19 de estos pacientes experimentaron síntomas de gripe (según datos de Pfizer, Inc.). Suponiendo que estas tabletas no influyen en los síntomas de la gripe, estime la probabilidad de que al menos 19 de las 863 personas experimenten síntomas de gripe. ¿Qué sugieren estos resultados acerca de los síntomas de gripe como una reacción adversa al fármaco?
23. **Teléfonos celulares y cáncer cerebral.** En un estudio de 420,095 usuarios de teléfono celular en Dinamarca, se encontró que 135 desarrollaron cáncer cerebral o del sistema nervioso. Suponiendo que los teléfonos celulares no tienen efecto alguno, existe una probabilidad de 0.000340 de que una persona desarrolle cáncer cerebral o del sistema nervioso. Por lo tanto, esperaríamos aproximadamente 143 casos de este tipo de cáncer en un grupo de 420,095 personas seleccionadas al azar. Estime la probabilidad de 135 o menos casos de este tipo de cáncer en un grupo de 420,095 personas elegidas al azar. ¿Qué sugieren estos resultados acerca de los reportes de los medios de comunicación que afirman que los teléfonos celulares causan cáncer cerebral o del sistema nervioso?
24. **Sobreventa de lugares en los vuelos.** Air America está considerando la nueva política de registrar hasta 400 personas en un avión que sólo tiene cupo para 350. (Estudios anteriores han revelado que sólo el 85% de los pasajeros registrados llegan para tomar el vuelo). Estime la probabilidad de que, si Air America registra a 400 personas, no existan suficientes asientos disponibles. ¿Es esta probabilidad lo suficientemente baja para ser funcional, o deberá modificarse la política?
25. **Identificación de discriminación por género.** Después de haber sido rechazada para un empleo, Kim Kelly se entera que la Bellevue Advertising Company contrató únicamente a 21 mujeres entre sus 62 nuevos empleados. También se entera de que el grupo de solicitantes es muy grande, con igual número de hombres y mujeres calificados. Ayúdela a presentar una acusación por discriminación, estimando la probabilidad de obtener 21 mujeres o menos cuando se contrata a 62 personas, suponiendo que no existe discriminación por género. ¿La probabilidad resultante realmente apoya una acusación como ésta?
26. **Dulces M&M: ¿el 20% son anaranjados?** Según Mars (la compañía de dulces), el 20% de todos los dulces sencillos M&M son anaranjados. El conjunto de datos 13 del apéndice B indica que de 100 M&M elegidos, 25 son anaranjados. Suponiendo que es correcta la afirmación de que el 20% de los dulces M&M son anaranjados, estime la probabilidad de seleccionar al azar 100 dulces M&M y obtener 25 o más que sean anaranjados. Con base en el resultado, ¿será infrecuente obtener 25 o más dulces M&M anaranjados cuando se seleccionan 100 al azar?
27. **Grupo sanguíneo.** El 45% de nosotros tiene sangre del tipo O, según datos del Great New York Program. El Providence Memorial Hospital está realizando una campaña de donación de sangre, ya que su abastecimiento de sangre del tipo O es bajo, y necesita 177 donadores de este grupo sanguíneo. Si 400 voluntarios donan sangre, estime la probabilidad de que el número de personas con sangre del tipo O sea al menos de 177. ¿Es probable que el grupo de 400 voluntarios sea suficiente?
28. **Muestreo de aceptación.** Algunas compañías verifican la calidad a través del método del muestreo de aceptación, por medio del cual se rechaza el lote completo de artículos

si una muestra aleatoria de un tamaño particular incluye más de un número especificado de defectos. La Dayton Machine Company compra tornillos de máquina en lotes de 5000 y rechaza un lote si, cuando se obtiene una muestra de 50, al menos dos son defectuosos. Estime la probabilidad de rechazar un lote si el proveedor está fabricando los tornillos con una tasa de defectos del 10%. ¿Es posible que el plan de verificación identifique la tasa inaceptable de defectos?

29. **Encuesta sobre clonación.** Una encuesta de Gallup reciente incluyó a 1012 adultos seleccionados al azar, a quienes se preguntó si “la clonación humana debe o no permitirse”. Según los resultados, el 89% de los encuestados indicaron que no debe permitirse la clonación. Un reportero de noticias desea determinar si estos resultados de encuesta constituyen una fuerte evidencia de que la mayoría (más del 50%) de las personas se oponen a ese tipo de clonación. Suponiendo que el 50% de todas las personas se oponen, estime la probabilidad de obtener al menos un 89% de oposición en una encuesta de 1012 personas seleccionadas al azar. Con base en el resultado, ¿existe fuerte evidencia que apoya la afirmación de que la mayoría se opone a ese tipo de clonación?
30. **Sesgo en la selección de miembros de un jurado.** En el condado de Orange, el 12% de las personas que pueden ser convocadas para formar parte de un jurado son zurdas. De 250 personas elegidas como parte de un jurado, 25 (o el 10%) son zurdas. Calcule la probabilidad de obtener a lo sumo 25 personas zurdas, suponiendo que son elegidas con un proceso diseñado para producir una tasa del 12% de zurdos. ¿Podemos concluir que este proceso de selección de miembros de un jurado discrimina a los zurdos?
31. **Detección de fraude con tarjetas de crédito.** La compañía Dynamic Credit emite tarjetas de crédito y utiliza programas de cómputo para detectar fraudes. Después de indagar los hábitos de gastos de un consumidor, se descubre que cobros mayores de \$100 constituyen el 35.8% de las transacciones de crédito. De los 30 cargos realizados este mes, 18 implican montos que exceden los \$100. ¿Se tratará de un patrón de gasto infrecuente que debe ser verificado? Explique.
32. **Detección de fraude.** Cuando trabajaba para el fiscal del distrito de Brooklyn, el investigador Robert Burton analizó el primer dígito de los montos de cheques expedidos por empresas sospechosas de fraude. De un total de 784 cheques, 479 tenían montos que terminaban en 5, aunque se esperaba que el 7.9% los cheques emitidos en el transcurso normal de transacciones terminaran en 5. ¿Existe una fuerte evidencia que indique que los montos de los cheques son significativamente diferentes de los montos que se esperarían en condiciones normales? Explique.

6-6 MÁS ALLÁ DE LO BÁSICO

33. **Ganar en la ruleta.** Marc Taylor planea hacer 200 apuestas, de \$1 cada una, al número 7 en la ruleta. Un triunfo es rentable con probabilidades de 35:1 y, en cualquier giro, existe una probabilidad de 1/38 de que el 7 sea el número ganador. De las 200 apuestas, ¿cuál es el número mínimo de triunfos necesarios para que Marc obtenga una ganancia? Estime la probabilidad de que Marc obtenga una ganancia.
34. **Reemplazo de televisores.** Los tiempos de reemplazo de televisores se distribuyen normalmente, con una media de 8.2 años y una desviación estándar de 1.1 años (según datos de “Getting Fixed”, *Consumer Reports*). Estime la probabilidad de que, de 250 televisores seleccionados al azar, al menos 15 tengan tiempos de reemplazo mayores a 10.0 años.
35. **Hits en béisbol.** Suponga que un jugador de béisbol batea de *hit* .350, de manera que su probabilidad de hacer un *hit* es de 0.350. (Ignore las complicaciones causadas por las bases por bolas). También suponga que sus intentos de *hit* son independientes unos de otros.
- Calcule la probabilidad de al menos 1 *hit* en 4 intentos, en un partido.
 - Suponiendo que este bateador sube a batir 4 veces cada juego, estime la probabilidad de obtener un total de al menos 56 *hits* en 56 juegos.

- c. Suponiendo que este bateador tiene la oportunidad de batear 4 veces en cada juego, estime la probabilidad de al menos 1 *hit* en cada uno de 56 juegos consecutivos (el récord de Joe DiMaggio en 1941).
 - d. ¿Cuál es el promedio mínimo de bateo que se requeriría para que la probabilidad del inciso c) sea mayor que 0.1?
36. **Sobreventa de lugares en vuelos.** Vertigo Airlines trabaja únicamente con reservaciones anticipadas y registra una tasa del 7% de personas que no se presentan. ¿Cuántas reservaciones podrían aceptarse para un avión con cupo para 250 pasajeros, si existe al menos una probabilidad de 0.95 de que todos los individuos que reservaron y se presenten puedan tener su lugar en el vuelo?
37. **Necesidad de aproximación normal.** En esta sección se incluyó la siguiente afirmación: “En realidad, casi todas las aplicaciones prácticas de la distribución de probabilidad binomial ahora se pueden manejar bien con un programa de cómputo o una calculadora TI-83/84 Plus”. Utilizando un programa estadístico de cómputo específico o una calculadora TI-83/84 Plus, identifique un caso en el que la tecnología no sea útil y, por lo tanto, se requiera hacer una aproximación normal de una distribución binomial.

6-7 Determinación de normalidad

Concepto clave Los siguientes capítulos incluyen algunos métodos estadísticos muy importantes que requieren que los datos muestrales se seleccionen al azar a partir de una población con una distribución *normal*. Esta sección presenta criterios para determinar si se cumplen los requisitos de una distribución normal. Los criterios incluyen la inspección visual de un histograma para ver si tiene forma de campana, la identificación de valores extremos y la construcción de una nueva gráfica denominada *gráfica cuantilar normal*.



Definición

Una **gráfica cuantilar normal** (o **gráfica de probabilidad normal**) es una gráfica de puntos (x, y) donde cada valor x proviene del conjunto original de datos muestrales, y cada valor y es la puntuación z correspondiente, que es un valor cuantilar esperado de la distribución normal estándar. (Véase el paso 3 en el siguiente procedimiento para conocer detalles sobre el cálculo de estas puntuaciones z).

Procedimiento para determinar si los datos se distribuyen normalmente

1. *Histograma:* Construya un histograma. Rechace la normalidad si el histograma difiere mucho de la forma de campana.
2. *Valores extremos:* Identifique valores extremos. Rechace la normalidad si existe más de un valor extremo. (La presencia de un solo valor extremo podría ser un error o el resultado de la variación por el azar, pero sea cuidadoso porque incluso un solo valor extremo podría tener un efecto importante en los resultados).
3. *Gráfica cuantilar normal:* Si el histograma es básicamente simétrico y existe a lo sumo un valor extremo, construya una *gráfica cuantilar normal*. Los siguientes pasos describen un procedimiento relativamente sencillo para construir una gráfica cuantilar normal, pero los diferentes paquetes estadísticos emplean diversos métodos. (STATDISK y la calculadora TI-83/84 Plus utilizan el procedimiento que aquí se describe). Este procedimiento es bastante confuso, por lo que generalmente utilizamos un programa de cómputo o una calculadora para generar la gráfica; al final de esta sección se incluyen las instrucciones para el uso de STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus.

- a. Primero ordene los datos del menor al mayor.
- b. Con una muestra de tamaño n , cada valor representa una proporción de $1/n$ de la muestra. Utilizando el tamaño muestral n conocido, identifique las áreas de $1/2n$, $3/2n$, $5/2n$, $7/2n$, etcétera. Éstas son las áreas acumulativas a la izquierda de los valores muestrales correspondientes.
- c. Utilice la distribución normal estándar (la tabla A-2, un programa de cómputo o una calculadora) para calcular las puntuaciones z correspondientes a las áreas acumulativas de la izquierda obtenidas en el paso b). (Se trata de las puntuaciones z que se esperan de una muestra distribuida normalmente).
- d. Asocie los valores originales de los datos ordenados con sus puntuaciones z correspondientes, calculadas en el paso c), después grafique los puntos (x, y) , donde cada x es un valor muestral original y y es la puntuación z correspondiente.
- e. Examine la gráfica cuantílica normal con los siguientes criterios: si los puntos no se acercan a una línea recta o si exhiben algún patrón sistemático diferente al de una línea recta, entonces parece que los datos provienen de una población que *no* tiene una distribución normal. Si el patrón de puntos se acerca razonablemente a una línea recta, entonces los datos pueden provenir de una población con distribución normal. (Estos criterios se emplean con escasa rigidez en el caso de muestras pequeñas, pero deben utilizarse de manera más estricta con muestras grandes).

Los pasos 1 y 2 son directos, pero ilustramos la construcción de una gráfica cuantílica normal (paso 3) en el siguiente ejemplo.

EJEMPLO Estaturas de hombres El conjunto de datos 1 del apéndice B incluye las estaturas (en pulgadas) de hombres elegidos al azar. Consideremos solamente las primeras cinco estaturas de hombres: 70.8, 66.2, 71.7, 68.7, 67.6. Con sólo cinco valores, un histograma no sería muy útil para revelar la distribución de los datos. En vez de ello, construya una gráfica cuantílica normal con estos cinco valores y determine si parece que provienen de una población distribuida normalmente.

SOLUCIÓN Los siguientes pasos corresponden a los listados en el procedimiento anterior para la construcción de una gráfica cuantílica normal.

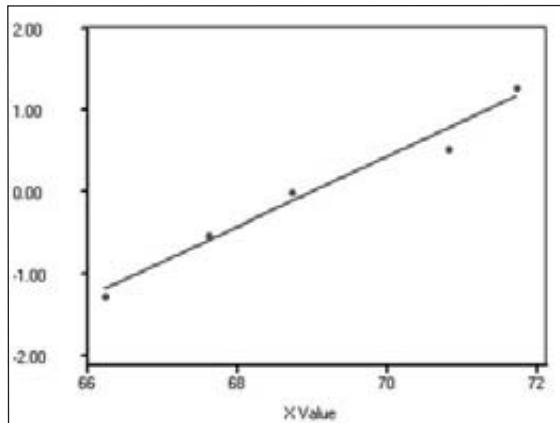
- a. Primero ordenamos los datos y obtenemos: 66.2, 67.6, 68.7, 70.8, 71.7.
- b. Con una muestra de tamaño $n = 5$, cada valor representa una proporción de $1/5$ de la muestra, por lo que procedemos a identificar las áreas acumulativas a la izquierda de los valores muestrales correspondientes. Estas áreas izquierdas acumulativas, que se expresan en general como $1/2n$, $3/2n$, $5/2n$, $7/2n$, etcétera, se convierten en las siguientes áreas específicas para este ejemplo, con $n = 5$: $1/10$, $3/10$, $5/10$, $7/10$ y $9/10$. Estas mismas áreas izquierdas acumulativas, expresadas en forma decimal, son 0.1 , 0.3 , 0.5 , 0.7 y 0.9 .
- c. Ahora buscamos en el cuerpo de la tabla A-2 las áreas izquierdas acumulativas de 0.1000 , 0.3000 , 0.5000 , 0.7000 y 0.9000 . Encontramos estas puntuaciones z correspondientes: -1.28 , -0.52 , 0 , 0.52 y 1.28 .
- d. Ahora apareamos las estaturas ordenadas originales con sus puntuaciones z correspondientes y obtenemos las siguientes coordenadas (x, y) , que están graficadas en la siguiente pantalla de STATDISK: $(66.2, -1.28)$, $(67.6, -0.52)$, $(68.7, 0)$, $(70.8, 0.52)$ y $(71.7, 1.28)$.



Los estados controlan las selecciones de lotería

Muchos estados tienen una lotería en la que los participantes seleccionan cuatro dígitos, como 1127 (el cumpleaños del autor). Si un jugador paga \$1 y selecciona la secuencia ganadora en el orden correcto, gana un premio de \$5000. Los estados observan las selecciones de números y, si una secuencia en particular se elige con demasiada frecuencia, se prohíbe a los participantes hacer más apuestas en ella. Las máquinas de lotería se controlan de tal manera que, una vez que una secuencia alcanza cierto nivel de ventas, la máquina no aceptará más esa secuencia. Esto evita que los estados paguen más de lo que reciben. Los críticos afirman que esta práctica es injusta. Según William Thompson, un experto en apuestas de la Universidad de Nevada en Las Vegas afirma que “ellos (los estados) quieren estar en el negocio del juego, pero no desean ser jugadores. Esto convierte al juego de números en una farsa”.

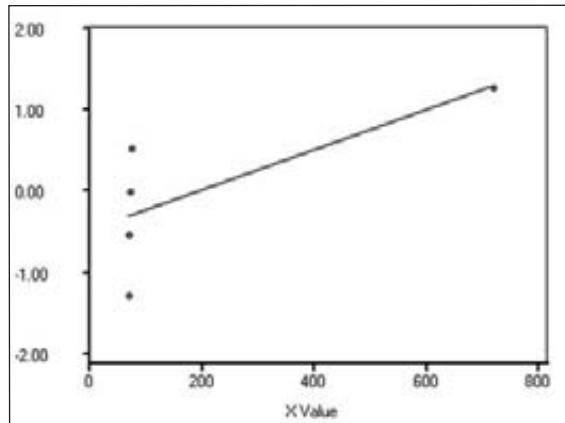
STATDISK



INTERPRETACIÓN Examinamos la gráfica cuantílica normal de la pantalla de STATDISK. Como los puntos parecen estar razonablemente cerca de una línea recta y no parece haber un patrón sistemático distinto al de una línea recta, concluimos que las cinco estaturas dadas parecen provenir de una población distribuida normalmente.

La siguiente pantalla de STATDISK presenta la gráfica cuantílica normal de los mismos datos del ejemplo anterior, con un cambio: el valor más grande de 71.7 se sustituyó por 717, el cual se convierte en un valor extremo. Observe cómo el cambio en ese valor afecta la gráfica. Esta gráfica cuantílica normal *no* produce puntos que se aproximen de manera razonable a un patrón de línea recta. La pantalla de STATDISK sugiere que los valores de 66.2, 67.6, 68.7, 70.8 y 717 provienen de una población con una distribución que *no* es normal.

STATDISK



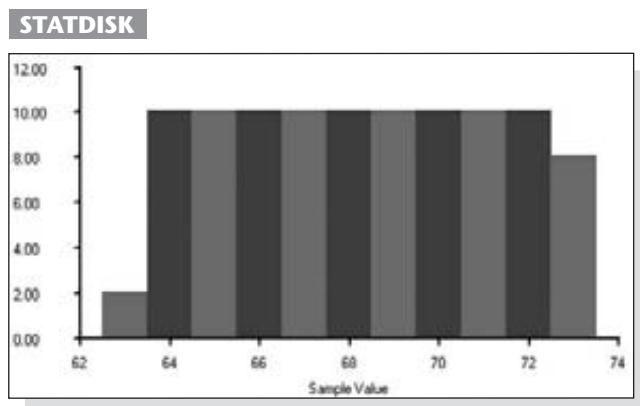
El siguiente ejemplo ilustra el uso de un histograma y de una gráfica cuantílica con un conjunto de datos más grande. Este tipo de conjuntos grandes de datos suelen requerir el uso de un programa de cómputo.

EJEMPLO Estaturas de hombres Las dos gráficas cuantilares normales anteriores se refieren a estaturas de hombres, pero sólo incluyen cinco valores muestrales. Considere las siguientes 100 estaturas de hombres, proporcionadas por un investigador a quien se pidió que seleccionara al azar a 100 hombres y que midiera sus estaturas.

63.3	63.4	63.5	63.6	63.7	63.8	63.9	64.0	64.1	64.2
64.3	64.4	64.5	64.6	64.7	64.8	64.9	65.0	65.1	65.2
65.3	65.4	65.5	65.6	65.7	65.8	65.9	66.0	66.1	66.2
66.3	66.4	66.5	66.6	66.7	66.8	66.9	67.0	67.1	67.2
67.3	67.4	67.5	67.6	67.7	67.8	67.9	68.0	68.1	68.2
68.3	68.4	68.5	68.6	68.7	68.8	68.9	69.0	69.1	69.2
69.3	69.4	69.5	69.6	69.7	69.8	69.9	70.0	70.1	70.2
70.3	70.4	70.5	70.6	70.7	70.8	70.9	71.0	71.1	71.2
71.3	71.4	71.5	71.6	71.7	71.8	71.9	72.0	72.1	72.2
72.3	72.4	72.5	72.6	72.7	72.8	72.9	73.0	73.1	73.2

SOLUCIÓN

Paso 1: Construya un histograma. La siguiente pantalla de STATDISK presenta el histograma de las 100 estaturas, el cual sugiere que tales estaturas *no* se distribuyen normalmente.



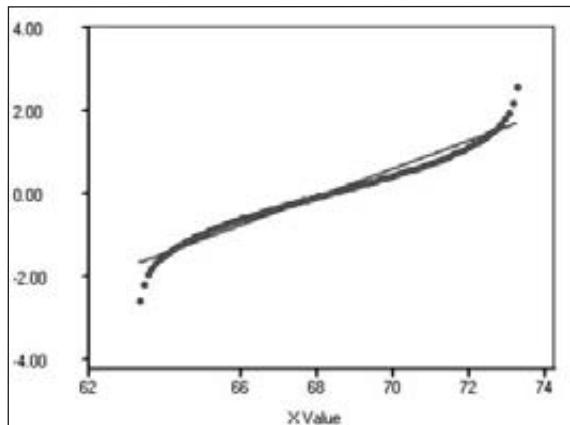
Paso 2: Identifique valores extremos. Al examinar la lista de las 100 estaturas, no encontramos datos que parezcan ser valores extremos.

Paso 3: Construya una gráfica cuantilar normal. La pantalla de STATDISK que aparece en la página 306 presenta la gráfica cuantilar normal. Al examinarla, se revela un patrón sistemático distinto de una línea recta, lo que sugiere que los datos no provienen de una población distribuida normalmente.

INTERPRETACIÓN Puesto que el histograma no tiene forma de campana y dado que la gráfica cuantilar normal revela un patrón diferente al de una línea recta, concluimos que las estaturas no parecen distribuirse de manera normal. Algunos de los procedimientos estadísticos de los capítulos posteriores requieren que los datos muestrales se distribuyan normalmente, pero ese requisito no

continúa

se cumple en este conjunto de datos. Se esperaría que las estaturas de 100 hombres elegidos al azar tuvieran una distribución aproximadamente normal, por lo que el investigador debe ser cuestionado. También podríamos examinar los datos con mayor detalle. Observe que los valores, al ordenarse, aumentan 0.1 de manera consistente. Ésta es otra evidencia de que las estaturas no son el resultado de mediciones obtenidas de hombres seleccionados al azar.

STATDISK


Transformación de datos Muchos conjuntos de datos tienen una distribución que no es normal, pero podemos *transformar* los datos para que los valores modificados tengan una distribución normal. Una transformación común consiste en reemplazar cada valor de x por $\log(x + 1)$. [En vez de utilizar $\log(x + 1)$, podríamos emplear una transformación más directa al reemplazar cada valor de x por $\log x$, pero el uso de $\log(x + 1)$ tiene algunas ventajas, incluyendo la propiedad de que si $x = 0$, entonces $\log(x + 1)$ se puede evaluar, en tanto que $\log x$ es indefinido]. Si la distribución de los valores $\log(x + 1)$ es normal, la distribución de los valores x se denomina **distribución log normal**. (Véase el ejercicio 22). Además de reemplazar cada valor x por $\log(x + 1)$, existen otras transformaciones, como reemplazar cada valor x por \sqrt{x} , o $1/x$ o x^2 . Además de obtener una distribución normal requerida cuando los valores de los datos originales no se distribuyen de manera normal, este tipo de transformaciones se pueden emplear para corregir otras deficiencias, como el requisito (que encontraremos en capítulos posteriores) de que distintos conjuntos de datos tengan la misma varianza.

A continuación presentamos unos comentarios finales acerca de los procedimientos empleados para determinar si los datos provienen de una población distribuida de manera normal:

- Si el requisito de una distribución normal no es demasiado estricto, el examen de un histograma y de los valores extremos podría ser todo lo que necesite para evaluar la normalidad.
- Las gráficas cuantilares normales pueden ser difíciles de construir, pero pueden generarse con una calculadora TI-83/84 Plus o con un programa de cómputo como STATDISK, SPSS, SAS, Minitab y Excel.
- Además de los procedimientos estudiados en esta sección, existen otros procedimientos más avanzados, como la chi cuadrada, la prueba de bondad de ajuste, la prueba de Kolmogorov-Smirnov y la prueba de Lilliefors.

Uso de la tecnología

STATDISK STATDISK puede utilizarse para generar una gráfica cuantilar normal, y el resultado es congruente con el procedimiento descrito en esta sección. Registre los datos en una columna de la ventana del editor de muestras (Sample Editor), después seleccione **Data** de la parte superior de la barra del menú principal, luego seleccione **Normal Quantile Plot**. Proceda a indicar el número

de columna para los datos y haga clic en **Evaluate**.

MINITAB Minitab puede emplearse para generar una gráfica de probabilidad normal similar a la descrita en esta sección. El procedimiento de Minitab es un poco diferente, pero la gráfica puede interpretarse utilizando los mismos criterios de esta sección. Es decir, los datos que se distribuyen de manera normal deben aproximarse a una línea recta, y los puntos no deben revelar un patrón distinto al de una línea recta. Primero anote los valores en la columna C1, después seleccione **Stat**, **Basic Statistics** y **Normality Test**. Introduzca C1 para la variable, después haga clic en **OK**.

EXCEL El complemento Data Desk XL puede utilizarse para generar una gráfica cuantilar normal similar a la descrita en esta sección. Primero registre los valores muestrales en la columna A, después haga clic en

DDXL. (Si DDXL no aparece en la barra del menú, instale el complemento Data Desk XL). Seleccione **Charts and Plots**, después seleccione la función de **Normal Probability Plot**. Haga clic en el ícono del lápiz para “Quantitative Variable”, luego ingrese los rangos de valores, tales como A1:A36. Presione **OK**.

TI-83/84 PLUS La calculadora TI-83/84 Plus puede utilizarse para generar una gráfica cuantilar normal, y el resultado es congruente con el procedimiento descrito en esta sección. Primero anote los datos muestrales en la lista L1, presione **2nd** y la tecla **Y =** (para **STAT PLOT**), después presione **ENTER**. Seleccione **ON**, el elemento “type”, que es el último del segundo renglón de opciones, luego **L1** para la lista de datos. Después de hacer todas las selecciones, presione **ZOOM** y luego **9**.



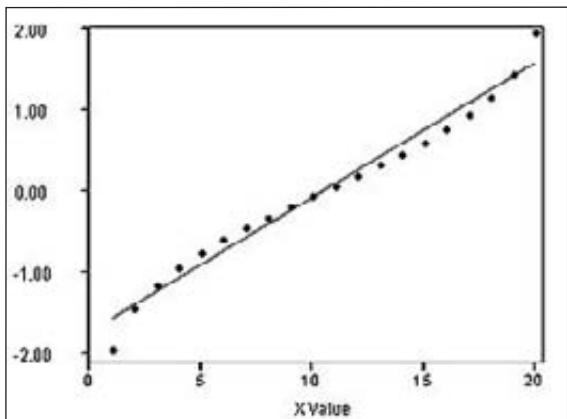
6-7 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

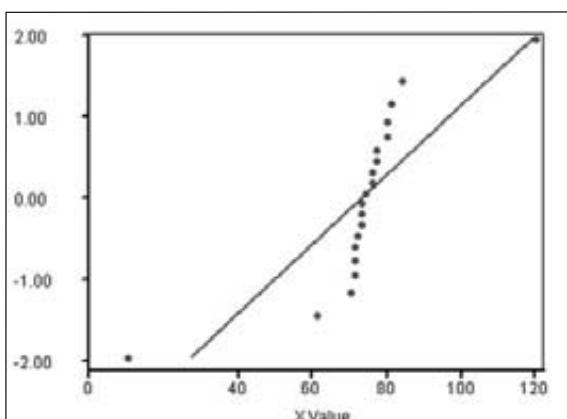
- Gráfica cuantilar normal.** ¿Cuál es el propósito de construir una gráfica cuantilar normal?
- Rechazo de normalidad.** Identifique dos características diferentes de una gráfica cuantilar normal, de tal manera que cada característica conduzca a la conclusión de que los datos no provienen de una población distribuida normalmente.
- Gráfica cuantilar normal.** Si usted selecciona al azar a 100 mujeres entre 21 y 30 años y luego construye una gráfica cuantilar normal con sus estaturas, describa cómo sería la gráfica que esperaría.
- Valor extremo.** Si usted tiene 49 valores de datos, elegidos al azar de una población distribuida normalmente, y también existe un 50º dato que representa un valor extremo, ¿este valor extremo destacará en la gráfica cuantilar normal o se ajustará aparentemente debido a que sólo es un valor entre 50?

Interpretación de gráficas cuantilares normales. En los ejercicios 5 a 8, examine la gráfica cuantilar normal y determine si describe datos de una población con una distribución normal.

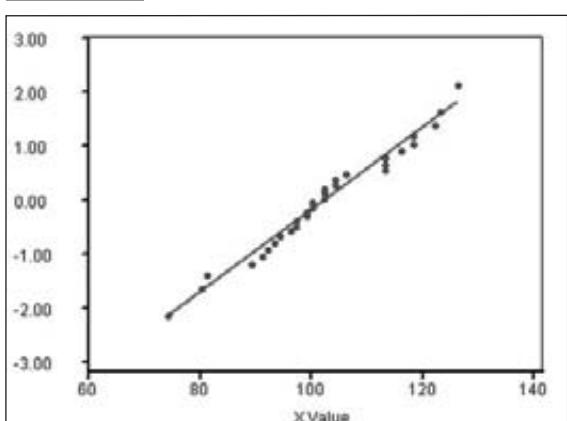
5. STATDISK



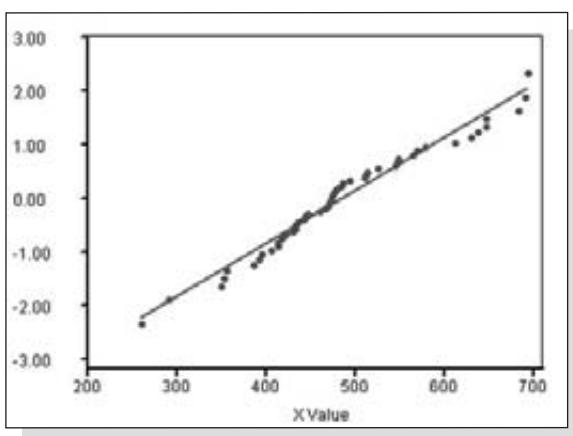
6. STATDISK



7. STATDISK



8. STATDISK



Determinación de normalidad. En los ejercicios 9 a 12, remítase al conjunto de datos indicado y determine si se satisface el requisito de una distribución normal. Suponga que este requisito es flexible, en el sentido de que la distribución poblacional no necesita ser exactamente normal, sino que debe tratarse de una distribución que sea básicamente simétrica y con una moda única.

9. **IMC.** Los valores de índice de masa corporal (IMC) medidos de una muestra de hombres, como aparecen en el conjunto de datos 1 del apéndice B.
10. **Pesos de peniques.** Los pesos de los llamados “peniques de trigo”, tal como se listan en el conjunto de datos 14 del apéndice B.
11. **Precipitación.** Las cantidades de precipitación, como aparecen en el conjunto de datos 8 del apéndice B.
12. **Temperaturas.** Las temperaturas medias diarias, tal como se listan en el conjunto de datos 9 del apéndice B.

Uso de la tecnología para construir gráficas cuantilares normales. En los ejercicios 13 a 16, utilice los datos del ejercicio indicado en esta sección. Utilice una calculadora TI-83/84 Plus o un programa de cómputo (como STATDISK, Minitab o Excel), capaz de generar gráficas cuantilares normales (o gráficas de probabilidad normal). Genere la gráfica y después determine si los datos provienen de una población distribuida normalmente.

- | | |
|------------------|------------------|
| 13. Ejercicio 9 | 14. Ejercicio 10 |
| 15. Ejercicio 11 | 16. Ejercicio 12 |
17. **Comparación de conjuntos de datos.** Con las estaturas de mujeres y los niveles de colesterol de mujeres, listados en el conjunto de datos 1 del apéndice B, analice los dos conjuntos de datos y determine si cada uno de ellos parece provenir de una población distribuida de manera normal. Compare los resultados y dé una posible explicación para cualquier diferencia notoria entre las dos distribuciones.
 18. **Comparación de conjuntos de datos.** Con los niveles de presión sanguínea sistólica y las anchuras del codo de mujeres, listados en el conjunto de datos 1 del apéndice B, analice los dos conjuntos de datos y determine si cada uno de ellos parece provenir de una población distribuida de manera normal. Compare los resultados y dé una posible explicación para cualquier diferencia notoria entre las dos distribuciones.

Construcción de gráficas cuantilares normales. En los ejercicios 19 y 20, utilice los valores dados e identifique las puntuaciones z correspondientes que se emplean para una gráfica cuantilar normal, después construya la gráfica cuantilar normal y determine si los datos parecen provenir de una población con distribución normal.

19. **Estaturas de los Lakers de Los Ángeles.** Utilice esta muestra de estaturas (en pulgadas) de los jugadores de la alineación estelar del equipo profesional de básquetbol de los Lakers de Los Ángeles: 85, 79, 82, 73, 78.
20. **Monitoreo del plomo en el aire.** Los días siguientes a la destrucción causada por los ataques terroristas del 11 de septiembre del 2001, se registraron las cantidades del plomo en el aire (en microgramos por metro cúbico), en el edificio 5 del World Trade Center y se obtuvieron los siguientes valores: 5.40, 1.10, 0.42, 0.73, 0.48, 1.10.

6-7 MÁS ALLÁ DE LO BÁSICO

21. **Uso de puntuaciones estándar.** Al construir una gráfica cuantilar normal suponga que, en vez de calcular las puntuaciones z por medio del procedimiento descrito en esta sección, cada valor en una muestra se convierte en su puntuación estándar correspondiente por medio de $z = (x - \bar{x})/s$. Si los puntos (x, z) se marcan en una gráfica, ¿puede usarse esta gráfica para determinar si la muestra proviene de una población distribuida normalmente? Explique.
22. **Distribución log normal.** Pruebe la normalidad de las siguientes duraciones de llamadas telefónicas (en segundos), luego reemplace cada valor x por $\log(x + 1)$ y pruebe la normalidad de los valores transformados. ¿Qué concluye?

31.5	75.9	31.8	87.4	54.1	72.2	138.1	47.9	210.6	127.7
160.8	51.9	57.4	130.3	21.3	403.4	75.9	93.7	454.9	55.1

Repasso

En el capítulo 5 estudiamos el concepto de distribuciones de probabilidad, pero sólo incluimos las distribuciones *discretas*. En este capítulo estudiamos las distribuciones de probabilidad *continuas* y nos enfocamos en su categoría más importante: las distribuciones normales, las cuales se utilizarán extensamente en los siguientes capítulos.

En la sección 6-2 observamos que, cuando se grafican, las distribuciones normales se aproximan a una forma de campana. El área total bajo la curva de densidad de una distribución normal es 1, de manera que existe una correspondencia conveniente entre áreas y probabilidades. Las áreas específicas pueden encontrarse por medio de la tabla A-2, de una calculadora TI-83/84 Plus o de un programa de cómputo. (No utilizamos la fórmula 6-1, que es la ecuación utilizada para definir la distribución normal).

En la sección 6-3 presentamos métodos importantes para trabajar con las distribuciones normales, incluyendo las que emplean la puntuación estándar $z = (x - \mu)/\sigma$ para resolver problemas como éstos:

- Dado que las puntuaciones de CI se distribuyen normalmente, con $\mu = 100$ y $\sigma = 15$, calcule la probabilidad de seleccionar al azar a un individuo con un CI por arriba de 90.
- Dado que las puntuaciones de CI se distribuyen normalmente, con $\mu = 100$ y $\sigma = 15$, calcule la puntuación de CI que separa al 85% inferior del 15% superior.

En la sección 6-4 analizamos el concepto de distribución muestral. La distribución muestral de la media es la distribución de probabilidad de medias muestrales, en la que todas las muestras tienen el mismo tamaño de muestra n . La distribución

muestral de la proporción es la distribución de probabilidad de proporciones muestrales, donde todas las muestras tienen el mismo tamaño de muestra n . En general, la distribución muestral de cualquier estadístico es la distribución de probabilidad de ese estadístico.

En la sección 6-5 presentamos los siguientes puntos importantes, asociados con el teorema del límite central:

1. Conforme el tamaño de muestra n se incrementa, la distribución de medias muestrales se aproxima a una distribución normal.
2. La media de las medias muestrales es la media poblacional μ .
3. La desviación estándar de las medias muestrales es σ/\sqrt{n} .

En la sección 6-6 señalamos que en ocasiones podemos aproximar una distribución de probabilidad binomial por medio de una distribución normal. Si $np \geq 5$ y $nq \geq 5$, la variable aleatoria binomial x se distribuye de manera aproximadamente normal, con una media $\mu = np$ y una desviación estándar dadas por $\sigma = \sqrt{npq}$. Puesto que la distribución de probabilidad binomial trata con datos discretos y la distribución normal trata con datos continuos, aplicamos la corrección por continuidad, que debe emplearse en aproximaciones normales de distribuciones binomiales.

Por último, en la sección 6-7 usamos un procedimiento para determinar si los datos muestrales parecen provenir de una población con distribución normal. Algunos de los métodos estadísticos que se estudiarán más adelante en este libro requieren, de forma flexible, de una población distribuida normalmente. En tales casos, es probable que lo único que se necesite sea el examen de un histograma y de los valores extremos. En otros casos, podrían necesitarse gráficas cuantilares normales ante el estricto requisito de que la población tenga una distribución normal.

Conocimientos estadísticos y pensamiento crítico

1. **Distribución normal.** ¿Qué es una distribución normal? ¿Qué es una distribución normal estándar?
2. **Distribución de medias muestrales.** Un proceso consiste en seleccionar al azar a 250 adultos, medir su la fuerza de agarre (sólo de la mano derecha) y calcular la media muestral. Suponiendo que este proceso se repite muchas veces, ¿qué factor importante conocemos acerca de la distribución de las medias muestrales resultantes?
3. **Muestra aleatoria simple.** Un investigador reunió una muestra grande de puntuaciones de CI de sus amigos y parientes. Él afirma que, puesto que su muestra es grande y la distribución de sus puntuaciones muestrales se aproxima mucho a la forma de campana de una distribución normal, su muestra es representativa de la población. ¿Es correcto este razonamiento?
4. **Teorema del límite central.** ¿Qué nos indica el teorema del límite central?

Ejercicios de repaso

1. **Errores de peso.** Una báscula está diseñada de tal manera que, cuando se pesan los artículos, los errores en los pesos indicados se distribuyen normalmente con una media de 0 g y una desviación estándar de 1 g. (Si la lectura de la báscula es demasiado baja, el error es negativo. Si la lectura de la báscula es demasiado alta, el error es positivo).
 - a. Si se elige al azar un artículo y se pesa, ¿cuál es la probabilidad de que haya un error entre -0.5 y 0.5 g?
 - b. Si se seleccionan 16 artículos al azar y se pesan, ¿cuál es la probabilidad de que la media de los errores caiga entre -0.5 y 0.5 g?
 - c. ¿Cuál es el percentil 90 de los errores?

- 2. Club Beanstalk de Boston.** El Club Beanstalk de Boston exige una estatura mínima de 74 pulgadas para los hombres. Las estaturas de los hombres se distribuyen normalmente, con una media de 69.0 in y una desviación estándar de 2.8 in (según datos de la National Health Survey).
- ¿Qué porcentaje de los hombres cubren el requisito de la estatura mínima?
 - Si se seleccionan cuatro hombres al azar, ¿cuál es la probabilidad de que su estatura media sea de al menos 74 in?
 - Si se modificara el requisito de la estatura mínima para los hombres, de manera que sólo el 10% de los hombres más altos pudieran ser miembros, ¿cuál sería la nueva estatura mínima?
- 3. Niveles altos de colesterol.** Los niveles de colesterol sérico de hombres entre 18 y 24 años de edad se distribuyen normalmente, con una media de 178.1 y una desviación estándar de 40.7. Las unidades son mg/100 mL y los datos están basados en la National Health Survey.
- Si se selecciona al azar a un hombre entre 18 y 24 años, calcule la probabilidad de que su nivel de colesterol sérico sea mayor de 260, valor considerado “moderadamente alto”.
 - Si se selecciona al azar a un hombre entre 18 y 24 años, calcule la probabilidad de que su nivel de colesterol sérico esté entre 170 y 200.
 - Si se seleccionan al azar 9 hombres entre 18 y 24 años, calcule la probabilidad de que su nivel medio de colesterol sérico esté entre 170 y 200.
 - La Providence Health Maintenance Organization desea establecer un criterio para recomendar cambios en la dieta, si los niveles de colesterol se encuentran dentro del 3% superior. ¿Cuál es el punto de corte para los hombres de 18 a 24 años?
- 4. Detección de sesgo por género en una pregunta de prueba** Al analizar una pregunta específica en una prueba de CI, se descubre que de las 20 personas que respondieron de forma incorrecta, 18 son mujeres. Dado que la prueba se aplicó al mismo número de hombres y mujeres, los cuales fueron elegidos cuidadosamente para que tuvieran la misma capacidad intelectual, calcule la probabilidad de que, de 20 respuestas incorrectas, al menos 18 provengan de mujeres. ¿El resultado constituye una fuerte evidencia de que la pregunta está sesgada a favor de los hombres?
- 5. Discriminación por género.** Cuando varias mujeres no son contratadas por la Tektronics Company, realizan una investigación y encuentran que, entre la gran cantidad de personas que solicitaron empleo, el 30% eran mujeres. Sin embargo, las 20 personas contratadas incluyen sólo 3 mujeres y 17 hombres. Calcule la probabilidad de seleccionar al azar 20 personas de un grupo grande de solicitantes (30% de las cuales son mujeres) y obtener 3 o menos mujeres. Con base en el resultado, ¿parece que la compañía está discriminando con base en el género?
- 6. Prueba de normalidad.** A continuación se incluyen las duraciones (en días) en que el presupuesto del estado de Nueva York ha tenido una demora durante 19 años recientes consecutivos. ¿Parece que tales duraciones provienen de una población distribuida normalmente? ¿Por qué?

4 4 10 19 18 48 64 1 4 68 67 103 125 13 125 34 124 45 44

Ejercicio de repaso acumulativo

- 1. Carbohidratos en la comida.** Se eligen al azar algunos artículos estándar de comida y se mide su contenido de carbohidratos (en gramos); los resultados se listan a continuación (según datos del Departamento de Agricultura de Estados Unidos). (Los artículos son 12 oz de café regular, una taza de leche entera con 3.3% de grasa, un huevo, un plátano, una rosquilla, una cucharada de mantequilla de maní, una zanahoria y 10 papas fritas).

0 12 1 27 24 3 7 10

- a. Calcule la media \bar{x} .
 - b. Calcule la mediana.
 - c. Calcule la desviación estándar s .
 - d. Calcule la varianza s^2 .
 - e. Convierta el valor de 3 g en una puntuación z .
 - f. Calcule el porcentaje real de estos valores muestrales que exceden 3 g.
 - g. Suponiendo una distribución normal, calcule el porcentaje de cantidades *poblacionales* que exceden 3 g. Utilice los valores muestrales de \bar{x} y s como estimados de μ y σ .
 - h. ¿Qué nivel de medición (nominal, ordinal, de intervalo, de razón) describe a este conjunto de datos?
 - i. Parece que las medidas están redondeadas al gramo más cercano, pero, ¿las cantidades exactas sin redondeo representan datos discretos o datos continuos?
 - j. ¿Tiene sentido utilizar la media muestral como un estimado del contenido de carbohidratos de los alimentos que consume el adulto estadounidense promedio? ¿Por qué?
- 2. Zurdos.** Según datos de la American Medical Association, el 10% de las personas son zurdas.
- a. Si se seleccionan tres personas al azar, calcule la probabilidad de que todas sean zurdas.
 - b. Si se seleccionan tres personas al azar, calcule la probabilidad de que al menos una de ellas sea zurda.
 - c. ¿Por qué no podemos resolver el problema del inciso b) con una aproximación normal de la distribución binomial?
 - d. Si se seleccionan al azar grupos de 50 personas, ¿cuál sería el número medio de individuos zurdos en estos grupos?
 - e. Si se seleccionan al azar grupos de 50 personas, ¿cuál sería la desviación estándar del número de personas zurdas en estos grupos?
 - f. ¿Sería infrecuente obtener 8 sujetos zurdos en un grupo de 50 personas seleccionadas al azar? ¿Por qué?

Actividades de cooperación en equipo

- 1. Actividad fuera de clase** Formen grupos de tres o cuatro estudiantes. En cada grupo, diseñen un procedimiento original para ilustrar el teorema del límite central. El objetivo principal es demostrar que cuando se seleccionan al azar muestras de una población, las medias de esas muestras tienden a distribuirse *normalmente*, sin importar la naturaleza de la distribución poblacional.
- 2. Actividad en clase** Formen grupos de tres o cuatro estudiantes. Utilicen una moneda para simular nacimientos y pidan a cada miembro del grupo que simule 25 nacimientos y registre el número de niñas simuladas. Combinen todos los resultados del grupo y registren n = número total de nacimientos y x = número de niñas. Con los grupos de n nacimientos, calculen la media y la desviación estándar del número de niñas. ¿Es frecuente o infrecuente el resultado simulado? ¿Por qué?

3. Actividad en clase Formen grupos de tres o cuatro estudiantes. Selecionen un conjunto de datos del apéndice B (excluyan los conjuntos de datos 1, 8, 9 y 14, que se utilizaron como ejemplos o ejercicios en la sección 6-7). Aplique los métodos de la sección 6-7 y construya un histograma y una gráfica cuantil normal; después determine si el conjunto de datos parece provenir de una población distribuida normalmente.

Proyecto tecnológico

Uno de los mejores proyectos tecnológicos para este capítulo consiste en verificar propiedades importantes del teorema de límite central descrito en la sección 6-5. Proceda de la siguiente manera:

1. Utilice una calculadora o un programa de cómputo para simular 100 lanzamientos de un dado. Seleccione un generador aleatorio que produzca los números enteros 1, 2, 3, 4, 5, 6, todos elegidos al azar. (Las instrucciones aparecen abajo).
2. Calcule y registre la media de los 100 resultados.
3. Repita los primeros dos pasos hasta obtener 50 medias muestrales.
4. Registre las 50 medias muestrales; luego, genere un histograma y estadísticos descriptivos para esas medias muestrales.
5. Sin generar realmente un histograma, ¿cuál es la forma aproximada del histograma para los 5000 lanzamientos simulados de un dado? ¿En qué difiere del histograma generado en el paso 4?
6. ¿Cuál es la media de las 50 medias muestrales? ¿En qué difiere de la media de muchos lanzamientos de un dado legal?
7. ¿Cuál es la desviación estándar de las 50 medias muestrales? ¿En qué difiere de la desviación estándar de los resultados que se obtienen cuando un solo dado se lanza muchas veces?
8. Describa con sus propias palabras la manera en que los resultados demuestran el teorema del límite central.

STATDISK: Seleccione **Data** de la barra del menú principal, después elija la opción de **Uniform Generator**. Introduzca 100 para el tamaño de la muestra, 1 para el valor mínimo, 6 para el valor máximo y utilice 0 decimales. **Sugerencia:** Utilice la misma ventana presionando el botón **Evaluate**. Una vez que se ha generado cada muestra, calcule la media muestral utilizando las funciones Copiar/Pegar para copiar los datos a la ventana Sample Editor.

Minitab:

Seleccione las opciones **Calc**, **Random Data** y luego **Integer**. Proceda a introducir 100 en el número de renglones y C1-C50 para las columnas donde se almacenarán los datos. (Esto le permitirá generar las 50 muestras en un solo paso). También introduzca un valor mínimo de 1 y un valor máximo de 6 y luego haga clic en **OK**. Calcule la media de cada columna utilizando **STAT**, **Basic Statistics**, **Display Descriptive Statistics**.

Excel:

Primero anote los valores 1, 2, 3, 4, 5, 6 en la columna A, y luego anote la expresión $=1/6$ en cada una de las primeras seis celdas de la columna B. (Esto define una distribución de probabilidad en las columnas A y B). Luego seleccione **Tools** de la barra del menú principal, después seleccione **Data Analysis** y **Random Number Generation**. Después de hacer clic en OK, utilice el cuadro de diálogo y anote 50 para el número de variables y 100 para la cantidad de números aleatorios; después seleccione "discrete" para el tipo de distribución. Introduzca **A1:B6** para "Value and Probability Input Range". Los datos deben aparecer. Calcule la media para cada columna utilizando **Tools**, **Data Analysis**, **Descriptive Statistics**.

Indique un rango de entrada de **A1:AX100** y haga clic en el recuadro **Summary Statistics**. Después de presionar el botón **OK**, deben aparecer todas las medias muestrales.

TI-83/84 Plus:

Presione **MATH**, luego seleccione **PRB** y después utilice **randInt** para registrar **randInt(1, 6, 100)**, lo cual generará 100 enteros entre 1 y 6, inclusive. Presione **STO→L1** para almacenar los datos en la lista L1. Ahora presione **STAT**, seleccione **Calc** y luego **1-Var Stats** y proceda a obtener los estadísticos descriptivos (incluyendo la media muestral) para la muestra generada.

De los datos a la decisión

Pensamiento crítico: Diseño de un asiento de avión

En este proyecto consideramos el problema de determinar la “distancia de asiento” que se muestra la figura 6-24a). Definimos la distancia de asiento como la longitud entre la parte trasera del cojín de un asiento y el asiento que está enfrente. La determinación de la distancia de asiento debe tomar en cuenta medidas del cuerpo humano. En específico, debemos considerar la “longitud desde el glúteo hasta la rodilla”, tal como se ilustra en la figura 6-24b). El establecimiento de la distancia de asiento de una aeronave es sumamente importante. Si la distancia de asiento es innecesariamente grande, es probable que se eliminen filas de

asientos. Se ha estimado que la eliminación de una sola fila de seis asientos puede costar alrededor de \$8 millones en el transcurso de la vida de la aeronave. Si la distancia de asiento es demasiado pequeña, los pasajeros podrían sentirse incómodos y preferirán volar en otra línea, o su seguridad podría verse comprometida debido a la limitación de su movilidad.

Para determinar la distancia de asiento de nuestro avión, utilizaremos datos reunidos previamente sobre mediciones de grandes grupos de personas. Los resultados de esas medidas se resumen en la tabla. Podemos utilizar los datos de la tabla para determinar la distancia de asiento requerida, pero debemos tomar algunas decisiones difíciles. Si queremos acomodar a *todos* los miembros de la población, tendremos una distancia de asiento tan costosa en

términos de un asiento reducido, que podría no ser viable económicamente. Una de las decisiones difíciles que debemos tomar es la siguiente: ¿Qué porcentaje de la población estamos dispuestos a excluir? Otra decisión necesaria es: ¿Qué cantidad de espacio adicional deseamos ofrecer para la comodidad y seguridad de los pasajeros? Utilice la información disponible para determinar la distancia de asiento. Identifique las opciones y las decisiones que se tomaron para esa determinación.

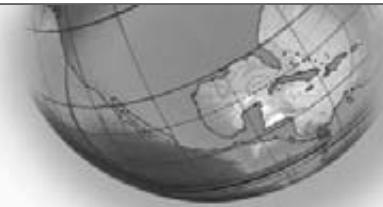
Longitud desde el glúteo hasta la rodilla (en pulgadas)

	Media	Desviación estándar	Distribución
Hombres	23.5 in	1.1 in	Normal
Mujeres	22.7 in	1.0 in	Normal

- Distancia del cojín trasero del asiento de enfrente
- Longitud desde el glúteo hasta la rodilla, más cualquier distancia adicional para brindar comodidad



Figura 6-24 Distancia de asiento y longitud desde el glúteo hasta la rodilla



Proyecto de Internet

Exploración del teorema del límite central

El teorema del límite central es uno de los resultados más importantes en estadística; también puede ser uno de los más sorprendentes. De manera informal, el teorema del límite central dice que la distribución normal está en todas partes. Sin importar qué distribución de probabilidad subyace en un experimento, existe una distribución correspondiente de medias que tendrá una forma aproximadamente normal.

La mejor manera para comprender y apreciar el teorema del límite central es verlo en acción.

El proyecto de Internet de este capítulo, que se encuentra en el sitio de Internet de *Estadística*,

<http://www.pearsoneducacion.net/triola>

le permitirá lograr lo anterior. Se le pedirá observar, interpretar y comentar una demostración del teorema del límite central como parte de un experimento con dados. Además, será guiado a través de una búsqueda en Internet para encontrar otras demostraciones como ésta.

La estadística en el trabajo

“Es posible ser un periodista y no sentirse cómodo con la estadística, pero definitivamente se estará limitado en lo que se puede hacer”.



Joel B. Obermayer

Reportero del *News & Observer*

Joel B. Obermayer escribe acerca de temas médicos y asuntos de salud para *The News & Observer*, un periódico que se distribuye en la región este de Carolina del Norte. Joel realiza reportes sobre la administración de la salud, asuntos de salud pública e investigaciones en centros médicos académicos, incluyendo Duke y la Universidad de Carolina del Norte en Chapel Hill.

¿Qué conceptos de estadística utiliza?

Utilizo ideas como la significancia estadística, porcentajes de error y probabilidad. No necesito hacer cosas increíblemente elaboradas, pero necesito sentirme muy cómodo con las matemáticas y con el planteamiento de preguntas acerca de ellas.

Utilizo la estadística para analizar investigaciones médicas y decidir si diferentes estudios son significativos, así como para definir la forma en que escribo acerca de eso. Principalmente, necesito ser capaz de leer estadísticos y comprenderlos, más que desarrollarlos. Empleo la estadística para plantear buenas preguntas y fundamentar los argumentos que escribo. También la utilizo para decidir si alguien está tratando de darme un punto de vista positivo sobre algo que puede ser cuestionable. Por ejemplo, en una ocasión una persona de una universidad local me envió un artículo sobre cremas milagrosas que, al parecer, ayudaban a bajar de peso disolviendo las células de grasa. Pues bien, yo dudé de que estas cremas funcionaran. El estudio no era muy bueno. Estaban tratando de hacer aseveraciones con base en un estudio que sólo incluyó a 11 personas. El investigador argumentó que 11 individuos eran suficientes para obtener buenas conclusiones empíricas sobre la salud. Eso no fue algo inusual. Las personas tratan de manipular a los medios de comunicación todo el tiempo. Los buenos estudios verificables, con buenas bases estadísticas verificables, ayudan a evitar la manipulación.

¿El uso que usted hace de la probabilidad y la estadística está aumentando, disminuyendo o permanece estable?

Está aumentando. El interés de la gente por nuevas terapias que pueden estar en la etapa de ensayo clínico se está incrementando, en parte debido al énfasis en investigaciones sobre el SIDA y en la obtención de nuevos fármacos que pronto serán aprobados y recetados a los pacientes. Es más importante que nunca que un escritor sobre asuntos médicos utilice la estadística para asegurarse de que los estudios realmente prueban lo que la gente de relaciones públicas dice que prueban.

¿Deben tener estudios de estadística quienes solicitan empleo?

Es posible ser un periodista y no sentirse cómodo con la estadística, pero definitivamente se estará limitado en lo que se puede hacer. Si usted escribe acerca de la eficacia de programas educativos financiados por el gobierno, o si escribe acerca de los peligros de contaminantes particulares del ambiente, necesitará utilizar la estadística.

En mi campo, los editores no suelen pensar sobre la estadística en los procesos de entrevista; se preocupan más por las habilidades de escritura. El conocimiento de la estadística es más importante para definir lo que se puede hacer una vez que se obtiene el empleo.



Estimaciones y tamaños de muestra

7



- 7-1 Panorama general**
- 7-2 Estimación de la proporción de una población**
- 7-3 Estimación de una media poblacional: σ conocida**
- 7-4 Estimación de una media poblacional: σ desconocida**
- 7-5 Estimación de la varianza de una población**

¿Funciona la terapia de contacto?

Muchos pacientes pagan de \$25 a \$50 por una sesión de terapia de contacto en la que el terapeuta coloca sus manos a unos centímetros del cuerpo del paciente, sin tener realmente contacto físico. El objetivo es curar una amplia variedad de problemas médicos, incluyendo cáncer, SIDA, asma, enfermedades cardíacas, dolores de cabeza, quemaduras y fracturas óseas. La teoría básica plantea que un terapeuta de contacto capacitado profesionalmente puede detectar un mal alineamiento en el campo de energía del paciente y generar un equilibrio energético que incrementa el proceso de curación.

Cuando Emily Rosa, una niña de nueve años, estaba en cuarto grado, eligió el tema de la terapia de contacto para el proyecto de una feria de ciencias y convenció a 21 terapeutas de contacto experimentados para que participaran en una prueba sencilla de su capacidad para detectar el campo de energía humana. Emily utilizó un cartón con dos agujeros para introducir las manos. Cada terapeuta de contacto pasaba sus dos manos a través de los agujeros, y Emily colocaba su mano por arriba de una de las manos del terapeuta; luego, se le pedía al terapeuta que identificara la mano que Emily había elegido. La niña lanzaba una moneda para seleccionar al azar la mano sobre la que colocaba la suya. Esta prueba se repitió 280 veces. Si los terapeutas de contacto realmente tenían la habilidad de percibir un campo energético humano, debían identificar la mano correcta mucho más del 50% de las veces. Si no tenían tal capacidad y sólo hacían conjjeturas, debían acertar alrededor del 50% de las veces. Emily obtuvo los siguientes resultados: de los 280 ensayos, los terapeutas de contacto identificaron la mano correcta 123 veces, es decir, tuvieron una tasa de

éxito del 44%. Emily, con la ayuda de su madre, un especialista en estadística y un médico, envió sus hallazgos para publicarlos en el prestigioso *Journal of the American Medical Association*. Después de una cuidadosa y detallada revisión del diseño experimental y de los resultados, se publicó el artículo “A Close Look at a Therapeutic Touch” (*Journal of the American Medical Association*, vol. 279, núm. 13). Emily se convirtió en la investigadora más joven en publicar un artículo en esa revista. Además, ganó el primer premio de la feria de ciencias por su proyecto.

Consideremos los principales resultados del proyecto de Emily. En los 280 ensayos, los terapeutas de contacto acertaron 123 veces. Tenemos una proporción muestral con $n = 280$ y $x = 123$. Los argumentos en contra de la validez del estudio podrían incluir la aseveración de que el número de ensayos es demasiado pequeño para ser significativo, o que los terapeutas de contacto tuvieron un mal día y que, debido al azar, no tuvieron tanto éxito como la población de todos los terapeutas de contacto. En este capítulo analizaremos estos temas.

También es importante señalar que el proyecto de Emily Rosa fue relativamente sencillo. Recuerde que ella realizó el estudio cuando cursaba el cuarto grado de primaria. Su proyecto es el tipo de actividad que cualquier estudiante de un curso de introducción a la estadística podría llevar a cabo. Después de comprender los conceptos que se enseñan en el curso de introducción a la estadística típico, los estudiantes tienen la habilidad para realizar trabajos significativos e importantes.

7-1 Panorama general

En este capítulo comenzaremos a trabajar con el verdadero núcleo de la estadística inferencial al utilizar datos muestrales para hacer inferencias acerca de poblaciones. En específico, usaremos datos muestrales para hacer estimaciones de parámetros de población. Por ejemplo, el problema del capítulo se refiere a terapeutas de contacto que identificaron correctamente el campo de energía humano únicamente en el 44% de 280 ensayos. Con base en el estadístico muestral de 44%, estimaremos el porcentaje de aciertos para toda la población de los terapeutas de contacto.

Las dos aplicaciones principales de la estadística inferencial implican el uso de datos muestrales para **1.** estimar el valor de un parámetro de una población y **2.** probar alguna aseveración (o hipótesis) acerca de una población. En este capítulo presentamos métodos para estimar valores de estos importantes parámetros de población: proporciones, medias y varianzas. También presentamos métodos para determinar los tamaños de muestra necesarios para estimar estos parámetros. En el capítulo 8 estudiaremos los métodos básicos para probar aseveraciones (o hipótesis) que se hacen acerca de un parámetro de una población.

7-2 Estimación de la proporción de una población

Concepto clave En esta sección presentamos métodos importantes donde se utiliza una proporción muestral para estimar el valor de una proporción poblacional con un *intervalo de confianza*. También presentamos métodos para calcular el tamaño muestral necesario para estimar una proporción poblacional. En esta sección se explican conceptos generales que se utilizan en las siguientes secciones y en los capítulos posteriores, por lo que es muy importante su comprensión.

Una estrategia de estudio: El tiempo dedicado a esta sección estará bien invertido, ya que presentamos el concepto de intervalo de confianza, un concepto general que se aplicará en las siguientes secciones del capítulo. Sugerimos que primero lea esta sección con el único objetivo de tratar de comprender qué son los intervalos de confianza, para qué sirven y por qué son indispensables. Segundo, trate de desarrollar la habilidad de construir estimaciones del intervalo de confianza de las proporciones de una población. Tercero, aprenda a interpretar correctamente un intervalo de confianza. Cuarto, lea la sección una vez más y trate de comprender la teoría subyacente. Usted tendrá mucho más éxito si comprende lo que está haciendo, en vez de aplicar a ciegas y de manera mecánica los pasos para obtener una respuesta que podría o no tener sentido.

En esta sección sólo consideraremos casos en los que es posible utilizar la distribución normal para aproximar la distribución muestral de proporciones de muestra. Los siguientes requisitos se aplican a los métodos de esta sección.

Requisitos

1. La muestra es aleatoria simple.
2. Las condiciones para la distribución binomial se satisfacen. Esto es, hay un número fijo de ensayos, los ensayos son independientes, hay dos categorías de resultados y las probabilidades permanecen constantes para cada ensayo. (Véase la sección 5-3).
3. Existen al menos 5 éxitos y al menos 5 fracasos. (Cuando p y q se desconocen, estimamos sus valores utilizando la proporción muestral, de manera que este

requisito es una forma de verificar que $np \geq 5$ y $nq \geq 5$ se cumplan para que la distribución normal sea una aproximación adecuada para la distribución binomial. Además, existen procedimientos para tratar situaciones en las que la distribución normal no es una aproximación adecuada. Véase el ejercicio 51).

Notación para proporciones

p = proporción de la *población*

$\hat{p} = \frac{x}{n}$ = proporción muestral de *x éxitos* en una muestra de tamaño n

$\hat{q} = 1 - \hat{p}$ = proporción muestral de *fracasos* en una muestra de tamaño n

Proporción, probabilidad y porcentaje Esta sección se enfoca en la proporción poblacional p , aunque también podemos trabajar con probabilidades o porcentajes. Cuando trabaje con un porcentaje, expréselo en forma decimal. (Por ejemplo, exprese el 44% como 0.44, de manera que $\hat{p} = 0.44$). Si desea estimar una proporción poblacional con un solo valor, el mejor estimado es \hat{p} . Puesto que \hat{p} consiste en un solo valor, se denomina estimado puntual.

Definición

Un **estimado puntual** es un valor individual (o punto) que se usa para aproximar un parámetro de población.

La proporción muestral \hat{p} es el mejor estimado puntual en la proporción poblacional p .

Usamos \hat{p} como el estimado puntual de p , ya que no está sesgado y es el más consistente de los estimadores que podrían usarse. No está sesgado en el sentido de que la distribución de las proporciones muestrales tiende a concentrarse alrededor del valor de p ; esto es, las proporciones muestrales \hat{p} no tienden sistemáticamente a subestimar ni a sobreestimar p . (Véase la sección 6-4). La proporción muestral \hat{p} es el estimador más consistente en el sentido de que la desviación estándar de las proporciones muestrales tiende a ser menor que la desviación estándar de cualquier otro estimador sin sesgo.



EJEMPLO Tasa de éxito de la terapia de contacto En el problema del capítulo señalamos que, en 280 ensayos con terapeutas de contacto, la mano correcta fue elegida en 123 ensayos, de manera que la tasa de éxito es $\hat{p} = 123/280 = 0.44$. Utilice estos resultados de prueba para calcular el mejor estimado puntual de la proporción de todas las selecciones correctas que resultarían si se pusiera a prueba a todos los terapeutas de contacto.

SOLUCIÓN Puesto que la proporción muestral es el mejor estimado puntual de la proporción poblacional, concluimos que el mejor estimado puntual de p es 0.44. (Utilizando el sentido común, podríamos examinar el diseño del experimento y concluir que la verdadera proporción poblacional es 0.5, pero si sólo utilizamos los resultados muestrales encontramos que 0.44 es el mejor estimado de la proporción poblacional p).



Muestra pequeña

El Children's Defense Fund se constituyó con el fin de promover el bienestar de los niños. El grupo publicó *Children Out of School in America*, donde se informó que, en una zona, el 37.5% de los jóvenes de 16 y 17 años ya no asistían a la escuela. Esta estadística recibió mucha cobertura por parte de los medios de comunicación masiva, pero se basaba en una muestra de sólo 16 jóvenes. Otra estadística se basó en una muestra de sólo tres estudiantes. (Véase "Firsthand Report: How Flawed Statistics Can Make an Ugly Picture Look Even Worse", *American School Board Journal*, vol. 162).

¿Por qué necesitamos intervalos de confianza?

En el ejemplo anterior vimos que 0.44 era el *mejor* estimado puntual de la proporción poblacional p , pero no tenemos indicación precisa de qué tan *bueno* es nuestro mejor estimado. Como el estimado puntual tiene el grave defecto de no revelar nada acerca de qué tan bueno es, los especialistas en estadística han diseñado ingeniosamente otro tipo de estimado: el *intervalo de confianza* o *estimado del intervalo*, que consiste en un rango (o un intervalo) de valores en vez de un solo valor.

Definición

Un **intervalo de confianza** (o **estimado del intervalo**) es un rango (o un intervalo) de valores que se usa para estimar el valor real de un parámetro de población. El intervalo de confianza suele abreviarse como IC.

Un intervalo de confianza se asocia con un nivel de confianza, como 0.95 (o 95%). El nivel de confianza nos da la tasa de éxitos del procedimiento que se utiliza para construir el intervalo de confianza. El nivel de confianza suele expresarse como la probabilidad o área $1 - \alpha$ (alfa griega minúscula). El valor de α es el complemento del *nivel de confianza*. Para un nivel de confianza de 0.95 (o 95%), $\alpha = 0.05$. Para un nivel de confianza de 0.99 (o 99%), $\alpha = 0.01$.

Definición

El **nivel de confianza** es la probabilidad $1 - \alpha$ (a menudo expresada como el valor de porcentaje equivalente), que es la proporción de veces que el intervalo de confianza realmente contiene el parámetro de población, suponiendo que el proceso de estimación se repite un gran número de veces. El nivel de confianza también se llama **grado de confianza** o **coeficiente de confianza**.

Las opciones más comunes para el nivel de confianza son 90% (con $\alpha = 0.10$), 95% (con $\alpha = 0.05$) y 99% (con $\alpha = 0.01$). La opción del 95% es la más común puesto que provee un buen equilibrio entre precisión (reflejada en el ancho del intervalo de confianza) y confiabilidad (expresada por el nivel de confianza).

A continuación se presenta un ejemplo de un intervalo de confianza basado en los datos muestrales de 280 ensayos de terapeutas de contacto, donde en el 44% de los ensayos se identifica correctamente la mano elegida:

El intervalo de confianza estimado de 0.95 (o 95%) de la proporción poblacional p es $0.381 < p < 0.497$.

Interpretación de un intervalo de confianza

Debemos ser cuidadosos para interpretar los intervalos de confianza correctamente. Existe una interpretación correcta y muchas diferentes y creativas interpretaciones erróneas del intervalo de confianza $0.381 < p < 0.497$.

Correcta: “Tenemos una confianza del 95% de que el intervalo de 0.381 a 0.497 realmente contiene el valor verdadero de p ”. Esto significa que si seleccionamos muchas muestras diferentes de tamaño 280 y construimos los intervalos de confianza correspondientes, el 95% de ellos incluirían realmente el valor de la proporción poblacional p . (Note que en esta interpretación correcta, el nivel del 95% se refiere a la tasa de éxitos del *proceso*, utilizada para estimar la proporción, y no a la proporción de la población en sí).

Errónea: “Existe un 95% de probabilidades de que el valor real de p esté entre 0.381 y 0.497”.

Para cualquier punto específico en el tiempo, una población tiene un valor fijo y constante de p , un intervalo de confianza construido a partir de una muestra que incluye o no a p . De manera similar, si un bebé acaba de nacer y el médico está por anunciar su género, es incorrecto decir que existe una probabilidad de 0.5 de que sea una niña; el bebé es o no una niña, y no hay una probabilidad implicada. Una proporción poblacional p es como el bebé que acaba de nacer: el valor de p es fijo, de manera que los límites del intervalo de confianza contienen o no a p . Por eso es incorrecto decir que existe un 95% de probabilidades de que p se localice entre valores tales como 0.381 y 0.497.

Un nivel de confianza del 95% nos dice que el *proceso* que estamos usando, a la larga, dará por resultado límites del intervalo de confianza que contienen la proporción real de la población el 95% del tiempo. Suponga que la proporción real de todas las identificaciones correctas de la mano por parte de los terapeutas de contacto es $p = 0.5$. Entonces, el intervalo de confianza obtenido de los datos muestrales no incluiría la proporción poblacional, ya que la proporción poblacional real de 0.5 no se encuentra entre 0.381 y 0.497. Esto se ilustra en la figura 7-1, la cual señala los intervalos de confianza típicos que resultan de 20 muestras diferentes. Con un 95% de confianza, esperamos que 19 de las 20 muestras den por resultado intervalos de confianza que contienen el valor real de p ; la figura 7-1 ilustra esto con 19 de los intervalos de confianza que contienen p , mientras un intervalo de confianza no contiene p .

Uso de intervalos de confianza para hacer comparaciones *Advertencia:* Los intervalos de confianza pueden usarse de manera informal para comparar conjuntos de datos diferentes, pero el traslape de intervalos de confianza no debe usarse para elaborar conclusiones formales y finales acerca de la igualdad de las proporciones. (Véase “On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals” de Schenker y Gentleman, *American Statistician*, vol. 55, núm. 3).

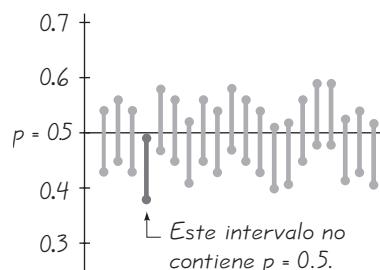


Figura 7-1

Intervalos de confianza de 20 muestras diferentes

Valores críticos

Los métodos de esta sección y muchos de los demás estadísticos que se encuentran en los capítulos siguientes incluyen el uso de una puntuación z estándar, que permite distinguir entre estadísticos muestrales que tienen probabilidades de ocurrir y aquellos que son improbables. Una puntuación z de este tipo se llama *valor crítico* (definido abajo). Los valores críticos se basan en las siguientes observaciones.

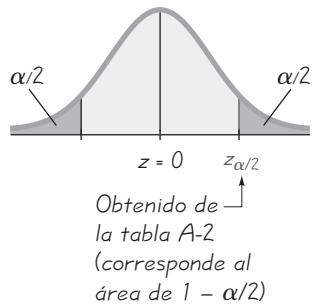


Figura 7-2

Valor crítico $z_{\alpha/2}$ en la distribución normal estándar

1. Sabemos, desde la sección 6-6, que en ciertas condiciones, la distribución muestral de las proporciones muestrales puede aproximarse por una distribución normal, como en la figura 7-2.
2. Las proporciones muestrales tienen una probabilidad relativamente pequeña (denotada por α) de caer en una de las colas sombreadas de gris oscuro de la figura 7-2.
3. Denotando el área de cada cola sombreada como $\alpha/2$, vemos que existe una probabilidad total de α de que una proporción muestral caiga en cualquiera de las dos colas sombreadas de gris oscuro.
4. Por la regla de los complementos (capítulo 4), existe una probabilidad de $1 - \alpha$ de que una proporción muestral caiga dentro de la región interior sombreada de gris claro de la figura 7-2.
5. La puntuación z que separa la región de la cola derecha generalmente se denota por $z_{\alpha/2}$ y se conoce como *valor crítico*, puesto que está en la frontera que separa las proporciones muestrales que tienen probabilidad de ocurrir de aquellas que no tienen probabilidad de ocurrir.

Estas observaciones se formalizan con la notación y definición siguientes.

Notación para el valor crítico

El valor crítico $z_{\alpha/2}$ es el valor z positivo que está en la frontera vertical que separa una área de $\alpha/2$ en la cola derecha de la distribución normal estándar. (El valor de $-z_{\alpha/2}$ está en la frontera vertical para el área de $\alpha/2$ en la cola izquierda). El subíndice $\alpha/2$ es simplemente un recordatorio de que la puntuación z separa una área de $\alpha/2$ en la cola derecha de la distribución normal estándar.

Definición

Un **valor crítico** es el número en la línea límitrofe que separa estadísticos muestrales que tienen mayor probabilidad de ocurrir de aquellos que no tienen probabilidad de ocurrir. El número $z_{\alpha/2}$ es un valor crítico, una puntuación z con la propiedad de que separa una área de $\alpha/2$ en la cola derecha de la distribución normal estándar (véase la figura 7-2).

EJEMPLO Cálculo de un valor crítico Calcule el valor crítico $z_{\alpha/2}$ que corresponde a un nivel de confianza del 95%.

SOLUCIÓN *Advertencia:* Para calcular el valor crítico z de un nivel de confianza del 95% *no* busque 0.95 en el cuerpo de la tabla A-2. Un nivel de confianza del 95% corresponde a $\alpha = 0.05$. Observe la figura 7-3, donde mostramos que el área en cada una de las colas sombreadas de gris oscuro es $\alpha/2 = 0.025$. Calculamos $z_{\alpha/2} = 1.96$ señalando que toda el área a su izquierda debe ser $1 - 0.025$, o 0.975. Podemos remitirnos a la tabla A-2 y

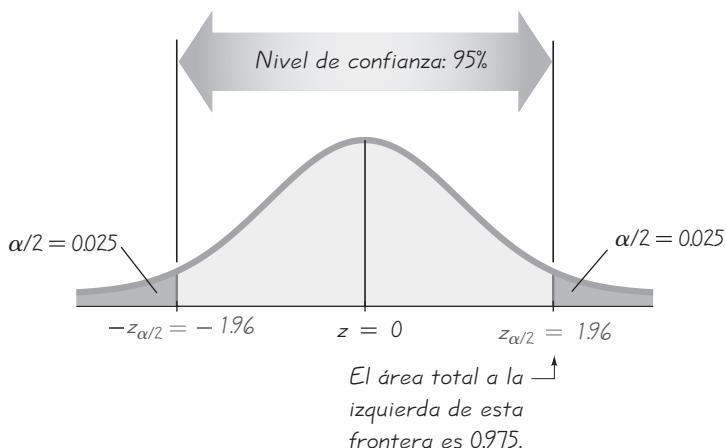


Figura 7-3 Cálculo de $z_{\alpha/2}$ para un nivel de confianza del 95%

encontrar que el área de 0.9750 (que se encuentra en el *cuerpo* de la tabla) corresponde exactamente a una puntuación z de 1.96. Para un nivel de confianza del 95%, el valor crítico es, por consiguiente, $z_{\alpha/2} = 1.96$. Para calcular la puntuación z crítica para un nivel de confianza del 95%, busque 0.9750 en el cuerpo la tabla A-2, no 0.95.

El ejemplo anterior mostró que un nivel de confianza del 95% da por resultado un valor crítico de $z_{\alpha/2} = 1.96$. Éste es el valor crítico más común y se lista junto con otros dos valores comunes en la siguiente tabla.

Nivel de confianza	α	valor crítico, $z_{\alpha/2}$
90%	0.10	1.645
95%	0.05	1.96
99%	0.01	2.575

Margen de error

Cuando reunimos un conjunto de datos muestrales, como los datos sobre la terapia de contacto de Emily Rosa en el problema del capítulo (donde el 44% de los 280 ensayos correspondieron a identificaciones correctas), podemos calcular la proporción muestral \hat{p} y esta proporción muestral suele ser diferente de la proporción poblacional p . La diferencia entre la proporción muestral y la proporción de la población se considera un error. Ahora definimos el *margen de error E* como sigue.

Definición

Cuando se utilizan los datos de una muestra aleatoria simple para estimar una proporción poblacional p , el **margin de error**, denotado por E , es la diferencia máxima probable (con probabilidad $1 - \alpha$) entre la proporción muestral \hat{p} observada y el valor real de la proporción poblacional p . El margen de error E también se llama *error máximo del estimado* y se calcula multiplicando el valor crítico por la desviación estándar de las proporciones muestrales, como se indica en la fórmula 7-1.

Fórmula 7-1

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

margen de error para proporciones



Falsificación de datos

El glosario del censo del 2000 define la *falsificación de datos (curbstoning)* como “la práctica por medio de la cual un censor fabrica un cuestionario para una vivienda, sin visitarla”. La falsificación de datos ocurre cuando un censor se sienta en la acera (o en cualquier otro lado) y llena las formas inventando las respuestas. Puesto que estos datos no son reales, afectan la validez del censo. En varios estudios se ha investigado la magnitud de la falsificación; uno de ellos reveló que aproximadamente el 4% de los censores realizan esta práctica al menos en una ocasión.

Los métodos de la sección 7-2 suponen que los datos muestrales se reunieron de una forma apropiada, así que si gran parte de los datos se obtuvieron a través de falsificaciones, entonces los estimados de los intervalos de confianza resultantes podrían tener muchos errores.



Intervalo de confianza (o estimado de intervalo) para la proporción poblacional p

$$\hat{p} - E < p < \hat{p} + E \quad \text{donde} \quad E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

El intervalo de confianza suele expresarse en los siguientes formatos equivalentes:

$$\hat{p} \pm E$$

o

$$(\hat{p} - E, \hat{p} + E)$$

En el capítulo 4, cuando las probabilidades se expresaban en forma decimal, redondeábamos a tres dígitos significativos. Aquí utilizamos esa misma regla de redondeo.

Regla de redondeo para estimados de intervalo de confianza de p

Redondee los límites del intervalo de confianza para p a tres dígitos significativos.

Con base en los resultados anteriores, podemos resumir el procedimiento para construir un estimado del intervalo de confianza de una proporción poblacional p como sigue.

Procedimiento para construir un intervalo de confianza para p

1. Verifique que los supuestos requeridos se cumplen (la muestra es aleatoria simple, las condiciones para la distribución binomial se satisfacen y existen al menos 5 éxitos y al menos 5 fracasos).
2. Remítase a la tabla A-2 y encuentre el valor crítico $z_{\alpha/2}$ que corresponde al nivel de confianza deseado. (Por ejemplo, si el nivel de confianza es del 95%, el valor crítico es $z_{\alpha/2} = 1.96$).
3. Evalúe el margen de error $E = z_{\alpha/2} \sqrt{\hat{p}\hat{q}/n}$.
4. Utilizando el valor del margen de error E calculado y el valor de la proporción muestral \hat{p} , calcule los valores de $\hat{p} - E$ y $\hat{p} + E$. Sustituya esos valores en el formato general para el intervalo de confianza:

$$\hat{p} - E < p < \hat{p} + E$$

o

$$\hat{p} \pm E$$

o

$$(\hat{p} - E, \hat{p} + E)$$

5. Redondee los límites del intervalo de confianza resultantes a tres dígitos significativos.



EJEMPLO Tasa de éxito de la terapia de contacto En el problema del capítulo señalamos que los terapeutas de contacto participaron en 280 pruebas de su capacidad para percibir el campo de energía humana. En cada ensayo se pidió a un terapeuta que identificara la mano que estaba debajo de la mano de Emily Rosa. De los 280 ensayos, los terapeutas acertaron en 123 ocasiones. Los resultados muestrales son $n = 280$ y $\hat{p} = 123/280 = 0.439286$. (En vez de utilizar 0.44 para la proporción muestral, empleamos decimales adicionales para que los cálculos posteriores no se vean afectados por un error de redondeo).

- Calcule el margen de error E que corresponde a un nivel de confianza del 95%.
- Calcule el estimado del intervalo de confianza del 95% de la proporción poblacional p .
- Con base en los resultados, ¿qué podemos concluir acerca de la eficacia de la terapia de contacto?

SOLUCIÓN

REQUISITO ✓ Primero debemos verificar que se satisfacen los requisitos necesarios. (Antes, en esta sección indicamos los requisitos para utilizar una distribución normal como aproximación de una distribución binomial). Dado el diseño del experimento, es razonable suponer que la muestra es aleatoria simple. Las condiciones para un experimento binomial se satisfacen, ya que existe un número fijo de ensayos (280), los ensayos son independientes (porque el resultado de un ensayo no afecta la probabilidad del resultado de otro ensayo), hay dos categorías de resultados (correcto e incorrecto) y la probabilidad de acertar permanece constante. Asimismo, con 123 identificaciones correctas en 280 ensayos, se obtienen 157 identificaciones incorrectas, de manera que el número de éxitos (123) y el número de fracasos (157) son de al menos 5. Completamos con éxito la verificación de los requisitos. ✓

- El margen de error se calcula usando la fórmula 7-1 con $z_{\alpha/2} = 1.96$ (como se calculó en el ejemplo anterior), $\hat{p} = 0.439286$, $\hat{q} = 1 - 0.439286 = 0.560714$ y $n = 280$.

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96 \sqrt{\frac{(0.439286)(0.560714)}{280}} = 0.058133$$

- Construir el intervalo de confianza es bastante fácil ahora que tenemos los valores de \hat{p} y E . Simplemente sustituimos esos valores para obtener este resultado:

$$\hat{p} - E < p < \hat{p} + E$$

$$0.439286 - 0.058133 < p < 0.439286 + 0.058133$$

$$0.381 < p < 0.497 \text{ (redondeado a tres dígitos significativos)}$$

Este mismo resultado podría expresarse en el formato de 0.439 ± 0.058 o $(0.381, 0.497)$. Si queremos el intervalo de confianza del 95% para el *porcentaje* real de la población, podemos expresar el resultado como $38.1\% < p < 49.7\%$. Este intervalo de confianza suele reportarse con una afirmación como ésta: “Se estima que la tasa de éxito es del 44%, con un margen de error de más o menos 6 puntos porcentuales”. Tal enunciado es una expresión verbal de ese formato para el intervalo de confianza: $44\% \pm 6\%$. El nivel de confianza debe reportarse también, aunque rara vez se hace en los medios de comunicación. En general, los medios de comunicación usan un nivel de confianza del 95%, pero omiten cualquier referencia a él.

continúa

c. Para interpretar los resultados, observe que el hecho de tratar de adivinar habría producido alrededor de un 50% de identificaciones correctas. Si los terapeutas de contacto realmente tienen la habilidad de percibir el campo de energía humano, su tasa de éxito debería ser mayor al 50% por una cantidad significativa. Sin embargo, los terapeutas de contacto tuvieron una tasa de éxito más baja de lo que se esperaría con el lanzamiento de una moneda. He aquí una aseveración del artículo publicado en el *Journal of the American Medical Association*: “El fracaso (de los terapeutas de contacto) para demostrar la afirmación fundamental de la TC (terapia de contacto) es evidencia irrefutable de que las aseveraciones de la TC (terapia de contacto) no tienen fundamento y de que su uso profesional no está justificado”. Con base en los resultados del proyecto de Emily Rosa para la feria de ciencias, parece que la terapia de contacto no es efectiva.

Fundamentos del margen de error Puesto que la distribución de proporciones muestrales es aproximadamente normal (ya que ambas condiciones $np \geq 5$ y $nq \geq 5$ se satisfacen), podemos utilizar los resultados de la sección 6-6 para concluir que μ y σ están dadas por $\mu = np$ y $\sigma = \sqrt{npq}$. Estos dos parámetros pertenecen a n ensayos, pero los convertimos a una base por ensayo dividiendo entre n como sigue:

$$\text{Medida de proporciones muestrales: } \mu = \frac{np}{n} = p$$

$$\text{Desviación estándar de proporciones muestrales: } \sigma = \frac{\sqrt{npq}}{n} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}$$

El primer resultado tal vez parezca trivial, puesto que ya estipulamos que la proporción real de la población es p . El segundo resultado no es trivial y resulta útil para describir el margen de error E , pero reemplazamos el producto pq por $\hat{p}\hat{q}$ porque no conocemos todavía el valor de p (ya que es el valor que tratamos de estimar). La fórmula 7-1 para el margen de error refleja el hecho de que \hat{p} tiene una probabilidad de $1 - \alpha$ de estar dentro de $z_{\alpha/2}\sqrt{pq/n}$ de p . El intervalo de confianza para p , como se dio previamente, refleja el hecho de que existe una probabilidad de $1 - \alpha$ de que \hat{p} difiera de p menos que el margen de error $E = z_{\alpha/2}\sqrt{pq/n}$.



Determinación del tamaño muestral

Suponga que queremos reunir datos muestrales con el objetivo de estimar alguna proporción de la población. ¿Cómo sabemos cuántos elementos muestrales deben obtenerse? Si tomamos la expresión para el margen de error E (fórmula 7-1), y luego despejamos n , obtenemos la fórmula 7-2, la cual requiere que \hat{p} sea un estimado de la proporción poblacional p ; pero si no se conoce un estimado como éste (como suele ser el caso), reemplazamos \hat{p} por 0.5 y reemplazamos \hat{q} por 0.5, con el resultado que se da en la fórmula 7-3.

Tamaño muestral para la estimación de la proporción p

$$\text{Cuando se conoce un estimado } \hat{p}: \quad \text{Fórmula 7-2} \quad n = \frac{[z_{\alpha/2}]^2 \hat{p} \hat{q}}{E^2}$$

$$\text{Cuando se desconoce el estimado } \hat{p}: \quad \text{Fórmula 7-3} \quad n = \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2}$$

Regla de redondeo para determinar el tamaño muestral

Para asegurar que el tamaño muestral requerido sea al menos tan grande como debe ser, si el tamaño muestral calculado no es un número entero, redondee al siguiente número entero *mayor*.

Utilice la fórmula 7-2 cuando se puedan hacer estimaciones razonables de \hat{p} usando muestras previas, un estudio piloto o el conocimiento experto de alguna persona. Cuando esto no sea posible, utilice la fórmula 7-3. Observe que las fórmulas 7-2 y 7-3 no incluyen el tamaño de la población N , de manera que el tamaño de la población es irrelevante. (*Excepción:* Cuando el muestreo es sin reemplazo y procede de una población finita relativamente pequeña. Véase el ejercicio 49).

EJEMPLO Tamaño muestral para una encuesta por correo electrónico

Las formas en las que nos comunicamos se han visto afectadas drásticamente por el uso de máquinas contestadoras telefónicas, máquinas de fax, correo de voz y correo electrónico. Suponga que un sociólogo quiere determinar el porcentaje actual de hogares en Estados Unidos que utilizan el correo electrónico. ¿Cuántos hogares deben encuestarse para tener una confianza del 95% de que el porcentaje muestral es erróneo por no más de 4 puntos porcentuales?

- Utilice el siguiente resultado de un estudio pionero: en 1997, el 16.9% de los hogares estadounidenses usaban correo electrónico (según datos de *The World Almanac and Book of Facts*).
- Suponga que no tenemos información previa que sugiera un posible valor de \hat{p} .

SOLUCIÓN

- El estudio previo sugiere que $\hat{p} = 0.169$, entonces $\hat{q} = 0.831$ (calculado de $\hat{q} = 1 - 0.169$). Con un nivel de confianza del 95%, tenemos $\alpha = 0.05$, entonces $z_{\alpha/2} = 1.96$. Además, el margen de error es $E = 0.04$ (el equivalente decimal de “cuatro puntos porcentuales”). Puesto que tenemos un valor estimado de \hat{p} , usamos la fórmula 7-2 como sigue:

$$\begin{aligned} n &= \frac{[z_{\alpha/2}]^2 \hat{p} \hat{q}}{E^2} = \frac{[1.96]^2 (0.169)(0.831)}{0.04^2} \\ &= 337.194 = 338 \quad (\text{redondeado}) \end{aligned}$$

Debemos encuestar al menos 338 hogares seleccionados al azar.

- Como en el inciso a), nuevamente utilizamos $z_{\alpha/2} = 1.96$ y $E = 0.04$, pero sin conocimiento previo de \hat{p} (o \hat{q}), usamos la fórmula 7-3 como sigue.

$$\begin{aligned} n &= \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2} = \frac{[1.96]^2 \cdot 0.25}{0.04^2} \\ &= 600.25 = 601 \quad (\text{redondeado}) \end{aligned}$$

INTERPRETACIÓN Para tener una confianza del 95% de que nuestro porcentaje muestral está dentro de cuatro puntos porcentuales del porcentaje verdadero para todos los hogares, debemos seleccionar al azar y encuestar 601 hogares. Comparando este resultado con el tamaño muestral de 338 calculado en el inciso a), podemos ver que si no tenemos conocimiento de un estudio previo, se requiere

continúa

una muestra más grande para obtener los mismos resultados que cuando se puede estimar el valor de \hat{p} . Pero ahora recurramos al sentido común: sabemos que el uso del correo electrónico está creciendo tan rápidamente que el estimado de 1997 es muy antiguo para ser de utilidad. En la actualidad, mucho más del 16.9% de los hogares estadounidenses utilizan correo electrónico. Siendo realistas, necesitamos una muestra mayor que 338 hogares. Suponiendo que en realidad no conocemos la tasa actual de uso de correo electrónico, deberíamos seleccionar al azar 601 hogares. Con 601 hogares, tendremos una confianza del 95% de que estamos dentro de cuatro puntos porcentuales del porcentaje verdadero de hogares que usan correo electrónico.

Errores comunes Cuando calcule el tamaño muestral, trate de evitar los siguientes dos errores comunes: **1.** No cometa el error de utilizar $E = 4$ como el margen de error correspondiente a “cuatro puntos porcentuales”. **2.** Asegúrese de sustituir la puntuación z crítica por $z_{\alpha/2}$. Por ejemplo, si usted trabaja con una confianza del 95%, asegúrese de reemplazar $z_{\alpha/2}$ por 1.96. No cometa el error de reemplazar $z_{\alpha/2}$ por 0.95 o 0.05.

Tamaño de la población Muchas personas creen de manera errónea que el tamaño de la muestra debe ser algún porcentaje de la población; sin embargo, la fórmula 7-3 indica que el tamaño de la población es irrelevante. (En realidad, el tamaño de la población se utiliza algunas veces, pero sólo en casos en los que hacemos un muestreo sin reemplazo de una población relativamente pequeña. Véase el ejercicio 49). Las encuestas suelen utilizar tamaños muestrales en el rango de 1000 a 2000 y, aunque encuestas como éstas quizás implican un porcentaje muy pequeño de la población total, logran generar resultados bastante buenos.

Cálculo del estimado puntual y de E desde un intervalo de confianza Algunas veces queremos comprender mejor un intervalo de confianza que podría haberse obtenido de un artículo de una revista, o que podría haberse generado por medio de programas de cómputo o una calculadora. Si ya conocemos los límites del intervalo de confianza, la proporción muestral \hat{p} y el margen de error E se calculan como sigue:

Estimado puntual de p :

$$\hat{p} = \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2}$$

Margen de error:

$$E = \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2}$$

EJEMPLO El artículo “High-Dose Nicotine Patch Therapy” de Dale, Hurt *et al.* (*Journal of the American Medical Association*, vol. 274, núm. 17) incluye esta afirmación: “De los 71 sujetos, el 70% se abstuvo de fumar durante 8 semanas (intervalo de confianza [IC] del 95%, del 58% al 81%)”. Utilice esta afirmación para calcular el estimado puntual \hat{p} y el margen de error E .

SOLUCIÓN De la afirmación dada, vemos que el intervalo de confianza del 95% es $0.58 < p < 0.81$. El estimado puntual \hat{p} es el valor medio entre los límites superior e inferior del intervalo de confianza, de manera que obtenemos

$$\hat{p} = \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2}$$

$$= \frac{0.81 + 0.58}{2} = 0.695$$

El margen de error se calcula como sigue:

$$E = \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2}$$

$$= \frac{0.81 - 0.58}{2} = 0.115$$

Intervalos de confianza con el mejor desempeño

Nota importante: Los ejercicios de esta sección se basan en el intervalo de confianza que describimos antes y no en los intervalos de confianza que se describen a continuación.

El intervalo de confianza descrito en esta sección tiene el formato que generalmente se presenta en los cursos de introducción a la estadística, pero no tiene un desempeño tan bueno como otros intervalos de confianza. El *intervalo de confianza Wald ajustado* se desempeña mejor en el sentido de que su probabilidad de contener la verdadera proporción de población p es más cercana al nivel de confianza que se utiliza. El intervalo de confianza Wald ajustado emplea el siguiente procedimiento sencillo: se suma 2 al número de éxitos x , se suma 2 al número de fracasos (de manera que el número de ensayos n aumenta en 4), y luego se calcula el intervalo de confianza de la misma forma descrita en esta sección. Por ejemplo, si utilizamos los métodos de esta sección con $x = 10$ y $n = 20$, obtenemos el siguiente intervalo de confianza del 95%: $0.281 < p < 0.719$. Con $x = 10$ y $n = 20$ aplicamos el intervalo de confianza Wald ajustado permitiendo que $x = 12$ y $n = 24$ para obtener el siguiente intervalo de confianza: $0.300 < p < 0.700$. La probabilidad de que el intervalo de confianza $0.300 < p < 0.700$ incluya a p se acerca más al 95% que la probabilidad de que $0.281 < p < 0.719$ incluya a p .

Otro intervalo de confianza que tiene mejor desempeño que el descrito en esta sección y que el intervalo de confianza Wald ajustado es el *intervalo de confianza de la puntuación de Wilson*. Este intervalo posee el límite inferior del intervalo de confianza que se indica abajo, y el límite superior del intervalo de confianza se expresa al cambiar el signo menos por un signo más. (Es fácil ver por qué este enfoque no se utiliza mucho en los cursos de introducción a la estadística). Utilizando $x = 10$ y $n = 20$, el intervalo de confianza de la puntuación de Wilson es $0.299 < p < 0.701$.

$$\frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} - z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z_{\alpha/2}^2}{4n}}}{1 + \frac{z_{\alpha/2}^2}{n}}$$

Para revisar una explicación sobre estos y otros intervalos de confianza para p , véase “Approximation is Better than ‘Exact’ for Interval Estimation of Binomial Proportions”, de Agresti y Coull, *American Statistician*, vol. 52, núm. 2.

Uso de la tecnología para intervalos de confianza

STATDISK Seleccione **Analysis**, luego **Confidence Intervals**, después **Proportion One Sample** y proceda a ingresar los elementos que se le piden.

MINITAB Seleccione **Stat**, **Basic Statistics**, luego **1 Proportion**. En el cuadro de diálogo, haga clic en el botón **Summarized Data**. También haga clic en el botón de **Options**, ingrese el nivel de confianza deseado (el predeterminado es del 95%). En vez de utilizar una aproximación normal, el procedimiento predeterminado de Minitab consiste en determinar los límites del intervalo de confianza por medio de un método exacto.

Para usar el método de aproximación normal presentado en esta sección, haga clic en **Options** y luego en el cuadro con la frase: “Use test and interval based on normal distribution”.

EXCEL Utilice Data Desk XL, que es un complemento de este libro. Primero ingrese el número de éxitos en la celda A1, después ingrese el número total de ensayos en la celda B1. Haga clic en **DDXL** y seleccione **Confidence Intervals**, luego seleccione **Summ 1 Var Prop Interval** (que es la forma abreviada de “intervalo de confianza para una proporción utilizando datos resumidos para una variable”). Haga clic en el ícono que muestra un lápiz para “Num successes” e ingrese A1. Haga clic en el ícono con forma de lápiz para “Num trials” e ingrese B1. Haga clic en **OK**. En el cuadro de

diálogo, seleccione el nivel de confianza y luego haga clic en **Compute Interval**.

TI-83/84 PLUS Oprima **STAT**, seleccione **TESTS**, luego seleccione **1-Prop-ZInt** y proceda a ingresar los elementos que se piden. La siguiente pantalla corresponde al resultado del ejemplo del intervalo de confianza presentado en esta sección.

TI-83/84 Plus

1-PropZInt
(.38115, .49742)
 $\hat{p} = .4392857143$
 $n = 280$

Uso de la tecnología para la determinación del tamaño muestral

STATDISK Seleccione **Analysis**, después **Sample Size Determination**, luego

Estimate Proportion. Proceda a ingresar los elementos que se le piden en el cuadro de diálogo.

La determinación del tamaño muestral no está disponible como una función incluida en Minitab, Excel o la calculadora TI-83/84 Plus.



7-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Valor crítico.** ¿Qué es un valor crítico en una distribución normal?
- 2. Margen de error.** ¿Qué es un margen de error?
- 3. Intervalo de confianza.** Al entrevistar a 500 personas, obtenemos 200 respuestas afirmativas a una pregunta específica, de manera que se estima que la proporción de respuestas afirmativas de toda la población es de 0.4. Dado que tenemos el valor estimado de 0.4, ¿por qué necesitaríamos un intervalo de confianza? Es decir, ¿qué información adicional brinda el intervalo de confianza?
- 4. Muestreo.** Un estudiante encuesta a 100 compañeros de clase y les pregunta si tienen deudas pendientes. Después de calcular la proporción muestral de esta muestra de $n = 100$ sujetos, ¿se pueden utilizar los métodos de esta sección para estimar la proporción de todos los adultos que tienen deudas pendientes? ¿Por qué?

Cálculo de valores críticos. En los ejercicios 5 a 8, calcule el valor crítico $z_{\alpha/2}$ que corresponde al nivel de confianza dado.

5. 99% 6. 90%
7. 98% 8. 99.5%
9. Exprese el intervalo de confianza $0.220 < p < 0.444$ en la forma de $\hat{p} \pm E$.
10. Exprese el intervalo de confianza $0.600 < p < 0.800$ en la forma de $\hat{p} \pm E$.
11. Exprese el intervalo de confianza $(0.206, 0.286)$ en la forma de $\hat{p} \pm E$.
12. Exprese el intervalo de confianza 0.337 ± 0.050 en la forma de $\hat{p} - E < p < \hat{p} + E$.

Interpretación de los límites del intervalo de confianza. En los ejercicios 13 a 16, utilice los límites de intervalo de confianza dados para calcular el estimado puntual \hat{p} y el margen de error E .

13. $(0.868, 0.890)$ 14. $0.325 < p < 0.375$
15. $0.607 < p < 0.713$ 16. $0.0144 < p < 0.0882$

Cálculo del margen de error. En los ejercicios 17 a 20, suponga que una muestra se utiliza para estimar una proporción poblacional p . Calcule el margen de error E que corresponde al estadístico y al nivel de confianza dados.

17. $n = 500, x = 200$, 95% de confianza
18. $n = 1200, x = 800$, 99% de confianza
19. 98% de confianza; el tamaño muestral es 1068, de los cuales el 25% son éxitos.
20. 90% de confianza; el tamaño muestral es 2107, de los cuales el 65% son éxitos.

Construcción de intervalos de confianza. En los ejercicios 21 a 24, use los datos muestrales y el nivel de confianza para construir el intervalo de confianza estimado de la proporción poblacional p .

21. $n = 500, x = 200$, 95% de confianza
22. $n = 1200, x = 800$, 99% de confianza
23. $n = 1068, x = 267$, 98% de confianza
24. $n = 4500, x = 2925$, 90% de confianza

Determinación del tamaño muestral. En los ejercicios 25 a 28, utilice los datos para calcular el tamaño muestral mínimo requerido para estimar una proporción o porcentaje de una población.

25. Margen de error: 0.020; nivel de confianza: 95%; \hat{p} y \hat{q} desconocidas
26. Margen de error: 0.050; nivel de confianza: 99%; \hat{p} y \hat{q} desconocidas
27. Margen de error: tres puntos porcentuales; nivel de confianza: 95%; de un estudio previo, \hat{p} se estima por el equivalente decimal del 27%.
28. Margen de error: cinco puntos porcentuales; nivel de confianza: 90%; de un estudio previo, \hat{p} se estima por el equivalente decimal del 65%.
29. **Selección del género.** El Genetics and IVF Institute realizó una prueba clínica del método XSORT, diseñado para incrementar la probabilidad de concebir una niña. Para cuando se escribía este libro, ya habían nacido 325 bebés de padres que utilizaron el método XSORT y 295 de ellos eran niñas. Utilice los datos muestrales para construir un estimado del intervalo de confianza del 99% del porcentaje de niñas nacidas de padres que utilizaron el método XSORT. Con base en el resultado, ¿parece que el método XSORT es efectivo? ¿Por qué?

- 30. Selección del género.** El Genetics and IVF Institute realizó una prueba clínica del método YSORT, diseñado para incrementar la probabilidad de concebir un hijo varón. Para cuando se escribía este libro, ya habían nacido 51 bebés de padres que utilizaron el método YSORT y 39 eran varones. Utilice los datos muestrales para construir un estimado del intervalo de confianza del 99% del porcentaje de niños nacidos de padres que utilizaron el método YSORT. Con base en el resultado, ¿parece que el método YSORT es efectivo? ¿Por qué?
- 31. Posposición de la muerte.** Una hipótesis interesante y del dominio público dice que los individuos pueden posponer temporalmente su muerte para estar presentes en una fiesta o en un suceso importante como un cumpleaños. En un estudio de este fenómeno, se descubrió que la semana previa y la semana posterior al Día de Acción de Gracias hubo un total de 12,000 muertes, y que 6062 de ellas ocurrieron la semana anterior al Día de Acción de Gracias (según datos de "Holidays, Birthdays, and Postponement of Cancer Death", de Young y Hade, *Journal of the American Medical Association*, vol. 292, núm. 24). Construya un estimado del intervalo de confianza del 95% para la proporción de muertes ocurridas la semana anterior al Día de Acción de Gracias, con respecto a las muertes totales durante la semana previa y la semana posterior al Día de Acción de Gracias. Con base en el resultado, ¿existe algún indicador de que la gente pueda posponer temporalmente su muerte para estar presente el Día de Acción de Gracias? ¿Por qué?
- 32. Negligencia médica.** Un problema importante que enfrentan los estadounidenses es la gran cantidad de demandas por negligencia médica y los gastos que éstas generan. En un estudio de 1228 demandas por negligencia médica elegidas al azar, se descubrió que 856 de ellas se retiraron o se rechazaron posteriormente (según datos de la Physician Insurers Association of America). Construya un estimado del intervalo de confianza del 99% de la proporción de demandas por negligencia médica que fueron retiradas o rechazadas. Al parecer, ¿la mayoría de estas demandas se retiraron o se rechazaron?
- 33. Genética mendeliana.** Cuando Mendel realizó sus famosos experimentos genéticos con guisantes, una muestra de vástagos consistió en 428 plantas de guisantes verdes y 152 de guisantes amarillos.
- Calcule un estimado del intervalo de confianza del 95% del porcentaje de plantas de guisantes amarillos.
 - Con base en su teoría genética, Mendel esperaba que el 25% de los vástagos dieran guisantes amarillos. Dado que el porcentaje de vástagos de guisantes amarillos no es el 25%, ¿contradicen los resultados la teoría de Mendel? ¿Por qué?
- 34. Respuestas de encuesta confusas.** En una encuesta de 1002 personas, 701 dijeron que habían votado en una elección presidencial reciente (según datos del ICR Research Group). Los registros de votos mostraron que el 61% de las personas con derecho a voto realmente votaron.
- Calcule un estimado del intervalo de confianza del 99% de la proporción de personas que dijeron haber votado.
 - ¿Son consistentes los resultados de encuesta con los votos reales del 61%? ¿Por qué?
- 35. Teléfonos celulares y cáncer.** Un estudio de 420,095 daneses usuarios de teléfono celular encontró que 135 de ellos desarrollaron cáncer cerebral o del sistema nervioso. Antes de este estudio del uso de teléfono celular, se encontró que la tasa de ese tipo de cáncer era de 0.0340% para aquellos que no usan teléfonos celulares. Los datos son del *Journal of the National Cancer Institute*.
- Utilice los datos muestrales para construir un estimado del intervalo de confianza del 95% para el porcentaje de usuarios de teléfono celular que desarrollan cáncer del cerebro o del sistema nervioso.
 - ¿Parecen tener los usuarios de teléfono celular una tasa de cáncer cerebral o del sistema nervioso diferente de la tasa de cáncer de este tipo entre aquellos que no usan teléfonos celulares? ¿Por qué?

36. Encuesta sobre clonación. Una reciente encuesta Gallup incluyó a 1012 adultos elegidos al azar, a quienes se preguntó “si se debe permitir o no la clonación de seres humanos”. Los resultados mostraron que 901 de los encuestados dijeron que no debe permitirse la clonación. Un reportero de noticias desea determinar si estos resultados de encuesta constituyen una fuerte evidencia de que la mayoría de las personas (más del 50%) se oponen a este tipo de clonación. Construya un estimado del intervalo de confianza del 99% de la proporción de adultos que considera que no debe permitirse la clonación de seres humanos. Con base en el resultado, ¿existen fuertes evidencias que apoyen la afirmación de que la mayoría de la gente se opone a este tipo de clonación?

37. Sesgo en la selección de integrantes de jurado. En el caso *Castaneda contra Partida*, se descubrió que en el condado de Hidalgo, Texas, durante un periodo de 11 años 870 personas habían sido elegidas para formar parte del gran jurado, y que el 39% de ellas eran mexico-estadounidenses. Utilice los datos muestrales para construir un estimado del intervalo de confianza del 99% de la proporción de miembros del gran jurado que eran mexico-estadounidenses. Dado que, entre la población elegible para fungir como juez, el 79.1% estaba integrado por mexico-estadounidenses, ¿parece que el proceso de selección de los jueces tuvo algún sesgo en contra de los mexico-estadounidenses? ¿Por qué?

38. Detección de fraude. Cuando trabajaba para el fiscal de distrito de Brooklyn, el investigador Robert Burton analizó los primeros dígitos de montos de cheques propiedad de empresas sospechosas de fraude. De 784 cheques, en el 61% el primer dígito era 5. Construya un estimado del intervalo de confianza del 99% de la proporción de cheques en los que el primer dígito era 5. Cuando los cheques se emiten en el transcurso normal de transacciones honestas, se espera que el 7.9% de ellos tengan montos en los que el primer dígito sea 5. ¿Qué sugiere el intervalo de confianza?

39. Hogares con teléfono. En 1920 sólo el 35% de los hogares estadounidenses tenían teléfono, pero ahora la proporción es mucho más alta. Una encuesta reciente de 4276 hogares elegidos al azar reveló que 94% de ellos cuentan con teléfono (según datos del U.S. Census Bureau). Utilice los resultados de la encuesta y construya un estimado del intervalo de confianza del 99% de la proporción de hogares con teléfonos. Dado que la encuesta sólo incluye a 4276 hogares de un total de 115 millones, ¿tenemos realmente evidencia suficiente para decir que el porcentaje de hogares con teléfono en la actualidad es mayor que la tasa del 35% en 1920?

40. Compras por Internet. En una encuesta Gallup se entrevistó a 1025 adultos seleccionados al azar, y el 29% de ellos dijeron que utilizaban Internet para hacer compras al menos unas cuantas veces al año.

- Calcule el estimado puntual del *porcentaje* de adultos que utilizan Internet para hacer compras.
- Calcule un estimado del intervalo de confianza del 99% del *porcentaje* de adultos que utilizan Internet para hacer compras.
- Si una tienda al detalle tradicional desea estimar el porcentaje de adultos que compran por Internet para determinar el impacto máximo de los compradores por Internet sobre sus ventas, ¿qué porcentaje de personas que compran por Internet debe considerarse?

Determinación del tamaño de la muestra. En los ejercicios 41 a 44, calcule el tamaño de muestra mínimo requerido para estimar una proporción o un porcentaje poblacional.

41. Tamaño de muestra para compras en Internet. Muchos estados están considerando con cuidado los pasos que les ayudarían a recolectar impuestos por ventas en artículos comprados a través de Internet. ¿Cuántas transacciones de ventas seleccionadas al azar deben registrarse para determinar el porcentaje que se lleva a cabo a través de Internet? Suponga que queremos tener una confianza del 99% de que el porcentaje muestral está dentro de dos puntos porcentuales del porcentaje real de la población para todas las transacciones de ventas.

- 42. Tamaño de muestra para descarga de canciones.** La industria de la música debe adaptarse a la creciente práctica de los consumidores de descargar canciones en vez de comprar CD. Por eso es importante estimar la proporción de canciones que se descargan actualmente. ¿Cuántas adquisiciones de canciones elegidas al azar se deben considerar para determinar el porcentaje que se obtuvo a través de descargas? Suponga que deseamos tener una confianza del 95% de que el porcentaje de la muestra está dentro de un punto porcentual del porcentaje real de la población de canciones que se descargan.
- 43. Nitrógeno en neumáticos.** Una campaña reciente se diseñó para convencer a los propietarios de automóviles de que debían llenar sus neumáticos con nitrógeno en vez de aire. A un costo de aproximadamente \$5 por neumático, se supone que el nitrógeno tiene la ventaja de escaparse con una rapidez mucho menor que el aire, por lo que la presión ideal se puede mantener de manera más consistente. Antes de gastar grandes sumas en la publicidad del nitrógeno, sería pertinente realizar una encuesta para determinar el porcentaje de propietarios de automóviles que estarían dispuestos a pagar por el nitrógeno. ¿Cuántos propietarios de automóviles elegidos al azar se deben encuestar? Suponga que deseamos tener una confianza del 98% de que el porcentaje de la muestra está dentro de tres puntos porcentuales del porcentaje real de todos los dueños de automóviles que estarían dispuestos a pagar por el nitrógeno.
- 44. Techo corredizo y bolsas de aire laterales.** Toyota ofrece la opción de un paquete de techo corredizo y bolsas de aire laterales para su modelo Corolla. Este paquete cuesta \$1400 (el precio facturado es de \$1159). Suponga que antes de ofrecer la opción de este paquete, Toyota desea determinar el porcentaje de compradores del Corolla que estarían dispuestos a pagar \$1400 adicionales por el techo corredizo y las bolsas laterales. ¿Cuántos compradores del Corolla se deben encuestar si deseamos tener una confianza del 95% de que el porcentaje muestral está dentro de cuatro puntos porcentuales del porcentaje real de todos los compradores del Corolla?

Uso de conjuntos de datos del apéndice B. En los ejercicios 45 a 48 utilice el conjunto de datos indicado del apéndice B.

- 45. Dulces M&M azules.** Remítase al conjunto de datos 13 del apéndice B y calcule la proporción muestral de dulces M&M que son azules. Use este resultado para crear un estimado del intervalo de confianza del 95% del porcentaje de la población de dulces M&M azules. ¿Es congruente el resultado con la tasa del 24% que reporta Mars, el fabricante de dulces?
- 46. Consumo de alcohol y tabaco en películas para niños.** Remítase al conjunto de datos 5 del apéndice B.
- Construya un estimado del intervalo de confianza del 95% para el porcentaje de películas infantiles de dibujos animados que presentan consumo de tabaco.
 - Construya un estimado del intervalo de confianza del 95% para el porcentaje de películas infantiles de dibujos animados que presentan consumo de alcohol.
 - Compare los resultados anteriores. ¿Aparecen en un alto porcentaje el tabaco o el alcohol en las películas de dibujos animados infantiles?
 - Utilizando los resultados de los incisos a) y b) como medidas de la descripción de hábitos no saludables, ¿qué característica importante de los datos no se incluye?
- 47. Precipitación en Boston.** Remítase al conjunto de datos 10 del apéndice B y considere que los días con valores de precipitación diferentes de 0 son días con precipitación. Construya un estimado del intervalo de confianza del 95% para la proporción de miércoles con precipitación y también construya un estimado del intervalo de confianza del 95% para la proporción de domingos con precipitación. Compare los resultados. ¿Hay mayor precipitación en alguno de esos días?
- 48. Exactitud de temperaturas pronosticadas.** Remítase al conjunto de datos 8 del apéndice B. Construya un estimado del intervalo de confianza del 95% para la proporción de días con una temperatura máxima real que difiera en más de 2° de la temperatura máxi-

ma pronosticada un día antes. Luego, construya un estimado del intervalo de confianza del 95% para la proporción de días con una temperatura máxima real que difiera en más de 2° de la temperatura máxima pronosticada cinco días antes. Compare los resultados.

7-2 MÁS ALLÁ DE LO BÁSICO

- 49. Uso del factor de corrección para una población finita.** En esta sección se presentaron las fórmulas 7-2 y 7-3, que se utilizan para determinar el tamaño muestral. En ambos casos se supuso que la población es infinita o muy grande, y que se realiza un muestreo con reemplazo. Cuando tenemos una población relativamente pequeña, con tamaño N , y el muestreo se hace sin reemplazo, modificamos E para incluir el *factor de corrección por población finita* que se presenta aquí y despejamos n para obtener así el resultado que se da a continuación. Utilice este resultado para repetir el ejercicio 44, suponiendo que limitamos nuestra población a 1250 compradores del Corolla de Toyota en una región.

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n} \sqrt{\frac{N-n}{N-1}}} \quad n = \frac{N\hat{p}\hat{q}[z_{\alpha/2}]^2}{\hat{p}\hat{q}[z_{\alpha/2}]^2 + (N-1)E^2}$$

- 50. Intervalo de confianza unilateral.** Un *intervalo de confianza unilateral* para p se expresa como $p < \hat{p} + E$ o $p > \hat{p} - E$, donde el margen de error E se modifica reemplazando $z_{\alpha/2}$ por z_{α} . Si Air America quiere reportar un rendimiento de puntualidad de al menos x por ciento con un 95% de confianza, construya el intervalo de confianza unilateral apropiado y luego calcule el porcentaje en cuestión. Suponga que una muestra aleatoria simple de 750 vuelos incluye 630 que son puntuales.

- 51. Intervalo de confianza de una muestra pequeña.** Existen tablas especiales disponibles para encontrar intervalos de confianza para proporciones que implican números pequeños de casos, cuando no se puede usar la aproximación por distribución normal. Por ejemplo, dado $x = 3$ éxitos en $n = 8$ ensayos, el intervalo de confianza del 95% que se encuentra en *Standard Probability and Statistics Tables and Formulae* (CRC Press) es $0.085 < p < 0.755$. Calcule el intervalo de confianza que resultaría si usted usara la distribución normal erróneamente como una aproximación de la distribución binomial. ¿Los resultados son razonablemente cercanos?

- 52. Interpretación de límites de intervalo de confianza.** Suponga que se modifica una moneda para que favorezca las caras, y que de 100 lanzamientos se obtienen 95 caras. Calcule el estimado del intervalo de confianza del 99% para la proporción de caras que se obtendrán con esta moneda. ¿Qué es poco común en los resultados obtenidos a través de los métodos de esta sección? ¿Sugiere el sentido común una modificación del intervalo de confianza resultante?

- 53. Regla de tres.** Suponga que en n ensayos de un experimento binomial no hay ningún éxito. De acuerdo con la *Regla de tres*, tenemos un 95% de confianza de que la proporción real de la población tiene una frontera superior de $3/n$. (Véase “A Look at the Rule of Three”, de Jovanovic y Levy, *American Statistician*, vol. 51, núm. 2).

- a. Si en n ensayos independientes no se obtiene ningún éxito, ¿por qué no podemos calcular los límites del intervalo de confianza usando los métodos que se describen en esta sección?
- b. Si se trata a 20 pacientes con un fármaco y no hay reacciones adversas, ¿cuál es la frontera superior del 95% para p , la proporción de todos los pacientes que experimentan reacciones adversas a este fármaco?

- 54. Exactitud de encuesta.** Un artículo del *New York Times* acerca de los resultados de una encuesta afirma: “En teoría, en 19 casos de 20, los resultados de una encuesta como ésta no deben diferir por más de un punto porcentual en cualquier dirección de lo que podría obtenerse entrevistando a todos los votantes de Estados Unidos”. Calcule el tamaño muestral sugerido por esta afirmación.

7-3 Estimación de una media de población: σ conocida

Concepto clave En la sección 7-2 presentamos el estimado puntual y el intervalo de confianza como herramientas para el uso de una proporción muestral con el fin de estimar una proporción poblacional; esta sección presenta métodos para utilizar datos muestrales y calcular un estimado puntual y un estimado del intervalo de confianza de una media poblacional. Un requisito fundamental en esta sección es que, además de tener datos muestrales, también conocemos σ , la desviación estándar de la población. También presentamos un método para calcular el tamaño muestral que se necesitaría para estimar una media poblacional.

Requisitos

1. La muestra es aleatoria simple. (Todas las muestras del mismo tamaño tienen la misma probabilidad de ser seleccionadas).
2. El valor de la desviación estándar poblacional σ es conocido.
3. Cualquiera o ambas de estas condiciones se satisfacen: la población está normalmente distribuida o $n > 30$.

Conocimiento de σ Los requisitos anteriores incluyen el conocimiento de la desviación estándar poblacional σ , pero la siguiente sección presenta métodos para estimar una media poblacional sin conocer el valor de σ .

Requisito de normalidad Los requisitos incluyen la propiedad de que la población se distribuya normalmente o que $n > 30$. Si $n \leq 30$, la población no necesita tener una distribución exactamente normal, sino aproximadamente normal. Podemos considerar que el requisito de normalidad se satisface si no hay valores extremos y si un histograma de los datos muestrales no se aleja mucho de la forma de campana. (Se dice que los métodos de esta sección son *robustos*, es decir, no se ven muy afectados si los datos se alejan de la normalidad, siempre y cuando no se alejen demasiado).

Requisitos del tamaño muestral Esta sección utiliza la distribución normal como la distribución de medias muestrales. Si la población original no está distribuida normalmente, entonces decimos que las medias de muestras con tamaño $n > 30$ tienen una distribución que puede aproximarse a una distribución normal. La condición de que el tamaño muestral sea $n > 30$ se usa por lo regular como directriz, pero no es posible identificar un tamaño muestral mínimo específico que sea suficiente para todos los casos. El tamaño muestral mínimo realmente depende de cuánto se desvía la distribución de la población de una distribución normal. Tamaños muestrales de 15 a 30 son adecuados si la población parece tener una distribución que no es lejana a la normal, pero algunas otras poblaciones tienen distribuciones que son extremadamente diferentes de la normal y pueden necesitarse tamaños muestrales de 50, 100 o más. Usaremos el criterio simplificado de $n > 30$ como justificación para el tratamiento de la distribución de medias muestrales como una distribución normal.

En la sección 7-2 vimos que la proporción muestral \hat{p} es el mejor estimado puntual de la proporción poblacional p . Por razones similares, la media muestral \bar{x} es el mejor estimado puntual de la media poblacional μ .

La media muestral \bar{x} es el mejor estimado puntual de la media de la población.

Por lo regular la media muestral \bar{x} brinda el mejor estimado, por las siguientes dos razones:

1. Para todas las poblaciones, la media muestral \bar{x} es un **estimador sin sesgo** de la media poblacional μ , lo que significa que la distribución de medias muestrales tiende a concentrarse alrededor del valor de la media poblacional μ . [Es decir, las medias muestrales no tienden sistemáticamente a sobreestimar el valor de μ , ni tienden sistemáticamente a subestimar el valor de μ , sino que tienden a coincidir con este valor (como se ilustra en la sección 6-4).]
2. Para muchas poblaciones, la distribución de las medias muestrales \bar{x} tiende a ser más consistente (con *menos variación*) que la distribución de otros estadísticos muestrales.

EJEMPLO Pulso cardiaco de mujeres El pulso cardiaco de las personas es sumamente importante. Sin él, ¿dónde estaríamos? El conjunto de datos 1 del apéndice B incluye pulsos cardíacos (en latidos por minuto) de mujeres seleccionadas al azar; los estadísticos son los siguientes: $n = 40$, $\bar{x} = 76.3$ y $s = 12.5$. Utilice esta muestra para calcular el mejor estimado puntual de la media poblacional μ de los pulsos cardíacos de todas las mujeres.

SOLUCIÓN Para los datos muestrales, $\bar{x} = 76.3$. Como la media muestral \bar{x} es el mejor estimado puntual de la media poblacional μ , concluimos que el mejor estimado puntual de los pulsos cardíacos de todas las mujeres es 76.3.

Intervalos de confianza

Aunque un estimado puntual es el *mejor* valor individual para estimar un parámetro poblacional, no nos da ninguna indicación precisa de qué tan *bueno* es este mejor estimado. Sin embargo, un intervalo de confianza nos ofrece información que nos permite comprender mejor la exactitud del estimado. El intervalo de confianza se asocia con un nivel de confianza, como 0.95 (o 95%). El nivel de confianza nos da la tasa de éxitos del procedimiento que se utiliza para construir el intervalo de confianza. Como se describió en la sección 7-2, α es el complemento del nivel de confianza. Para un nivel de confianza de 0.95 (o 95%), $\alpha = 0.05$. Para un nivel de confianza de 0.99 (o 99%), $\alpha = 0.01$.

Margen de error Cuando reunimos un conjunto de datos muestrales, como los datos de los 40 pulsos de mujeres que se incluyen en el conjunto de datos 1 del apéndice B, podemos calcular la media muestral \bar{x} y esa media muestral por lo regular es diferente de la media poblacional μ . La diferencia entre la media muestral y la media poblacional es un error. En la sección 6-5 vimos que σ/\sqrt{n} es la desviación estándar de las medias muestrales. Utilizando σ/\sqrt{n} y la notación $z_{\alpha/2}$ que se presentó en la sección 7-2, ahora podemos usar el *margen de error E* que se expresa como sigue:

$$\text{Fórmula 7-4} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad \text{margen de error para la media (con base en } \sigma \text{ conocida)}$$

La fórmula 7-4 refleja el hecho de que la distribución del muestreo de la media muestral \bar{x} es *exactamente* una distribución normal con media μ y desviación



Estimación del tamaño de las poblaciones silvestres

La Ley Nacional de Administración de los Bosques de Estados Unidos protege a las especies en peligro de extinción, entre las que destaca el búho moteado del norte; la ley impidió que la industria silvícola talara vastas regiones de árboles en el noroeste del Pacífico. Se pidió a biólogos y a especialistas en estadística que analizaran el problema, y ellos concluyeron que estaban disminuyendo las tasas de supervivencia y los tamaños de las poblaciones de los búhos hembra, quienes desempeñan un papel importante en la supervivencia de la especie. Los biólogos y especialistas en estadística también estudiaron el salmón en los ríos Snake y Columbia del estado de Washington, así como los pingüinos en Nueva Zelanda. En el artículo "Sampling Wildlife Populations" (*Chance*, vol. 9, núm. 2), los autores Brian Manly y Lyman McDonald comentan que, en estudios de esta clase, "los biólogos ganan habilidades de modelamiento, que son una característica distinta de la buena estadística. Por su parte, los especialistas en estadística aprenden a competir en la realidad de los problemas, ya que los biólogos los introducen en asuntos cruciales".



Los números de serie de tanques capturados revelan el tamaño de la población

Durante la Segunda Guerra Mundial, especialistas en espionaje del bando de los aliados querían determinar el número de tanques que Alemania estaba produciendo. Las técnicas de espionaje tradicionales produjeron resultados poco confiables, pero los especialistas en estadística obtuvieron estimaciones exactas al analizar los números de serie de los tanques capturados. Por ejemplo, los registros muestran que Alemania realmente produjo 271 tanques en junio de 1941. La estimación basada en los números de serie fue de 244, en tanto que los métodos de espionaje tradicionales dieron como resultado una estimación extrema de 1550. (Véase “An Empirical Approach to Economic Intelligence in World War II”, de Ruggles y Brodie, *Journal of the American Statistical Association*, vol. 42).

estándar σ/\sqrt{n} , siempre y cuando la población tenga una distribución normal con media μ y desviación estándar σ . Si la población no está distribuida normalmente, las muestras grandes producen medias muestrales con una distribución que *se aproxima* a la normal. (La fórmula 7-4 requiere del conocimiento de la desviación estándar poblacional σ , pero la sección 7-4 presentará un método para calcular el margen de error E cuando se desconoce σ).

Utilizando el margen de error E , ahora podemos identificar el intervalo de confianza para la media poblacional μ (si se satisfacen los requisitos de esta sección). Los tres formatos que suelen usarse para expresar el intervalo de confianza se presentan en el siguiente cuadro.

Estimación del intervalo de confianza de la media poblacional μ (con σ conocida)

$$\bar{x} - E < \mu < \bar{x} + E \quad \text{donde} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

o

$$\bar{x} \pm E$$

o

$$(\bar{x} - E, \bar{x} + E)$$

Definición

Los dos valores $\bar{x} - E$ y $\bar{x} + E$ se llaman **límites del intervalo de confianza**.

Procedimiento para construir un intervalo de confianza para μ (con σ conocida)

1. Verifique que los supuestos requeridos se satisfagan. (Tenemos una muestra aleatoria simple, σ es conocida, y la población parece estar distribuida normalmente o $n > 30$).
2. Remítase a la tabla A-2 y calcule el valor crítico $z_{\alpha/2}$ correspondiente al nivel de confianza deseado. (Por ejemplo, si el nivel de confianza es del 95%, el valor crítico es $z_{\alpha/2} = 1.96$).
3. Evalúe el margen de error $E = z_{\alpha/2} \cdot \sigma/\sqrt{n}$.
4. Utilizando el valor calculado del margen de error E y el valor de la media muestral \bar{x} , calcule los valores de $\bar{x} - E$ and $\bar{x} + E$. Sustituya esos valores en el formato general para el intervalo de confianza:

$$\bar{x} - E < \mu < \bar{x} + E$$

o

$$\bar{x} \pm E$$

o

$$(\bar{x} - E, \bar{x} + E)$$

5. Redondee los valores resultantes usando la siguiente regla de redondeo.

Regla de redondeo para intervalos de confianza utilizados para estimar μ

1. Cuando utilice el *conjunto de datos original* para construir un intervalo de confianza, redondee los límites del intervalo de confianza a un decimal más del que se usa para el conjunto de datos original.
2. Cuando el conjunto de datos original se desconoce y sólo se utiliza el *resumen de estadísticos* (n, \bar{x}, s), redondee los límites del intervalo de confianza al mismo número de espacios decimales utilizados para la media muestral.

Interpretación de un intervalo de confianza Al igual que en la sección 7-2, debemos ser cuidadosos para interpretar correctamente los intervalos de confianza. Después de obtener un estimado del intervalo de confianza de la media poblacional μ , como un intervalo de confianza del 95% de $72.4 < \mu < 80.2$, existe una interpretación correcta y muchas interpretaciones erróneas.

Correcta: “Tenemos una confianza del 95% de que el intervalo de 72.4 a 80.2 realmente contiene el valor verdadero de μ ”. Esto significa que si seleccionamos muchas muestras diferentes del mismo tamaño y construimos los intervalos de confianza correspondientes, a la larga, el 95% de éstos contendrían realmente el valor de μ . (Como en la sección 7-2, esta interpretación correcta se refiere a la tasa de éxitos del *proceso* que se usa para estimar la media poblacional).

Errónea: Puesto que μ es una constante fija, sería incorrecto decir que “existe un 95% de probabilidades de que μ se localice entre 72.4 y 80.2”. También sería incorrecto decir que “el 95% de todos los valores de los datos están entre 72.4 y 80.2” y que “el 95% de las medias muestrales caen entre 72.4 y 80.2”.

EJEMPLO Pulsos cardíacos de mujeres Para la muestra de pulsos cardíacos de mujeres en el conjunto de datos 1 del apéndice B, tenemos $n = 40$ y $\bar{x} = 76.3$, y la muestra es aleatoria simple. Suponga que sabemos que σ es 12.5. Utilice un nivel de confianza del 95% y calcule lo siguiente:

- a. El margen de error E .
- b. El intervalo de confianza para μ .

SOLUCIÓN

REQUISITO ✓ Primero debemos verificar que se cumplan los requisitos. La muestra es aleatoria simple. Se supone que conocemos el valor de σ (12.5). Con $n > 30$, se satisface el requisito de que “la población se distribuye normalmente o $n > 30$ ”. Por lo tanto, los requisitos se cumplen y podemos continuar con los métodos de esta sección. ✓

- a. El nivel de confianza del 0.95 implica que $\alpha = 0.05$, entonces $z_{\alpha/2} = 1.96$ (como se mostró en un ejemplo de la sección 7-2). El margen de error E se calcula usando la fórmula 7-4 como sigue. Los lugares decimales adicionales se usan para minimizar los errores de redondeo en el intervalo de confianza calculado en el inciso b).

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{12.5}{\sqrt{40}} = 3.8737901$$

continúa

- b. Con $\bar{x} = 76.3$ y $E = 3.8737901$, construimos el intervalo de confianza como sigue:

$$\begin{aligned}\bar{x} - E &< \mu < \bar{x} + E \\ 76.3 - 3.8737901 &< \mu < 76.3 + 3.8737901 \\ 72.4 &< \mu < 80.2 \quad (\text{redondeado a un decimal como en } \bar{x})\end{aligned}$$

INTERPRETACIÓN Este resultado también podría expresarse como 76.3 ± 3.9 o como $(72.4, 80.2)$. Con base en la muestra con $n = 40$, $\bar{x} = 76.3$, y suponiendo que σ es 12.5, el intervalo de confianza para la media de la población μ es $72.4 < \mu < 80.2$ y este intervalo tiene un nivel de confianza de 0.95. Esto significa que si seleccionamos muchas muestras diferentes de 40 mujeres y construimos los intervalos de confianza como lo hicimos aquí, el 95% de ellos incluirían realmente el valor de la media poblacional μ .

Fundamentos del intervalo de confianza La idea básica que subyace en la construcción de intervalos de confianza se relaciona con el teorema del límite central, que indica que si reunimos muestras aleatorias simples de una población distribuida normalmente, las medias muestrales se distribuyen de manera normal, con media μ y desviación estándar σ/\sqrt{n} . Si reunimos muestras aleatorias simples de tamaño $n > 30$ de cualquier población, la distribución de medias muestrales es aproximadamente normal, con media μ y desviación estándar σ/\sqrt{n} . El formato del intervalo de confianza es realmente una variación de la ecuación que ya se usó con el teorema del límite central. En la expresión $z = (\bar{x} - \mu_{\bar{x}})/\sigma_{\bar{x}}$, sustituya $\sigma_{\bar{x}}$ por σ/\sqrt{n} , sustituya $\mu_{\bar{x}}$ por μ , y entonces despeje μ para obtener

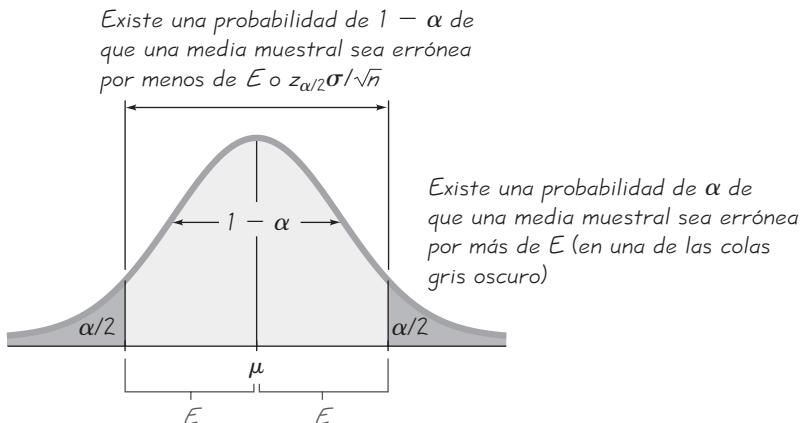
$$\mu = \bar{x} - z \frac{\sigma}{\sqrt{n}}$$

Se usan los valores positivo y negativo para los resultados de z en los límites del intervalo de confianza que estamos empleando.

Consideremos el caso específico de un nivel de confianza del 95%, de manera que $\alpha = 0.05$ y $z_{\alpha/2} = 1.96$. Para este caso, hay una probabilidad de 0.05 de que una media muestral esté a más de 1.96 desviaciones estándar (o $z_{\alpha/2} \cdot \sigma/\sqrt{n}$, lo que denotamos como E) de la media poblacional μ . Por otra parte, existe una probabilidad del 0.95 de que una media muestral esté dentro de 1.96 desviaciones estándar (o $z_{\alpha/2} \cdot \sigma/\sqrt{n}$) a partir de μ (véase la figura 7-4). Si la media muestral \bar{x} está dentro de $z_{\alpha/2} \cdot \sigma/\sqrt{n}$ de la media poblacional μ , entonces μ debe estar entre $\bar{x} - z_{\alpha/2} \cdot \sigma/\sqrt{n}$ y $\bar{x} + z_{\alpha/2} \cdot \sigma/\sqrt{n}$; esto se expresa en el formato general de nuestro intervalo de confianza (con $z_{\alpha/2} \cdot \sigma/\sqrt{n}$, denotado como E): $\bar{x} - E < \mu < \bar{x} + E$.

Determinación del tamaño muestral requerido para estimar μ

Ahora examinaremos esta importante pregunta: cuando planeamos reunir una muestra aleatoria simple de datos que se usarán para estimar una media poblacional μ , ¿cuántos valores muestrales deben obtenerse? Por ejemplo, suponga que queremos estimar el peso medio de pasajeros de líneas aéreas (un valor importante por razones de seguridad). ¿Cuántos pasajeros deben seleccionarse al azar y pesarse? La determinación del tamaño de una muestra aleatoria simple es un aspecto muy importante,

**Figura 7-4**

Distribución de medias muestrales con σ conocida

puesto que muestras que son innecesariamente grandes desperdician tiempo y dinero, en tanto que muestras muy pequeñas conducen a resultados deficientes.

Si empezamos con la expresión para el margen de error E (fórmula 7-4) y despejamos el tamaño muestral n , obtenemos lo siguiente.

Tamaño muestral para estimar la media μ

Fórmula 7-5

$$n = \left[\frac{z_{\alpha/2}\sigma}{E} \right]^2$$

donde $z_{\alpha/2}$ = puntuación z crítica basada en el nivel de confianza deseado
 E = margen de error deseado
 σ = desviación estándar poblacional

La fórmula 7-5 es relevante puesto que indica que el tamaño muestral no depende del tamaño de la población (N); el tamaño muestral depende del nivel de confianza deseado, del margen de error deseado y del valor de la desviación estándar σ . (Véase el ejercicio 40 para casos en los que se selecciona una muestra relativamente grande sin reemplazo a partir de una población finita).

El tamaño muestral debe ser un número entero, ya que representa el número de valores muestrales que deben encontrarse. Sin embargo, la fórmula 7-5 suele dar un resultado que no es un número entero, de manera que utilizamos la siguiente regla de redondeo. (Esta regla se basa en el principio de que cuando es necesario redondear, el tamaño de muestra requerido debe redondearse *hacia arriba* para que sea al menos adecuadamente grande en oposición a un tamaño ligeramente más pequeño).

Regla de redondeo para el tamaño muestral n

Cuando se calcula el tamaño muestral n , si el uso de la fórmula 7-5 no produce un número entero, siempre *incrementa* el valor de n al siguiente número entero *mayor*.

Cómo manejar σ desconocida al calcular el tamaño muestral Cuando se aplica la fórmula 7-5, existe un dilema práctico: la fórmula requiere que sustitu yamos algún valor de la desviación estándar poblacional σ , pero en realidad ésta suele desconocerse. Cuando se determina un tamaño muestral requerido (sin construir un intervalo de confianza), existen algunos procedimientos que pueden funcionar para este problema:

1. Use la regla práctica del intervalo (véase la sección 3-3) para estimar la desviación estándar como sigue: $\sigma \approx \text{rango}/4$. (Con una muestra de 87 valores o más, seleccionada al azar de una población normalmente distribuida, el rango/4 nos da un valor que es mayor que o igual a σ al menos el 95% de las veces. Véase “Using the Sample Range as a Basis for Calculating Sample Size in Power Calculations”, de Richard Browne, *American Statistician*, vol. 55, núm. 4).
2. Realice un estudio piloto empezando por el proceso de muestreo. Comience el proceso de muestreo y, utilizando los primeros valores, calcule la desviación estándar muestral s y úsela en lugar de σ . Entonces, el valor estimado de σ puede mejorar conforme se obtienen más datos muestrales, y de este modo es posible refinar el tamaño muestral.
3. Estime el valor de σ utilizando los resultados de algún otro estudio hecho con anterioridad.

Asimismo, algunas veces podemos ser creativos en nuestro uso de otros resultados conocidos. Por ejemplo, por lo regular las pruebas de CI están diseñadas para que la media sea 100 y la desviación estándar sea 15. Los profesores de estadística tienen puntuaciones de CI con una media mayor que 100 y una desviación estándar menor que 15 (puesto que son un grupo más homogéneo que las personas seleccionadas al azar de la población general). No conocemos el valor específico de σ para los profesores de estadística, pero podemos calcular con seguridad usando $\sigma = 15$. Utilizar un valor de σ que sea mayor que el valor real producirá un tamaño muestral mayor del necesario, pero utilizar un valor de σ que sea muy pequeño daría por resultado un tamaño muestral inadecuado. *Cuando se calcula el tamaño muestral n , cualquier error siempre debe ser conservador, en el sentido de que haga a n muy grande y no muy pequeña.*

EJEMPLO Puntuaciones de CI de profesores de estadística Supongamos que queremos estimar la puntuación media del CI de la población de profesores de estadística. ¿Cuántos profesores de estadística deben seleccionarse al azar para efectuar pruebas de CI, si queremos tener una confianza del 95% de que la media muestral estará dentro de 2 puntos de CI de la media poblacional?

SOLUCIÓN Los valores que requiere la fórmula 7-5 se calculan como sigue:

$z_{\alpha/2} = 1.96$ (Esto se calcula convirtiendo el nivel de confianza del 95% a $\alpha = 0.05$, y luego calculando la puntuación crítica z como se describe en la sección 7-2).

$E = 2$ (Puesto que queremos que la media muestral esté dentro de dos puntos de CI de μ , el margen de error deseado es 2).

$\sigma = 15$ (Véase el análisis en el párrafo que está antes de este ejemplo).

Con $z_{\alpha/2} = 1.96$, $E = 2$ y $\sigma = 15$, utilizamos la fórmula 7-5 como sigue:

$$n = \left[\frac{z_{\alpha/2}\sigma}{E} \right]^2 = \left[\frac{1.96 \cdot 15}{2} \right]^2 = 216.09 = 217 \quad (\text{redondeado hacia arriba})$$

INTERPRETACIÓN Entre los miles de profesores de estadística, necesitamos obtener una muestra aleatoria simple de al menos 217 de ellos y luego necesitamos obtener sus puntuaciones de CI. Con una muestra aleatoria simple de sólo 217 profesores de estadística, tendremos un nivel de confianza del 95% de que la media muestral \bar{x} está dentro de dos puntos de CI de la media poblacional μ real.

Si estamos dispuestos a obtener resultados menos exactos utilizando un margen de error más grande, como por ejemplo 4, el tamaño muestral disminuye a 54.0225, el cual se redondea *hacia arriba* a 55. La duplicación del margen de error hace que el tamaño muestral requerido disminuya a un cuarto de su valor original. Por el contrario, dividir a la mitad el margen de error cuadriplica el tamaño muestral. En consecuencia, si usted desea resultados más exactos, el tamaño muestral debe incrementarse sustancialmente. Ya que los muestreos grandes por lo regular requieren de más tiempo y dinero, con frecuencia existe la necesidad de realizar un balance entre el tamaño de la muestra y el margen de error E .

Uso de la tecnología

Intervalos de confianza Vea al final de la sección 7-4 los procedimientos del intervalo de confianza que se aplican a los métodos de

esta sección, así como también los de la sección 7-4. STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus se pueden usar para calcular intervalos de confianza cuando queremos estimar la media de una población y se satisfacen todos los supuestos de esta sección (incluido el valor conocido de σ).

Determinación del tamaño muestral Los cálculos para el tamaño muestral no se incluyen en la calculadora TI-83/84 Plus, Minitab o Excel. El procedimiento de STATDISK para determinar el tamaño muestral requerido para estimar una media poblacional μ se describe a continuación.

STATDISK Seleccione **Analysis** de la parte superior de la barra del menú principal, luego seleccione **Sample Size Determination**, seguido por **Estimate Mean**. Ahora debe ingresar el nivel de confianza (como 0.95) y el error E . También ingrese la desviación estándar poblacional σ , si se conoce. Además existe una opción que le permite ingresar el tamaño poblacional N , suponiendo que usted está haciendo el muestreo sin reemplazo de una población finita. (Véase el ejercicio 40).

7-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Intervalo de confianza.** Con base en datos muestrales se obtiene el siguiente intervalo de confianza del 95%: $2.5 < \mu < 6.0$. Escriba un enunciado que interprete el intervalo de confianza de manera correcta.
- Estimador sin sesgo.** Una de las características de la media muestral que la convierte en un buen estimador de una media poblacional μ es que es un estimador sin sesgo. ¿Qué significa que un estadístico sea un estimador sin sesgo de un parámetro poblacional?



- 3. Intervalo de confianza.** Un fabricante de juegos para parques de diversiones necesita el estimado de un intervalo de confianza de la fuerza que se puede ejercer cuando un visitante presiona un freno de seguridad. Incapaz de encontrar datos, obtiene una muestra midiendo la fuerza de 100 estudiantes de preparatoria que participan en una feria de ciencias. ¿El estimado de confianza resultante será un buen estimado de la fuerza media de la población de todos los visitantes potenciales? ¿Por qué?
- 4. Tamaño de muestra.** Un investigador calcula el tamaño muestral necesario para estimar la fuerza que pueden ejercer las piernas de las personas en juegos de parques de diversiones y obtiene el tamaño muestral de 120. Si el investigador no puede obtener una muestra aleatoria y en vez de ello debe confiar en una muestra de conveniencia integrada por sus amigos y parientes, ¿será posible compensar y obtener buenos resultados utilizando un tamaño muestral mucho más grande?

Cálculo de valores críticos. En los ejercicios 5 a 8, encuentre el valor crítico $z_{\alpha/2}$ que corresponda al nivel de confianza dado.

- | | |
|--------|--------|
| 5. 95% | 6. 96% |
| 7. 92% | 8. 99% |

Verificación de requisitos y cálculo del margen de error. En los ejercicios 9 a 12, calcule el margen de error $E = z_{\alpha/2}\sigma/\sqrt{n}$, si se satisfacen los requisitos necesarios. Si estos requisitos no se satisfacen, determine que el margen de error no puede calcularse por medio de los métodos de esta sección.

9. El nivel de confianza es del 95%, el tamaño muestral es $n = 100$ y $\sigma = 15$.
10. El nivel de confianza es del 95%, el tamaño muestral es $n = 9$ y se desconoce σ .
11. El nivel de confianza es del 99%, el tamaño muestral es $n = 9$, $\sigma = 15$ y la población original está distribuida normalmente.
12. El nivel de confianza es de 99%, el tamaño muestral es $n = 12$, se desconoce σ y la población original está distribuida normalmente.

Cálculo de un intervalo de confianza. En los ejercicios 13 a 16, utilice el intervalo de confianza y los datos muestrales indicados para calcular un intervalo de confianza para estimar la media poblacional μ .

13. Salarios de graduados universitarios que tomaron un curso de estadística en la universidad: confianza del 95%; $n = 41$, $\bar{x} = \$67,200$, y se sabe que σ es $\$18,277$.
14. Las velocidades de conductores multados en una zona con límite de velocidad de 55 mi/h: confianza del 95%; $n = 90$, $\bar{x} = 66.2$ mi/h, y se sabe que σ es 3.4 mi/h.
15. Calificaciones de crédito de FICO (Fair, Isacc and Company) de solicitantes de tarjetas de crédito: confianza del 99%; $n = 70$, $\bar{x} = 688$ y se sabe que σ es 68.
16. Cantidades perdidas por jugadores que tomaron un autobús a un casino de Atlantic City: confianza del 99%; $n = 40$, $\bar{x} = \$189$ y se sabe que σ es $\$87$.

Cálculo del tamaño muestral. En los ejercicios 17 a 20, use el margen de error, el nivel de confianza y la desviación estándar poblacional s indicados para calcular el tamaño de muestra mínimo requerido para estimar una media poblacional m desconocida.

17. Margen de error: 0.5 pulgadas, nivel de confianza: 95%, $\sigma = 2.5$ pulgadas.
18. Margen de error: 0.25 segundos, nivel de confianza: 99%, $\sigma = 5.4$ segundos.
19. Margen de error: \$1, nivel de confianza: 90%, $\sigma = \$12$.
20. Margen de error: 1.5 mm, nivel de confianza: 95%, $\sigma = 8.7$ mm.

Interpretación de resultados. En los ejercicios 21 a 24, remítase a los resultados de Minitab que se presentan abajo, sobre un intervalo de confianza del 95% generado con los métodos de esta sección. Los resultados muestrales provienen del uso de una muestra aleatoria de las velocidades de conductores multados en una sección de la carretera interestatal 95, en Connecticut.

MINITAB

Variable	N	Mean	StDev	SE Mean	95% CI
Speed	81	67.3849	3.3498	0.3722	(66.6554, 68.1144)

21. Identifique el valor del estimado puntual de la media poblacional μ .
22. Exprese el intervalo de confianza en el formato de $\bar{x} - E < \mu < \bar{x} + E$.
23. Exprese el intervalo de confianza en el formato de $\bar{x} \pm E$.
24. Escriba una afirmación que interprete el intervalo de confianza del 95%.
25. **Tiempo para obtener un título de licenciatura.** En un estudio sobre el tiempo que un estudiante requiere para obtener un título universitario, se selecciona al azar a 80 estudiantes y se descubre que tienen una media de 4.8 años (según datos del National Center for Education Statistics). Suponiendo que $\sigma = 2.2$ años, construya un estimado de un intervalo de confianza de la media poblacional. ¿El intervalo de confianza resultante contradice el hecho de que el 39% de los estudiantes obtienen su título universitario en cuatro años?
26. **Edades de motociclistas muertos en accidentes.** Un estudio de las edades de motociclistas muertos en accidentes incluye la selección aleatoria de 150 conductores con una media de 37.1 años (según datos del Insurance Institute for Highway Safety). Suponiendo que $\sigma = 12.0$ años, construya un estimado de un intervalo de confianza del 99% de la media de la edad de todos los motociclistas muertos en accidentes. Si los límites del intervalo de confianza no incluyen edades menores de 20 años, ¿eso significa que los motociclistas menores de 20 años rara vez mueren en accidentes?
27. **Percepción del tiempo.** Alumnos del autor, seleccionados al azar, participaron en un experimento con el fin de poner a prueba su habilidad para determinar el transcurso de 1 minuto (o 60 segundos). Cuarenta estudiantes produjeron una media muestral de 58.3 segundos. Suponiendo que $\sigma = 9.5$ segundos, construya un estimado de un intervalo de confianza del 95% de la media poblacional de todos los estudiantes estadística. Con base en el resultado, ¿es probable que sus estimados tengan una media que se acerque razonablemente a 60 segundos?
28. **Niveles de cotinina en fumadores.** Cuando las personas fuman, la nicotina que absorben se convierte en cotinina, la cual puede medirse. Una muestra de 40 fumadores tiene una media del nivel de cotinina de 172.5. Suponga que se sabe que σ es 119.5, calcule un estimado del intervalo de confianza del 90% de la media del nivel de cotinina para todos los fumadores. ¿Qué aspecto de este problema no es realista?
29. **Niveles de presión sanguínea.** Cuando 14 estudiantes de segundo año de medicina del Bellevue Hospital midieron la presión sanguínea de la misma persona, obtuvieron los resultados que se listan abajo. Suponiendo que se sabe que la desviación estándar poblacional es de 10 mmHg, construya un estimado de un intervalo de confianza del 95% de la media poblacional. De manera ideal, ¿cuál debe ser el intervalo de confianza en esta situación?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

30. **El mamífero más pequeño del mundo.** El mamífero más pequeño del mundo es el murciélagos abejorro, también conocido como murciélagos nariz de cochino (o *Craseonycteris thonglongyai*). Estos animales apenas alcanzan el tamaño de un abejorro grande. A continuación se listan los pesos (en gramos) de una muestra de estos murciélagos. Suponiendo que los pesos de todos estos murciélagos tienen una desviación

estándar de 0.30 gramos, construya un estimado de un intervalo de confianza del 95% de su peso medio. Utilice el intervalo de confianza para determinar si esta muestra de murciélagos proviene de la misma población con una media conocida de 1.8 gramos.

1.7 1.6 1.5 2.0 2.3 1.6 1.6 1.8 1.5 1.7 2.2 1.4 1.6 1.6 1.6

31. **Pesos de monedas de 25 centavos del apéndice B.** Utilice los pesos de las monedas de 25 centavos, acuñadas después de 1964, que se incluyen en el conjunto de datos 14 del apéndice B. Suponiendo que las monedas se acuñan de tal manera que tienen pesos con una desviación estándar poblacional de 0.068 gramos, utilice la muestra de pesos para construir un estimado del intervalo de confianza del 99% del peso medio. Las especificaciones en Estados Unidos exigen que las monedas pesen entre 5.443 gramos y 5.897 gramos. ¿Qué sugiere el intervalo de confianza acerca del proceso de fabricación?
32. **Errores de pronóstico del apéndice B.** Remítase al conjunto de datos 8 del apéndice B y reste cada temperatura máxima real de la temperatura máxima que se pronosticó un día antes. El resultado es una lista de errores. Suponiendo que todos los errores tienen una desviación estándar de 2.5° , construya un estimado del intervalo de confianza del 95% para la media de todos estos errores. ¿Qué sugiere el resultado acerca de la exactitud de las temperaturas pronosticadas?

Cálculo del tamaño de muestra. En los ejercicios 33 a 38, calcule el tamaño de muestra indicado.

33. **Tamaño muestral para la media del CI de estudiantes de estadística.** La prueba Weschler de CI está diseñada de tal forma que, para la población de adultos normales, la media es 100 y la desviación estándar es 15. Calcule el tamaño muestral necesario para estimar la media de la puntuación de CI de estudiantes de estadística. Queremos tener una confianza del 95% de que nuestra media muestral está dentro de dos puntos de CI de la media real. La media para esta población es claramente mayor que 100. Tal vez la desviación estándar de esta población es menor que 15, porque se trata de un grupo con menor variación que un grupo seleccionado al azar de la población general; por lo tanto, si usamos $\sigma = 15$, estamos siendo conservadores al emplear un valor que hará que el tamaño de la muestra sea al menos tan grande como se necesite. Suponga entonces que $\sigma = 15$ y determine el tamaño de muestra requerido.
34. **Tamaño muestral de pesos de monedas de 25 centavos.** La Tyco Video Game Corporation encontró que está perdiendo ingresos por las fichas que se usan en sus juegos de video. Las máquinas deben ser ajustadas para aceptar monedas sólo si caen dentro de límites prefijados. Para ajustar estos límites, debe estimarse la media del peso de monedas de un cuarto de dólar en circulación. Se pesará una muestra de monedas de un cuarto de dólar para determinar la media. ¿Cuántas monedas de un cuarto de dólar debemos seleccionar al azar y pesar si queremos tener un nivel de confianza del 99% de que la media muestral está dentro de 0.025 g de la media real de la población de todas las monedas de un cuarto de dólar? Con base en los resultados de una muestra de monedas de un cuarto de dólar, podemos estimar que la desviación estándar de la población es 0.068 g.
35. **Tamaño muestral para la dieta Atkins.** Usted desea estimar la pérdida media de peso de las personas, un año después de utilizar la dieta Atkins. ¿Cuántas personas sometidas a la dieta se deben encuestar si deseamos tener una confianza del 95% de que la media muestral de la pérdida de peso está dentro de 0.25 lb de la media poblacional real? Suponga que sabemos que la desviación estándar poblacional es de 10.6 lb (según datos de “Comparison of the Atkins, Ornish, Weight Watchers and Zone Diets for Weight Loss and Heart Disease Risk Reduction”, de Dansinger *et al.*, *Journal of the American Medical Association*, vol. 293, núm. 1).
36. **Tamaño muestral de uso de la televisión.** Nielsen Media Research desea estimar la media de la cantidad de tiempo (en minutos) que los alumnos universitarios que estudian tiempo completo dedican a ver la televisión cada día de la semana. Calcule el tamaño muestral necesario para estimar esta media con un margen de error de 15 minutos. Suponga que se desea un nivel de confianza del 96%. También suponga que un estudio piloto indicó que la desviación estándar se estima en 112.2 min.

- 37. Tamaño muestral utilizando la regla práctica del intervalo.** Usted acaba de ser contratado por la división de marketing de General Motors para estimar la media de la cantidad de dinero que se gasta actualmente en la compra de automóviles nuevos en Estados Unidos. Primero use la regla práctica del intervalo para hacer un estimado burdo de la desviación estándar de las cantidades gastadas. Es razonable suponer que el rango típico de cantidades va desde \$12,000 hasta alrededor de \$70,000. Luego use esa desviación estándar estimada para determinar el tamaño muestral que corresponde a un nivel de confianza del 95% y a un margen de error de \$100. ¿Es práctico el tamaño muestral? Si no es así, ¿qué se debe cambiar para obtener un tamaño muestral práctico?
- 38. Tamaño muestral utilizando datos muestrales.** Usted quiere estimar la media del pulso de adultos varones. Remítase al conjunto de datos 1 en el apéndice B y calcule el pulso máximo y mínimo para varones, luego utilice estos valores con la regla práctica del intervalo para estimar σ . ¿Cuántos adultos varones debe usted seleccionar al azar y examinar si quiere tener un nivel de confianza del 95% de que la media muestral del pulso está dentro de 2 latidos (por minuto) de la media poblacional μ real? Si en vez de usar la regla práctica del intervalo se emplea la desviación estándar de los pulsos de varones del conjunto de datos 1 como un estimado de σ , ¿es muy diferente el tamaño muestral requerido? ¿Qué tamaño muestral parece estar más cerca del tamaño muestral correcto?

7-3 MÁS ALLÁ DE LO BÁSICO

- 39. Intervalo de confianza con factor de corrección por población finita.** El error estándar de la media es σ/\sqrt{n} , siempre y cuando el tamaño de la población sea infinito. Si el tamaño de la población es finito y se denota como N , entonces el factor de corrección $\sqrt{(N-n)/(N-1)}$ debe usarse siempre y cuando $n > 0.05N$. Este factor de corrección multiplica el margen de error E dado en la fórmula 7-4, de manera que el margen de error es como se indica abajo. Calcule el intervalo de confianza del 95% para la media de 250 puntuaciones de CI, si una muestra de 35 de esas puntuaciones produce una media de 110. Suponga que $\sigma = 15$.

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

- 40. Tamaño muestral con factor de corrección por población finita.** En la fórmula 7-4 para el margen de error E , suponemos que la población es infinita, que estamos realizando un muestreo con reemplazo, o que la población es muy grande. Si tenemos una población relativamente pequeña y hacemos el muestreo sin reemplazo, debemos modificar E para incluir un *factor de corrección por población finita*, para que el margen de error sea como el que se indica en el ejercicio 39, donde N es el tamaño de la población. En esta expresión del margen de error se despeja n para obtener

$$n = \frac{N\sigma^2(z_{\alpha/2})^2}{(N-1)E^2 + \sigma^2(z_{\alpha/2})^2}$$

Repita el ejercicio 33, suponiendo que los estudiantes de estadística se seleccionan al azar y sin reemplazo, de una población de $N = 200$ estudiantes de estadística.

7-4 Estimación de la media poblacional: σ desconocida

Concepto clave En esta sección se presentan métodos para construir un estimado del intervalo de confianza de una media poblacional cuando *no se conoce* la desviación estándar. (En la sección 7-3 se presentaron métodos para estimar μ cuando se conoce σ). Cuando se desconoce σ , se utiliza la *distribución t de Student* (en vez de la distribución normal), suponiendo que ciertos requisitos (los cuales se señalan abajo) se satisfacen. Como generalmente se desconoce σ en circunstancias reales, los métodos de esta sección son muy realistas y prácticos, y se utilizan con frecuencia.

Requisitos

1. La muestra es aleatoria simple.
2. La muestra proviene de una población distribuida normalmente o $n > 30$.

Como en la sección 7-3, el requisito de una población distribuida normalmente no es estricto. Por lo regular, podemos considerar que la población está distribuida normalmente después de usar los datos muestrales para confirmar que no existen valores extremos y que el histograma tiene una forma que no es muy lejana a la de una distribución normal. Además, al igual que en la sección 7-3, el requisito de que el tamaño muestral sea $n > 30$ suele usarse como directriz, pero el tamaño muestral mínimo realmente depende de cuánto se aleja la distribución de la población de la distribución normal. [Si se sabe que una población se distribuye normalmente, la distribución de medias muestrales \bar{x} es *exactamente* una distribución normal con media μ y desviación estándar σ/\sqrt{n} ; si la población no está distribuida normalmente, muestras grandes ($n > 30$) producen medias muestrales con una distribución que es *aproximadamente* normal, con media μ y desviación estándar σ/\sqrt{n}].

Al igual que en la sección 7-3, la media muestral \bar{x} es el mejor estimado puntual (o estimado de un solo valor) de la media poblacional μ .

La media muestral \bar{x} es el mejor estimado puntual de la media poblacional μ .

He aquí el aspecto clave de esta sección: si σ no se conoce, pero los requisitos anteriores se satisfacen, utilizamos la *distribución t de Student* (en vez de la distribución normal), que desarrolló William Gosset (1876-1937). Gosset fue un empleado de la cervecería Guinness Brewery que necesitaba una distribución que pudiera utilizarse con muestras pequeñas. La cervecería irlandesa donde trabajaba no permitía la publicación de resultados de investigaciones, entonces Gosset publicó bajo el seudónimo de *Student*. (En aras de la investigación y para servir a sus lectores, el autor visitó la cervecería Guinness Brewery y probó una muestra del producto. ¡Qué comprometido!).

Puesto que no conocemos el valor de σ , lo estimamos con el valor de la desviación estándar muestral s , pero esto introduce otra fuente de falta de confiabilidad, en especial con las muestras pequeñas. Para mantener un intervalo de confianza en algún nivel deseado, como el 95%, compensamos esta falta de confiabilidad adicional haciendo más ancho el intervalo de confianza: utilizamos valores críticos $t_{\alpha/2}$ (de una distribución *t de Student*), los cuales son más grandes que los valores críticos de $z_{\alpha/2}$ de la distribución normal.

Distribución *t de Student*

Si una población tiene una distribución normal, entonces la distribución de

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

es una **distribución *t de Student*** para todas las muestras de tamaño n . La distribución *t de Student*, conocida a menudo como **distribución *t***, se utiliza para calcular valores críticos denotados por $t_{\alpha/2}$.

Pronto analizaremos algunas de las propiedades importantes de la distribución *t*, pero antes presentamos los componentes necesarios para la construcción de intervalos

de confianza. Comencemos con el valor crítico denotado por $t_{\alpha/2}$. Un valor de $t_{\alpha/2}$ se puede encontrar en la tabla A-3 localizando el número apropiado de *grados de libertad* en la columna izquierda y avanzando por el renglón correspondiente hasta encontrar el número que aparece directamente abajo del área adecuada en la parte superior.

Definición

El número de **grados de libertad** para un conjunto de datos muestrales recolectados es el número de valores muestrales que pueden variar después de haber impuesto ciertas restricciones a todos los valores de los datos.

Por ejemplo, si 10 estudiantes tienen puntuaciones de examen con una media de 80, podemos asignar con libertad valores a las primeras 9 puntuaciones, pero la décima puntuación se calcula. La suma de las 10 puntuaciones debe ser 800, entonces la décima puntuación debe ser igual a 800 menos la suma de las primeras 9 puntuaciones. Puesto que esas primeras 9 puntuaciones pueden seleccionarse con libertad para adoptar cualquier valor, decimos que existen 9 grados de libertad disponibles. Para las aplicaciones de esta sección, el número de grados de libertad es simplemente el tamaño muestral menos 1.

$$\text{grados de libertad} = n - 1$$

EJEMPLO Cálculo de un valor crítico Una muestra de tamaño $n = 23$ es una muestra aleatoria simple seleccionada de una población distribuida normalmente. Calcule el valor crítico $t_{\alpha/2}$ correspondiente a un nivel de confianza del 95%.

SOLUCIÓN Puesto que $n = 23$, el número de grados de libertad está dado por $n - 1 = 22$. Utilizando la tabla A-3, localizamos el renglón 22 con respecto a la columna de la extrema izquierda. Al igual que en la sección 7-2, un nivel de confianza del 95% corresponde a $\alpha = 0.05$, de manera que encontramos los valores listados en la columna para una *área de 0.05 en dos colas*. El valor correspondiente al renglón para 22 grados de libertad y la columna para una área de 0.05 en dos colas es 2.074; entonces $t_{\alpha/2} = 2.074$.

Ahora que sabemos cómo encontrar valores críticos denotados por $t_{\alpha/2}$, podemos describir el margen de error E de este intervalo de confianza.

Margen de error E para la estimación de μ (con σ desconocida)

Fórmula 7-6

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

donde $t_{\alpha/2}$ tiene $n - 1$ grados de libertad. La tabla A-3 lista valores de $t_{\alpha/2}$.



Extractos de una circular del Departamento del Transporte

Los siguientes extractos de una circular del departamento de transporte de Estados Unidos ataúan algunos de los requisitos de exactitud para el equipo de navegación empleado en aviones. Observe el uso del intervalo de confianza. “El total de las contribuciones de error del equipo a bordo, combinado con los errores técnicos de vuelo correspondientes incluidos en la lista, no debe exceder lo siguiente, con un nivel de confianza del 95% (2-sigma), durante un periodo igual al ciclo de actualización”. “El sistema de vías y rutas aéreas de Estados Unidos tiene anchuras de protección de ruta que se utilizan en un sistema VOR con una exactitud de ± 4.5 grados con base en una probabilidad del 95%”.



Intervalo de confianza para la estimación de μ (con σ desconocida)

$$\bar{x} - E < \mu < \bar{x} + E$$

donde

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

El siguiente procedimiento utiliza el margen de error anterior en la construcción de estimados del intervalo de confianza de μ .

Procedimiento para construir un intervalo de confianza para μ (con σ desconocida)

1. Verifique que los requisitos se satisfacen. (Tenemos una muestra aleatoria simple y la población parece estar distribuida normalmente o $n > 30$).
2. Utilizando $n - 1$ grados de libertad, remítase a la tabla A-3 y encuentre el valor crítico $t_{\alpha/2}$ que corresponde al nivel de confianza deseado. (Para el nivel de confianza, remítase al “área en dos colas”).
3. Evalúe el margen de error $E = t_{\alpha/2} \cdot s / \sqrt{n}$.
4. Utilizando el valor del margen de error E calculado y el valor de la media muestral \bar{x} , calcule los valores de $\bar{x} - E$ y $\bar{x} + E$. Sustituya estos valores en el formato general para el intervalo de confianza:

$$\bar{x} - E < \mu < \bar{x} + E$$

o

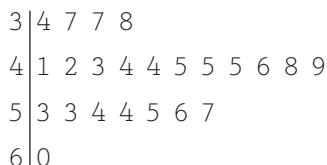
$$\bar{x} \pm E$$

o

$$(\bar{x} - E, \bar{x} + E)$$

5. Redondee los límites del intervalo de confianza resultantes. Si utiliza el conjunto original de datos, redondee a un decimal más del que se usa para el conjunto original de datos. Si utiliza un resumen de estadísticos (n, \bar{x}, s), redondee los límites del intervalo de confianza al mismo número de lugares decimales utilizados para la media muestral.

Gráfica de tallo y hojas de las edades



EJEMPLO Construcción de un intervalo de confianza En el diagrama de tallo y hojas que aparece al margen, se incluyen las edades de solicitantes que no lograron un ascenso (según datos de “Debating the Use of Statistical Evidence in Allegations of Age Discrimination”, de Barry y Boland, *American Statistician*, vol. 58, núm. 2). Existe el tema más importante de si ciertos solicitantes fueron víctimas de discriminación por edad, pero por ahora nos enfocaremos en el simple aspecto de utilizar esos valores como una muestra con el propósito de estimar la media de una población más grande. Suponga que la muestra es aleatoria simple y utilice los datos muestrales con un nivel de confianza del 95% para calcular lo siguiente:

- a. El margen de error E
- b. El intervalo de confianza para μ

SOLUCIÓN

REQUISITO ✓ Primero debemos verificar que los dos requisitos para esta sección se satisfacen. Estamos suponiendo que la muestra es aleatoria simple. Ahora revisamos el requisito de que “la población se distribuya normalmente o $n > 30$ ”. Puesto que $n = 23$, debemos verificar que la distribución sea aproximadamente normal. La forma de la gráfica de tallo y hojas sugiere una distribución normal. Además, una gráfica cuantil normal confirma que los datos

muestrales provienen de una población con una distribución aproximadamente normal. Por consiguiente, los requisitos se satisfacen y procedemos con los métodos de esta sección. ✓

- El nivel de confianza de 0.95 implica que $\alpha = 0.05$, de manera que $t_{\alpha/2} = 2.074$ (utilice la tabla A-3 con $gl = n - 1 = 22$, como se mostró en el ejemplo anterior). Después de encontrar que los estadísticos muestrales son $n = 23$, $\bar{x} = 47.0$ y $s = 7.2$, el margen de error E se calcula utilizando la fórmula 7-6 como sigue. Se utilizan decimales adicionales para minimizar los errores de redondeo en el intervalo de confianza calculado en el inciso b).

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 2.074 \cdot \frac{7.2}{\sqrt{23}} = 3.11370404$$

- Con $\bar{x} = 47.0$ y $E = 3.11370404$, construimos el intervalo de confianza de la siguiente manera:

$$\bar{x} - E < \mu < \bar{x} + E$$

$$47.0 - 3.11370404 < \mu < 47.0 + 3.11370404$$

$$43.9 < \mu < 50.1 \quad (\text{redondeado a un decimal más que los datos originales})$$

INTERPRETACIÓN Este resultado también podría expresarse en la forma de 47.0 ± 3.1 o $(43.9, 50.1)$. Con base en los resultados muestrales dados, tenemos una confianza del 95% de que los límites de 43.9 años y 50.1 años realmente contienen el valor de la media poblacional μ .

Ahora listamos las propiedades importantes de la distribución t que utilizamos en esta sección.

Propiedades importantes de la distribución t de Student

- La distribución t de Student es diferente para distintos tamaños de muestra. (Véase la figura 7-5 para los casos $n = 3$ y $n = 12$).
- La distribución t de Student tiene la misma forma de campana simétrica que la distribución normal estándar, pero refleja una mayor variabilidad (con distribuciones más amplias) de lo que se espera con muestras pequeñas.

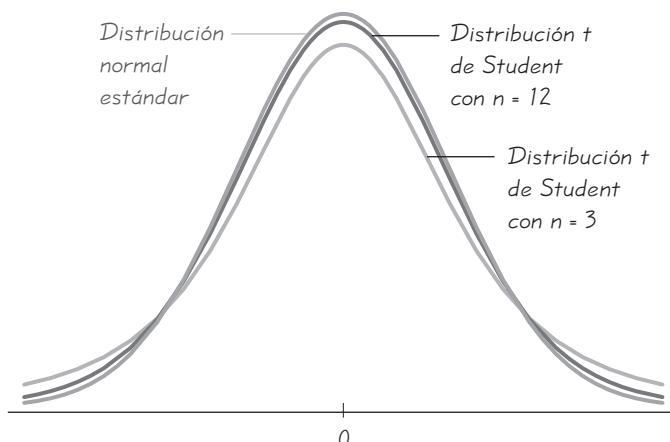


Figura 7-5

Distribuciones t de Student para $n = 3$ y $n = 12$

La distribución t de Student tiene la misma forma y simetría general de la distribución normal estándar, pero refleja una mayor variabilidad de lo que se espera con muestras pequeñas.

3. La distribución t de Student tiene una media de $t = 0$ (así como la distribución normal estándar tiene una media de $z = 0$).
4. La desviación estándar de la distribución t de Student varía con el tamaño muestral, pero es mayor que 1 (a diferencia de la distribución normal estándar, que tiene $\sigma = 1$).
5. Conforme el tamaño muestral n se hace más grande, la distribución t de Student se acerca más a la distribución normal estándar.

Elección de la distribución apropiada

En ocasiones es difícil decidir entre utilizar la distribución normal estándar z o la distribución t de Student. El diagrama de flujo de la figura 7-6 y la tabla 7-1 resumen los aspectos clave a considerarse cuando se construyen intervalos de confianza para estimar μ , la media poblacional. En la figura 7-6 o en la tabla 7-1, note que si tenemos una muestra pequeña ($n \leq 30$) obtenida de una distribución que difiere drásticamente de una distribución normal, no podemos usar los métodos descritos en este capítulo. Una alternativa es utilizar métodos no paramétricos (véase el capítulo 13); otra alternativa es usar el método de *bootstrap* por computadora. En ambos enfoques no se hacen supuestos acerca de la población original. El método *bootstrap* se describe en el proyecto tecnológico al final del capítulo.

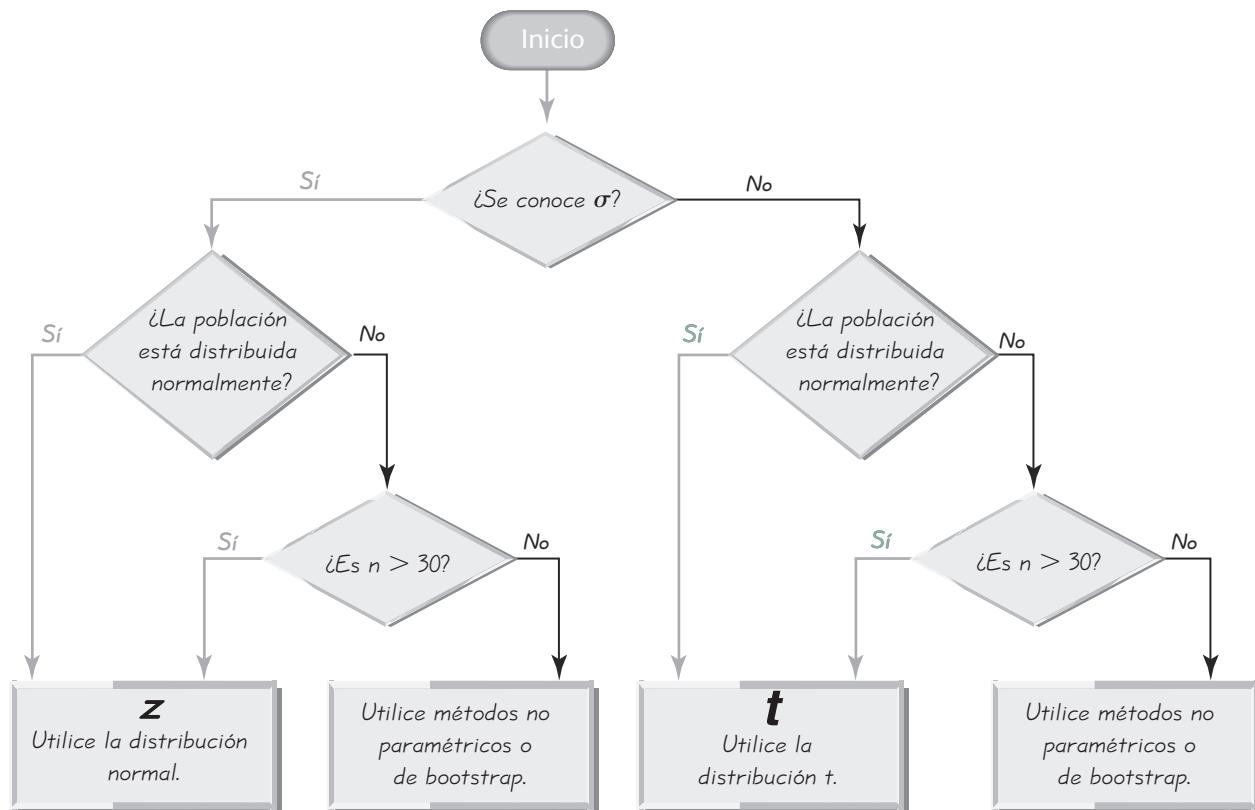


Figura 7-6 Elección entre z y t

Tabla 7-1 Elección entre z y t

Método	Condiciones
Utilice la distribución normal (z).	σ conocida y población distribuida normalmente o σ conocida y $n > 30$
Utilice la distribución t .	σ desconocida y población distribuida normalmente o σ desconocida y $n > 30$
Utilice un método no paramétrico o de <i>bootstrap</i> .	La población no está distribuida normalmente y $n \leq 30$

Notas:

- 1. Criterios para decidir si la población está distribuida normalmente:** La población no necesita ser exactamente normal, pero debe tener una apariencia un tanto simétrica, con una moda y sin valores extremos.
- 2. Tamaño muestral $n > 30$:** Éste es un lineamiento que se usa regularmente, pero tamaños muestrales de 15 a 30 son adecuados si la población parece tener una distribución normal y no existen valores extremos. Para algunas distribuciones poblacionales que estén extremadamente alejadas de la normal, puede requerirse que el tamaño muestral sea mayor de 50 o aun de 100.

El siguiente ejemplo se enfoca en elegir la aproximación correcta utilizando los métodos de esta sección y de la sección 7-3.

EJEMPLO Selección de distribuciones Suponiendo que usted planea construir un intervalo de confianza para la media poblacional μ , utilice los datos para determinar si el margen de error E debe calcularse utilizando un valor crítico de $z_{\alpha/2}$ (de la distribución normal), un valor crítico de $t_{\alpha/2}$ (de la distribución t) o ninguno de estos (de manera que los métodos de la sección 7-3 y de esta sección no son viables).

- a. $n = 150$, $\bar{x} = 100$, $s = 15$, y la población tiene una distribución sesgada.
- b. $n = 8$, $\bar{x} = 100$, $s = 15$, y la población tiene una distribución normal.
- c. $n = 8$, $\bar{x} = 100$, $s = 15$, y la población tiene una distribución muy sesgada.
- d. $n = 150$, $\bar{x} = 100$, $\sigma = 15$, y la distribución está sesgada. (Esta situación casi nunca ocurre).
- e. $n = 8$, $\bar{x} = 100$, $\sigma = 15$, y la distribución está extremadamente sesgada. (Esta situación casi nunca ocurre).

SOLUCIÓN Remítase a la figura 7-6 o a la tabla 7-1 para determinar lo siguiente:

- a. Puesto que la desviación estándar poblacional σ no se conoce y la muestra es grande ($n > 30$), el margen de error se calcula usando $t_{\alpha/2}$ en la fórmula 7-6.
- b. Puesto que la desviación estándar poblacional σ no se conoce y la población está distribuida normalmente, el margen de error se calcula usando $t_{\alpha/2}$ en la fórmula 7-6.



Estimación de azúcar en las naranjas

En Florida, los miembros de la industria de los cítricos usan profusamente métodos estadísticos. Una aplicación específica tiene que ver con la forma en que se paga a los agricultores por las naranjas que se usan para elaborar jugo de naranja. Cuando llega un camión cargado con naranjas, primero se pesa la carga en la planta receptora, luego se elige al azar una muestra de una docena de naranjas. La muestra se pesa, se exprime y se mide la cantidad de azúcar que contiene el jugo. Con base en los resultados de la muestra, se estima la cantidad total de azúcar contenida en toda la carga del camión. El pago por la carga de naranjas se basa en la estimación de la cantidad de azúcar, ya que las naranjas más dulces son más valiosas que las menos dulces, aunque las cantidades de jugo sean iguales.

continúa



Estimados para mejorar el censo

En el censo del decenio no se cuenta a todas las personas y es posible que algunas de ellas se cuenten más de una vez. Existen métodos estadísticos para mejorar los conteos de población con ajustes en cada condado de cada estado. Algunos argumentan que la Constitución especifica que el censo debe ser una “enumeración real” que no permite ajustes. Una norma de la Suprema Corte prohíbe el uso de conteos de población ajustados por las repercusiones que esto tendría en la asignación de escaños en el Congreso, pero una disposición reciente establecida por una Corte Federal de apelación ordenó que se permitieran los conteos ajustados, incluso si no pueden utilizarse para ese propósito. Según la Associated Press, “el Census Bureau ha dejado abierta la posibilidad de utilizar datos ajustados para financiamientos federales en el futuro”, de manera que el uso de métodos estadísticos poderosos, con el tiempo, produciría una mejor asignación de fondos federales y estatales.

- c. Puesto que la muestra es pequeña y la población no tiene una distribución normal, el margen de error E no debe calcularse usando un valor crítico de $z_{\alpha/2}$ o $t_{\alpha/2}$. No se aplican los métodos de la sección 7-3 ni de esta sección.
- d. Puesto que la desviación estándar poblacional σ se conoce y la muestra es grande ($n > 30$), el margen de error se calcula usando $z_{\alpha/2}$ en la fórmula 7-4.
- e. Puesto que la población no está distribuida normalmente y la muestra es pequeña ($n \leq 30$), el margen de error E no debe calcularse usando un valor crítico de $z_{\alpha/2}$ o $t_{\alpha/2}$. No se aplican los métodos de la sección 7-3 ni de esta sección.

EJEMPLO Intervalo de confianza para pesos al nacer En un estudio de los efectos sobre los bebés del consumo de cocaína durante el embarazo, se obtuvieron los siguientes datos de pesos al nacer: $n = 190$, $\bar{x} = 2700$ g, $s = 645$ g (según datos de “Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure”, de Singer *et al.*, *Journal of the American Medical Association*, vol. 291, núm. 20). El diseño del estudio justifica el supuesto de que la muestra puede tratarse como una muestra aleatoria simple. Utilice los datos muestrales para construir un estimado del intervalo de confianza del 95% de μ , el peso medio al nacer de todos los bebés hijos de madres que consumieron cocaína durante el embarazo.

SOLUCIÓN

REQUISITO ✓ Primero debemos verificar que los requisitos se cumplan. Se trata de una muestra aleatoria simple. Puesto que $n = 190$, se satisface el requisito de que “la población se distribuye normalmente o $n > 30$ ”. Por lo tanto, los requisitos se cumplen. (Este es el paso 1 del procedimiento de cinco pasos que se describió antes, y ahora podemos continuar con los pasos restantes). ✓

Paso 2: El valor crítico es $t_{\alpha/2} = 1.972$. En la tabla A-3 encontramos que el valor crítico corresponde a $n - 1 = 189$ grados de libertad (columna izquierda de la tabla A-3) y una área en dos colas de 0.05. (Puesto que la tabla A-3 no incluye $gl = 189$, utilizamos el valor crítico más cercano de 1.972. Si utilizamos un programa de cómputo encontraremos que un valor crítico más exacto es 1.973, de manera que la aproximación que se hace aquí es bastante buena).

Paso 3: *Calcule el margen de error E:* El margen de error $E = 2.97355$ se calcula utilizando la fórmula 7-6 como se indica a continuación, con decimales adicionales para minimizar el error de redondeo en el intervalo de confianza calculado en el paso 4.

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 1.972 \cdot \frac{645}{\sqrt{190}} = 92.276226$$

Paso 4: *Calcule el intervalo de confianza:* Ahora podemos calcular el intervalo de confianza utilizando $\bar{x} = 2700$ g y $E = 92.276226$, como se indica a continuación:

$$\begin{aligned} \bar{x} - E < \mu &< \bar{x} + E \\ 2700 - 92.276226 &< \mu < 2700 + 92.276226 \\ 2607.7238 &< \mu < 2792.2762 \end{aligned}$$

Paso 5: Redondee los límites del intervalo de confianza. Como la media muestral se redondea a un número entero, redondee los límites del intervalo de confianza para obtener este resultado: $2608 < \mu < 2792$.

INTERPRETACIÓN Con base en los datos muestrales, tenemos una confianza del 95% de que los límites de 2608 g y 2792 g realmente contienen el valor del peso medio al nacer. Ahora podemos comparar este resultado con un intervalo de confianza construido para los pesos al nacer de niños cuyas madres no consumieron cocaína. (Véase el ejercicio 17).



Cálculo del estimado puntual y de E a partir de un intervalo de confianza

Más adelante en esta sección describiremos cómo pueden utilizarse las calculadoras y los programas de cómputo para calcular un intervalo de confianza. Un uso común requiere que usted ingrese un nivel de confianza y estadísticos muestrales, y la pantalla indica los límites del intervalo de confianza. La media muestral \bar{x} es el valor intermedio de estos límites; el margen de error E es la mitad de la diferencia entre estos límites (ya que el límite superior es $\bar{x} + E$ y el límite inferior es $\bar{x} - E$, y la distancia que los separa es $2E$).

Estimado puntual de μ :

$$\bar{x} = \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2}$$

Margen de error:

$$E = \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2}$$

EJEMPLO Edades de polizones Al analizar las edades de todos los polizones del *Queen Mary* (según datos de Cunard Line), se obtiene la pantalla de Minitab que se muestra abajo. Utilice el intervalo de confianza dado para calcular el estimado puntual \bar{x} y el margen de error E . Trate los valores como datos muestrales seleccionados al azar de una población grande.

Minitab

95.0% CI
(24.065, 27.218)

SOLUCIÓN En los siguientes cálculos, los resultados se redondean a un decimal, que es un espacio decimal adicional más del redondeo utilizado para la lista original de edades.

$$\begin{aligned}\bar{x} &= \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2} \\ &= \frac{27.218 + 24.065}{2} = 25.6 \text{ años}\end{aligned}$$

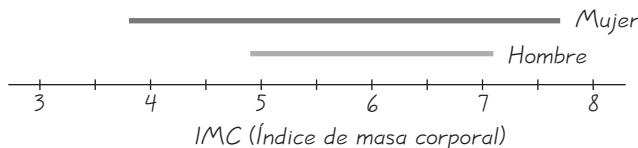
$$\begin{aligned}E &= \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2} \\ &= \frac{27.218 - 24.065}{2} = 1.6 \text{ años}\end{aligned}$$

Estimación del tamaño de multitudes

Existen métodos complejos para analizar el tamaño de una multitud. Se pueden emplear fotografías aéreas y medidas de densidad demográfica con una exactitud bastante razonable. Sin embargo, los reportes de estimaciones del tamaño de multitudes a menudo son simples conjeturas. Después de que los Medias Rojas de Boston ganaron la Serie Mundial por primera vez en 86 años, las autoridades de la ciudad de Boston estimaron que a la fiesta de celebración acudieron 3.2 millones de fanáticos. La policía de Boston dio un estimado de alrededor de un millón, pero aceptó que este cálculo se basaba en conjeturas de los comandantes de la policía. Un análisis fotográfico produjo un estimado de alrededor de 150,000. El profesor Farouk El-Baz de la Universidad de Boston utilizó imágenes del U.S. Geological Survey para llegar a un estimado de casi 400,000. El físico Bill Donnelly del MIT dijo que “es un problema serio que la gente sólo indique un número cualquiera. Esto significa que otros tantos asuntos no se están investigando de manera cuidadosa”.

Figura 7-7

Índices de masa corporal (IMC) de hombres y mujeres



Uso de los intervalos de confianza para describir, explorar o comparar datos

En algunos casos, podríamos utilizar un intervalo de confianza para lograr el objetivo final de estimar el valor de un parámetro poblacional. En otros casos, un intervalo de confianza podría ser una de varias herramientas para describir, explorar o comparar conjuntos de datos. En la figura 7-7 se presentan gráficas de intervalos de confianza para los índices de masa corporal (IMC) de una muestra de mujeres y hombres (véase el conjunto de datos 1 del apéndice B). Puesto que los intervalos de confianza se traslanan, no parece haber una diferencia significativa entre la media del índice IMC de mujeres y hombres.

Uso de la tecnología

Los siguientes procedimientos se aplican a intervalos de confianza para estimar una media μ e incluyen los intervalos de confianza descritos en la sección 7-3, así como los intervalos de confianza presentados en esta sección. Antes de utilizar programas de cómputo o una calculadora para generar un intervalo de confianza, asegúrese de revisar que los requisitos se satisfagan. Consulte los requisitos listados casi al principio de esta sección y de la sección 7-3.

STATDISK Primero debe calcular el tamaño muestral n , la media muestral \bar{x} y la desviación estándar muestral s . (Vea el procedimiento del STATDISK descrito en la sección 3-3). Seleccione **Analysis** de la barra del menú principal, luego **Confidence Intervals** y después **Population Mean**. Proceda a ingresar los elementos en el cuadro de diálogo; luego, haga clic en el botón **Evaluate**. El intervalo de confianza aparecerá en la pantalla.



MINITAB Minitab Release 14 ahora le permite utilizar ya sea el resumen de estadísticos n , \bar{x} y s o una lista de los valores muestrales originales. Seleccione **Stat** y **Basic Statistics**. Si no se conoce σ , seleccione **1-sample t** e ingrese el resumen de estadísticos o ingrese **C1** en el recuadro ubicado en la parte superior derecha. (Si se conoce σ , seleccione **1-sample Z** e ingrese el resumen de estadísticos o ingrese **C1** en el recuadro ubicado en la parte superior derecha. También ingrese el valor de σ en el cuadro "Standard Deviation" o "Sigma"). Utilice el botón **Options** para ingresar el nivel de confianza.

EXCEL Utilice el programa complementario Data Desk XL que es un complemento de este libro. Haga clic en **DDXL** y seleccione **Confidence Intervals**. Dentro de las opciones para tipo de función, seleccione **1 Var t Interval** si se desconoce σ . (Si se conoce σ , seleccione **1 Var z Interval**). Haga clic en el ícono con forma de lápiz e ingrese el rango de datos, como A1:A12 si usted tiene 12 valores listados en la columna A. Haga clic en **OK**. En el cuadro de diálogo, seleccione el nivel de confianza. (Si está utilizando 1 Var z Interval, también ingrese el valor de σ). Haga clic en **Compute Interval** y el intervalo de confianza aparecerá en la pantalla.

No se recomienda el uso de la herramienta de Excel para calcular intervalos de confian-

za, ya que supone que se conoce σ , y usted debe calcular primero el tamaño muestral n y la desviación estándar muestral s (que puede calcularse usando **fxx**, **Statistical**, **ST-DEV**). En vez de generar el intervalo de confianza completo con límites específicos, esta herramienta calcula sólo el margen de error E . Luego usted debe restar este resultado a \bar{x} y sumarlo a \bar{x} para poder identificar los límites reales del intervalo de confianza. Para utilizar esta herramienta cuando se conoce σ , haga clic en **fxx**, seleccione la categoría de funciones **Statistical** y luego seleccione el elemento de **CONFIDENCE**. En el cuadro de diálogo, ingrese el valor de α (llamado *nivel de significancia*), la desviación estándar y el tamaño muestral. El resultado será el valor del margen de error E .

TI-83/84 PLUS La calculadora TI-83/84 Plus sirve para generar intervalos de confianza para valores muestrales originales guardados en una lista, o bien, usted puede utilizar un resumen de los estadísticos n , \bar{x} y s . Ingrese los datos en la lista L1 o tenga disponible el resumen de los estadísticos, luego presione la tecla **STAT**. Ahora seleccione **TEST** y elija **TInterval** si no se conoce σ (elija **ZInterval** si se conoce σ). Despues de efectuar los ingresos requeridos, la pantalla de la calculadora incluirá el intervalo de confianza en el formato $(\bar{x} - E, \bar{x} + E)$.

Advertencia: Como en las secciones 7-2 y 7-3, los intervalos de confianza pueden usarse de manera informal para comparar diferentes conjuntos de datos, pero el traslape de intervalos de confianza no debe usarse para obtener conclusiones formales ni finales acerca de la igualdad de medias. En capítulos posteriores se incluyen procedimientos para decidir si dos poblaciones tienen medias iguales; esos métodos no tendrán las dificultades asociadas con las comparaciones que se basan en el traslape de intervalos de confianza.

No utilice el traslape de intervalos de confianza como base para obtener conclusiones formales acerca de la igualdad de medias.

7-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **¿Qué está mal?** Una nota en *USA Today* señaló que “los consumidores gastarán un promedio estimado de \$483 en mercancía” por el reinicio de las clases. Se reportó que el valor se basa en una encuesta de 8453 consumidores, y el margen de error es “ ± 1 punto porcentual”. ¿Qué está incorrecto en esta información?
2. **Intervalo de confianza.** El *Newport Chronicle* publicó un reporte que afirmaba que, con base en una muestra de hogares, la factura fiscal media es de \$4626, con un margen de error de \$591. Exprese el intervalo de confianza en el formato de $\bar{x} - E < \mu < \bar{x} + E$.
3. **Interpretación de un intervalo de confianza.** Utilizando los niveles de presión sanguínea sistólica de 40 hombres que se listan en el conjunto de datos 1 del apéndice B, obtenemos el siguiente intervalo de confianza del 99%: $114.4 < \mu < 123.4$. Redacte una aseveración que interprete correctamente ese intervalo de confianza.
4. **Verificación de requisitos.** Suponga que deseamos construir un estimado de un intervalo de confianza para las cantidades de precipitación que caen los lunes en Boston, y que planeamos utilizar las cantidades que aparecen en el conjunto de datos 10 del apéndice B. Al examinar esas cantidades vemos que, de los 52 lunes, 33 tienen cantidades de 0. Con base en esa observación, ¿parece que las cantidades de precipitación que caen los lunes se distribuyen normalmente? Suponiendo que la muestra se puede considerar aleatoria simple, ¿podemos utilizar los métodos de esta sección para construir un estimado de un intervalo de confianza para la media poblacional? ¿Por qué?

Uso de la distribución correcta. En los ejercicios 5 a 12, realice una de las siguientes acciones, según sea apropiado: a) calcule el valor crítico $z_{\alpha/2}$, b) calcule el valor crítico $t_{\alpha/2}$, c) determine que no se aplica ni la distribución normal ni la distribución *t*.

5. 95%; $n = 12$; se desconoce σ ; la población parece estar distribuida normalmente.
6. 99%; $n = 15$; se desconoce σ ; la población parece estar distribuida normalmente.
7. 99%; $n = 4$; se conoce σ ; la población parece estar muy sesgada.
8. 95%; $n = 50$; se conoce σ ; la población parece estar muy sesgada.
9. 90%; $n = 200$; se desconoce σ ; la población parece estar distribuida normalmente.
10. 98%; $n = 16$; $\sigma = 5.0$; la población parece estar muy sesgada.
11. 98%; $n = 18$; $\sigma = 21.5$; la población parece estar distribuida normalmente.
12. 90%; $n = 33$; se desconoce σ ; la población parece estar distribuida normalmente.

Cálculo de intervalos de confianza. En los ejercicios 13 y 14, utilice el nivel de confianza y los datos muestrales indicados para calcular a) el margen de error y b) el intervalo de confianza para la media poblacional m . Suponga que la población tiene una distribución normal.

13. Peso perdido por una dieta de Weight Watchers: 95% de confianza; $n = 40$, $\bar{x} = 3.0$ kg, $s = 4.9$ kg.
14. Periodo de vida de una computadora de escritorio: 99% de confianza; $n = 21$, $\bar{x} = 6.8$ años, $s = 2.4$ años.

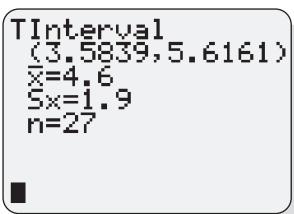
Interpretación de la pantalla de la calculadora. En los ejercicios 15 y 16 utilice los datos y la imagen de la pantalla de la calculadora TI-83/84 Plus correspondiente para expresar el intervalo de confianza en el formato de $\bar{x} - E < \mu < \bar{x} + E$. Además, escriba un enunciado que interprete el intervalo de confianza.

15. Puntuaciones de CI de estudiantes de estadística: 95% de confianza; $n = 25$, $\bar{x} = 118.0$, $s = 10.7$.

Minitab

N	Mean	StDev	SE Mean	95% CI
25	118.000	10.700	2.140	(113.583, 122.417)

TI-83/84 Plus



16. Periodo de vida de teléfonos celulares: 99% de confianza; $n = 27$, $\bar{x} = 4.6$ años, $s = 1.9$ años. (Véase la pantalla de la calculadora TI-83/84 Plus que aparece al margen).

Construcción de intervalos de confianza. En los ejercicios 17 a 26, construya el intervalo de confianza.

17. **Pesos al nacer.** Una muestra aleatoria de los pesos al nacer de 186 bebés tiene una media de 3103 g y una desviación estándar de 696 g (según datos de “Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure”, de Singer *et al.*, *Journal of the American Medical Association*, vol. 291, núm. 20). Estos bebés son hijos de madres que no consumieron cocaína durante el embarazo. Construya un estimado del intervalo de confianza del 95% del peso medio al nacer de todos estos bebés. Compare el resultado con el intervalo de confianza obtenido en el ejemplo de esta sección sobre los pesos al nacer de hijos de madres que consumieron cocaína durante el embarazo. Al parecer, ¿el consumo de cocaína afecta el peso que tiene un bebé al nacer?
18. **Temperatura media corporal.** El conjunto de datos 2 del apéndice B incluye 106 temperaturas corporales, para las cuales $\bar{x} = 98.20^\circ\text{F}$ y $s = 0.62^\circ\text{F}$. Utilizando los estadísticos de la muestra, construya un estimado del intervalo de confianza del 99% para la temperatura media corporal de todos los seres humanos saludables. ¿Los límites del intervalo de confianza incluyen los 98.6°F ? ¿Qué sugiere la muestra acerca del uso de 98.6°F como la temperatura corporal media?
19. **Temperaturas pronosticadas y reales.** El conjunto de datos 8 del apéndice B incluye una lista de temperaturas máximas reales y la lista correspondiente del pronóstico de temperaturas máximas para tres días. Si la diferencia para cada día se obtiene restando la temperatura máxima pronosticada para tres días de la temperatura máxima real, el resultado es una lista de 35 valores con una media de -1.3° y una desviación estándar de 4.7° .
 - a. Construya un estimado del intervalo de confianza del 99% para la media de la diferencia entre todas las temperaturas máximas reales y las temperaturas máximas pronosticadas para tres días.
 - b. ¿El intervalo de confianza incluye 0° ? Si un meteorólogo afirma que el pronóstico de temperaturas máximas para tres días tiende a ser muy alto puesto que la diferencia media de la muestra es -1.3° ¿parece ser válida esa afirmación? ¿Por qué?
20. **Ritmos cardiacos al trabajar con la pala.** Ya que las muertes por problemas cardíacos parecen incrementarse después de las fuertes nevadas, se diseñó un experimento para comparar las demandas cardíacas al retirar nieve con una pala con las de aquellos que utilizan un aparato eléctrico para retirarla. Diez sujetos retiraron la nieve del terreno

usando ambos métodos y se registraron sus frecuencias cardíacas máximas (en latidos por minuto) durante ambas actividades. Se obtuvieron los siguientes resultados (según datos de “Cardiac Demands of Heavy Snow Shoveling”, de Franklin *et al.*, *Journal of the American Medical Association*, vol. 273, núm. 11):

Frecuencias cardíacas máximas al retirar la nieve con pala: $n = 10$, $\bar{x} = 175$, $s = 15$

Frecuencias cardíacas máximas al usar un aparato eléctrico para retirar la nieve:

$n = 10$, $\bar{x} = 124$, $s = 18$

- Calcule el estimado del intervalo de confianza del 95% de la media poblacional para aquellas personas que retiran la nieve con la ayuda de una pala.
- Calcule el estimado del intervalo de confianza del 95% de la media poblacional de aquellas personas que usan el aparato eléctrico para retirar la nieve.
- Si usted fuera un médico preocupado por las muertes debidas a problemas cardíacos provocados por el paleo manual de nieve, ¿qué valor individual del intervalo de confianza del inciso a) sería de mayor interés?
- Compare los intervalos de confianza de los incisos a) y b) e interprete lo que encontró.

- 21. Control del plomo en el aire.** A continuación se listan las cantidades de plomo medidas (en microgramos por metro cúbico o $\mu\text{g}/\text{m}^3$) en el aire. La Environmental Protection Agency estableció un estándar de calidad del aire para el plomo de $1.5 \mu\text{g}/\text{m}^3$. Las medidas que se presentan abajo se registraron en el edificio 5 del World Trade Center en diferentes días, inmediatamente después de la destrucción causada por los ataques terroristas del 11 de septiembre de 2001. Despues del colapso de los dos edificios del World Trade Center hubo una gran preocupación por la calidad del aire. Utilice los valores dados para construir un estimado del intervalo de confianza del 95% para la cantidad media de plomo en el aire. ¿Hay algo en este conjunto de datos que sugiera que el intervalo de confianza tal vez no sea muy bueno? Explique.

5.40 1.10 0.42 0.73 0.48 1.10

- 22. Construcción de un intervalo de confianza.** La gráfica de tallo y hojas que se presenta abajo incluye las edades de solicitantes que lograron un ascenso (según datos de “Debating the Use of Statistical Evidence in Allegations of Age Discrimination”, de Barry y Boland, *American Statistician*, vol. 58, núm. 2). Suponga que la muestra es aleatoria simple y construya un estimado del intervalo de confianza del 95% para la edad media de todas estas personas exitosas. Compare el resultado con el intervalo de confianza para las edades de los individuos que no lograron el ascenso (véase el ejemplo en esta sección).



- 23. Calificación de crédito.** Cuando los consumidores solicitan crédito, su crédito se califica con puntuaciones FICO (Fair, Isaac, and Company). A continuación se presentan calificaciones de crédito de una muestra de solicitantes de préstamos para adquirir un automóvil. Utilice los datos muestrales con el fin de construir un intervalo de confianza del 99% para la media de la calificación FICO de todos los solicitantes de crédito. Si un banco requiere una calificación de crédito de al menos 620 puntos para un préstamo destinado a adquirir un automóvil, al parecer, ¿casi todos los solicitantes tendrán calificaciones de crédito adecuadas?

661 595 548 730 791 678 672 491 492 583 762 624 769 729 734 706

- 24. El mamífero más pequeño del mundo.** El mamífero más pequeño del mundo es el murciélagos abejorro, también conocido como murciélagos nariz de cochino (o *Cra- seonycteris thonglongyai*). Estos murciélagos apenas alcanzan el tamaño de un abejorro grande. A continuación se presentan los pesos (en gramos) de una muestra de esos

murciélagos. Construya un estimado del intervalo de confianza del 95% de su peso medio. ¿Difieren mucho los límites del intervalo de confianza de los límites de 1.56 y 1.87 que se obtienen al suponer que se sabe que σ es 0.30 g?

1.7 1.6 1.5 2.0 2.3 1.6 1.6 1.8 1.5 1.7 2.2 1.4 1.6 1.6 1.6

- 25. Estimación de contaminación de automóviles.** En una muestra de siete automóviles, se prueban las emisiones de óxido de nitrógeno de cada uno (en gramos por milla) y se obtienen los siguientes resultados: 0.06, 0.11, 0.16, 0.15, 0.14, 0.08, 0.15 (según datos de la Environmental Protection Agency). Suponiendo que esta muestra es representativa de los automóviles en uso, construya un estimado del intervalo de confianza del 98% para la cantidad media de emisiones de óxido de nitrógeno para todos los automóviles. Si la Environmental Protection Agency exige que las emisiones de óxido de nitrógeno sean menores que 0.165 g/mi, ¿podemos concluir con seguridad que se cubre este requisito?
- 26. Anchura de cráneos.** Las anchuras máximas de muestras de cráneos egipcios masculinos del año 4000 a.C. y del año 150 d.C. son (según datos de *Ancient Races of the Thebaid*, de Thomson y Randall-Maciver):

4000 a.C.: 131 119 138 125 129 126 131 132 126 128 128 131
150 d.C.: 136 130 126 126 139 141 137 138 133 131 134 129

Los cambios en el tamaño de la cabeza con el paso del tiempo sugieren un mestizaje con individuos de otras regiones. Utilice intervalos de confianza para determinar si el tamaño de la cabeza parece haber cambiado del año 4000 a.C. al 150 d.C. Explique su resultado.

Conjuntos de datos del apéndice B. En los ejercicios 27 y 28, utilice los conjuntos de datos del apéndice B.

- 27. Pulso.** Una doctora quiere desarrollar criterios para determinar si el pulso de un paciente es anormal y desea determinar si hay diferencias significativas entre hombres y mujeres. Utilice los pulsos muestrales del conjunto de datos 1 del apéndice B.
- Construya un estimado del intervalo de confianza del 95% del pulso medio de los hombres.
 - Construya un estimado del intervalo de confianza del 95% del pulso medio de las mujeres.
 - Compare los resultados anteriores. ¿Podemos concluir que las medias poblacionales para hombres y para mujeres son diferentes? ¿Por qué?
- 28. Comparación de Coca-Cola regular y dietética.** Remítase al conjunto de datos 12 del apéndice B y utilice los datos muestrales.
- Construya un estimado del intervalo de confianza del 95% para el peso medio de la bebida de cola en latas de Pepsi regular.
 - Construya un estimado del intervalo de confianza del 95% para el peso medio de la bebida de cola en latas de Pepsi dietética.
 - Compare los resultados de los incisos a) y b) e interprételos. ¿Parece haber alguna diferencia? Si es así, identifique una razón de esta diferencia.

7-4 MÁS ALLÁ DE LO BÁSICO

- 29. Efecto de un valor extremo.** Pruebe el efecto de un valor extremo como sigue: utilice los datos muestrales del ejercicio 22 para calcular un estimado del intervalo de confianza del 95% de la media poblacional, después de cambiar el primer valor de 54 años por el de 540 años. Este valor no es realista, pero un error de este tipo puede ocurrir fácilmente durante un proceso de captura de datos. ¿Se altera mucho el intervalo de confianza cuando se cambia el valor de 54 años por el de 540 años? ¿Los límites del intervalo de confianza son sensibles a los valores extremos? ¿Cómo debe usted manejar los valores extremos cuando se encuentran en conjuntos de datos muestrales que se usarán para la construcción de intervalos de confianza?

- 30. Método alternativo.** La figura 7-6 y la tabla 7-1 resumen la decisión tomada al elegir entre las distribuciones normal y t . Un método alternativo que se incluye en algunos libros de texto (pero que casi nunca se incluye en revistas científicas) se basa en el siguiente criterio: sustituya la desviación estándar muestral s por σ siempre que $n > 30$, y luego proceda como si se conociera σ . Suponga que para una muestra aleatoria simple, $n = 35$, $\bar{x} = 50.0$ y $s = 10.0$, luego construya estimados del intervalo de confianza del 95% de μ utilizando el método de esta sección y el método alternativo. Compare los resultados.
- 31. Factor de corrección por población finita.** Si se selecciona una muestra aleatoria simple de tamaño n sin reemplazo de una población finita de tamaño N , y el tamaño muestral es mayor que el 5% del tamaño de la población ($n > 0.05N$), se pueden obtener mejores resultados utilizando el factor de corrección por población finita, el cual implica multiplicar el margen de error E por $\sqrt{(N - n)/(N - 1)}$. Para la muestra de 100 pesos de dulces M&M del conjunto de datos 13 en el apéndice B, obtenemos $\bar{x} = 0.8565$ g y $s = 0.0518$ g. Primero construya un estimado del intervalo de confianza del 95% de μ , suponiendo que la población es grande, y luego construya un estimado del intervalo de confianza del 95% del peso medio de dulces M&M que se encuentran en la bolsa llena de donde se tomó la muestra. La bolsa llena tiene 465 dulces M&M. Compare los resultados.
- 32. Uso de la distribución incorrecta.** Suponga que se selecciona una muestra aleatoria simple pequeña de una población distribuida normalmente, para la que σ es desconocida. La construcción de un intervalo de confianza debe utilizar la distribución t , pero ¿cómo se afecta el intervalo de confianza si se usa la distribución normal incorrectamente, en vez de la distribución t ?
- 33. Intervalo de confianza para muestra de tamaño $n = 1$.** Cuando una nave espacial dirigida por la NASA llega a Marte, los astronautas encuentran a un solo adulto marciano que mide 12.0 pies de estatura. Es razonable suponer que las estaturas de todos los marcianos se distribuyen normalmente.
- Los métodos de este capítulo requieren información acerca de la variación de una variable. Si sólo está disponible un valor muestral, ¿puede éste darnos alguna información acerca de la variación de la variable?
 - Al utilizar los métodos de esta sección, ¿qué pasa cuando usted trata de usar la estatura individual en la construcción de un intervalo de confianza del 95%?
 - Con base en el artículo “An Effective Confidence Interval for the Mean with Samples of Size One and Two”, de Wall, Boen y Tweedie (*American Statistician*, vol. 55, núm. 2), se calcula un intervalo de confianza del 95% para μ (utilizando métodos que no se analizan en este libro) con una muestra de tamaño $n = 1$ seleccionada al azar de una población distribuida normalmente, y se expresa como $x \pm 9.68|x|$. Utilice este resultado para construir un intervalo de confianza del 95% empleando el valor muestral individual de 12.0 ft, y expréselo en la forma de $\bar{x} - E < \mu < \bar{x} + E$. Con base en el resultado ¿es posible que algún otro marciano seleccionado al azar mida 50 ft de estatura?

Estimación de la varianza 7-5 poblacional

Concepto clave En esta sección presentamos métodos para **1.** calcular un intervalo de confianza de una desviación estándar o una varianza poblacional y **2.** determinar el tamaño muestral requerido para estimar una desviación estándar o una varianza poblacional. En esta sección se presenta la distribución chi cuadrada, la cual se utiliza para calcular un estimado de un intervalo de confianza de σ o de σ^2 .

Requisitos

1. La muestra es aleatoria simple.
2. La población debe tener valores distribuidos normalmente (aun si la muestra es grande).

El supuesto de una población distribuida normalmente se mencionó en secciones anteriores, pero este requisito es mucho más importante aquí. Para los métodos de esta sección, los alejamientos de una distribución normal pueden generar errores muy graves. En consecuencia, el requisito de tener una distribución normal es mucho más estricto, y debemos revisar la distribución de los datos construyendo histogramas y gráficas cuantilares normales, como se describe en la sección 6-7.

Cuando consideramos estimados de proporciones y medias, utilizamos las distribuciones normal y *t* de Student. Cuando desarrollamos estimados de varianzas o desviaciones estándar utilizamos otra distribución, conocida como la distribución chi cuadrada. Examinaremos características importantes de esta distribución antes de proceder con el desarrollo de intervalos de confianza.

Distribución chi cuadrada

En una población distribuida normalmente con varianza σ^2 , suponga que seleccionamos al azar muestras independientes de tamaño n y, para cada muestra, calculamos la varianza muestral s^2 (que es el cuadrado de la desviación estándar muestral s). El estadístico muestral $\chi^2 = (n - 1)s^2/\sigma^2$ tiene una distribución llamada **distribución chi cuadrada**.

Distribución chi cuadrada

Fórmula 7-7

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

donde n = tamaño muestral

s^2 = varianza muestral

σ^2 = varianza poblacional

Denotamos chi cuadrada por χ^2 , que se pronuncia “ji cuadrada”. Para calcular valores críticos de la distribución chi cuadrada, remítase a la tabla A-4. La distribución chi cuadrada se determina por el número de grados de libertad y en este capítulo usamos $n - 1$ grados de libertad.

$$\text{grados de libertad} = n - 1$$

En capítulos posteriores encontraremos situaciones en las que los grados de libertad no son $n - 1$, por lo que no debemos hacer la generalización incorrecta de que el número de grados de libertad es siempre $n - 1$.

Propiedades de la distribución del estadístico chi cuadrada

1. La distribución chi cuadrada no es simétrica, a diferencia de las distribuciones normal y *t* de Student (véase la figura 7-8). (Conforme el número de grados de

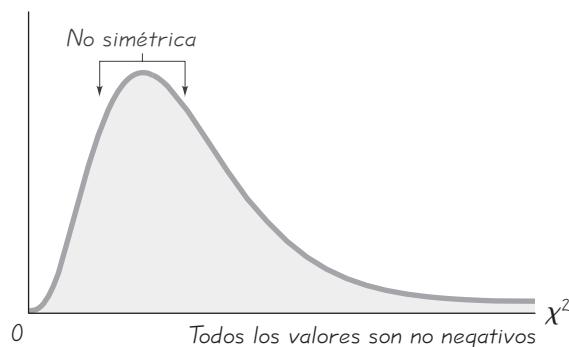


Figura 7-8
Distribución chi cuadrada

libertad se incrementa, la distribución se vuelve más simétrica, como ilustra la figura 7-9).

2. Los valores de chi cuadrada pueden ser cero o positivos, pero no pueden ser negativos (véase la figura 7-8).
3. La distribución chi cuadrada es diferente para cada número de grados de libertad (véase la figura 7-9), y en esta sección el número de grados de libertad está dado por $gl = n - 1$. Conforme el número de grados de libertad se incrementa, la distribución chi cuadrada se approxima a una distribución normal.

Puesto que la distribución chi cuadrada es sesgada y no simétrica, el intervalo de confianza no se ajusta al formato de $s^2 \pm E$ y debemos hacer cálculos separados para los límites de confianza superior e inferior. Si se utiliza la tabla A-4 para calcular valores críticos, observe su siguiente característica:

En la tabla A-4 cada valor crítico de χ^2 corresponde a una área que se encuentra en el renglón superior de la tabla, y esa área representa la región acumulativa localizada a la derecha del valor crítico.

La tabla A-2 para la distribución normal estándar proporciona áreas acumulativas de la *izquierda*, pero la tabla A-4 para la distribución chi cuadrada provee áreas acumulativas de la *derecha*.

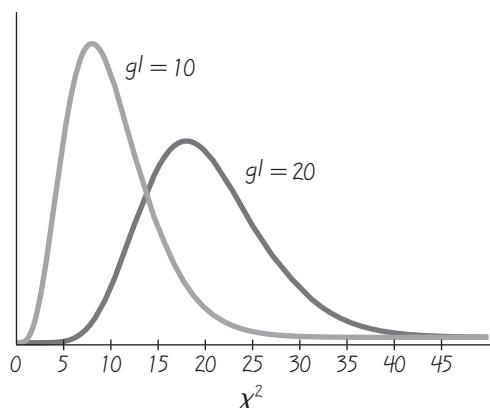


Figura 7-9
Distribución chi cuadrada para $gl = 10$ y $gl = 20$

EJEMPLO Valores críticos Calcule los valores críticos de χ^2 que determinan las regiones críticas que contienen una área de 0.025 en cada cola. Suponga que el tamaño muestral relevante es 10, de manera que el número de grados de libertad es $10 - 1$, o 9.

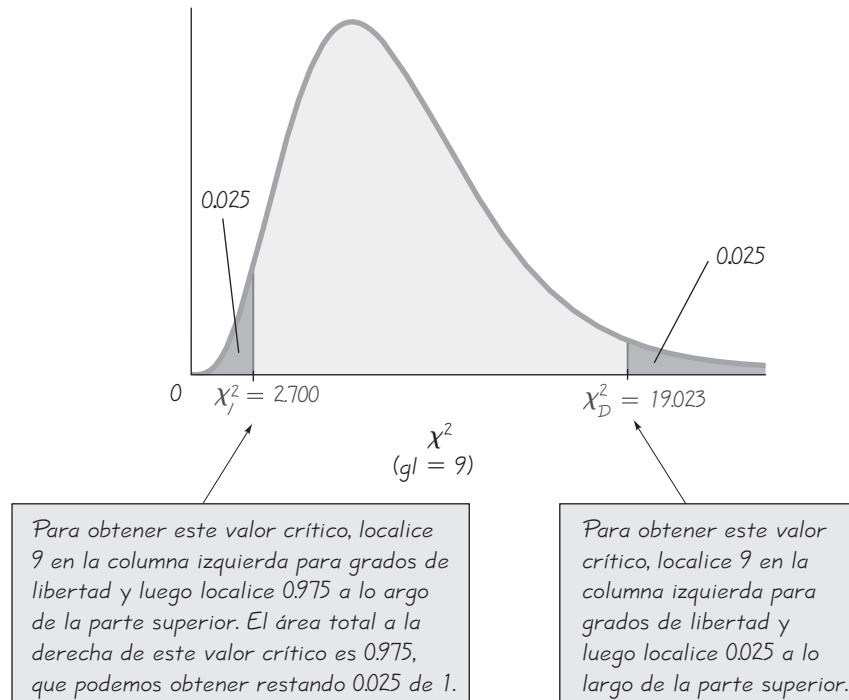


Figura 7-10 Valores críticos de la distribución chi cuadrada

SOLUCIÓN Vea la figura 7-10 y remítase a la tabla A-4. El valor crítico a la derecha ($\chi^2 = 19.023$) se obtiene de manera directa localizando 9 en la columna de grados de libertad a la izquierda y 0.025 a lo largo de la parte superior. El valor crítico de $\chi^2 = 2.700$ a la izquierda otra vez corresponde a 9 en la columna de grados de libertad, pero debemos localizar 0.975 (que se calcula al restar 0.025 de 1) a lo largo de la parte superior, puesto que los valores en el renglón superior son siempre *áreas a la derecha* del valor crítico. Remítase a la figura 7-10 y vea que el área total a la derecha de $\chi^2 = 2.700$ es 0.975. La figura 7-10 nos indica que, para una muestra de 10 valores tomados de una población distribuida normalmente, el estadístico chi cuadrada $(n - 1)s^2/\sigma^2$ tiene una probabilidad de 0.95 de caer dentro de los valores críticos de chi cuadrada de 2.700 y 19.023.

Observe que, cuando se obtienen valores críticos de χ^2 de la tabla A-4, los números de grados de libertad son enteros consecutivos del 1 al 30, seguidos por 40, 50, 60, 70, 80, 90 y 100. Cuando un número de grados de libertad (por ejemplo 52) no se encuentra en la tabla, generalmente se utiliza el valor crítico más cercano. Por ejemplo, si el número de grados de libertad es 52, remítase a la tabla A-4 y utilice 50 grados de libertad. (Si el número de grados de libertad está exactamente a la mitad de dos valores de la tabla, como por ejemplo 55, simplemente calcule la media de los dos valores χ^2). Para números de grados de libertad mayores de 100,

use la ecuación que se incluye en el ejercicio 27, una tabla más detallada o un programa de cómputo de estadística.

Estimadores de σ^2

En la sección 6-4 señalamos que las varianzas muestrales s^2 tienden a coincidir con (o centrarse en) el valor de la varianza poblacional σ^2 , por lo que decimos que s^2 es un *estimador sin sesgo* de σ^2 . Es decir, las varianzas muestrales s^2 no tienden sistemáticamente a sobreestimar el valor de σ^2 , ni tampoco tienden sistemáticamente a subestimar σ^2 . En vez de ello, tienden a coincidir con el valor de la propia σ^2 . Además, los valores de s^2 tienden a producir errores más pequeños por estar más cercanos a σ^2 que otras medidas de variación sin sesgo. Por estas razones, generalmente se utiliza s^2 para estimar σ^2 . [Sin embargo, existen otros estimadores de σ^2 que podrían considerarse mejores que s^2 . Por ejemplo, aun cuando $(n - 1)s^2/(n + 1)$ es un estimador sesgado de σ^2 , tiene la propiedad muy deseable de minimizar la media de los cuadrados de los errores y, por lo tanto, tiene una mayor probabilidad de acercarse a σ^2 . Véase el ejercicio 28].

La varianza muestral s^2 es el mejor estimado puntual de la varianza poblacional σ^2 .

Puesto que s^2 es un estimador sin sesgo de σ^2 , esperaríamos que s fuera un estimador sin sesgo de σ , pero no es así. (Véase la sección 5.4). Sin embargo, si el tamaño muestral es grande, el sesgo es tan pequeño que podemos utilizar s como un estimado de σ razonablemente bueno. Aunque s es un estimado sesgado, se usa con frecuencia como un estimado puntual de σ .

La desviación estándar muestral s suele utilizarse como un estimado puntual de σ (aunque es un estimado sesgado).

Si bien s^2 es el mejor estimado puntual de σ^2 , no existe una indicación de qué tan bueno es en realidad. Para compensar esta deficiencia, desarrollamos un estimado de intervalo (o intervalo de confianza) que es más informativo.

Intervalo de confianza (o estimado de intervalo) para la varianza poblacional σ^2

$$\frac{(n - 1)s^2}{\chi_D^2} < \sigma^2 < \frac{(n - 1)s^2}{\chi_I^2}$$

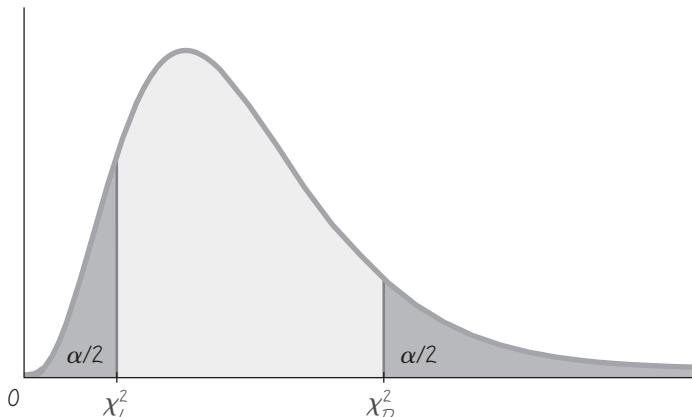
Esta expresión se utiliza para calcular un intervalo de confianza para la varianza σ^2 , pero un intervalo de confianza (o un estimado de intervalo) para la desviación estándar σ se calcula tomando la raíz cuadrada de cada componente, como se indica abajo.

$$\sqrt{\frac{(n - 1)s^2}{\chi_D^2}} < \sigma < \sqrt{\frac{(n - 1)s^2}{\chi_I^2}}$$

Las notaciones χ_D^2 y χ_I^2 en las expresiones anteriores se describen como sigue. (Observe que algunos otros libros de texto utilizan $\chi_{\alpha/2}^2$ en vez de χ_D^2 y utilizan $\chi_{1-\alpha/2}^2$ en vez de χ_I^2).

Figura 7-11**Distribución chi cuadrada con valores críticos χ_I^2 y χ_D^2**

Los valores críticos χ_I^2 y χ_D^2 separan las áreas extremas correspondientes a varianzas muestrales que son improbables (con probabilidad α).

**Notación**

Con una área total de α dividida por igual entre las dos colas de una distribución chi cuadrada, χ_I^2 denota el valor crítico de la cola izquierda y χ_D^2 denota el valor crítico de la cola derecha (como se ilustra en la figura 7-11).

Con base en los resultados anteriores, podemos resumir el procedimiento para construir un estimado del intervalo de confianza de σ o σ^2 como sigue.

Procedimiento para construir un intervalo de confianza para σ o σ^2

1. Verifique que los requisitos se satisfagan. (La muestra es aleatoria simple y un histograma o gráfica cuantilar normal sugiere que la población tiene una distribución que es muy cercana a la distribución normal).
2. Utilizando $n - 1$ grados de libertad, remítase a la tabla A-4 y encuentre los valores críticos χ_D^2 y χ_I^2 correspondientes al nivel de confianza deseado.
3. Evalúe los límites del intervalo de confianza superior e inferior utilizando el siguiente formato para el intervalo de confianza:

$$\frac{(n - 1)s^2}{\chi_D^2} < \sigma^2 < \frac{(n - 1)s^2}{\chi_I^2}$$

4. Si se desea un estimado del intervalo de confianza de σ , calcule la raíz cuadrada de los límites del intervalo de confianza superior e inferior y cambie σ^2 por σ .
5. Redondee los límites del intervalo de confianza resultantes. Si se utiliza el conjunto original de datos, redondee a un decimal más del que se usa para el conjunto original de datos. Si se utiliza la desviación estándar o varianza muestrales, redondee los límites del intervalo de confianza al mismo número de espacios decimales.

Advertencia: Los intervalos de confianza pueden usarse de manera informal para comparar conjuntos diferentes de datos, pero el traslape de intervalos de confianza no debe usarse para obtener conclusiones formales ni finales acerca de la igualdad de las varianzas o de las desviaciones estándar. En capítulos posteriores se incluirán procedimientos para decidir si dos poblaciones tienen varianzas o desviaciones estándar iguales, y esos métodos no tendrán las

deficiencias asociadas con comparaciones basadas en el traslape de los intervalos de confianza.

No utilice el traslape de intervalos de confianza como base para obtener conclusiones definitivas acerca de la igualdad de varianzas o desviaciones estándar.

EJEMPLO Intervalo de confianza para pesos de monedas de 1 centavo

En la actualidad las monedas de 1 centavo de dólar se acuñan con una desviación estándar de 0.0165 g (de acuerdo con el conjunto de datos 14 del apéndice B). Se prueba un nuevo equipo con la intención de mejorar la calidad al reducir la variación. Se obtiene una muestra aleatoria simple de 10 monedas de 1 centavo acuñadas con el equipo nuevo. Una gráfica cuantílica normal y un histograma indican que los pesos provienen de una población distribuida normalmente, y la muestra tiene una desviación estándar de 0.0125 g. Utilice los resultados muestrales para construir un estimado de un intervalo de confianza del 95% de σ , la desviación estándar de los pesos de monedas de 1 centavo fabricadas con el equipo nuevo. Con base en los resultados, ¿parece que el equipo nuevo sirve para reducir la variación de los pesos?

SOLUCIÓN

REQUISITO ✓ Primero verificamos si se satisfacen los requisitos. Se indicó que se trata de una muestra aleatoria simple. Con base en las descripciones del histograma y de la gráfica cuantílica normal, también se satisface el requisito de una distribución normal. Por lo tanto, ambos requisitos se cumplen. (Esta verificación de los requisitos es el paso 1 en el proceso del cálculo de un intervalo de confianza de σ , de manera que procedemos al paso 2). ✓

Paso 2: Utilizando $n - 1$ grados de libertad, ahora calculamos los valores críticos de χ^2 . El tamaño muestral es $n = 10$, de manera que hay 9 grados de libertad. Nos remitimos a la tabla A-4, al renglón correspondiente a 9 grados de libertad, y revisamos las columnas con áreas de 0.975 y 0.025. (Para un nivel de confianza del 95%, dividimos $\alpha = 0.05$ entre las dos colas de la distribución chi cuadrada y nos remitimos a los valores de 0.975 y 0.025 a lo largo del renglón superior de la tabla A-4). Los valores críticos de χ^2 son $\chi_I^2 = 2.700$ y $\chi_D^2 = 19.023$ (véase la figura 7-10).

Paso 3: Usando los valores críticos de 2.700 y 19.023, la desviación estándar muestral de $s = 0.0125$ y el tamaño muestral de 10, construimos el intervalo de confianza del 95% al evaluar lo siguiente:

$$\frac{(10 - 1)(0.0125)^2}{19.023} < \sigma^2 < \frac{(10 - 1)(0.0125)^2}{2.700}$$

Paso 4: La evaluación de la expresión anterior produce $0.0000739237 > \sigma^2 < 0.000520833$. El cálculo de la raíz cuadrada de cada parte (antes de redondear) y el redondeo posterior con cuatro decimales produce $0.0086 \text{ g} < \sigma < 0.0228 \text{ g}$.

INTERPRETACIÓN Con base en este resultado, tenemos una confianza del 95% de que los límites de 0.0086 g y 0.0228 g contienen el valor real de σ . Observe que este intervalo incluye la desviación estándar de 0.0165 g para los

continúa

pesos de las monedas que se fabrican actualmente. No parece que el nuevo equipo reduzca significativamente la variación. Aun cuando la desviación estándar de la muestra (0.0125 g) es menor que la desviación estándar actual de 0.0165 g, no es lo suficientemente baja para ser significativa. Con base en los datos disponibles, parece que el equipo nuevo no es efectivo.

El intervalo de confianza $0.0086 < \sigma < 0.0228$ también se expresa como $(0.0086, 0.0228)$, pero la forma de $s \pm E$ no puede utilizarse porque el intervalo de confianza no tiene s en su parte central.

Fundamentos Ahora explicamos por qué los intervalos de confianza para σ y σ^2 tienen las formas que acabamos de dar. Si obtenemos muestras de tamaño n de una población con varianza σ^2 , la distribución de los valores $(n - 1)s^2/\sigma^2$ será como se observa en la figura 7-11. Para una muestra aleatoria simple, existe una probabilidad de $1 - \alpha$ de que el estadístico $(n - 1)s^2/\sigma^2$ quede entre los valores críticos de χ_I^2 y χ_D^2 . En otras palabras (y símbolos), existe una probabilidad de $1 - \alpha$ de que las dos expresiones siguientes sean verdaderas:

$$\frac{(n - 1)s^2}{\sigma^2} < \chi_D^2 \quad \text{y} \quad \frac{(n - 1)s^2}{\sigma^2} > \chi_I^2$$

Si multiplicamos las dos desigualdades anteriores por σ^2 y dividimos cada desigualdad entre el valor crítico de χ^2 apropiado, veremos que las dos desigualdades pueden expresarse en las formas equivalentes:

$$\frac{(n - 1)s^2}{\chi_D^2} < \sigma^2 \quad \text{y} \quad \frac{(n - 1)s^2}{\chi_I^2} > \sigma^2$$

Estas últimas dos desigualdades se combinan en una desigualdad:

$$\frac{(n - 1)s^2}{\chi_D^2} < \sigma^2 < \frac{(n - 1)s^2}{\chi_I^2}$$

Determinación del tamaño muestral

Los procedimientos para calcular el tamaño muestral necesario para estimar σ^2 son mucho más complejos que los procedimientos que se vieron antes para las medias y las proporciones. En vez de utilizar procedimientos muy complicados, usaremos la tabla 7-2. STATDISK también provee tamaños muestrales. Si usa STATDISK, seleccione **Analysis**, **Sample Size Determination** y luego **Estimate St Dev**. Minitab, Excel y la calculadora TI-83/84 Plus no dan tamaños muestrales de este tipo.

EJEMPLO Deseamos estimar σ , la desviación estándar de todas las temperaturas corporales. Queremos tener una confianza del 95% de que nuestro estimado está dentro del 10% del valor real de σ . ¿Qué tan grande debe ser la muestra? Suponga que la población está distribuida normalmente.

SOLUCIÓN En la tabla 7-2 podemos ver que un 95% de confianza y un error del 10% para σ corresponden a una muestra de tamaño 191. Debemos seleccionar al azar 191 valores de la población de temperaturas corporales.

Tabla 7-2

Tamaño muestral para σ^2		Tamaño muestral para σ	
Para tener una confianza del 95% de que s^2 está dentro del	del valor de σ^2 , el tamaño muestral n debe ser al menos	Para tener una confianza del 95% de que s está dentro del	del valor de σ , el tamaño muestral n debe ser al menos
1%	77,207	1%	19,204
5%	3,148	5%	767
10%	805	10%	191
20%	210	20%	47
30%	97	30%	20
40%	56	40%	11
50%	37	50%	7
Para tener una confianza del 99% de que s^2 está dentro del		Para tener una confianza del 99% de que s está dentro del	
Para tener una confianza del 99% de que s^2 está dentro del	del valor de σ^2 , el tamaño muestral n debe ser al menos	Para tener una confianza del 99% de que s está dentro del	del valor de σ , el tamaño muestral n debe ser al menos
1%	133,448	1%	33,218
5%	5,457	5%	1,335
10%	1,401	10%	335
20%	368	20%	84
30%	171	30%	37
40%	100	40%	21
50%	67	50%	13

Uso de la tecnología para intervalos de confianza

STATDISK Primero obtenga los estadísticos descriptivos y verifique que la distribución sea normal utilizando un histograma o una gráfica cuantil normal. Después, seleccione **Analysis** del menú principal, luego

seleccione **Confidence Intervals** y **Population StDev**. Proceda a ingresar los datos requeridos.

MINITAB Con Minitab Release 14 ingrese los datos en la columna C1, haga clic en **Stat**, luego en **Basic Statistics** y seleccione **Graphical Summary** (o **Display Descriptive Statistics** en versiones anteriores). Ingrese C1 en el recuadro de Variables. (En versiones anteriores de Minitab, haga clic en **Graphs** y luego en **Graphical Summary**). Ingrese el nivel de confianza y presione **OK**. Los resultados incluirán un intervalo de confianza para la desviación estándar.

EXCEL Utilice DDXL. Seleccione **Confidence Intervals**, haga clic en el icono en forma de lápiz e ingrese el rango de celdas con los datos muestrales, como A1:A16. Seleccione un nivel de confianza y presione **OK**.

TI-83/84 PLUS La calculadora TI-83/84 Plus no brinda intervalos de confianza para σ ni para σ^2 directamente, pero se puede emplear el programa **S2INT**, el cual fue escrito por Michael Loyd de Henderson State University y se encuentra disponible para su descarga en www.pearsoneducacion.net/triola. El programa S2INT usa el programa ZZINNEWT, por lo que también debe instalarse. Después de almacenar los programas en la calculadora, presione la tecla **PRGM**, seleccione **S2INT** y proceda a ingresar la varianza muestral s^2 , el tamaño muestral n y el nivel de confianza (como 0.95). Presione la tecla **ENTER** y espere un momento hasta que aparezcan los límites del intervalo de confianza para σ^2 . Calcule la raíz cuadrada de los límites del intervalo de confianza si desea una estimación de σ .



7-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Interpretación de intervalos de confianza.** Consultamos las estaturas de mujeres del conjunto de datos 1 en el apéndice B y utilizamos la desviación estándar de la muestra ($s = 2.741$ in) para obtener el siguiente estimado del intervalo de confianza del 95% para la desviación estándar de las estaturas de todas las mujeres: $2.25 \text{ in} < \sigma < 3.52 \text{ in}$. Redacte un enunciado que interprete correctamente ese intervalo de confianza.
- Expresión de intervalos de confianza.** El intervalo de confianza del ejercicio 1 también se expresa como $(2.25, 3.52)$, pero no como 2.885 ± 0.635 . Dado que 2.885 ± 0.635 da por resultado los valores 2.25 y 3.52, ¿por qué no podemos expresar el intervalo de confianza como 2.885 ± 0.635 ?
- Interpretación de intervalos de confianza.** Para cada una de las 50 entidades de Estados Unidos, un investigador obtiene una muestra aleatoria de las deudas de tarjeta de crédito y calcula la media para obtener 50 valores representativos. Luego, utiliza las 50 medias muestrales para construir un estimado del intervalo de confianza de σ . ¿El resultado es un estimado de la desviación estándar de las deudas de tarjeta de crédito de todos los estudiantes estadounidenses de estadística? ¿Por qué?
- Estimadores sin sesgo.** ¿Qué es un estimador sin sesgo? ¿La varianza muestral es un estimador sin sesgo de la varianza poblacional? ¿La desviación estándar muestral es un estimador sin sesgo de la desviación estándar poblacional?

Cálculo de valores críticos. En los ejercicios 5 a 8, calcule los valores críticos χ_I^2 y χ_D^2 correspondientes al nivel de confianza y tamaño muestral dados.

5. 95%; $n = 27$

6. 95%; $n = 7$

7. 99%; $n = 41$

8. 90%; $n = 91$

Cálculo de un intervalo de confianza. En los ejercicios 9 a 12, utilice el nivel de confianza y los datos muestrales indicados para calcular un intervalo de confianza para la desviación estándar poblacional σ . En cada caso, suponga que se obtuvo una muestra aleatoria simple de una población que tiene una distribución normal.

- Salarios de graduados universitarios que tomaron un curso de estadística en la universidad: confianza del 95%; $n = 41$, $\bar{x} = \$67,200$, $s = \$18,277$.
- Las velocidades de conductores multados en una zona con límite de velocidad de 55 mi/h: confianza del 95%; $n = 90$, $\bar{x} = 66.2$ mi/h, $s = 3.4$ mi/h.
- Calificaciones de crédito de FICO (Fair, Isacc and Company) de solicitantes de tarjetas de crédito: confianza del 99%; $n = 70$, $\bar{x} = 688$, $s = 68$.
- Cantidades perdidas por jugadores que tomaron un autobús a un casino de Atlantic City: confianza del 99%; $n = 40$, $\bar{x} = \$189$, $s = \$87$.

Determinación del tamaño muestral. En los ejercicios 13 a 16, suponga que cada muestra es aleatoria simple y que se obtuvo de una población distribuida normalmente.

- Calcule el tamaño muestral mínimo que se necesita para tener una confianza del 95% de que la desviación estándar muestral s está dentro del 5% de σ .
- Calcule el tamaño muestral mínimo que se necesita para tener una confianza del 95% de que la desviación estándar muestral s está dentro del 20% de σ .
- Calcule el tamaño muestral mínimo que se necesita para tener una confianza del 99% de que la varianza muestral está dentro del 10% de la varianza poblacional. ¿Resulta práctico un tamaño muestral como éste para la mayoría de los casos?

16. Calcule el tamaño muestral mínimo que se necesita para tener una confianza del 95% de que la varianza muestral está dentro del 30% de la varianza poblacional.

Cálculo de intervalos de confianza. En los ejercicios 17 a 24, suponga que cada muestra es aleatoria simple y que se obtuvo de una población con una distribución normal.

17. **Pesos al nacer.** En un estudio de los efectos sobre los bebés que tiene el consumo de cocaína durante el embarazo, se obtuvieron los siguientes datos muestrales de pesos al nacer: $n = 190$ g, $\bar{x} = 2700$ g, $s = 645$ g (según datos de “Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure”, de Singer *et al.*, *Journal of American Medical Association*, vol. 291, núm. 20). Utilice los datos muestrales para construir un estimado del intervalo de confianza del 95% para la desviación estándar de todos los pesos al nacer de hijos de madres que consumieron cocaína durante el embarazo. (Como la tabla A-4 incluye un máximo de 100 grados de libertad, y se requieren 189, utilice los siguientes valores críticos obtenidos de STATDISK: $\chi^2_L = 152.8222$ y $\chi^2_D = 228.9638$). Con base en el resultado, ¿parece que la desviación estándar difiere de la desviación estándar de 696 g de los pesos al nacer de hijos de madres que no consumieron cocaína durante el embarazo?
18. **Acuñación de monedas de 25 centavos.** En la actualidad, las monedas de 25 centavos de dólar se acuñan con pesos que tienen una media de 5.670 g y una desviación estándar de 0.062 g. Se está probando un nuevo equipo con la intención de mejorar la calidad al reducir la variación. Se obtiene una muestra aleatoria simple de 24 monedas acuñadas con el equipo nuevo; la muestra tiene una desviación estándar de 0.049 g. Utilice los resultados muestrales para construir un estimado del intervalo de confianza del 95% de σ , la desviación estándar de los pesos de las monedas de 25 centavos acuñadas con el equipo nuevo. Con base en el intervalo de confianza, ¿parece que el nuevo equipo produce una desviación estándar claramente menor que la desviación estándar de 0.062 g del antiguo equipo? Con base en los resultados, ¿parece que el nuevo equipo sirve para reducir la variación de los pesos?
19. **Temperatura corporal.** El conjunto de datos 2 del apéndice B incluye 106 temperaturas corporales, donde $\bar{x} = 98.20^\circ\text{F}$ $s = 0.62^\circ\text{F}$. Utilice los estadísticos muestrales para construir un estimado del intervalo de confianza del 99% para la desviación estándar de la temperatura corporal de todos los seres humanos saludables. Con base en el resultado, ¿podemos concluir con seguridad que la desviación estándar poblacional es menor que 2.10°F ? (Si la desviación estándar poblacional es 2.10 o mayor, la variación es lo suficientemente grande para que la media muestral de 98.20°F no difiera de 98.6°F por una cantidad significativa).
20. **Ritmos cardiacos al trabajar con pala.** Puesto que, al parecer, las muertes por problemas cardíacos se incrementan después de fuertes nevadas, se diseñó un experimento para comparar las demandas cardíacas al retirar la nieve con la ayuda de una pala con las que se registran al utilizar una máquina eléctrica para remover la nieve. Diez sujetos retiraron montones de nieve utilizando ambos métodos, y se registraron sus frecuencias cardíacas máximas (latidos por minuto) durante ambas actividades. Se obtuvieron los resultados que se muestran a continuación (según datos de “Cardiac Demands of Heavy Snow Shoveling”, de Franklin *et al.*, *Journal of American Medical Association*, vol. 273, núm. 11). Frecuencias cardíacas máximas de retirar la nieve con una pala: $n = 10$, $\bar{x} = 175$, $s = 15$. Frecuencias cardíacas máximas utilizando un aparato eléctrico para retirar la nieve: $n = 10$, $\bar{x} = 124$, $s = 18$.
- Construya un estimado del intervalo de confianza del 95% de la desviación estándar poblacional σ para las personas que retiran la nieve con una pala.
 - Construya un estimado del intervalo de confianza del 95% de la desviación estándar poblacional σ de las personas que usan el aparato eléctrico para retirar la nieve.
 - Compare los resultados. Al parecer, ¿la variación difiere en los dos grupos?

- 21. Calificación de crédito.** Cuando los consumidores solicitan un crédito, su crédito se califica utilizando puntuaciones FICO (Fair, Isaac, and Company). A continuación se presentan las calificaciones de crédito de una muestra de solicitantes de préstamos para adquirir un automóvil. Utilice los datos muestrales para construir un intervalo de confianza del 99% para la desviación estándar de las calificaciones FICO de todos los solicitantes de crédito.

661 595 548 730 791 678 672 491 492 583 762 624 769 729 734 706

- 22. El mamífero más pequeño del mundo.** El mamífero más pequeño del mundo es el murciélagos abejorro, también conocido como murciélagos nariz de cochino (o *Crassonycteris thonglongyai*). Estos murciélagos apenas alcanzan el tamaño de un abejorro grande. A continuación se presentan los pesos (en gramos) de una muestra de esos murciélagos. Construya un estimado del intervalo de confianza del 95% de la desviación estándar de los pesos de todos esos murciélagos.

1.7 1.6 1.5 2.0 2.3 1.6 1.6 1.8 1.5 1.7 2.2 1.4 1.6 1.6 1.6

- 23. Control de plomo en el aire.** En la lista de abajo se incluyen cantidades de plomo medidas en el aire (en microgramos por metro cúbico o $\mu\text{g}/\text{m}^3$). La Environmental Protection Agency estableció un estándar de calidad del aire para el plomo de $1.5 \mu\text{g}/\text{m}^3$. Las mediciones que se presentan abajo se registraron en el edificio 5 del World Trade Center en diferentes días posteriores a la destrucción causada por los ataques terroristas del 11 de septiembre de 2001. Después del colapso de los dos edificios del World Trade Center, hubo una considerable preocupación acerca de la calidad del aire. Utilice los valores dados para construir un estimado del intervalo de confianza del 95% de la desviación estándar de las cantidades de plomo en el aire. ¿Hay algo en este conjunto de datos que sugiera que el intervalo de confianza no es muy bueno? Explique.

5.40 1.10 0.42 0.73 0.48 1.10

- 24. a. Comparación de filas de espera.** Los valores listados son tiempos de espera (en minutos) de clientes en el Jefferson Valley Bank, donde los clientes se forman en una sola fila atendida por tres ventanillas. Construya un intervalo de confianza del 95% para la desviación estándar poblacional σ .

6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7

- b.** Los valores listados son tiempos de espera (en minutos) de clientes en el Bank of Providence, donde los clientes se forman en una de tres filas atendidas por tres ventanillas distintas. Construya un intervalo de confianza del 95% para la desviación estándar poblacional σ .

4.2 5.4 5.8 6.2 6.7 7.7 7.7 8.5 9.3 10.0

- c.** Interprete los resultados obtenidos en los incisos *a*) y *b*). ¿Los intervalos de confianza sugieren una diferencia en la variación de los tiempos de espera? ¿Qué sistema parece ser mejor: el sistema de una sola fila o el de filas múltiples?

- 25. Datos del apéndice B sobre el índice de masa corporal (IMC).** Remítase al conjunto de datos 1 en el apéndice B y utilice los datos muestrales.

- a.** Construya un estimado del intervalo de confianza del 99% de la desviación estándar de los índices de masa corporal para hombres.
- b.** Construya un estimado del intervalo de confianza del 99% de la desviación estándar de los índices de masa corporal para mujeres.
- c.** Compare e interprete los resultados.

- 26. Datos del apéndice B sobre pesos de monedas de 25 centavos.** Remítase al conjunto de datos 14 del apéndice B y utilice los datos muestrales.

- a.** Construya un estimado del intervalo de confianza del 99% de la desviación estándar para los pesos de monedas de 25 centavos acuñadas después de 1964.

- b. Construya un estimado del intervalo de confianza del 99% de la desviación estándar para los pesos de monedas de 25 centavos de plata fabricadas antes de 1964.
- c. Compare e interprete los resultados.

7-5 MÁS ALLÁ DE LO BÁSICO

- 27. Cálculo de valores críticos.** En la construcción de intervalos de confianza para σ o σ^2 , utilizamos la tabla A-4 para encontrar los valores críticos χ_I^2 y χ_D^2 , pero esta tabla sólo se aplica a casos en los que $n \leq 101$, de manera que el número de grados de libertad es 100 o menor. Para números de grados de libertad más grandes, podemos aproximar χ_I^2 y χ_D^2 utilizando

$$\chi^2 = \frac{1}{2} [\pm z_{\alpha/2} + \sqrt{2k - 1}]^2$$

donde k es el número de grados de libertad y $z_{\alpha/2}$ es la puntuación crítica z que se describió por primera vez en la sección 7-2. Se utilizó STATDISK para calcular los valores críticos de 189 grados de libertad, con un nivel de confianza del 95%, y esos valores críticos se incluyen en el ejercicio 17. Utilice la aproximación que se presenta aquí para calcular los valores críticos y compare los resultados con los obtenidos por medio de STATDISK.

- 28. Cálculo del mejor estimador.** Señalamos que los valores de s^2 tienden a producir errores más pequeños al estar más cerca más de σ^2 que otras medidas de variación sin sesgo. Ahora consideremos el estimador sesgado de $(n - 1)s^2/(n + 1)$. Dada la población de valores $\{2, 3, 7\}$, utilice el valor de σ^2 y nueve muestras diferentes posibles de tamaño $n = 2$ (con un muestreo hecho con reemplazo) para lo siguiente.
- a. Calcule s^2 para cada una de las nueve muestras y luego obtenga el error $s^2 - \sigma^2$ para cada muestra; después eleve al cuadrado los errores y calcule la media de esos cuadrados. El resultado es el valor del error del cuadrado medio.
 - b. Calcule $(n - 1)s^2/(n + 1)$ para cada una de las nueve muestras y luego obtenga el error $(n - 1)s^2/(n + 1) - \sigma^2$ para cada muestra; después eleve al cuadrado los errores y calcule la media de los cuadrados. El resultado es el error del cuadrado medio.
 - c. El error del cuadrado medio se puede emplear para medir qué tan cercano es un estimador al parámetro poblacional. ¿Qué estimador es mejor para producir el menor error del cuadrado medio? ¿Ese estimador es sesgado o no sesgado?

Repaso

Las dos actividades principales de la estadística inferencial son la estimación de parámetros poblacionales y la prueba de aseveraciones que se hacen acerca de parámetros poblacionales. En este capítulo estudiamos métodos básicos para calcular *estimados* de proporciones, medias y varianzas poblacionales, y desarrollamos procedimientos para calcular:

- el estimado puntual
- el intervalo de confianza
- el tamaño muestral requerido

Analizamos el estimado puntual (o estimado de un solo valor) y obtuvimos las siguientes conclusiones:

- Proporción: el mejor estimado puntual de p es \hat{p} .
- Media: el mejor estimado puntual de μ es \bar{x} .
- Variación: el valor de s suele emplearse como un estimado puntual de σ , aun cuando éste es un estimado sesgado. Además, s^2 es el mejor estimado puntual de σ^2 .

Puesto que los estimados puntuales anteriores consisten en valores individuales, tienen la grave desventaja de no revelar qué tan buenos son; por eso, suelen utilizarse intervalos de confianza (o estimados de intervalo) como estimados más reveladores y útiles. También consideramos formas que permiten determinar los tamaños muestrales necesarios para estimar parámetros dentro de márgenes de error dados. En este capítulo también se estudiaron las distribuciones t de Student y chi cuadrada. Debemos ser cuidadosos para utilizar la distribución de probabilidad correcta para cada conjunto de circunstancias. En este capítulo se emplearon los siguientes criterios para seleccionar la distribución apropiada:

Intervalo de confianza para la proporción p : Utilice la distribución *normal* (considerando que los supuestos requeridos se satisfacen y que haya al menos 5 éxitos y al menos 5 fracasos para que se pueda usar la distribución normal como aproximación de la distribución binomial).

Intervalo de confianza para μ : Véase la figura 7-6 o la tabla 7-1 para elegir entre la distribución *normal* o t (o concluir que no se aplica ninguna).

Intervalo de confianza para σ o σ^2 : Utilice la distribución *chi cuadrada* (considerando que los supuestos requeridos se satisfacen).

Para aplicar los procedimientos del intervalo de confianza y el tamaño muestral de este capítulo, es muy importante verificar que los supuestos requeridos se satisfacen. Si no es así, entonces no podemos utilizar los métodos de este capítulo y tal vez necesitemos emplear otros métodos, como el método *bootstrap* que se describe en el proyecto tecnológico al final de este capítulo, o métodos no paramétricos, como los que se analizan en el capítulo 13.

En este capítulo también se examinó el concepto de valores críticos. Es importante aprender a obtener valores críticos para la distribución normal, t y chi cuadrada, ya que los emplearemos en los siguientes capítulos.

Conocimientos estadísticos y pensamiento crítico

- Valores críticos.** Al trabajar con una distribución normal, se obtiene el valor crítico de $z = 1.96$ para un nivel de confianza del 95%. ¿Cuál es la relación entre $z = 1.96$ y el nivel de confianza del 95%?
- Intervalo de confianza.** Al tratar de estimar el peso medio de la basura desechada por las familias en una semana, obtenemos los siguientes resultados muestrales: $n = 62$, $\bar{x} = 27.44$ lb, $s = 12.46$ lb. Sabiendo que la media muestral es un estimador sin sesgo de la media poblacional, concluimos de manera correcta que nuestro mejor estimado de μ es 27.44 lb. ¿Por qué entonces necesitamos un intervalo de confianza? ¿Qué nos indica el intervalo de confianza que está faltando en el estimado de 27.44 lb?
- Margen de error.** Un periódico reporta resultados de una encuesta y afirma que “el 65% de los encuestados se mostraron a favor de la propuesta, con un margen de error de ± 3 puntos porcentuales”. ¿Qué intervalo de confianza sugiere esa afirmación?
- Interpretación de un intervalo de confianza.** Utilice el intervalo de confianza calculado en el ejercicio 3. Por lo general, los medios de comunicación no reportan el nivel de confianza, pero suponiendo que el intervalo de confianza se calculó utilizando un nivel de confianza del 95%, redacte un enunciado que interprete correctamente el intervalo de confianza.

Ejercicios de repaso

- Política de servicio de alcohol.** En una encuesta de Gallup de 1004 adultos, el 93% indicó que los restaurantes y bares deberían negar el servicio a los clientes que han bebido demasiado. Construya un estimado del intervalo de confianza del 95% para el porcentaje de todos los adultos que consideran que los restaurantes y bares deben negar el servicio a los clientes que han bebido demasiado. Redacte un enunciado que interprete el intervalo de confianza.
- Determinación del tamaño de muestra para una encuesta.** Considere la encuesta descrita en el ejercicio 1. Si usted planea realizar una nueva encuesta para confirmar que el porcentaje continúa siendo correcto, ¿cuántos adultos seleccionados al azar deberá encuestar si quiere un nivel de confianza del 95% de que el margen de error es de cuatro puntos porcentuales?
- Máquina expendedora.** Las especificaciones para máquinas expendedoras, establecidas por la empresa Newton Machine Company, requieren que las máquinas despachen cantidades de café con una media de 12 oz. A continuación se incluyen las cantidades de café (en onzas) elegidas al azar de diferentes máquinas. Utilice esos resultados muestrales para construir un intervalo de confianza del 95% para la cantidad media de café de todas las tazas servidas. ¿Este intervalo de confianza sugiere que las máquinas funcionan adecuadamente? ¿Hay algo más en los datos que sugiera que existe un problema?

$$\begin{array}{ccccccc} 11.5 & 10.8 & 9.7 & 13.0 & 11.5 & 11.1 \\ 13.2 & 11.1 & 11.1 & 12.6 & 6.8 & 11.4 \\ 9.0 & 10.6 & 10.1 & 10.8 \end{array} \left. \begin{array}{l} n = 16 \\ \bar{x} = 10.89 \\ s = 1.56 \end{array} \right\}$$

- Intervalo de confianza para σ .** Utilice los mismos datos muestrales del ejercicio 3 para construir un estimado del intervalo de confianza del 95% para σ . Se están considerando nuevas especificaciones para controlar la variación de las cantidades de café que se despacha. Se desea que casi todas las cantidades despachadas estén dentro de 0.5 oz de 12 oz y, utilizando la regla práctica del intervalo, esto sugiere que la desviación estándar debe ser de 0.25 oz. Con base en el intervalo de confianza, ¿el valor de 0.25 oz es un valor factible de la desviación estándar poblacional? ¿La máquina requiere modificaciones para reducir la variación?
- Tamaño de muestra.** Usted fue contratado por un consorcio de granjeros lecheros para realizar una encuesta sobre el consumo de leche.
 - Si usted desea estimar el porcentaje de adultos que beben leche diariamente, ¿cuántos adultos debe encuestar si desea tener una confianza del 95% de que su porcentaje muestral tiene un error no mayor de dos puntos porcentuales?
 - Si usted desea estimar la cantidad media de leche que consumen diariamente los adultos, ¿cuántos adultos debe encuestar si desea tener una confianza del 95% de que su media muestral tiene un error no mayor de 0.5 oz? (De acuerdo con los resultados de un estudio piloto, suponga que $\sigma = 8.7$ oz).
 - Si usted planea obtener los estimados descritos en los incisos a) y b) con una sola encuesta que incluye a varias preguntas, ¿cuántas personas debe encuestar?
- Estimación de tiempo de propiedad de automóviles.** Un distribuidor de partes automotrices de la NAPA quiere información acerca de cuánto tiempo planean conservar sus automóviles los propietarios. Una muestra aleatoria simple de 25 propietarios de automóviles dio por resultado $\bar{x} = 7.01$ años y $s = 3.74$ años (según datos de una encuesta de Roper). Suponga que la muestra se obtuvo de una población distribuida normalmente.
 - Calcule un estimado del intervalo de confianza del 95% de la media poblacional.
 - Calcule un estimado del intervalo de confianza del 95% de la desviación estándar poblacional.

continúa

- c. Si pasaron varios años y usted quiere realizar una nueva encuesta para estimar la cantidad media del tiempo que los propietarios de automóviles planean conservar sus autos, ¿cuántos propietarios de automóviles seleccionados al azar debe encuestar? Suponga que desea tener una confianza del 99% de que la media muestral está dentro de 0.25 años (o tres meses) de la media poblacional, y también suponga que $\sigma = 3.74$ años (de acuerdo con el último resultado).
- d. Cuando se realiza la encuesta descrita en el inciso c), usted descubre que el proceso de encuesta puede simplificarse a un costo sustancialmente reducido si utiliza una base de datos disponibles de personas que compraron un automóvil General Motors durante los 10 últimos años. ¿Se obtendrían buenos resultados de esta población?
- 7. Tabaquismo y educación universitaria.** La industria tabacalera supervisa de cerca todas las encuestas relacionadas con el tabaquismo. Una encuesta reveló que, de 785 sujetos elegidos al azar que completaron cuatro años de estudios universitarios, el 18.3% fuma (según datos de la American Medical Association).
- Construya un intervalo de confianza del 98% para el porcentaje real de fumadores que se registra entre todas las personas que completaron cuatro años de estudios universitarios.
 - Con base en los resultados del inciso a), ¿parece que la tasa de tabaquismo de los individuos que completaron cuatro años de estudios universitarios difiere sustancialmente de la tasa del 27% que se registra entre la población general?
- 8. Costos hospitalarios de accidentes.** Se realizó un estudio para estimar los costos hospitalarios de las víctimas de accidentes que usaban cinturones de seguridad. Veinte casos elegidos al azar tienen una distribución con forma de campana, con una media de \$9,004 y una desviación estándar de \$5,629 (según datos del Departamento del Transporte de EUA).
- Construya un intervalo de confianza del 99% para la media de todos esos costos.
 - Si usted fuera el gerente de una compañía de seguros que ofrece tasas bajas a los conductores que usan el cinturón de seguridad y quisiera un estimado conservador para el escenario del peor caso, ¿qué cantidad debería utilizar como el posible costo hospitalario para una víctima de accidente que usa cinturón de seguridad?

Ejercicios de repaso acumulativo

- 1. Análisis de pesos de supermodelos.** Algunas veces se critica a las supermodelos porque sus bajos pesos fomentan hábitos de alimentación no saludables entre las mujeres jóvenes. A continuación se presentan los pesos (en libras) de nueve supermodelos seleccionadas al azar.

125 (Taylor)	119 (Auermann)	128 (Schiffer)	128 (MacPherson)
119 (Turlington)	127 (Hall)	105 (Moss)	123 (Mazza)
115 (Hume)			

Calcule lo siguiente:

- a. Media
- b. Mediana
- c. Moda
- d. Punto medio del rango
- e. Rango
- f. Varianza
- g. Desviación estándar
- h. Q_1
- i. Q_2
- j. Q_3
- k. ¿Cuál es el nivel de medición de estos datos (nominal, ordinal, de intervalo, de razón)?
- l. Construya una gráfica de cuadro para los datos.
- m. Construya un intervalo de confianza del 99% para la media poblacional.
- n. Construya un intervalo de confianza del 99% para la desviación estándar σ .
- o. Calcule el tamaño muestral necesario para estimar la media del peso de todas las supermodelos, con una confianza del 99% de que la media muestral tenga un error

continúa

- que no rebase las 2 lb. Utilice la desviación estándar muestral s del inciso g) como un estimado de la desviación estándar poblacional σ .
- p. Cuando se seleccionan al azar mujeres de la población general, sus pesos están distribuidos normalmente con una media de 143 lb y una desviación estándar de 29 lb (según datos de la National Health Examination Survey). Con base en los valores muestrales dados, ¿parece que los pesos de las supermodelos son sustancialmente menores que los pesos de mujeres seleccionadas al azar? Explique.
2. **Trastorno recesivo del cromosoma X.** Un experto en genética determinó que ciertas parejas tienen una probabilidad de 0.25 de que cualquiera de sus hijos presente un trastorno recesivo del cromosoma X.
- Calcule la probabilidad de que entre 200 de esos descendientes, al menos 65 presenten el trastorno recesivo del cromosoma X.
 - Un estudio posterior de 200 nacimientos reales reveló que 65 de los hijos presentaron el trastorno recesivo del cromosoma X. Con base en estos resultados muestrales construya un intervalo de confianza del 95% para la proporción de todos esos descendientes que presentan el trastorno.
 - Con base en los incisos a) y b), ¿parece que la determinación del experto de una probabilidad de 0.25 es correcta? Explique.
3. **Estimación de la asistencia a parques temáticos.** Cada año se gastan miles de millones de dólares en parques temáticos propiedad de Disney, Universal Studios, Sea World, Busch Gardens y otros. Se encuestó a 111 viajeros que realizaron una visita a un parque temático, así como a 1122 individuos que hicieron viajes sin visitar un parque temático (según datos de la Travel Industry Association of America)
- Calcule el estimado puntual del porcentaje de todas las personas que visitaron un parque temático cuando hicieron un viaje.
 - Calcule un estimado del intervalo de confianza del 95% del porcentaje de todas las personas que visitaron un parque temático cuando hicieron un viaje.
 - La encuesta se realizó con personas que viajaron, pero no se dio información acerca del porcentaje de personas que hicieron viajes de placer. Si usted quiere estimar el porcentaje de adultos que hacen un viaje de placer en un año, ¿a cuántas personas debe entrevistar si quiere tener una confianza del 99% de que su porcentaje muestral está dentro de 2.5 puntos porcentuales del porcentaje correcto de la población?

Actividades de cooperación en equipo

1. **Actividad fuera de clase** Reúna datos muestrales y utilice los métodos de este capítulo para construir estimados de intervalos de confianza de parámetros poblacionales. Aquí están algunas sugerencias de parámetros:
- La proporción de estudiantes de su universidad que pueden identificar correctamente al presidente, al vicepresidente y al secretario de Estado.
 - La proporción de estudiantes de su universidad que son mayores de 18 años y que están registrados en el padrón electoral.
 - La media de la edad de los estudiantes de tiempo completo en su universidad.
 - La proporción de vehículos motorizados en su región que son automóviles.
2. **Actividad en clase** Sin utilizar ningún aparato de medición, cada estudiante debe dibujar una línea de lo que considere que son 3 pulgadas de largo y otra línea de 3 centímetros de largo. Después utilicen reglas para medir y registren las longitudes de las líneas. Calcule las medias y las desviaciones estándar de los dos conjuntos de longitudes. Utilice los datos muestrales para construir un intervalo de confianza para la longitud de la línea considerada de

3 in, y luego haga lo mismo para la longitud de la línea considerada de 3 cm. ¿Los límites del intervalo de confianza realmente contienen la longitud correcta? Compare los resultados. ¿Los cálculos de la línea de 3 in parecen ser más exactos que los de la línea de 3 cm?

- 3. Actividad en clase** Suponga que un método de selección del género puede afectar la probabilidad de que un bebé sea niña, de manera que la probabilidad es 1/4. Cada estudiante debe simular 20 nacimientos al sacar 20 naipes de un mazo barajado. Reemplace cada carta después de sacarla y vuelva a barajar. Considere que los corazones son niñas y que el resto de las cartas son niños. Después de hacer 20 selecciones y registrar “el género” de los bebés, construya un estimado del intervalo de confianza de la proporción de niñas. Al parecer, ¿el resultado sirve para identificar el valor real de la proporción poblacional? (Si no dispone de una baraja, utilice algún otro recurso para simular los nacimientos, como el generador de números aleatorios de una calculadora o los dígitos de números telefónicos o de números del seguro social).
- 4. Actividad fuera de clase** En grupos de tres o cuatro estudiantes, acudan a la biblioteca y reúnan la muestra de las fechas de edición de libros (según las fechas de los derechos de autor). Planee y describa el plan de muestreo, ejecute el procedimiento de muestreo y luego use los resultados para construir un estimado del intervalo de confianza de la antigüedad media de todos los libros de la biblioteca.
- 5. Actividad en clase** Cada estudiante debe anotar un estimado de la edad del presidente actual de Estados Unidos. Todos los estimados se reúnen y se calcula la media y la desviación estándar muestrales. Después, se utilizan los resultados muestrales para construir un intervalo de confianza. ¿Los límites del intervalo de confianza incluyen la edad correcta del presidente?
- 6. Actividad en clase** Se debe diseñar un proyecto de clase para realizar una prueba en la que cada estudiante beba un poco de Coca-Cola y un poco de Pepsi. Luego se pide a cada estudiante que identifique cuál de ellas es Coca-Cola. Después de reunir todos los resultados, analice la afirmación de que la tasa de éxito es mejor que la tasa que se esperaría al hacer conjeturas.
- 7. Actividad en clase** Cada estudiante debe estimar la longitud del salón de clases. Los valores deben basarse

en estimaciones visuales, sin tomar medidas. Una vez que se han reunido las estimaciones, construya un intervalo de confianza y luego mida la longitud del salón. ¿El intervalo de confianza contiene la longitud real del salón de clases? ¿Existe una “sabiduría colectiva” por medio de la cual la media de la clase es aproximadamente igual a la longitud real del salón?

- 8. Actividad en clase** Divida la clase en grupos de tres o cuatro estudiantes. Examine una revista reciente como *Time* o *Newsweek* y calcule la proporción de páginas que incluyen anuncios. Con base en los resultados, construya un estimado del intervalo de confianza del 95% para el porcentaje de todas las páginas que tienen anuncios. Compare los resultados con otros grupos.
- 9. Actividad en clase** Divida la clase en grupos de dos estudiantes. Primero calcule el tamaño muestral requerido para estimar la proporción de veces que una moneda cae en cara al lanzarla, suponiendo que desea tener una confianza del 80% de que la proporción muestral está dentro de 0.08 de la verdadera proporción poblacional. Luego lance una moneda las veces necesarias y registre sus resultados. ¿Qué porcentaje de estos intervalos de confianza debe contener realmente el valor verdadero de la proporción poblacional, que sabemos que es $p = 0.5$? Verifique este último resultado comparando su intervalo de confianza con los intervalos de confianza calculados en los otros grupos.
- 10. Actividad fuera de clase** Identifique un tema de interés general y coordínense entre todos los miembros de la clase para realizar una encuesta. En vez de utilizar una encuesta “científica” con principios sólidos de selección aleatoria, utilice una muestra de conveniencia consistente en individuos que estén al alcance, como amigos, parientes y otros estudiantes. Analice e interprete los resultados. Identifique la población, describa las desventajas de utilizar una muestra de conveniencia y trate de determinar en qué diferiría de una muestra de sujetos seleccionados al azar de la población.
- 11. Actividad fuera de clase** Cada estudiante debe encontrar un artículo en una revista científica que incluya un intervalo de confianza como los que analizamos en este capítulo. Redacte un breve informe que describa el intervalo de confianza y su papel en el contexto del artículo.

Proyecto tecnológico

Muestreo repetido bootstrap El método *bootstrap* se emplea para construir intervalos de confianza en situaciones en las que los métodos tradicionales no pueden (o no deben) utilizarse. Por ejemplo, la siguiente muestra de 10 valores se seleccionó al azar de una población con una distribución que se aleja mucho de la normal, por lo que no es posible utilizar alguno de los métodos que requieren

una distribución normal.

2.9 564.2 1.4 4.7 67.6 4.8 51.3 3.6 18.0 3.6

Si queremos utilizar los datos muestrales anteriores para construir un estimado de un intervalo de confianza de la media poblacional μ , observamos que la muestra es pequeña y que existe un valor extremo. El método *bootstrap*, que no

necesita establecer supuestos de la población original, por lo regular requiere de una computadora para construir una población *bootstrap* replicando (duplicando) una muestra muchas veces. Podemos extraer de la muestra con reemplazo, creando así una aproximación de la población original. De esta forma, “estiramos” la muestra *bootstrap* (término que significa literalmente “punta de la bota”) para simular la población original. Utilizando los datos muestrales anteriores, construya un estimado del intervalo de confianza del 95% de la media poblacional μ con el método *bootstrap*.

Se puedan emplear varias tecnologías para este procedimiento. (Si utiliza el programa estadístico STATDISK que se encuentra en el CD que acompaña este libro, ingrese los 10 valores muestrales en la columna 1 de la ventana de datos, luego seleccione **Analysis** del menú principal y después **Bootstrap Resampling**).

- a. Desarrolle 500 muestras nuevas, cada una de tamaño 10, seleccionando 10 valores con reemplazo de los 10 valores muestrales dados arriba.
- b. Calcule las medias de las 500 muestras *bootstrap* generadas en el inciso a).
- c. Ordene las 500 medias.

- d. Obtenga los percentiles $P_{2.5}$ y $P_{97.5}$ para las medias ordenadas que resultan del paso anterior. ($P_{2.5}$ es la media de los valores decimosegundo y decimotercero de la lista ordenada de medias; $P_{97.5}$ es la media de los valores en los lugares 487 y 488 de la lista ordenada de medias). Identifique el intervalo de confianza resultante sustituyendo los valores para $P_{2.5}$ y $P_{97.5}$ en $P_{2.5} < \mu < P_{97.5}$. ¿Contiene este intervalo de confianza el valor real de μ , que es 148?

Ahora utilice el método *bootstrap* para calcular un intervalo de confianza del 95% para la desviación estándar poblacional σ . [Utilice los mismos pasos listados arriba, pero especifique *desviación estándar* (Standard Deviation) en vez de media (Mean)]. ¿El procedimiento *bootstrap* ofrece un intervalo de confianza para σ que contiene a 232.1, demostrando que el método *bootstrap* es efectivo?

Existe un paquete de cómputo especial diseñado específicamente para métodos de muestreo repetido *bootstrap*: Resampling Stats, comercializado por Resampling Stats, Inc., 612 N, Jackson St., Arlington, VA, 22201; teléfono: (703)522-2713.

De los datos a la decisión

Pensamiento crítico: ¿Qué nos indican los resultados de la encuesta “cámara vigilante”

Las encuestas se han convertido en un componente importante de la vida de los ciudadanos actuales. Las encuestas afectan directamente los programas de televisión que vemos, los productos que compramos, los líderes políticos que elegimos y la ropa que usamos. Como las encuestas ahora forman parte integral de nuestras vidas, es importante que cada ciudadano tenga la habilidad de interpretar los resultados de encuestas. En este ejercicio nos enfocamos en esta técnica de investigación.

El *Star Tribune*, un periódico de Saint Paul, Minneapolis, patrocinó una encuesta diseñada para revelar opiniones acerca de la “cámara vigilante”, que consiste en cámaras ubicadas para descubrir a los conductores que no respetan la luz roja del semáforo. Las cámaras toman fotografías de las placas de los automóviles que se pasan el alto, y los propietarios de los automóviles reciben multas de tránsito por correo. El periódico patrocinó la encuesta como respuesta a una legislación pendiente en Minnesota que aprobaría el uso de cámaras para expedir multas de tránsito.

(Agradezco a Beth Hentges por dar la información del periódico). Los encuestadores entrevistaron a 829 adultos de Minnesota y encontraron que el 51% se oponía a la legislación de la cámara vigilante.

Análisis de los datos

1. Utilice los resultados de la encuesta para construir un estimado del intervalo de confianza para el porcentaje de todos los habitantes de Minnesota que se oponen a la ley de la cámara vigilante.
2. Dado que el 51% de los 829 habitantes de Minnesota encuestados se oponen a la ley de la cámara vigilante, explique por qué sería o no adecuado que un periódico hiciera la siguiente afirmación: “Con base en los resultados de una encuesta reciente, la mayoría de los habitantes de Minnesota se oponen a la ley de la cámara vigilante”.
3. Una crítica común de las encuestas es que consideran únicamente un porcentaje muy pequeño de la población y, por lo tanto, no son exactas. ¿Una muestra de sólo 829 personas, tomada de una población de 3.4 millones de habitantes, es demasiado pequeña? Escriba una explicación de por qué el tamaño muestral de 829 es adecuado o demasiado pequeño.

4. En referencia a otra encuesta, el presidente de una compañía escribió a la Associated Press acerca de una encuesta nacional de 1223 sujetos. He aquí lo que escribió:

Cuando usted o cualquier otra persona trata de decirme a mí y a mis socios que 1223 personas representan esas opiniones y gustos aquí en Estados Unidos, ¡me pongo furioso! ¡Cómo se atreve! Cuando usted o cualquier otra persona me dice que 1223 personas representan a Estados Unidos, me parece increíble e injusto y creo que debería prohibirse.

Más adelante, el escritor de esa carta afirma que, puesto que el tamaño muestral de 1223 personas representa a 120 millones de individuos, entonces su carta representa a 98,000 individuos (120 millones divididos entre 1223) que comparten la misma perspectiva. ¿Coincide usted con esta afirmación? Redacte la respuesta que apoye o refute esta aseveración.





Proyecto de Internet

Intervalos de confianza

Los intervalos de confianza en este capítulo ilustran un aspecto importante de la ciencia de la estimación estadística. A saber, las estimaciones basadas en datos muestrales se hacen con ciertos grados de confianza. En el proyecto de Internet de este capítulo, usted utilizará intervalos de confianza para hacer una afirmación acerca de la temperatura del lugar donde vive. Visite el sitio de Internet de este libro de texto:

<http://www.pearsoneducacion.net/triola>

Localice el proyecto para este capítulo. Ahí encontrará las instrucciones sobre el uso de Internet para obtener datos de temperatura recolectados por la estación meteorológica más cercana a su casa. Con esos datos a la mano, usted construirá intervalos de confianza para las temperaturas durante diferentes períodos y tratará de obtener conclusiones acerca de los cambios de temperatura en su área. Además, aprenderá más acerca de la relación entre confianza y probabilidad.

La estadística en el trabajo

"Para la investigación y la enseñanza en el campo de la ecología, el comportamiento animal y la ecotoxicología, el conocimiento de la estadística es esencial para obtener un buen trabajo y conservarlo".



Joanna Burger

Profesora distinguida de biología en Rutgers University y miembro del Environmental and Occupational Health Sciences Institute.

Joanna Burger es docente, hace investigación y forma parte de muchos comités ambientales nacionales e internacionales que tratan con especies en peligro de extinción, contaminantes de la vida silvestre, los efectos de químicos en el comportamiento animal y los efectos de las personas en los ecosistemas.

¿Qué conceptos de la estadística utiliza en su trabajo?

Utilizo diversos métodos estadísticos, incluidos los paramétricos y los no paramétricos. Sin un firme entendimiento de la estadística yo no sería capaz de probar si los factores ambientales afectan el éxito reproductivo. Utilizo la estadística para probar hipótesis que planteo al observar animales dentro de sus ambientes naturales. Mientras que la observación nos conduce a formular hipótesis, sólo es posible responder a las preguntas mediante el uso de experimentos bien diseñados y ensayos estadísticos. Para la investigación y la enseñanza en el campo de la ecología, el comportamiento animal y la ecotoxicología, el conocimiento de la estadística es esencial para obtener un buen trabajo y conservarlo.

¿Podría dar un ejemplo específico de cómo ha usado usted la estadística en el pasado?

La estadística es muy útil para identificar factores importantes que influyen en el comportamiento animal. Los pájaros anidan en hábitat particulares, pero nos preguntamos si anidan de manera aleatoria o si eligen sitios específicos para hacer sus nidos. Esto es importante puesto que la conservación requiere conocer las necesidades de los animales para poder crear, proteger y/o manejar ese hábitat. Yo probé la hipótesis de que las golondrinas de mar comunes estaban seleccionando islas de pantanos salados específicos. Al comparar estadísticamente un amplio rango de factores ambientales (como son la altitud, el tamaño del territorio, el tipo y la densidad de la vegetación) en todas las islas con el mismo conjunto de factores en las islas que las golondrinas de mar usan para anidar, pudimos demostrar que estos animales en realidad seleccionan un conjunto de características muy específicas.

Aunque existen más de 250 islas en la bahía donde se realizó el estudio, sólo 36 reúnen los criterios que usan las golondrinas de mar. Estas aves eligen islas que son lo bastante altas para evitar las mareas en las tormentas de verano, pero lo bastante bajas para que los depredadores no puedan sobrevivir durante el invierno. Las islas que son lo suficientemente altas para evitar las mareas de las tormentas de invierno pueden tener poblaciones viables de depredadores, como zorros y mapaches, que se comerían los huevos y los polluelos de las golondrinas.

¿El conocimiento de la estadística es esencial para su trabajo?

Una firme comprensión de la estadística es absolutamente esencial para realizar investigación con humanos y animales. Con el uso de pruebas de hipótesis y análisis de regresión múltiple, es posible comenzar a identificar y evaluar los factores que afectan comportamientos, tales como la conducta de pesca y de consumo de las personas, el comportamiento de alimentación de las aves costeras y la construcción de nidos de las aves marinas.

En términos de estadística, ¿qué recomendaría a los aspirantes a obtener un empleo en su campo?

Toda persona que quiera trabajar en el campo de la biología de la conservación, la ecotoxicología, el comportamiento animal o ecológico necesita poseer una amplia gama de habilidades estadísticas. Sería recomendable tomar dos o tres cursos, incluyendo estadística general, regresión y métodos no paramétricos. La naturaleza de cada problema y las características de los datos determinan la estadística que se requiere, y uno no debe limitarse por una carencia de conocimiento de la estadística.



Prueba de hipótesis

8



- 8-1 Panorama general**
- 8-2 Fundamentos de la prueba de hipótesis**
- 8-3 Prueba de una aseveración respecto de una proporción**
- 8-4 Prueba de una aseveración respecto de una media: σ conocida**
- 8-5 Prueba de una aseveración respecto de una media:
 σ desconocida**
- 8-6 Prueba de una aseveración respecto de una desviación
estándar o de una varianza**

¿Cuál es la mejor manera de conseguir empleo?

Después de obtener un título universitario, los graduados se enfrentan a la importante tarea de conseguir empleo. Las opciones incluyen establecer redes de contacto por medio de amigos y parientes, buscar anuncios en periódicos, hacer solicitudes personalmente, realizar una búsqueda de empleo por Internet, recurrir a una agencia de empleo o a un reclutador, y consultar una agencia universitaria de colocación. Algunos de estos medios tienen mayores posibilidades de resultar fructíferos que otros. En específico, se ha visto que las redes de contacto son uno de los métodos más eficaces. Por medio de los contactos, el individuo que busca un empleo se vincula con otros e intercambia información a través de una red informal de personas.

Una encuesta reciente incluyó a 703 sujetos elegidos al azar, los cuales tenían un empleo. De ellos, el 61% dijo que había conseguido el trabajo por medio del contacto

con amigos y parientes (según datos de Taylor Nelson Sofres Intersearch). Parece que otro método eficaz son los anuncios de los periódicos, ya que el 16% de los encuestados obtuvieron el empleo por este medio. ¿Los resultados de esta encuesta son fundamento suficiente para publicar un artículo periodístico con el título “La mayoría de los empleados encuentran trabajo gracias a redes de contacto”? Algunos podrían argumentar que, aunque el 61% es mayor que el 50%, esta encuesta sólo incluye a 703 personas entre millones de empleados, por lo que no ofrece una justificación suficiente para la aseveración de que la mayoría de los empleados consiguen trabajo por medio de redes de contacto. Sin embargo, ¿los resultados de la encuesta ofrecen justificaciones suficientes para esta aseveración? En este capítulo se presentan métodos convencionales para poner a prueba de manera formal y objetiva este tipo de aseveraciones.

8-1 Panorama general

Las dos actividades principales de la estadística inferencial son el uso de datos para **1. estimar** un parámetro poblacional (como se hizo en el capítulo 7), y **2. probar** una hipótesis o afirmación con respecto a un parámetro poblacional (como se hará en este capítulo).

Definición

En estadística, una **hipótesis** es una aseveración o afirmación acerca de una propiedad de una población.

Una **prueba de hipótesis** (o **prueba de significancia**) es un procedimiento estándar para probar una aseveración acerca de una propiedad de una población.

Los siguientes son ejemplos de hipótesis que pueden someterse a prueba por medio de los procedimientos estudiados en este capítulo.

- **Negocios** El encabezado de una nota periodística afirma que la mayoría de los empleados consiguen trabajo por medio de redes de contactos.
- **Medicina** Investigadores médicos aseveran que la temperatura corporal media de adultos sanos no es igual a 98.6°F.
- **Seguridad de aeronaves** La Federal Aviation Administration afirma que el peso promedio de un pasajero de aeronave (con equipaje de mano) es mayor que las 185 libras de hace 20 años.
- **Control de calidad** Cuando se usa equipo nuevo para fabricar altímetros de aviones, los nuevos altímetros son mejores porque la variación en los errores es reducida y, por lo tanto, las lecturas son más consistentes. (En muchas industrias, la calidad de los bienes y servicios a menudo se puede mejorar al reducir la variación).

Los métodos que se presentan en este capítulo se basan en la regla del suceso infrecuente (sección 4-1) para la estadística inferencial, de manera que la repasaremos antes de continuar.

Regla del suceso infrecuente para la estadística inferencial

Si, bajo un supuesto dado, la probabilidad de un suceso observado particular es excepcionalmente pequeña, concluimos que el supuesto probablemente es incorrecto.

Siguiendo esta regla, probamos una aseveración analizando datos muestrales en un intento por distinguir entre resultados que pueden *ocurrir fácilmente por azar* y resultados cuya ocurrencia es *extremadamente improbable debido al azar*. Podemos explicar la ocurrencia de resultados extremadamente improbables al decir que en realidad ha ocurrido un suceso infrecuente o que el supuesto subyacente no es verdadero. Apliquemos este razonamiento en el siguiente ejemplo.

EJEMPLO Selección del género ProCare Industries, Ltd., alguna vez ofreció un producto llamado “Gender Choice”, el cual, según aseveraciones publicitarias, permitía a las parejas “incrementar hasta en un 80% sus posibilidades de tener una niña”. Suponga que realizamos un experimento con 100 parejas que desean tener niñas, y que las 100 parejas siguen el “sistema casero fácil de usar”

de Gender Choice, descrito en el paquete rosa diseñado para concebir niñas. Suponiendo que Gender Choice no tiene efecto alguno, y basados en el sentido común, sin un método estadístico formal, ¿qué debemos concluir acerca del supuesto de que Gender Choice no tiene efecto alguno, si 100 parejas lo utilizaron y tuvieron 100 bebés, de los cuales

- a. 52 fueron niñas?
- b. 97 fueron niñas?

SOLUCIÓN

- a. Generalmente esperamos que nazcan alrededor de 50 niñas por cada 100 nacimientos. El resultado de 52 niñas es cercano a 50, por lo que no debemos concluir que el producto Gender Choice es eficaz. El resultado de 52 niñas podría ocurrir fácilmente por azar, de manera que no existe evidencia suficiente para afirmar que Gender Choice sea eficaz.
- b. Es extremadamente improbable que el resultado de 97 niñas en 100 nacimientos suceda por azar. Nosotros podríamos explicar el nacimiento de 97 niñas de dos maneras: o se trata de un evento *extremadamente* infrecuente que ha ocurrido por azar, o Gender Choice es eficaz. La probabilidad extremadamente baja de que resulten 97 niñas sugiere que Gender Choice es eficaz.

El aspecto central del ejemplo anterior es que debemos concluir que el producto es eficaz sólo si obtenemos *significativamente* más niñas de las que esperaríamos normalmente. Aun cuando los resultados de 52 niñas y 97 niñas están “por arriba del promedio”, el resultado de 52 niñas no es significativo, mientras que el de 97 niñas es un resultado significativo.

Este breve ejemplo ilustra el método básico utilizado en la prueba de hipótesis. El método formal incluye una variedad de términos y condiciones convencionales incorporados en un procedimiento organizado. Le sugerimos que inicie el estudio de este capítulo con la lectura de las secciones 8-2 y 8-3, de manera informal, para tener una idea general de estos conceptos, y que después lea nuevamente la sección 8-2, ahora con mayor detenimiento, para familiarizarse con la terminología.

Fundamentos de la prueba 8-2 de hipótesis

Concepto clave En esta sección se presentan los componentes individuales de una prueba de hipótesis, y las siguientes secciones utilizan esos componentes en procedimientos detallados. Es necesario comprender el papel de los siguientes componentes: hipótesis nula, hipótesis alternativa, estadístico de prueba, región crítica, nivel de significancia, valor crítico, valor P , error tipo I y error tipo II. En la parte 1 de esta sección se presentan los conceptos básicos; es importante comprenderlos antes de considerar el concepto de potencia de una prueba, que se analiza en la parte 2.

Parte 1: Conceptos básicos de la prueba de hipótesis

A continuación se describen los objetivos de esta sección, los cuales debemos alcanzar antes de considerar el análisis de la potencia en la parte 2.

Objetivos de esta sección

- Dada una aseveración, identificar la hipótesis nula y la hipótesis alternativa, y expresar ambas de forma simbólica.



Detectores de mentiras

¿Por qué no se exige que todos los sospechosos de un crimen sean sometidos a la prueba del detector de mentiras para así prescindir de los juicios? El Council of Scientific Affairs de la American Medical Association afirma que “está establecido que la clasificación de los sujetos con base en la culpabilidad puede realizarse con una precisión del 75 al 97%, pero la tasa de falsos positivos suele ser lo suficientemente alta como para excluir el uso de esta prueba (del polígrafo) como único criterio para determinar la culpabilidad o inocencia”. Un “falso positivo” es una indicación de culpabilidad cuando el sujeto en realidad es inocente. Incluso con una precisión tan alta como del 97%, el porcentaje de resultados falsos positivos puede ser del 50%, de manera que la mitad de los sujetos inocentes aparecerían incorrectamente como culpables.

- Dados una aseveración y datos muestrales, calcular el valor del estadístico de prueba.
- Dado un nivel de significancia, identificar el valor (o los valores) crítico(s).
- Dado un valor del estadístico de prueba, identificar el valor P .
- Plantear la conclusión de una prueba de hipótesis en términos sencillos y sin tecnicismos.

Es recomendable estudiar el siguiente ejemplo hasta comprenderlo exhaustivamente. Una vez que lo logre, ya habrá captado el principal concepto de la estadística.

EJEMPLO Selección del género y probabilidad Vamos a referirnos de nuevo a los empaques color rosa de Gender Choice. ProCare Industries afirmaba que las parejas que usaban los productos de Gender Choice con empaque color rosa tendrían niñas en una proporción mayor al 50% o 0.5. Consideremos de nuevo un experimento en el que 100 parejas usan Gender Choice en un intento por concebir una niña y supongamos que los 100 bebés incluyen exactamente 52 niñas. Procederemos a formalizar parte del análisis, pero hay dos aspectos que pueden ser confusos:

1. **Suponga que $p = 0.5$:** Al tratar de determinar si 52 niñas en 100 nacimientos representan una evidencia de la eficacia de Gender Choice, suponemos que $p = 0.5$, de manera que podemos determinar si el resultado de 52 niñas puede ocurrir fácilmente por azar (sin efecto del tratamiento) o si es improbable que este resultado ocurra por azar (de manera que el tratamiento sea efectivo).
2. **Use $P(52 \text{ o más niñas})$:** Al determinar si el resultado de “52 niñas” puede ocurrir al azar, utilice la probabilidad de 52 o más niñas . [Repase la sección “Uso de las probabilidades para determinar resultados infrecuentes” de la sección 5-2, donde señalamos que “ x éxitos en n ensayos es un número *inusualmente alto* de éxitos si $P(x \text{ o más}) \leq 0.05$ ”].

En circunstancias normales, la proporción de niñas es $p = 0.5$, de manera que la aseveración de que Gender Choice es eficaz puede expresarse como $p > 0.5$. Respalдamos la aseveración de que $p > 0.5$ sólo si un resultado como el de 52 niñas es improbable (con una escasa probabilidad, como menor que o igual a 0.05). Si se utiliza una distribución normal como aproximación de la distribución binomial (véase la sección 6-6), encontramos que $P(52 \text{ o más niñas en } 100 \text{ nacimientos}) = 0.3821$. La figura 8-1 muestra que, con una probabilidad de 0.5, el resultado de 52 niñas en 100 nacimientos no es infrecuente, de manera que *no* rechazamos el azar como una explicación razonable. Concluimos que la proporción de niñas nacidas de parejas que usan Gender Choice *no* es significativamente mayor que el número que esperaríamos por el azar. He aquí los aspectos clave de este ejemplo:

- Aseveración: En las parejas que utilizan Gender Choice, la proporción de niñas es $p > 0.5$.
- Supuesto de trabajo: La proporción de niñas es $p = 0.5$ (sin efecto del Gender Choice).
- La muestra dio por resultado 52 niñas de entre 100 nacimientos, por lo tanto, la proporción muestral es $\hat{p} = 52/100 = 0.52$.
- Suponiendo que $p = 0.5$, empleamos una distribución normal como aproximación de la distribución binomial para calcular que $P(\text{al menos } 52 \text{ niñas en } 100 \text{ nacimientos}) = 0.3821$.

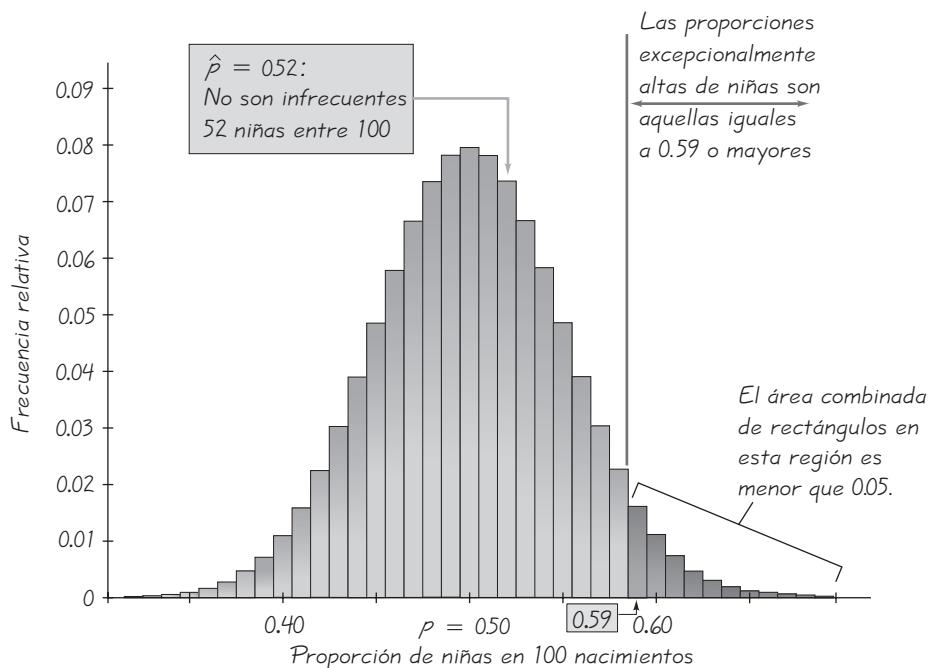


Figura 8-1 Distribución muestral de proporciones de niñas en 100 nacimientos

- Existen dos explicaciones posibles para el resultado de 52 niñas en 100 nacimientos: o bien ocurrió un suceso aleatorio (con una probabilidad de 0.3821), o la proporción de niñas nacidas de parejas que usan Gender Choice es mayor que 0.5. Como la probabilidad de obtener al menos 52 niñas por azar es tan alta (0.3821), consideramos que el azar es una explicación razonable. No existe evidencia suficiente para sustentar la aseveración de que Gender Choice es eficaz para concebir más niñas que lo esperado por el azar. (En realidad fue este tipo de análisis el que condujo a que Gender Choice fuera retirado del mercado).

En la sección 8-3 describiremos los pasos específicos que se utilizan en la prueba de hipótesis; antes de ello, describamos los componentes de una **prueba de hipótesis formal** o **prueba de significancia**. Estos términos suelen emplearse en una gran variedad de disciplinas cuando se requieren métodos estadísticos.

Componentes de una prueba de hipótesis formal

Hipótesis nula y alternativa

- La **hipótesis nula** (denotada por H_0) es la afirmación de que el valor de un parámetro de población (como una proporción, media o desviación estándar) es *igual a* un valor aseverado. Las siguientes son hipótesis nulas típicas del tipo considerado en este capítulo:

$$H_0: p = 0.5 \quad H_0: \mu = 98.6 \quad H_0: \sigma = 15$$

La hipótesis nula se prueba en forma directa, en el sentido de que suponemos que es verdadera, y llegamos a una conclusión para rechazar H_0 o no rechazar H_0 .

- La **hipótesis alternativa** (denotada por H_1 o H_a o H_A) es la afirmación de que el parámetro tiene un valor que, de alguna manera, difiere de la hipótesis nula. Para los métodos de este capítulo, la forma simbólica de la hipótesis alternativa debe emplear alguno de estos símbolos: $<$, $>$, o bien, \neq . A continuación se presentan nueve ejemplos diferentes de hipótesis alternativas que incluyen proporciones, medias y desviaciones estándar:

Proporciones: $H_1: p > 0.5$ $H_1: p < 0.5$ $H_1: p \neq 0.5$

Medias: $H_1: \mu > 98.6$ $H_1: \mu < 98.6$ $H_1: \mu \neq 98.6$

Desviaciones estándar: $H_1: \sigma > 15$ $H_1: \sigma < 15$ $H_1: \sigma \neq 15$

Nota sobre el uso del símbolo de igualdad en H_0 : Algunos libros de texto utilizan los símbolos \leq y \geq en la hipótesis nula H_0 , pero la mayoría de las revistas científicas emplean sólo el símbolo de igual para expresar igualdad. Realizamos la prueba de hipótesis suponiendo que la proporción, media o desviación estándar es *igual a* algún valor especificado, de manera que podemos trabajar con una sola distribución teniendo un valor específico.

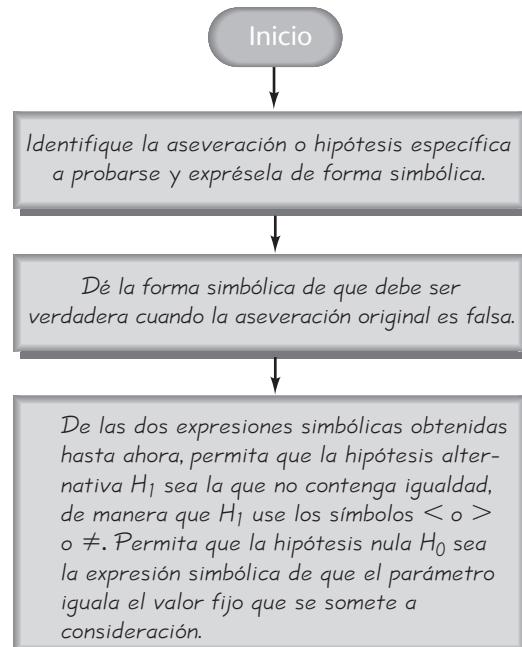
Nota sobre la formulación de sus propias aseveraciones (hipótesis): Si usted está realizando un estudio y desea emplear una prueba de hipótesis para *sustentar* su aseveración, ésta debe redactarse de tal manera que se convierta en la hipótesis alternativa. Esto quiere decir que su aseveración debe expresarse utilizando sólo estos símbolos: $<$, $>$, o bien, \neq). No puede utilizar una prueba de hipótesis para sustentar la aseveración de que algún parámetro es *igual a* algún valor especificado.

Por ejemplo, si usted ha creado un método de selección del género, que aumenta la probabilidad de concebir una niña, redacte su aseveración como $p > 0.5$, para que ésta pueda ser sustentada. (En el contexto de tratar de sustentar la meta de la investigación, la hipótesis alternativa en ocasiones se conoce como la *hipótesis de investigación*). Para el propósito de la prueba, usted supondrá que $p = 0.5$, pero usted esperará que $p = 0.5$ sea rechazada para que $p > 0.5$ se sostenga.

Nota sobre la identificación de H_0 y H_1 : La figura 8-2 resume los procedimientos para identificar las hipótesis nula y alternativa. Observe que la afirmación

Figura 8-2

Identificación de H_0 and H_1



original puede convertirse en la hipótesis nula, en la hipótesis alternativa o tal vez no corresponda con exactitud a ninguna de las dos.

Por ejemplo, en ocasiones probamos la validez de la aseveración de alguien más, como la afirmación de la Coca-Cola Bottling Company de que “la cantidad media de Coca-Cola en las latas es de al menos 12 onzas”. Esta afirmación puede expresarse en símbolos tales como $\mu \geq 12$. En la figura 8-2 vemos que si la aseveración original es falsa, entonces $\mu < 12$. La hipótesis alternativa se vuelve $\mu < 12$, pero la hipótesis nula es $\mu = 12$. Podremos considerar la aseveración original después de determinar si existe suficiente evidencia para rechazar la hipótesis nula de $\mu = 12$.

EJEMPLO Identificación de las hipótesis nula y alternativa Remítase a la figura 8-2 y utilice las aseveraciones para expresar las hipótesis nula y alternativa de forma simbólica.

- La proporción de empleados que consiguen trabajo por medio de una red de contactos es mayor que 0.5.
- El peso medio de los pasajeros de avión, con su equipaje de mano, es a lo sumo de 195 libras (la cifra que la Federal Aviation Administration difunde actualmente).
- La desviación estándar de las puntuaciones de CI de actores es igual a 15.

SOLUCIÓN Véase la figura 8-2, que muestra el procedimiento de los tres pasos.

- En el paso 1 de la figura 8-2, expresamos la aseveración dada como $p > 0.5$. En el paso 2 observamos que si $p > 0.5$ es falso, entonces $p \leq 0.5$ debe ser verdadero. En el paso 3 vemos que la expresión $p > 0.5$ no contiene igualdad, por lo que permitimos que la hipótesis alternativa H_1 sea $p > 0.5$, y dejamos que H_0 sea $p = 0.5$.
- En el paso 1 de la figura 8-2, expresamos “una media de a lo sumo 195 libras” en símbolos como $\mu \leq 195$. En el paso 2 observamos que si $\mu \leq 195$ es falso, entonces $\mu > 195$ debe ser verdadero. En el paso 3 vemos que la expresión $\mu > 195$ no contiene igualdad, por lo que permitimos que la hipótesis alternativa H_1 sea $\mu > 195$ y que H_0 sea $\mu = 195$.
- En el paso 1 de la figura 8-2 expresamos la aseveración dada como $\sigma = 15$. En el paso 2 observamos que si $\sigma = 15$ es falso, entonces $\sigma \neq 15$ debe ser verdadero. En el paso 3, permitimos que la hipótesis alternativa sea $\sigma \neq 15$, y que H_0 sea $\sigma = 15$.

Estadístico de prueba

- El **estadístico de prueba** es un valor que se utiliza para tomar la decisión sobre la hipótesis nula, y se calcula convirtiendo al estadístico muestral (como la proporción muestral \hat{p} , la media muestral \bar{x} , o la desviación estándar muestral s) en una puntuación (como z , t o χ^2), bajo el supuesto de que la hipótesis nula es verdadera. En este capítulo empleamos los siguientes estadísticos de prueba:

Estadístico de prueba para proporciones

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$



Un tamaño muestral grande no es lo suficientemente bueno

Los datos muestrales sesgados no deben emplearse para hacer inferencias, sin importar cuán grande sea la muestra. Por ejemplo, en *Women and Love: A Cultural Revolution in Progress*, Shere Hite basa sus conclusiones en 4500 respuestas que recibió después de enviar por correo 100,000 cuestionarios a diversos grupos de mujeres. Por lo general, una muestra aleatoria de 4500 sujetos da buenos resultados, pero la muestra de Hite está sesgada y ha sido criticada por considerarse que en ella tienen excesiva representación las mujeres con fuertes sentimientos acerca de los temas planteados. Como la muestra de Hite está sesgada, sus inferencias no son válidas, aun cuando el tamaño de muestra de 4500 pueda parecer lo suficientemente grande.

Estadístico de prueba para medias

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{o} \quad t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Estadístico de prueba para desviaciones estándar

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

El estadístico de prueba para una media usa la distribución normal o la distribución *t* de Student, dependiendo de los requisitos que se satisfagan. En este capítulo se utilizarán los mismos criterios descritos en la sección 7-4. (Véase la figura 7-6 y la tabla 7-1).

EJEMPLO Cálculo del estadístico de prueba Una encuesta de $n = 703$ empleados seleccionados al azar, reveló que el 61% ($\hat{p} = 0.61$) de ellos consiguieron trabajo por medio de una red de contactos. Calcule el valor del estadístico de prueba para la aseveración de que la mayoría de los empleados (más del 50%) consiguen trabajo por medio de una red de contactos. (En la sección 8-3 veremos que existen supuestos que deben verificarse. Para este ejemplo, suponga que se satisfacen los supuestos requeridos y concéntrese en el cálculo del estadístico de prueba indicado).

SOLUCIÓN El ejemplo anterior demostró que la aseveración da por resultado las siguientes hipótesis nula y alternativa: $H_0: p = 0.5$ y $H_1: p > 0.5$. Como trabajamos bajo el supuesto de que la hipótesis nula es verdadera, con $p = 0.5$, obtenemos el siguiente estadístico de prueba:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.61 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{703}}} = 5.83$$

INTERPRETACIÓN De capítulos previos sabemos que la puntuación *z* de 5.83 es “poco común” (porque es mayor que 2). Parece que, además de ser mayor que el 50%, el resultado muestral de 61% es *significativamente* mayor que el 50%. Observe la figura 8-3, donde demostramos que la proporción muestral de 0.61 (del 61%) cae dentro del rango de valores considerados significativos, ya que están tan arriba de 0.5 que no es probable que ocurran por azar (suponiendo que la proporción de la población es $p = 0.5$).

Región crítica, nivel de significancia, valor crítico y valor *P*

- La **región crítica** (o **región de rechazo**) es el conjunto de todos los valores del estadístico de prueba que pueden provocar que rechacemos la hipótesis nula. Por ejemplo, observe la región sombreada más oscura en la figura 8-3.
- El **nivel de significancia** (denotado por α) es la probabilidad de que el estadístico de prueba caiga en la región crítica, cuando la hipótesis nula es verdadera. Si el estadístico de prueba cae en la región crítica, rechazaremos la hipótesis nula, de manera que α es la probabilidad de cometer el error de rechazar la hipótesis nula cuando es verdadera. Se trata de la misma α presentada en

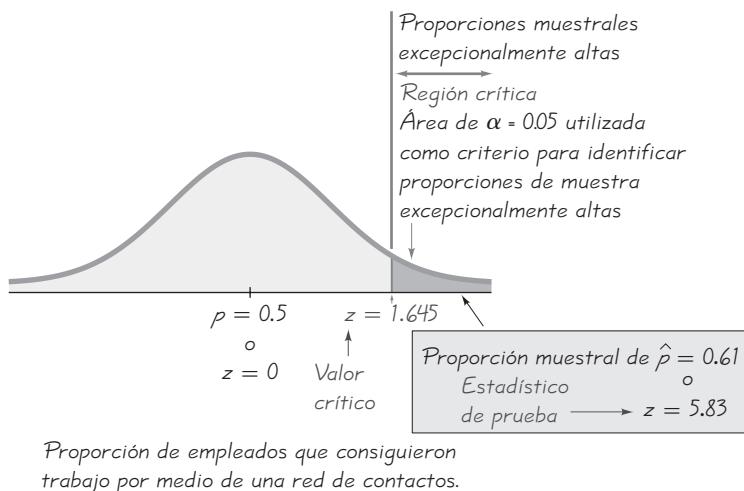


Figura 8-3 Región crítica, valor crítico y estadístico de prueba

la sección 7-2, donde definimos el nivel de confianza para un intervalo de confianza como la probabilidad $1 - \alpha$. Las opciones comunes para α son 0.05, 0.01 y 0.10, aunque la más común es 0.05.

- Un **valor crítico** es cualquier valor que separa la región crítica (donde rechazamos la hipótesis nula) de los valores del estadístico de prueba que no conducen al rechazo de la hipótesis nula. Los valores críticos dependen de la naturaleza de la hipótesis nula, de la distribución muestral que se aplique y del nivel de significancia α . Observe la figura 8-3, donde el valor crítico de $z = 1.645$ corresponde a un nivel de significancia de $\alpha = 0.05$. (Los valores críticos se estudiaron antes, en el capítulo 7).

EJEMPLO Cálculo de valores críticos Con un nivel de significancia de $\alpha = 0.05$, calcule los valores z críticos para cada una de las siguientes hipótesis alternativas (suponiendo que la distribución normal puede emplearse como aproximación de la distribución binomial):

- $p \neq 0.5$ (de manera que la región crítica está en *ambas* colas de la distribución normal)
- $p < 0.5$ (de manera que la región crítica está en la *cola izquierda* de la distribución normal)
- $p > 0.5$ (de manera que la región crítica está en la *cola derecha* de la distribución normal)

SOLUCIÓN

- Observe la figura 8-4a). Las colas sombreadas contienen una área total de $\alpha = 0.05$, por lo que cada cola contiene una área de 0.025. Empleando los métodos de la sección 6-2, los valores de $z = 1.96$ y $z = -1.96$ separan las regiones de la cola izquierda y la cola derecha. Por lo tanto, los valores críticos son $z = 1.96$ y $z = -1.96$.



Gane \$1,000,000 por sus poderes sobrenaturales

El mago James Randi instituyó una fundación educativa que ofrece un premio de \$1 millón a quien pueda demostrar poderes paranormales, sobrenaturales u ocultos. Cualquiera que posea un poder como el de adivinar el futuro, percepción extrasensorial (PES) o la habilidad para comunicarse con los muertos, puede ganar el premio si pasa ciertos procedimientos de prueba. Primero se realiza una prueba preliminar y después una formal, pero hasta ahora nadie ha aprobado la prueba preliminar. La prueba formal se diseñaría con métodos estadísticos sólidos, y probablemente incluiría un análisis con una prueba de hipótesis formal. Según la fundación, se consultan “especialistas competentes en estadística cuando se necesita evaluar los resultados o diseñar experimentos”. En la página de Internet de la fundación, randi.org, se puede encontrar información sobre la solicitud.

continúa

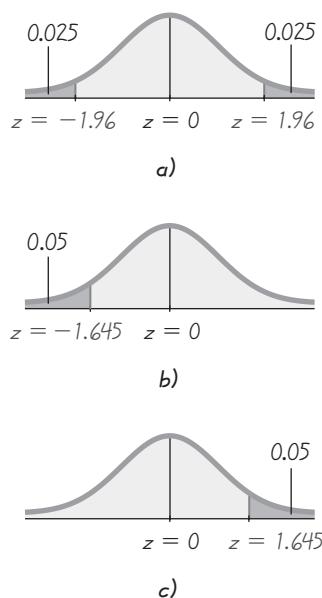


Figura 8-4
Cálculo de valores críticos

- b. Observe la figura 8-4b). Con una hipótesis alternativa de $p < 0.5$, la región crítica se encuentra en la cola izquierda. Con una área de cola izquierda de 0.05, se obtiene que el valor crítico es $z = -1.645$ (empleando los métodos de la sección 6-2).
- c. Observe la figura 8-4c). Con una hipótesis alternativa de $p > 0.5$, la región crítica está en la cola derecha. Con una área de cola derecha de 0.05, se obtiene que el valor crítico es $z = 1.645$ (empleando los métodos de la sección 6-2).

Dos colas, cola izquierda y cola derecha Las *colas* en una distribución son las regiones extremas limitadas por los valores críticos. Algunas pruebas de hipótesis incluyen dos colas, otras la cola derecha y otras la cola izquierda.

- **Prueba de dos colas:** La región crítica se encuentra en las dos regiones extremas (colas) bajo la curva [como en la figura 8-4a)].
- **Prueba de cola izquierda:** La región crítica se encuentra en la región extrema izquierda (cola) bajo la curva [como en la figura 8-4b)].
- **Prueba de cola derecha:** La región crítica se encuentra en la región extrema derecha (cola) bajo la curva [como en la figura 8-4c)].

En la prueba de dos colas, el nivel de significancia α está dividido equitativamente entre las dos colas que constituyen la región crítica. Por ejemplo, en una prueba de dos colas con un nivel de significancia de $\alpha = 0.05$, existe una área de 0.025 en cada una de las dos colas. En las pruebas de cola derecha o cola izquierda, el área de la región crítica en una cola es α (véase la figura 8-4).

Al examinar la hipótesis alternativa, podemos determinar si la prueba es de cola derecha, de cola izquierda o de dos colas. La cola corresponderá a la región crítica que contiene los valores que entrarán en conflicto, de manera significativa, con la hipótesis nula. En las figuras al margen se resume información útil (véase la figura 8-5), la cual indica que el signo de desigualdad de H_1 señala en la dirección de la región crítica. El símbolo \neq suele expresarse en lenguaje de programación como $<>$, y esto nos recuerda que una hipótesis alternativa, como $p \neq 0.5$ corresponde a una prueba de dos colas.

- El **valor P** (o **valor p** o **valor de probabilidad**) es la probabilidad de obtener un valor del estadístico de prueba que sea *al menos tan extremo* como el que representa a los datos muestrales, suponiendo que la hipótesis nula es verdadera. La hipótesis nula se rechaza si el valor P es muy pequeño, tanto como 0.05 o menos. Los valores P se calculan con el procedimiento resumido en la figura 8-6.

Decisiones y conclusiones

El procedimiento convencional de prueba de hipótesis requiere que siempre probemos la hipótesis nula, de manera que nuestra conclusión inicial siempre será una de las siguientes:

1. Rechazo de la hipótesis nula.
2. No rechazo de la hipótesis nula.

Criterio de decisión La decisión de rechazar o no rechazar la hipótesis nula suele realizarse por medio del método tradicional (o método clásico) de prueba de

hipótesis, el método del valor P , o bien, la decisión puede basarse en intervalos de confianza. En años recientes, el uso del método del valor P ha aumentado, junto con la inclusión de valores P en los resultados de programas de cómputo.

Método tradicional: *Rechace H_0* si el estadístico de prueba cae dentro de la región crítica.

No rechace H_0 si el estadístico de prueba no cae dentro de la región crítica.

Método del valor P : *Rechace H_0* si el valor de $P \leq \alpha$ (donde α es el nivel de significancia, tal como 0.05).

No rechace H_0 si el valor $P > \alpha$.

Otra opción: En vez de usar un nivel de significancia como $\alpha = 0.05$, simplemente identifique el valor P y deje la decisión al lector.

Intervalos de confianza: Como un estimado del intervalo de confianza de un parámetro de población contiene los valores posibles de tal parámetro, rechace la aseveración de que el parámetro de población tiene un valor que no está incluido en el intervalo de confianza.

EJEMPLO Cálculo de valores P Primero determine si las condiciones planteadas dan por resultado una prueba de cola derecha, de cola izquierda o de dos colas; después utilice la figura 8-6 para calcular el valor P , luego saque una conclusión acerca de la hipótesis nula.

- Se utiliza un nivel de significancia de $\alpha = 0.05$ para probar la aseveración de que $p > 0.25$, y los datos muestrales dan por resultado un estadístico de prueba de $z = 1.18$.
- Se utiliza un nivel de significancia de $\alpha = 0.05$ para probar la aseveración de que $p \neq 0.25$, y los datos muestrales dan por resultado un estadístico de prueba de $z = 2.34$.

SOLUCIÓN

- Con la aseveración de que $p > 0.25$, se trata de una prueba de cola derecha. Al utilizar la figura 8-6 para una prueba de cola derecha, vemos que el valor P es el área a la derecha del estadístico de prueba $z = 1.18$. Nos remitimos a la tabla A-2 y encontramos que el área a la derecha de $z = 1.18$ es 0.1190. El valor P de 0.1190 es mayor que el nivel de significancia $\alpha = 0.05$, por lo que no rechazamos la hipótesis nula. El valor P de 0.1190 es relativamente grande, lo que indica que los resultados muestrales podrían suceder fácilmente por azar.
- Con la aseveración de $p \neq 0.25$, se trata de una prueba de dos colas. Al utilizar la figura 8-6 para una prueba de dos colas, observamos que el valor P es *dos veces* el área a la derecha de $z = 2.34$. Nos remitimos a la tabla A-2 y encontramos que el área a la derecha de $z = 2.34$ es 0.0096, de manera que el valor $P = 2 \times 0.0096 = 0.0192$. El valor P de 0.0192 es menor o igual que el nivel de significancia, por lo que rechazamos la hipótesis nula. El pequeño valor P de 0.0192 indica que los resultados muestrales no podrían suceder por azar.

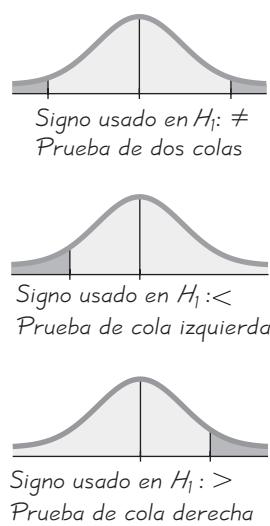


Figura 8-5
Pruebas de dos colas, cola izquierda y cola derecha

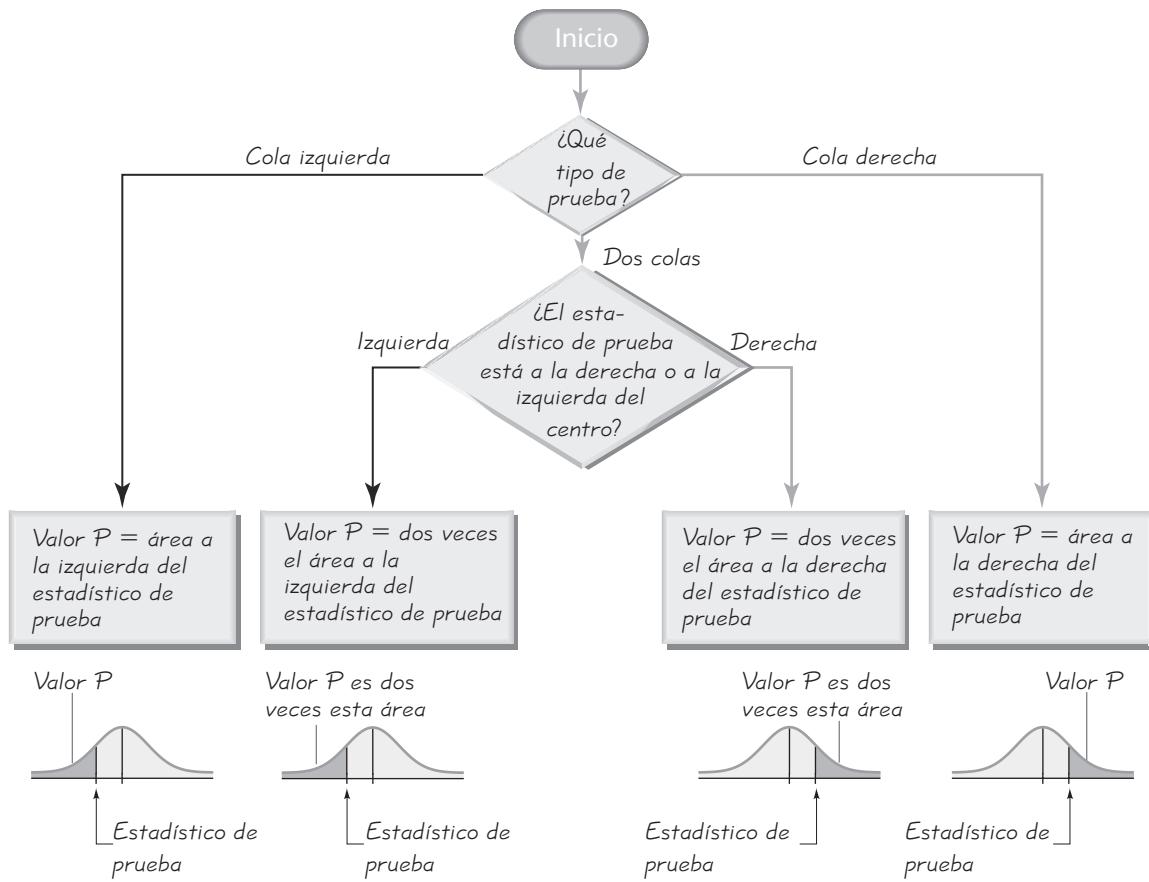


Figura 8-6 Procedimiento para el cálculo de los valores P

Redacción de la conclusión final La conclusión de rechazar o no la hipótesis nula es adecuada para aquellos que tenemos el buen juicio de tomar un curso de estadística, pero debemos emplear términos sencillos y sin tecnicismos al establecer el verdadero significado de la conclusión. La figura 8-7 resume un procedimiento para redactar la conclusión final. Observe que sólo un caso conduce a la indicación de que los datos muestrales en realidad *sustentan* la conclusión. Si desea sustentar alguna aseveración, exprésela de manera tal que se convierta en la hipótesis alternativa, y después espere que la hipótesis nula sea rechazada.

Aceptación/no rechazo Algunos libros de texto dicen “aceptar la hipótesis nula” en vez de “no rechazar la hipótesis nula”. Ya sea que usemos el término *aceptar* o *no rechazar*, debemos reconocer que *no estamos probando la hipótesis nula*; únicamente estamos diciendo que la evidencia muestral no es lo suficientemente fuerte como para justificar el rechazo de la hipótesis nula. (Cuando un jurado no encuentra evidencia suficiente para sentenciar a un sospechoso, emite un veredicto de no culpabilidad y no un veredicto de inocencia). El término *aceptar* es un poco confuso, ya que parece indicar incorrectamente que la hipótesis nula

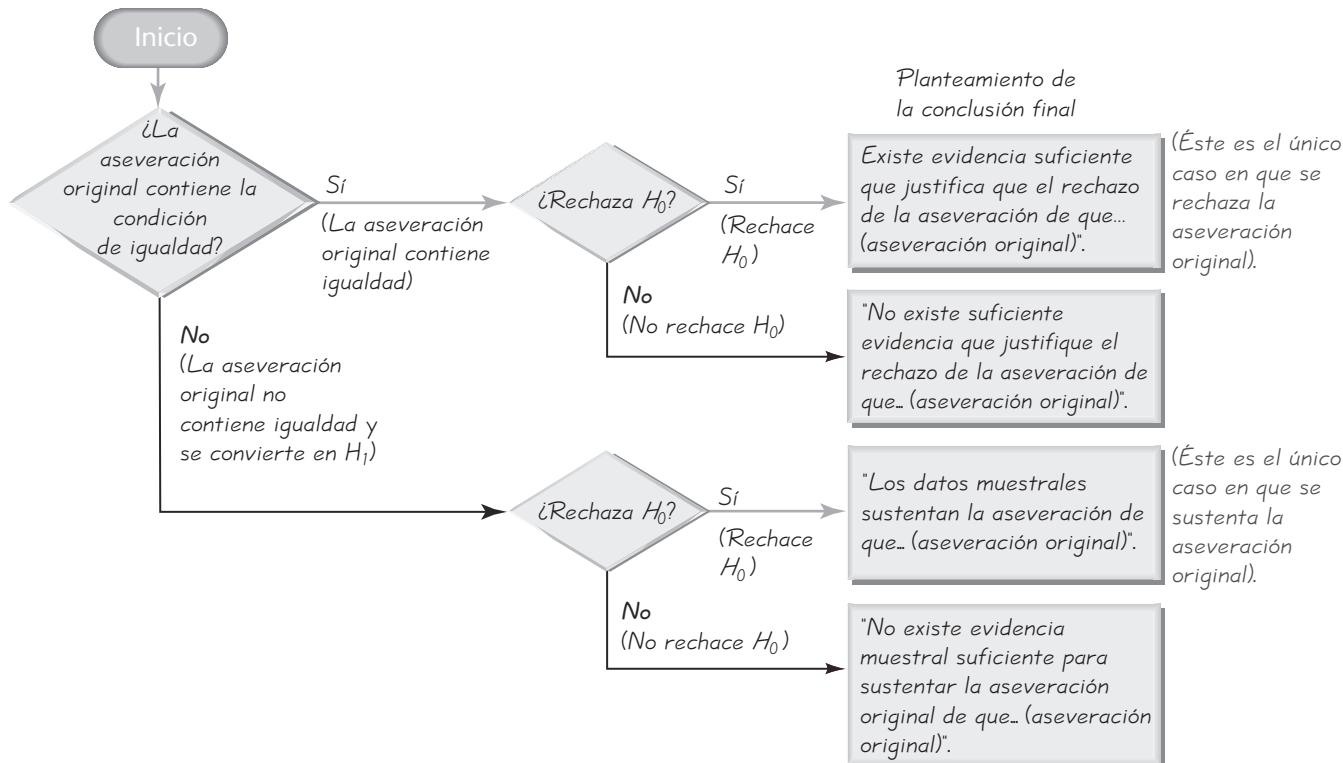


Figura 8-7 Redacción de la conclusión final

ha sido probada. (Es confuso decir que “existe evidencia suficiente para aceptar la hipótesis nula”). La frase *no rechazar* dice con mayor corrección que la evidencia disponible no es lo suficientemente fuerte para justificar el rechazo de la hipótesis nula. En este texto emplearemos la terminología *no rechazar la hipótesis nula*, en vez de *aceptar la hipótesis nula*.

Múltiples negativos Cuando se establece la conclusión final en términos no técnicos, es posible enunciar afirmaciones correctas con hasta tres términos negativos. (Ejemplo: “*No existe evidencia suficiente para justificar el rechazo* de la aseveración de que *no hay* diferencia entre 0.5 y la proporción poblacional”). Las conclusiones con demasiados términos negativos resultan confusas, por lo que es aconsejable volver a redactarlas en una forma comprensible, pero teniendo cuidado de no alterar el significado. Por ejemplo, en vez de decir que “*no existe evidencia suficiente para justificar el rechazo de la aseveración de que no existen diferencias entre 0.5 y la proporción poblacional*”, los siguientes serían mejores enunciados:

- No se rechaza la aseveración de que la proporción poblacional es igual a 0.5.
- Hasta no obtener evidencia más firme, continuamos suponiendo que la proporción poblacional es igual a 0.5.

		Verdadero estado de las cosas	
		La hipótesis nula es verdadera	La hipótesis nula es falsa
Decisión	Decidimos rechazar la hipótesis nula	Error tipo I (rechazo de una hipótesis nula verdadera) α	Decisión correcta
	Decidimos no rechazar la hipótesis nula	Decisión correcta	Error tipo II (no rechazo de una hipótesis nula falsa) β

EJEMPLO Redacción de la conclusión final Suponga que un reportero asevera que “más de la mitad” (más del 50%) de los empleados consiguen trabajo por medio de redes de contactos. Esta aseveración de $p > 0.5$ se convierte en la hipótesis alternativa, mientras que $p = 0.5$ se convierte en la hipótesis nula. Además, suponga que la evidencia muestral hace que rechacemos la hipótesis nula de $p = 0.5$. Enuncie la conclusión en términos sencillos y sin tecnicismos.

SOLUCIÓN Remítase a la figura 8-7. La aseveración original no incluye la condición de igualdad, y rechazamos la hipótesis nula. Por lo tanto, la redacción de la conclusión final debe ser la siguiente: “Los datos muestrales sustentan la aseveración de que la mayoría de los empleados consiguen trabajo por medio de redes de contactos”.

Errores tipo I y tipo II Cuando probamos una hipótesis nula, llegamos a la conclusión de rechazarla o no rechazarla. Tales conclusiones pueden ser correctas o incorrectas (incluso cuando hacemos todo correctamente). La tabla 8-1 resume los dos distintos tipos de errores que pueden cometerse, junto con los dos tipos de decisiones correctas. Distinguimos entre los dos tipos de errores denominándolos errores tipo I y tipo II.

- **Error tipo I:** El error de rechazar la hipótesis nula cuando en realidad es verdadera. Se utiliza el símbolo α (alfa) para representar la probabilidad de un error tipo I.
- **Error tipo II:** El error de no rechazar la hipótesis nula cuando en realidad es falsa. Se utiliza el símbolo β (beta) para representar la probabilidad de un error tipo II.

Debido a que los estudiantes suelen considerar difícil recordar cuál error es el tipo I y cuál es el error tipo II, recomendamos una herramienta mnemotécnica, como podría ser “revisión no refinada” (ReVisióN No ReFiNada”). Si utilizamos algunas de las consonantes de estas palabras podemos recordar que el error tipo I es RVN: rechazar verdadera nula (hipótesis), mientras que el error tipo II es NRFN: no rechazar falsa nula (hipótesis).

Notación

α (alfa) = probabilidad de un error tipo I (la probabilidad de rechazar la hipótesis nula cuando es verdadera)

β (beta) = probabilidad de un error tipo II (la probabilidad de no rechazar una hipótesis nula cuando es falsa)

EJEMPLO Identificación de errores tipo I y tipo II Suponga que estamos realizando una prueba de hipótesis de la aseveración de que $p < 0.5$. He aquí las hipótesis nula y alternativa

$$H_0: p = 0.5$$

$$H_1: p < 0.5$$

Escriba afirmaciones que identifiquen

- a. Un error tipo I.
- b. Un error tipo II.

SOLUCIÓN

- a. Un error tipo I se comete cuando se rechaza una hipótesis nula verdadera. El siguiente es un error tipo I: concluir que existe evidencia suficiente para sustentar $p < 0.5$, cuando en realidad $p = 0.5$.
- b. Un error tipo II se comete al no rechazar la hipótesis nula cuando es falsa. El siguiente es un error tipo II: no rechazar $p = 0.5$ (y, por lo tanto, no sustentar $p < 0.5$) cuando en realidad $p < 0.5$.

Control de los errores tipo I y tipo II: Un paso de nuestro procedimiento estándar para la prueba de hipótesis implica la selección del nivel de significancia α , que corresponde a la probabilidad de un error tipo I. Sin embargo, no seleccionamos β [P (error tipo II)]. Sería formidable si pudieramos tener siempre $\alpha = 0$ y $\beta = 0$, pero en realidad eso no es posible, así que debemos intentar manejar las probabilidades de los errores α y β . Matemáticamente, es posible demostrar que α , β y el tamaño de muestra n están relacionados, de manera que cuando usted elige o determina cualesquiera dos, el tercero se determina automáticamente. La práctica común en investigación y en la industria es seleccionar los valores de α y n , de manera que se determine el valor de β . Dependiendo de la gravedad de un error tipo I, trate de emplear el α más grande que pueda tolerar. Para errores tipo I con consecuencias más graves, seleccione valores más pequeños de β . Despues elija un tamaño de muestra n razonablemente grande, con base en consideraciones de tiempo, costo y otros factores relevantes. (La determinación del tamaño de muestra se estudió en el capítulo 7). Las siguientes consideraciones prácticas resultarán útiles:

1. Para cualquier α fija, un incremento en el tamaño de la muestra n causará un decremento en β . Es decir, una muestra grande disminuirá la posibilidad de que usted cometa el error de no rechazar la hipótesis nula cuando en realidad es falsa.
2. Para cualquier tamaño de muestra n fijo, un decremento en α causará un incremento en β . A la inversa, un incremento en α provocará un decremento en β .
3. Para disminuir tanto α como β , aumente el tamaño de la muestra.

Para que estas ideas abstractas tengan sentido, consideremos dulces M&M (producidos por Mars, Inc.) y tabletas de aspirina marca Bufferin (producidas por Bristol-Mayers Products).

- Se supone que el peso medio de los dulces M&M es de al menos 0.8535 g (de acuerdo con el peso indicado en la etiqueta del empaque).
- Se supone que las tabletas Bufferin tienen un peso medio de 325 mg de aspirina.

Puesto que los dulces M&M se emplean para disfrutarlos, mientras que las tabletas Bufferin son fármacos utilizados para el tratamiento de problemas de salud, estamos enfrentando dos niveles muy diferentes de gravedad. Para probar la aseveración de que $\mu = 0.8535$ g de los M&M, podríamos elegir $\alpha = 0.05$ y un tamaño muestral de $n = 100$; para probar la aseveración de que $\mu = 325$ mg de las tabletas Bufferin, podríamos elegir $\alpha = 0.01$ y un tamaño de muestra más grande de $n = 500$. Se elige un nivel de significancia α menor y un tamaño de muestra n más grande en virtud de las consecuencias más graves asociadas con la prueba de un fármaco comercial.

Prueba exhaustiva de hipótesis En esta sección describimos los componentes individuales utilizados en una prueba de hipótesis; las siguientes secciones combinarán tales componentes en procedimientos más exhaustivos. Podemos probar aseveraciones sobre parámetros de población utilizando el método del valor P que se resume en la figura 8-8, el método tradicional que se resume en la figura 8-9, o bien, podemos emplear un intervalo de confianza. En el caso de pruebas de hipótesis de dos colas, construya un intervalo de confianza con un nivel de confianza de $1 - \alpha$; pero para una prueba de hipótesis de una cola, con un nivel de significancia α , construya un intervalo de confianza con un nivel de confianza de $1 - 2\alpha$. (Véase la tabla 8-2 para los casos comunes). Después de construir el intervalo de confianza, use este criterio:

Un estimado de intervalo de confianza de un parámetro de población contiene los valores probables de tal parámetro. Por lo tanto, debemos rechazar la aseveración de que el parámetro de población tiene un valor que no está incluido en el intervalo de confianza. Advertencia: En algunos casos, una conclusión basada en un intervalo de confianza puede ser diferente de una conclusión basada en una prueba de hipótesis. Vea los comentarios en las secciones individuales.

Los ejercicios de esta sección incluyen componentes aislados de las pruebas de hipótesis, pero las siguientes secciones incluirán pruebas de hipótesis completas y exhaustivas.

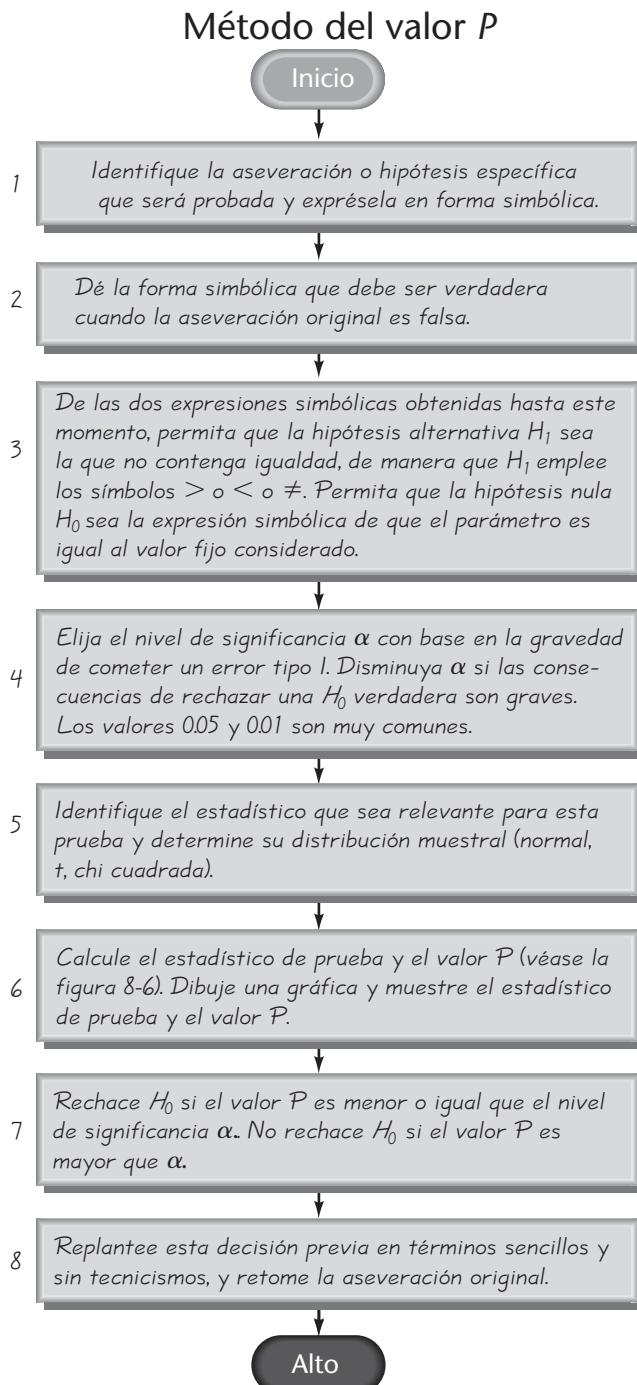
Parte 2: Más allá de lo básico de la prueba de hipótesis: La potencia de una prueba

Potencia de una prueba: Utilizamos β para denotar la probabilidad de no rechazar una hipótesis nula falsa (error tipo II). Se deduce que $1 - \beta$ es la probabilidad de rechazar una hipótesis nula falsa. Los especialistas en estadística se refieren a esta probabilidad como la *potencia* de una prueba, y con frecuencia la utilizan al evaluar la eficacia de la prueba para reconocer que una hipótesis nula es falsa.

Definición

La **potencia** de una prueba de hipótesis es la probabilidad $(1 - \beta)$ de rechazar una hipótesis nula falsa; se calcula utilizando un nivel de significancia α particular y un valor específico del parámetro de población que representa una alternativa al valor considerado como verdadero en la hipótesis nula. Es decir, la potencia de una prueba de hipótesis es la probabilidad de sustentar una hipótesis alternativa que es verdadera.

Figura 8-8 Método del valor P



Método del intervalo de confianza

Construya un intervalo de confianza con un nivel de confianza seleccionado de la misma forma que en la tabla 8-2.

Puesto que un estimado del intervalo de confianza de un parámetro de población contiene los probables valores de tal parámetro, rechace la aseveración de que el parámetro de población tiene un valor que no está incluido en el intervalo de confianza.

Figura 8-9 Método tradicional

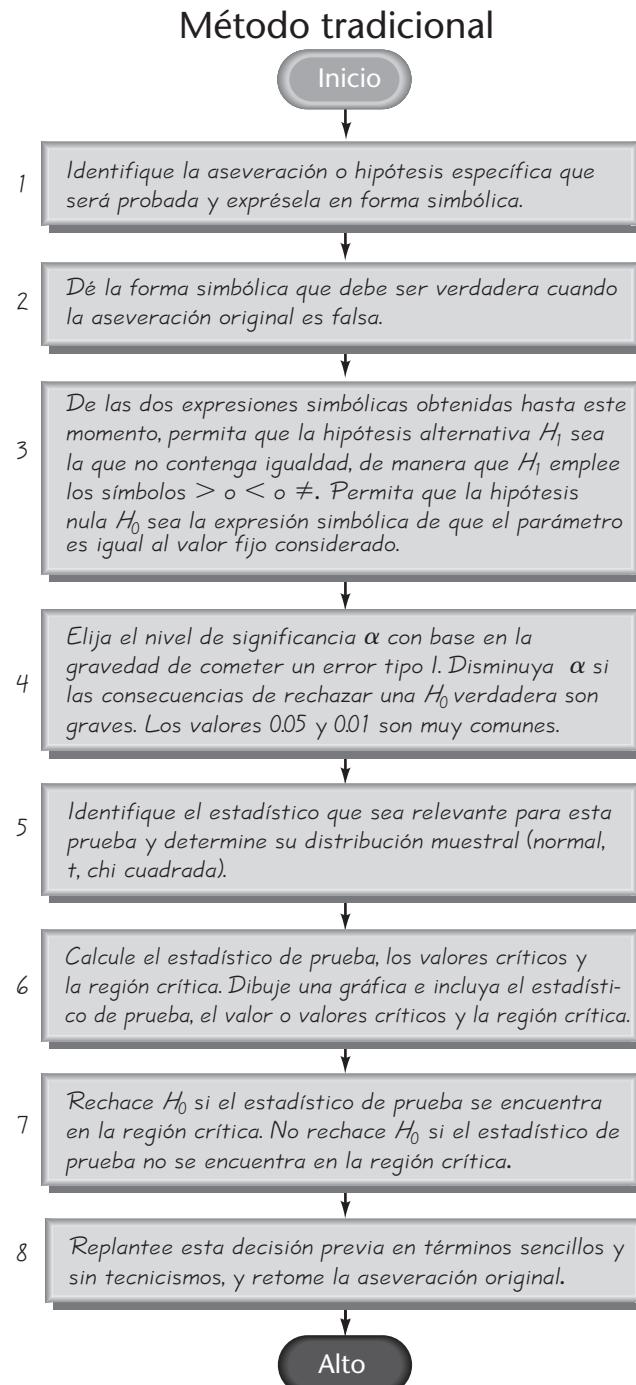


Tabla 8-2 Nivel de confianza para un intervalo de confianza

Nivel de significancia para la prueba de hipótesis	Prueba de dos colas	Prueba de una cola
0.01	99%	98%
0.05	95%	90%
0.10	90%	80%

Observe que en la definición anterior, la determinación de potencia requiere de un valor específico que es una alternativa al valor supuesto en la hipótesis nula. En consecuencia, una prueba de hipótesis puede tener muchos valores de potencia diferentes, dependiendo de los valores específicos elegidos como alternativas a la hipótesis nula.

EJEMPLO Potencia de una prueba de hipótesis Suponga que tenemos la hipótesis nula, la hipótesis alternativa, el nivel de significancia y los datos muestrales siguientes

$$H_0: p = 0.5 \quad H_1: p \neq 0.5 \quad \text{Nivel de significancia: } \alpha = 0.05$$

$$\begin{array}{lll} \text{Tamaño} & & \\ \text{muestral:} & n = 100 & \text{Proporción muestral: } \hat{p} = 0.57 \end{array}$$

Utilizando sólo estos componentes, podemos realizar una prueba de hipótesis completa. (El estadístico de prueba es $z = 1.4$, los valores críticos son $z = \pm 1.96$, el valor P es 0.1616, no rechazamos la hipótesis nula y concluimos que no existe evidencia suficiente para justificar el rechazo de la aseveración de que la proporción poblacional es igual a 0.5). Sin embargo, la determinación de la potencia requiere de otro elemento: un valor *específico* de p para utilizarlo como una alternativa al valor de $p = 0.5$ supuesto en la hipótesis nula. Si utilizamos los componentes de prueba anteriores, junto con distintos valores alternativos de p , obtenemos los siguientes ejemplos de valores de potencia. (Los valores de potencia se pueden calcular utilizando algunos programas de cómputo como Minitab, o bien, de manera manual. Como los cálculos de la potencia son complicados, en esta sección sólo el ejercicio 47 se refiere a la potencia, y éste incluye un procedimiento para su cálculo).

Valor alternativo específico de p	β	Poder de la prueba ($1 - \beta$)
0.3	0.013	0.987
0.4	0.484	0.516
0.6	0.484	0.516
0.7	0.013	0.987

INTERPRETACIÓN DE LA POTENCIA Con base en la lista anterior de valores de la potencia, observamos que esta prueba de hipótesis tiene una potencia de 0.987 (o 98.7%) de rechazo de $H_0: p = 0.5$, cuando la proporción poblacional p es en realidad 0.3. Es decir, si la verdadera proporción poblacional es igual a 0.3, existe una probabilidad del 98.7% de llegar a la conclusión correcta de rechazar la hipótesis nula falsa que plantea que $p = 0.5$. De manera similar, existe una probabilidad de 0.516 de rechazar $p = 0.5$, cuando el verdadero valor de p es 0.4. Parece lógico que esta prueba es más eficaz para rechazar la aseveración de $p = 0.5$ cuando la proporción poblacional es en realidad 0.3, que cuando ésta es en realidad 0.4. [Cuando se identifica animales que se supone son caballos, existe una mayor posibilidad de rechazar un elefante como caballo (a causa de la gran diferencia) que rechazar una mula como caballo]. En general, al aumentar la diferencia entre el valor paramétrico supuesto y el valor paramétrico real, se incrementa la potencia.

Potencia y diseño de experimentos Si se realiza una prueba de hipótesis con datos muestrales que sólo consisten en unas cuantas observaciones, la potencia será baja, pero ésta se incrementa conforme aumenta el tamaño muestral (y los otros componentes permanecen iguales). Además de incrementar el tamaño muestral, existen otras formas para aumentar la potencia, como incrementar el nivel de

significancia, usar un valor más extremo para el parámetro poblacional o disminuir la desviación estándar. Así como 0.05 es una elección común para el nivel de significancia, una potencia de al menos 0.80 es un requisito común para determinar que una prueba de hipótesis es efectiva. (Algunos especialistas en estadística argumentan que la potencia debe ser más alta, como 0.85 o 0.90).

Al diseñar un experimento, podríamos considerar cuál sería una diferencia importante entre el valor aseverado de un parámetro y su valor verdadero. Si se pone a prueba la eficacia de un método de selección del género, tal vez un cambio de 0.5 a 0.501 en la proporción de niñas sea poco importante, mientras que un cambio de 0.5 a 0.6 sí podría serlo. Este tipo de magnitudes de las diferencias afectan la potencia. Cuando se diseña un experimento, con frecuencia se puede plantear la meta de lograr una potencia de al menos 0.80 para determinar el tamaño muestral mínimo requerido. Por ejemplo, he aquí una aseveración similar a otra tomada de un artículo del *Journal of the American Medical Association*: “El diseño de prueba supuso que con un nivel de significancia de 0.05, se necesitarían 153 sujetos elegidos al azar para lograr una potencia del 80% con el fin de detectar una reducción de 0.5 a 0.4 en la tasa de enfermedades coronarias”. Antes de llevar a cabo el experimento, los investigadores determinaron que para lograr una potencia de al menos 0.80 necesitaban por lo menos 153 sujetos elegidos al azar. Ante factores tales como las tasas de deserción, los investigadores quizás necesiten un poco más de 153 sujetos.

8-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Prueba de que $p = 0.5$.** Un artículo periodístico afirma que “con base en una encuesta reciente, se demostró que el 50% de los conductores de camiones fuman”. ¿Qué es incorrecto en esta aseveración?
2. **Hipótesis nula.** El gerente de control de calidad de una empresa embotelladora de bebidas de cola afirma que “la cantidad media de bebida en nuestras latas es de al menos 12 onzas”. Para probar esta aseveración, exprese la hipótesis nula y la hipótesis alternativa en forma simbólica. ¿La aseveración original corresponde a la hipótesis nula, a la hipótesis alternativa o a ninguna de las dos?
3. **Estadístico de prueba y valor crítico.** ¿Cuál es la diferencia entre un estadístico de prueba y un valor crítico?
4. **Valor P .** Usted desarrolló un nuevo fármaco y asegura que disminuye el colesterol, de manera que expresa la aseveración como $\mu < 100$ (donde 100 es un índice estandarizado de colesterol). ¿Qué valor P preferiría obtener: 0.04 o 0.01? ¿Por qué?

Conclusiones sobre aseveraciones. En los ejercicios 5 a 8, tome una decisión sobre la aseveración dada. (No emplee procedimientos formales y cálculos exactos. Use sólo la regla del suceso infrecuente descrita en la sección 8-1 y haga estimaciones subjetivas para determinar si los sucesos son probables).

5. Aseveración: Al lanzar una moneda se obtienen más caras, y resultan 11 caras en 20 lanzamientos.
6. Aseveración: La mayoría de las personas entre 18 y 29 años edad votaron en la última elección presidencial, y las encuestas revelaron que, de 1000 votantes en ese rango de edad, votaron 170.
7. Aseveración: El pulso cardíaco medio (en latidos por minuto) de hombres es mayor que 60, y una muestra aleatoria de 400 hombres tiene una tasa de pulso cardíaco de 69.4.

- 8.** Aseveración: Estudiantes universitarios tienen puntuaciones de CI que varían menos que la población general con $\sigma = 15$, y una muestra aleatoria de 500 estudiantes universitarios arroja puntuaciones de CI con $s = 10.2$.

Identificación de H_0 y H_1 . En los ejercicios 9 a 16, examine la afirmación dada, después exprese la hipótesis nula H_0 y la hipótesis alternativa H_1 de manera simbólica. Asegúrese de emplear el símbolo correcto (μ , p , σ) para el parámetro indicado.

- 9.** Más del 25% de los usuarios de Internet pagan sus facturas en línea.
- 10.** La mayoría de los hogares tienen teléfono.
- 11.** El peso medio de mujeres que han ganado el título de Miss America es igual a 121 libras.
- 12.** La altura media de la rodilla de un hombre sentado es de 20.7 pulgadas.
- 13.** Las puntuaciones de CI de profesores universitarios tienen una desviación estándar menor que 15, que es la desviación estándar de la población general.
- 14.** Los profesores de preparatoria tienen ingresos con una desviación estándar menor que \$20,000.
- 15.** Los dulces M&M sencillos tienen un peso medio de al menos 0.8535 g.
- 16.** El porcentaje de empleados que consiguen trabajo por medio de su universidad no es mayor del 2%.

Cálculo de valores críticos. En los ejercicios 17 a 24, calcule los valores z críticos. En cada caso, suponga que se aplica la distribución normal.

- 17.** Prueba de dos colas; $\alpha = 0.05$.
- 18.** Prueba de dos colas; $\alpha = 0.01$.
- 19.** Prueba de cola derecha; $\alpha = 0.01$.
- 20.** Prueba de cola izquierda; $\alpha = 0.05$.
- 21.** $\alpha = 0.10$; H_1 es $p \neq 0.17$.
- 22.** $\alpha = 0.10$; H_1 es $p > 0.18$.
- 23.** $\alpha = 0.02$; H_1 es $p < 0.19$.
- 24.** $\alpha = 0.005$; H_1 es $p \neq 0.20$.

Cálculo de estadísticos de prueba. En los ejercicios 25 a 28, calcule el valor del estadístico de prueba z utilizando

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

- 25. Encuesta Gallup sobre el tabaquismo.** La aseveración es que la proporción de adultos que fumaron un cigarrillo la semana pasada es menor que 0.25, y los estadísticos de muestra incluyen $n = 1018$ sujetos, de los cuales 224 dicen que fumaron un cigarrillo la semana pasada.
- 26. Experimento de genética.** La aseveración es que la proporción de plantas de guisantes con vainas amarillas es igual a 0.25 (o 25%), y los estadísticos de muestra incluyen $n = 580$ plantas de guisantes, de las cuales 152 tienen vainas amarillas.
- 27. Encuesta Gallup sobre satisfacción laboral.** La aseveración es que más del 75% de los empleados están satisfechos con su trabajo, y los estadísticos de muestra incluyen a 580 adultos empleados, de los cuales 516 afirman sentirse satisfechos con su trabajo.

- 28. Demandas por negligencia médica.** La aseveración es que la mayoría de las demandas por negligencia médica se retiran o se rechazan, y una muestra aleatoria de 500 demandas incluye 349 que fueron retiradas o rechazadas.

Cálculo de valores P. En los ejercicios 29 a 36, utilice la información dada para calcular el valor P. (Sugerencia: Consulte la figura 8-6).

- 29.** El estadístico de prueba, en una prueba de cola derecha, es $z = 1.00$.
- 30.** El estadístico de prueba, en una prueba de cola izquierda, es $z = -2.00$.
- 31.** El estadístico de prueba, en una prueba de dos colas, es $z = 1.96$.
- 32.** El estadístico de prueba, en una prueba de dos colas, es $z = -0.50$.
- 33.** Con $H_1: p > 0.333$, el estadístico de prueba es $z = 1.50$.
- 34.** Con $H_1: p \neq 0.667$, el estadístico de prueba es $z = 2.05$.
- 35.** Con $H_1: p \neq 1/4$, el estadístico de prueba es $z = -1.75$.
- 36.** Con $H_1: p < 2/3$, el estadístico de prueba es $z = -0.45$.

Redacción de conclusiones. En los ejercicios 37 a 40, plantee la conclusión final en términos sencillos y sin tecnicismos. Asegúrese de retomar la aseveración original. (Sugerencia: Consulte la figura 8-7).

- 37.** Aseveración original: La proporción de hombres golfistas es menor que 0.5.
Conclusión inicial: Rechazar la hipótesis nula.
- 38.** Aseveración original: La proporción de mujeres entrenadoras es mayor al 50%.
Conclusión inicial: Rechazar la hipótesis nula.
- 39.** Aseveración original: La proporción de dulces M&M rojos difiere de 0.13.
Conclusión inicial: No rechazar la hipótesis nula.
- 40.** Aseveración original: La proporción de fumadores que tienen problemas para dormir es igual a 0.34.
Conclusión inicial: Rechazar la hipótesis nula.

Identificación de errores tipo I y tipo II. En los ejercicios 41 a 44, identifique el error tipo I y el error tipo II correspondiente a la hipótesis dada.

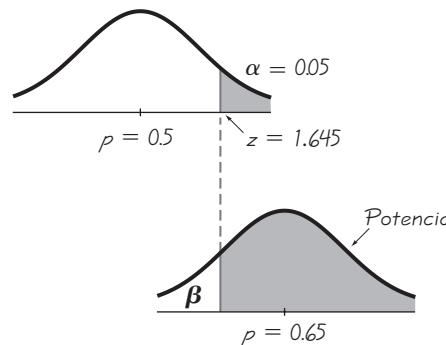
- 41.** La proporción de demandas por negligencia médica resueltas es de 0.25.
- 42.** La proporción de teléfonos privados en Nevada es de 0.524.
- 43.** La proporción de asesinatos aclarados por un arresto es de 0.62.
- 44.** La proporción de accidentes automovilísticos que ocurren a menos de una milla de casa del conductor es de 0.23.

8-2 MÁS ALLÁ DE LO BÁSICO

- 45. Prueba innecesaria.** Para probar la aseveración de que la mayoría de los estadounidenses adultos están en contra de la pena de muerte para una persona sentenciada por asesinato, se obtiene una muestra aleatoria de 491 adultos, y el 27% de ellos se manifiesta en contra de la pena de muerte (según datos de una encuesta Gallup). Calcule el valor P. ¿Por qué no es necesario seguir los pasos para realizar una prueba formal de hipótesis?
- 46. Nivel de significancia.** Si se rechaza una hipótesis nula con un nivel de significancia de 0.05, ¿también se rechaza con un nivel de significancia de 0.01? ¿Por qué?

47. Potencia de una prueba. Suponga que utiliza un nivel de significancia de $\alpha = 0.05$ para probar la aseveración de que $p > 0.5$ y que su muestra es aleatoria simple con tamaño $n = 64$.

- a. Suponiendo que la proporción poblacional verdadera es 0.65, calcule la potencia de la prueba, que es la probabilidad de rechazar la hipótesis nula cuando es falsa. (En el siguiente procedimiento, nos referimos a $p = 0.5$ como el valor “supuesto”, ya que éste se supuso en la hipótesis nula; nos referimos a $p = 0.65$ como el valor “alternativo”, puesto que éste es el valor de la proporción poblacional utilizado como alternativa a 0.5). Utilice el siguiente procedimiento y observe la figura que se muestra a continuación.



Paso 1: Use el nivel de significancia y calcule el valor o valores z críticos. (Para una prueba de cola derecha existe un solo valor z crítico que es positivo; para una prueba de cola izquierda existe un solo valor z crítico que es negativo; y para una prueba de dos colas tendremos un valor z crítico que es negativo, junto con otro valor z crítico que es positivo).

Paso 2: En la expresión que se muestra abajo para el estadístico de prueba, sustituya el valor supuesto de p (utilizado en la hipótesis nula). Evalúe $1 - p$ y sustituya ese resultado por el valor de q . También sustituya el valor (o valores) crítico(s) para z . Después resuelva para el estadístico muestral \hat{p} . (Si la prueba es de dos colas, sustituya el valor crítico de z que es positivo y luego resuelva para el estadístico de prueba \hat{p} . Luego sustituya el valor crítico de z que es negativo y resuelva para el estadístico de prueba \hat{p}). Así pues, una prueba de dos colas debe dar por resultado dos valores diferentes de \hat{p} .) El valor o valores resultantes de \hat{p} separan la región o las regiones donde la hipótesis nula se rechaza, de la región donde ésta no se rechaza.

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Paso 3: El cálculo de la potencia requiere un valor específico de p que se utilizará como alternativa al valor supuesto en la hipótesis nula. Identifique este valor alternativo de p (no el valor utilizado en la hipótesis nula), dibuje una curva normal con este valor alternativo al centro, y grafique el valor o los valores de \hat{p} obtenidos en el paso 2.

Paso 4: Remítase a la gráfica del paso 3 y calcule el área de la nueva región crítica limitada por el valor o los valores de \hat{p} que se obtuvieron en el paso 2. (*Advertencia:* cuando evalúe $\sqrt{pq/n}$, asegúrese de utilizar el valor alternativo de p y no el valor de p empleado para la hipótesis nula). Ésta es la probabilidad de rechazar la hipótesis nula, dado que el valor alternativo de p es el valor verdadero de la proporción poblacional. Puesto que se trata de la probabilidad de rechazar la hipótesis nula falsa, representa la *potencia* de la prueba.

- b. Calcule β , que es la probabilidad de *no* rechazar la hipótesis nula falsa. Es fácil determinar el valor de β calculando el complemento de la potencia.

- 48. Cálculo de tamaño muestral.** Un investigador planea hacer una prueba de hipótesis utilizando la hipótesis alternativa de $H_1: p < 0.4$, y piensa utilizar un nivel de significancia de $\alpha = 0.05$. Calcule el tamaño muestral que se requiere para lograr una potencia de al menos el 80% para detectar una reducción en p de 0.4 a 0.3. (*Este ejercicio es muy difícil. Sugerencia: Consulte el ejercicio 47.*)

Prueba de una aseveración respecto de una proporción

Concepto clave En esta sección se presentan procedimientos completos para probar una hipótesis (o aseveración) respecto de una proporción poblacional. Aquí se utilizan los componentes presentados en la sección 8-2 para el método del valor P , el método tradicional y el uso de intervalos de confianza. Además de poner a prueba aseveraciones respecto de proporciones poblacionales, podemos emplear los mismos procedimientos para poner a prueba aseveraciones sobre probabilidades o equivalentes decimales de porcentajes. Los métodos de esta sección utilizan la distribución normal como aproximación de la distribución de probabilidad binomial.

Los siguientes son ejemplos de los tipos de aseveraciones que podremos someter a prueba:

- Más del 50% de los empleados consiguen trabajo por medio de redes de contactos.
- Los sujetos que toman el fármaco Lipitor, que reduce el colesterol, experimentan dolores de cabeza en una proporción mayor que el 7% registrado entre quienes no toman Lipitor.
- El porcentaje de televidentes nocturnos que ven *The Late Show with David Letterman* es igual al 20%.

Las aseveraciones sobre una proporción poblacional suelen probarse utilizando una distribución normal como aproximación de la distribución binomial (véase la sección 6-6). En vez de utilizar exactamente los mismos métodos de la sección 6-6, empleamos una forma diferente, pero equivalente, del estadístico de prueba mostrado a continuación, y no incluimos la corrección por continuidad (debido a que su efecto tiende a ser muy pequeño en muestras grandes). Si no se satisfacen todos los requisitos, debemos utilizar otros métodos que no se describen en esta sección. En esta sección todos los ejemplos y ejercicios incluyen casos en que los requisitos se cumplen, de manera que la distribución muestral de proporciones muestrales puede aproximarse por medio de la distribución normal.

Prueba de aseveraciones sobre una proporción poblacional p

Requisitos

1. Las observaciones muestrales son una muestra aleatoria simple.
2. Se satisfacen las condiciones para una *distribución binomial*. (Existe un número fijo de ensayos independientes con probabilidades constantes, y cada ensayo tiene dos categorías de resultados de “éxito” y “fracaso”).
3. Se satisfacen las condiciones $np \geq 5$ y $nq \geq 5$, por lo tanto, la **distribución binomial de proporciones muestrales puede aproximarse con una distribución normal, con $\mu = np$ y $\sigma = \sqrt{npq}$** (como se describió en la sección 6-6). Observe que p es la proporción *supuesta* que se utiliza en la aseveración y no la proporción muestral.

continúa



Proceso de aprobación de un fármaco

Lograr la aprobación de la FDA para un fármaco nuevo es costoso y requiere de mucho tiempo. El proceso inicia con un **estudio de fase I**, en el que se prueba la seguridad del fármaco con un grupo pequeño de voluntarios (de 20 a 100). En la **fase II**, se prueba la eficacia del fármaco en ensayos aleatorios con un grupo más grande de sujetos (entre 100 y 300). Esta fase a menudo incluye sujetos asignados al azar a un grupo de tratamiento o a un grupo placebo. En la **fase III**, la meta consiste en comprender mejor la eficacia del fármaco, así como sus efectos adversos. En la fase III generalmente participan de 1000 a 3000 sujetos, y suele requerir varios años de pruebas. Lisa Gibbs escribió en la revista *Money* que “la industria (farmacéutica) afirma que por cada 5000 tratamientos que se someten a prueba, sólo cinco llegan a los ensayos clínicos y sólo uno termina en las farmacias”. Los estimados del costo total varían desde \$40 millones hasta \$1,500 millones.

Notación

n = tamaño de muestra o número de ensayos

$$\hat{p} = \frac{x}{n} \text{ (proporción muestral)}$$

p = proporción de la población (utilizada en la hipótesis nula)

$$q = 1 - p$$

Estadístico de prueba para probar una aseveración sobre una proporción

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Valores P: Utilice la distribución normal estándar (tabla A-2) y remítase a la figura 8-6.

Valores críticos: Utilice la distribución normal estándar (tabla A-2).



EJEMPLO Obtención de empleo por medio de redes de contactos El problema del capítulo incluyó los siguientes resultados de encuesta: de 703 empleados elegidos al azar, el 61% obtuvo trabajo por medio de redes de contactos. Utilice los datos muestrales, con un nivel de significancia de 0.05, para probar la aseveración de que la mayoría de los empleados (más del 50%) consiguen su trabajo por medio de redes de contactos. El siguiente es un resumen de la aseveración y de los datos muestrales:

Aseveración: La mayoría de los empleados consigue trabajo por medio de redes de contactos. Es decir, $p > 0.5$.

Datos muestrales: $n = 703$ y $\hat{p} = 0.61$

Daremos soluciones utilizando el método del valor P , el método tradicional y los intervalos de confianza. Sin embargo, antes de continuar debemos verificar que se satisfagan los requisitos necesarios.

REQUISITO ✓ Lea con atención los tres siguientes requisitos.

1. Al parecer, la encuesta se realizó utilizando métodos de muestreo adecuados, de manera que podemos suponer que la muestra es aleatoria simple.
2. Hay un número fijo (703) de ensayos independientes con dos categorías (el empleado consiguió trabajo por medio de redes de contactos o no). (Técnicamente, los ensayos no son independientes, pero pueden tratarse como tales utilizando el siguiente lineamiento presentado en la sección 5-3: “Cuando se hace un muestreo sin reemplazo, los sucesos pueden manejarse como si fueran independientes si el tamaño muestral no es mayor que el 5% del tamaño poblacional. Es decir, $n \leq 0.05N$ ”.)
3. Los requisitos $np \geq 5$ y $nq \geq 5$ se satisfacen con $n = 703$, $p = 0.5$ y $q = 0.5$. [Obtenemos $np = (703)(0.5) = 351.5 \geq 5$ y $nq = (703)(0.5) = 351.5 \geq 5$.]

Con los tres requisitos satisfechos, ahora procedemos a realizar una prueba de hipótesis formal. A continuación se ejemplifican el método del valor P , el método tradicional y el uso de intervalos de confianza. ✓

El método del valor P

Cuando ponemos a prueba la aseveración $p > 0.5$ del ejemplo anterior, los siguientes pasos corresponden al método del valor P de la prueba de hipótesis, que se resumen en la figura 8-8.

- Paso 1: La aseveración original en forma simbólica es $p > 0.5$.
- Paso 2: El opuesto de la aseveración original es $p \leq 0.5$.
- Paso 3: De las dos expresiones simbólicas anteriores, la expresión $p > 0.5$ no contiene igualdad, por lo que se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de que p iguala el valor fijo de 0.5. Por consiguiente, podemos expresar H_0 y H_1 de la siguiente manera:

$$H_0: p = 0.5$$

$$H_1: p > 0.5$$

- Paso 4: En ausencia de cualquier circunstancia especial, seleccionamos $\alpha = 0.05$ para el nivel de significancia.
- Paso 5: Para esta prueba se usa la distribución normal y, en virtud de que estamos probando una aseveración acerca de una proporción poblacional p , el estadístico de prueba \hat{p} es relevante y la distribución muestral de las proporciones muestrales \hat{p} se aproxima por medio de una distribución normal.
- Paso 6: El estadístico de prueba es $z = 5.83$, que se calcula de la siguiente manera:

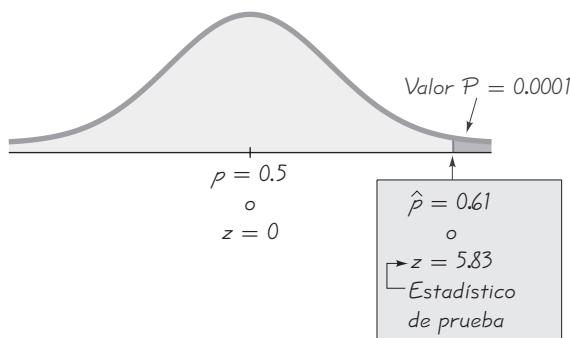
$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.61 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{703}}} = 5.83$$

Ahora obtenemos el valor P utilizando el siguiente procedimiento, que se muestra en la figura 8-6:

Prueba de cola derecha:	Valor P = área a la derecha del estadístico de prueba z
Prueba de cola izquierda:	Valor P = área a la izquierda del estadístico de prueba z
Prueba de dos colas:	Valor P = <i>dos veces</i> el área de la región extrema limitada por el estadístico de prueba z

Puesto que la prueba de hipótesis que estamos realizando es de cola derecha con un estadístico de prueba de $z = 5.83$, el valor P es el área a la derecha de $z = 5.83$. Si nos remitimos a la tabla A-2, observamos que para los valores de $z = 3.50$ y mayores utilizamos 0.0001 para el área acumulativa a la derecha del estadístico de prueba. Por lo tanto, el valor P es 0.0001. La figura 8-10 muestra el estadístico de prueba y el valor P para este ejemplo.

- Paso 7: Puesto que el valor P de 0.0001 es menor o igual que el nivel de significancia $\alpha = 0.05$, rechazamos la hipótesis nula.
- Paso 8: Concluimos que existe suficiente evidencia muestral para sustentar la aseveración de que la mayoría de los empleados consiguen trabajo por medio de redes de contactos. (Consulte la figura 8-7 para ayudarse a redactar esta conclusión final).

Figura 8-10**Método del valor P** 

El método tradicional

El método tradicional de prueba de hipótesis se resume en la figura 8-9, y los pasos 1 al 5 son iguales a los pasos 1 al 5 del método del valor P , como se puede ver arriba. Así que continuamos con el paso 6 del método tradicional.

Paso 6: El estadístico de prueba es $z = 5.83$, como vimos en el método del valor P .

Ahora calculamos el valor crítico (en vez del valor P). Se trata de una prueba de cola derecha, de manera que el área de la región crítica es una área de $\alpha = 0.05$ en la cola derecha. Si nos remitimos a la tabla A-2 y aplicamos los métodos de la sección 6-2, obtenemos que el valor crítico de $z = 1.645$ se encuentra en el límite de la región crítica. Observe la figura 8-3 de la página 395, que muestra la región crítica, el valor crítico y el estadístico de prueba.

Paso 7: Como el estadístico de prueba se localiza dentro de la región crítica, rechazamos la hipótesis nula.

Paso 8: Concluimos que existe suficiente evidencia muestral para sustentar la aseveración de que la mayoría de los empleados consiguen trabajo por medio de redes de contactos. (Consulte la figura 8-7 para ayudarse a redactar esta conclusión final).

Método del intervalo de confianza

La aseveración de $p > 0.5$ puede probarse con un nivel de significancia de 0.05, construyendo un intervalo de confianza del 90%. (Véase la tabla 8-2. En general, para las pruebas de hipótesis de dos colas, construya un intervalo de confianza con un nivel de confianza de $1 - \alpha$; pero las pruebas de hipótesis de una cola, con nivel de significancia α , requieren un intervalo de confianza con un nivel de confianza de $1 - 2\alpha$).

Ahora utilicemos el método del intervalo de confianza para probar la aseveración de $p > 0.5$, con datos muestrales que consisten en $n = 703$ y $\hat{p} = 0.61$ (como en el problema del capítulo). Si deseamos un nivel de significancia de $\alpha = 0.05$ en una prueba de cola derecha, empleamos un nivel de confianza del 90% con los métodos de la sección 7-2 para obtener este resultado: $0.580 < p < 0.640$. Como tenemos una confianza del 90% de que el valor verdadero de p está contenido dentro de los límites de 0.580 y 0.640, tenemos evidencia suficiente para sustentar la aseveración de que $p > 0.5$.

Advertencia: Cuando se prueban aseveraciones sobre una proporción poblacional, el método tradicional y el método del valor P son equivalentes en el sentido de que siempre dan los mismos resultados, pero el método del intervalo de confianza es un poco diferente. Tanto el método tradicional como el método del valor P utilizan

la misma desviación estándar basada en la *proporción aseverada* p , pero el intervalo de confianza emplea una desviación estándar estimada con base en la *proporción muestral* \hat{p} . Como consecuencia, es posible que en algunos casos los métodos tradicional y del valor P de prueba de una aseveración sobre una proporción produzcan una conclusión diferente a la del método del intervalo de confianza (véase el ejercicio 29). Una buena estrategia consiste en usar un intervalo de confianza para estimar una proporción poblacional, pero usar el método del valor P o el método tradicional para probar una hipótesis.

Cuando pruebe una aseveración sobre una proporción poblacional p , tenga cuidado en identificar correctamente la proporción muestral \hat{p} . En ocasiones la proporción muestral \hat{p} está dada directamente, pero en otros casos debe calcularse. Observe los siguientes ejemplos.

Afirmación dada	Cálculo de \hat{p}
El 10% de los automóviles deportivos observados son rojos.	\hat{p} está dada directamente: $\hat{p} = 0.10$
96 hogares encuestados tienen servicio de televisión por cable y 54 no lo tienen.	\hat{p} debe calcularse utilizando $\hat{p} = x/n$. $\hat{p} = \frac{x}{n} = \frac{96}{96 + 54} = 0.64$

EJEMPLO Experiments genéticos de Mendel Cuando Gregor Mendel realizó sus famosos experimentos de hibridación con guisantes, uno de ellos dio por resultado descendencia que consistía en 428 plantas de guisantes con vainas verdes y 152 plantas de guisantes con vainas amarillas. Según la teoría de Mendel, $1/4$ de los vástagos de guisantes debían tener vainas amarillas. Utilice un nivel de significancia de 0.05, con el método del valor P , para probar la aseveración de que la proporción de vástagos de guisantes con vainas amarillas es igual a $1/4$.

SOLUCIÓN

REQUISITO ✓ 1. Con base en el diseño del experimento, es razonable suponer que se trata de una muestra aleatoria simple. 2. Existe un número fijo ($428 + 152 = 580$) de ensayos independientes con dos categorías (vainas verdes o vainas amarillas). 3. Los requisitos de $np \geq 5$ y $nq \geq 5$ se satisfacen con $n = 580$, $p = 1/4$ y $q = 3/4$. [Obtenemos $np = (580)(1/4) = 145 \geq 5$ y $nq = (580)(3/4) = 435 \geq 5$]. Con los tres requisitos satisfechos, ahora podemos realizar una prueba formal de hipótesis. ✓

Usamos el método del valor P , resumido en la figura 8-8 de la sección 8-2. Observe que $n = 428 + 152 = 580$, $\hat{p} = 152/580 = 0.262$, y, para los fines de la prueba, suponemos que $p = 1/4 = 0.25$.

Paso 1: La aseveración original dice que la proporción de plantas de guisantes con vainas amarillas es igual a $1/4$. Expresamos esto en forma simbólica como $p = 0.25$.

Paso 2: El opuesto de la aseveración original es $p \neq 0.25$

Paso 3: Como $p \neq 0.25$ no contiene igualdad, se convierte en H_1 . Obtenemos

$$H_0: p = 0.25 \quad (\text{hipótesis nula y aseveración original})$$

$$H_1: p \neq 0.25 \quad (\text{hipótesis alternativa})$$

Paso 4: El nivel de significancia es $\alpha = 0.05$.

continúa

Paso 5: Puesto que la aseveración implica a la proporción p , el estadístico relevante para esta prueba es la proporción muestral \hat{p} , y la distribución muestral de proporciones muestrales se aproxima por medio de la distribución normal.

Paso 6: El estadístico de prueba de $z = 0.67$ se calcula de la siguiente manera:

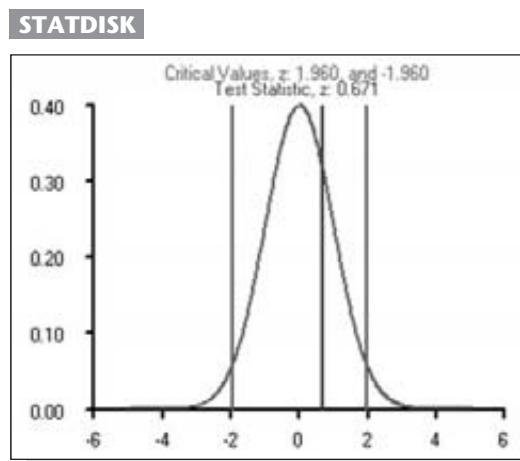
$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.262 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{580}}} = 0.67$$

Remítase a la figura 8-6 para el procedimiento del cálculo del valor P . La figura 8-6 indica que para esta prueba de dos colas, con el estadístico de prueba localizado a la derecha del centro (puesto que $z = 0.67$ es positivo), el valor P es el *doble* del área a la derecha del estadístico de prueba. En la tabla A-2, $z = 0.67$ tiene una área de 0.7486 a su izquierda, de manera que el área a la derecha de $z = 0.67$ es $1 - 0.7486 = 0.2514$, que duplicamos para obtener 0.5028.

Paso 7: Puesto que el valor P de 0.5028 es mayor que el nivel de significancia de 0.05, no rechazamos la hipótesis nula.

INTERPRETACIÓN Los métodos de prueba de hipótesis nunca nos permiten sustentar una aseveración de igualdad, de manera que no podemos concluir que la proporción de plantas de guisantes con vainas verdes sea igual a $1/4$. He aquí la conclusión correcta: no existe evidencia suficiente para justificar el rechazo de la aseveración de que $1/4$ de los vástagos de guisantes tienen vainas amarillas.

Método tradicional: Si repitiéramos el ejemplo anterior con el método tradicional de prueba de hipótesis, veríamos que en el paso 6 los valores críticos son $z = -1.96$ y $z = 1.96$. En el paso 7 no rechazaríamos la hipótesis nula, ya que el estadístico de prueba $z = 0.67$ no caería dentro de la región crítica. Observe la siguiente pantalla de STATDISK. Llegaríamos a la misma conclusión del método del valor P : no existe evidencia suficiente para justificar el rechazo de la aseveración de que $1/4$ de los vástagos de guisantes tienen vainas amarillas.



Método del intervalo de confianza: Si repitiéramos el ejemplo anterior con el método del intervalo de confianza, obtendríamos el siguiente intervalo de confianza del 95%: $0.226 < p < 0.298$. Puesto que los límites del intervalo de confianza contienen el valor aseverado de 0.25, concluimos que no existe evidencia suficiente

que justifique el rechazo de la aseveración de que 1/4 de los vástagos de guisantes tienen vainas amarillas. En este caso, el método del valor P , el método tradicional y el método del intervalo de confianza conducen a la misma conclusión. En otros casos relativamente raros, el método del valor P y el método tradicional podrían llevarnos a una conclusión diferente de la obtenida por medio del método del intervalo de confianza.

Fundamentos del estadístico de prueba: El estadístico de prueba empleado en esta sección se justifica señalando que cuando se usa la distribución normal para aproximar la distribución binomial, utilizamos $\mu = np$ y $\sigma = \sqrt{npq}$ para obtener

$$z = \frac{x - \mu}{\sigma} = \frac{x - np}{\sqrt{npq}}$$

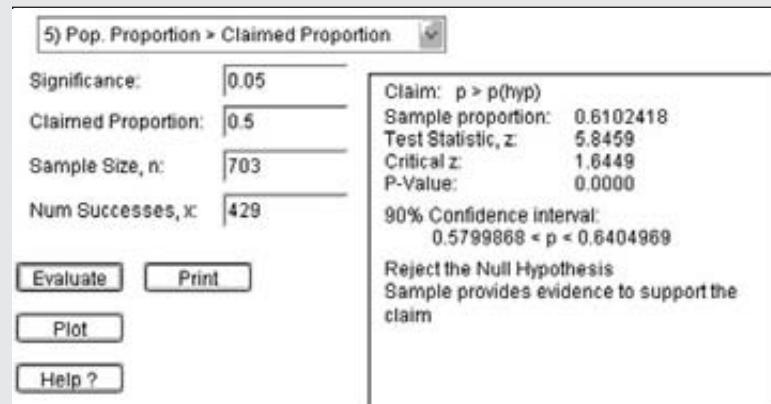
Empleamos la expresión anterior en la sección 6-6, junto con una corrección por continuidad, pero cuando se prueban aseveraciones sobre una proporción poblacional, hacemos dos modificaciones. Primero, no empleamos la corrección por continuidad porque su efecto suele ser muy pequeño para las muestras grandes que estamos considerando. Además, en vez de utilizar la expresión anterior para calcular el estadístico de prueba, empleamos una expresión equivalente obtenida al dividir el numerador y el denominador entre n , y sustituimos x/n por el símbolo \hat{p} para obtener el estadístico de prueba que estamos usando. El resultado final es que el estadístico de prueba es simplemente la misma puntuación estándar (de la sección 3-4) de $z = (x - \mu)/\sigma$, pero modificada para la notación binomial.

Uso de la tecnología

STATDISK Seleccione **Analysis, Hypothesis Testing, Proportion-One Sample**, después proceda a introducir los datos en el cuadro de diálogo. Vea la imagen que se presenta abajo, sobre el primer ejemplo de esta sección.

MINITAB Seleccione **Stat, Basic Statistics, 1 Proportion**, luego haga clic en el botón de “Summarized data”. Introduzca el tamaño de muestra y el número de éxitos,

STATDISK



después haga clic en **Options** y proceda a introducir los datos en el cuadro de diálogo. Para el nivel de confianza, introduzca el complemento del nivel de significancia. Para el valor de “test proportion”, ingrese la proporción empleada en la hipótesis nula. Para “alternative”, seleccione el formato usado para la hipótesis alternativa. En vez de usar una aproximación normal, el procedimiento predeterminado de Minitab consiste en determinar el valor P empleando un método exacto. Para utilizar el método de aproximación normal presentado en esta sección, haga clic en el

botón **Options** y luego en el recuadro que dice “Use test and interval based on normal distribution”.

EXCEL Primero introduzca el número de éxitos en la celda A1 e introduzca el número total de ensayos en la celda B1. Utilice el complemento Data Desk XL haciendo clic en **DDXL**, luego selecciones **Hypothesis Test**. En la función de teclear opciones, seleccione **Summ 1 Var Prop Test** (para probar una proporción aseverada usando datos resumidos de una variable). Haga clic en el ícono del lápiz en “Num successes” e introduzca A1. Haga clic en el ícono del lápiz en “Num trials” e introduzca B1. Haga clic en **OK**. Siga los cuatro pasos listados en el cuadro de diálogo. Después de marcar **Compute** en el paso 4, obtendrá el valor P , el estadístico de prueba y la conclusión.

T1-83/84 PLUS Presione **STAT**, seleccione **TESTS** y luego seleccione **1-PropZ-Test**. Introduzca el valor aseverado de la proporción de población para p_0 , luego introduzca los valores de x y n y después seleccione el tipo de prueba. Resalte **Calculate** y luego presione la tecla **ENTER**.

8-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Distribución.** Suponiendo que se satisfacen los requisitos listados en esta sección, ¿qué distribución se utiliza para probar una aseveración sobre una proporción poblacional? ¿Por qué?
- Proporción muestral.** Cuando se plantea una pregunta en una encuesta, 40 individuos responden *sí*, 60 de ellos responden *no* y no existen otras respuestas. ¿Cuál es la proporción muestral de respuestas *afirmativas* y qué notación se utiliza para representarla?
- Muestreo.** America Online realiza una encuesta en la que pide a usuarios de Internet que respondan una pregunta. De 96,772 respuestas, 76,885 son afirmativas. ¿Es válido usar esos resultados de muestra para probar la aseveración de que la mayoría de los integrantes de la población general responden “*sí*”?
- Método del valor *P*.** Se obtiene un valor *P* de 0.00001 al utilizar datos muestrales para probar la aseveración de que la mayoría de los accidentes automovilísticos ocurren a 5 millas de la casa del conductor. Interprete este valor *P* en el contexto de esta prueba de hipótesis. Es decir, ¿qué nos indica el valor *P*?

En los ejercicios 5 a 8, identifique los valores indicados o interprete la imagen de resultados.

- Experimentos de hibridación de Mendel.** En uno de los famosos experimentos de Mendel sobre la hibridación, se obtuvieron 8023 vástagos de guisantes y el 24.94% de ellos presentaban flores verdes. El resto tenía flores blancas. Considere una prueba de hipótesis que utiliza un nivel de significancia de 0.05 para probar la aseveración de que las plantas de guisantes con flores verdes se presentan en una proporción del 25%.
 - ¿Cuál es el estadístico de prueba?
 - ¿Cuáles son los valores críticos?
 - ¿Cuál es el valor *P*?
 - ¿Cuál es la conclusión?
 - ¿Se podría utilizar una hipótesis para “probar” que el porcentaje de plantas de guisantes con flores verdes es del 25%, como se aseveró?
- Encuesta de empleados.** En una encuesta de 703 empleados seleccionados al azar, el 15.93% obtuvo el trabajo a través de anuncios del periódico (según datos de Taylor Nelson Sofres Intereach). Considere una prueba de hipótesis que utiliza un nivel de significancia de 0.05 para probar la aseveración de que menos del 20% de los empleados obtuvo su trabajo a través de anuncios del periódico.
 - ¿Cuál es el estadístico de prueba?
 - ¿Cuál es el valor crítico?
 - ¿Cuál es el valor *P*?
 - ¿Cuál es la conclusión?
 - Con base en los resultados anteriores, ¿podemos concluir que el 15.93% es significativamente menor que el 20% de todo este tipo de pruebas de hipótesis?
- Interpretación de la pantalla de resultados.** Cuando se seleccionaron al azar 109,857 arrestos por crímenes federales, se encontró que 31,969 de ellos eran crímenes relacionados con drogas. Cuando se puso a prueba la aseveración de que más del 29% de los crímenes federales eran crímenes por drogas, se obtuvieron los resultados de la calculadora TI-83/84 Plus que aparecen al margen. Utilice estos resultados para probar la aseveración dada.
- Porcentaje de usuarios del teléfono.** Una encuesta reciente de 4276 hogares seleccionados al azar, reveló que 4019 de ellos tenían teléfono (según datos del U.S. Census Bureau). Se utilizó Minitab para probar la aseveración de que el porcentaje de hogares ahora es mayor que la tasa del 35% que se registró en 1920. Los resultados de Minitab se presentan en la siguiente página. La proporción actual de 4019/4276 (o 94%) parece ser significativamente mayor que la tasa del 35% de 1920; sin embargo, ¿existe evidencia suficiente para sustentar esta aseveración? Utilice un nivel de significancia de 0.01.

TI-83/84 Plus

```
1-PropZTest
PROP>.29
z=.7345177804
P=.2313165397
P=.29100558
n=109857
```

Minitab Resultados para el ejercicio 8

Test of $p = 0.35$ vs $p > 0.35$						
Sample	X	N	Sample p	95%	Z-Value	P-Value
				Lower Bound		
1	4019	4276	0.939897	0.933919	80.87	0.000

Prueba de aseveraciones sobre proporciones. En los ejercicios 9 a 24, pruebe la aseveración dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor P o el valor (o valores) crítico(s), la conclusión sobre la hipótesis nula y la conclusión final referente a la aseveración original. Utilice el método del valor P , a menos que su profesor especifique otra cosa.

9. **Selección del género para niñas.** El Genetics and IVF Institute llevó a cabo un ensayo clínico del método XSORT, diseñado para incrementar la probabilidad de concebir una niña. Mientras se escribía este libro, ya habían nacido 325 bebés de padres que utilizaron el método XSORT, y 295 de ellos fueron niñas. Utilice los datos muestrales con un nivel de significancia de 0.01 para probar la aseveración de que, con este método, la probabilidad de que un bebé sea niña es mayor que 0.5. ¿Parece que el método funciona?
10. **Selección del género para niños.** El Genetics and IVF Institute llevó a cabo un ensayo clínico del método YSORT, diseñado para incrementar la probabilidad de concebir un hijo varón. Mientras se escribía este libro, ya habían nacido 51 bebés de padres que utilizaron el método YSORT, y 39 de ellos fueron niños. Utilice los datos muestrales con un nivel de significancia de 0.01 para probar la aseveración de que, con este método, la probabilidad de que un bebé sea niño es mayor que 0.5. ¿Parece que el método funciona?
11. **Accidentes automovilísticos.** En un estudio de 11,000 accidentes automovilísticos, se descubrió que 5720 de ellos ocurrieron a 5 millas de casa del conductor (según datos de Progressive Insurance). Utilice un nivel de significancia de 0.01 para probar la aseveración de que más del 50% de los accidentes automovilísticos ocurren dentro de 5 millas de distancia de la casa del conductor. ¿Los resultados son cuestionables porque se basan en una encuesta patrocinada por una compañía de seguros?
12. **Viajes por medio de Internet.** De 734 usuarios de Internet elegidos al azar, se descubrió que 360 de ellos usan Internet para hacer planes de viaje (según datos de una encuesta Gallup). Utilice un nivel de significancia de 0.01 para probar la aseveración de que, de los usuarios de Internet, menos del 50% lo utiliza para hacer planes de viaje. ¿Los resultados son importantes para los agentes de viajes?
13. **Porcentaje de usuarios de correo electrónico.** La tecnología está cambiando de forma drástica la forma en que nos comunicamos. En 1997 una encuesta de 880 hogares estadounidenses reveló que 149 de ellos emplean el correo electrónico (según datos de *The World Almanac and Book of Facts*). Utilice los resultados de esta muestra para probar la aseveración de que más del 15% de los hogares estadounidenses emplean el correo electrónico. Use un nivel de significancia de 0.05. ¿Sería válida la conclusión aún hoy? ¿Por qué?
14. **Prueba de drogas a solicitantes de empleo.** En 1990 el 5.8% de quienes solicitaban empleo no pasaban la prueba de drogas. Con un nivel de significancia de 0.01, pruebe la aseveración de que el porcentaje que no pasa la prueba ahora es menor, si en una muestra actual de 1520 solicitantes de empleo hay 58 individuos que no pasan la prueba (según datos de la American Management Association). ¿Sugiere el resultado que en la actualidad un menor número de solicitantes consumen drogas?
15. **Teléfonos celulares y cáncer.** En un estudio de 420,095 usuarios daneses de teléfonos celulares, 135 sujetos desarrollaron cáncer cerebral o del sistema nervioso (según datos del *Journal of the National Cancer Institute*, reportados en *USA Today*). Pruebe la

aseveración, antes generalizada, de que estos tipos de cáncer se ven afectados por el uso de teléfonos celulares. Es decir, pruebe la aseveración de que los usuarios de teléfonos celulares desarrollan cáncer cerebral o del sistema nervioso en un porcentaje diferente al de 0.0340% registrado entre quienes no utilizan teléfonos celulares. Como este tema es de gran importancia, utilice un nivel de significancia de 0.005. ¿Deberían preocuparse los usuarios de teléfonos celulares acerca del cáncer cerebral o del sistema nervioso?

16. **Prueba de la eficacia de los parches de nicotina.** Un estudio realizado a fumadores que intentaban dejar el hábito con terapia de parches de nicotina reveló que 39 de ellos continuaban fumando un año después de iniciado el tratamiento y 32 habían dejado de fumar (según datos de “High-Dose Nicotine Patch Therapy”, de Dale *et al.*, *Journal of the American Medical Association*, vol. 274, núm. 17). Utilice un nivel de significancia de 0.10 para probar la aseveración de que, de los fumadores que intentan dejar el cigarrillo, la mayoría continúa fumando un año después de iniciar el tratamiento. ¿Sugieren estos resultados que la terapia de parches de nicotina es ineficaz?
17. **Precisión del verificador de precios de una tienda.** En un estudio de verificadores de precios, se verificaron 1234 artículos y se encontró que 20 de ellos tenían un sobreprecio y 1214 no lo tenían (según datos de “UPC Scanner Pricing Systems: Are They Accurate?” de Goodstein, *Journal of Marketing*, vol. 58). Emplee un nivel de significancia de 0.05 para probar la aseveración de que con los verificadores de precios, el 1% de las ventas tienen un sobreprecio. (Antes de que se utilizaran los verificadores de precios, se estimaba que el porcentaje de sobreprecio era de alrededor del 1%). Con base en estos resultados, ¿parece que los verificadores de precio ayudan a los consumidores a evitar los sobreprecios?
18. **Posposición de la muerte.** Una hipótesis interesante y generalizada dice que los individuos pueden posponer temporalmente su muerte para estar presentes en una festividad o en un suceso importante como un cumpleaños. En un estudio de este fenómeno, se descubrió que hubo 6062 muertes la semana previa al Día de Acción de Gracias, y 5938 muertes la semana posterior a esta festividad (según datos de “Holidays, Birthdays, and Postponement of Cancer Death” de Young y Hade, *Journal of the American Medical Association*, vol. 292, núm. 24). Si la gente puede posponer su muerte para después del Día de Acción de Gracias, entonces la proporción de muertes la semana anterior debe ser menor que 0.5. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la proporción de muertes durante la semana anterior al Día de Acción de Gracias es menor que 0.5. Con base en el resultado, ¿existe algún indicador de que la gente pue-
da posponer temporalmente su muerte para estar presente el Día de Acción de Gracias?
19. **Encuesta sobre la conducta de beber.** Una encuesta Gallup reciente de 976 adultos elegidos al azar reveló que 312 de ellos nunca beben. Utilice esos resultados de encuesta para probar la aseveración de que menos de 1/3 de todos los adultos nunca beben. Utilice un nivel de significancia de 0.05. Asimismo, examine la siguiente redacción de la pregunta real y determine si producirá respuestas honestas: “¿Con qué frecuencia, si acaso, consume bebidas alcohólicas como licor, vino o cerveza: todos los días, algunas veces por semana, una vez por semana, menos de una vez por semana, sólo en ocasiones especiales como Año Nuevo o festividades, o nunca?”
20. **Tabaquismo.** En una encuesta Gallup de 1018 adultos, se descubrió que el 22% fumó cigarrillos la semana anterior. Utilice un nivel de significancia de 0.05 para probar la aseveración de que menos del 25% de los adultos fumaron durante la semana anterior. ¿La conclusión cambiaría si, en vez de una encuesta Gallup, los resultados se hubieran obtenido de una encuesta por Internet, en la que se hubiera pedido a usuarios de este medio que respondieran?
21. **Viajes por avión.** En una encuesta Gallup de 1125 adultos, se encontró que el 47% nunca o casi nunca realiza un viaje aéreo. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el porcentaje de adultos que nunca o casi nunca realizan un viaje aéreo es igual al 50%. Dado que los sujetos encuestados accedieron voluntariaamente a responder una pregunta sobre su costumbre de volar, ¿es posible que los resultados muestrales tengan una cantidad considerable de error?

- 22. Sesgo en la selección de integrantes de jurado.** En el caso *Castaneda contra Partida*, se descubrió que en el condado de Hidalgo, Texas, durante un periodo de 11 años 807 personas habían sido elegidas para integrar el gran jurado, y que el 39% de ellas eran mexicano-estadounidenses. De las personas que podían ser seleccionadas para un gran jurado, el 79.1% eran mexicano-estadounidenses. Utilice un nivel de significancia de 0.01 para probar la aseveración de que el proceso de selección está sesgado en contra de los mexicano-estadounidenses.
- 23. Prueba de las reacciones adversas de Clarinex.** Clarinex es un fármaco utilizado para tratar el asma. En pruebas clínicas de este fármaco, 1655 pacientes fueron tratados con dosis de 5 mg, y el 2.1% de ellos experimentaron fatiga (según datos de Schering Corporation). Utilice un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de usuarios de Clarinex que experimentan fatiga es mayor que la tasa del 1.2% registrada entre quienes no utilizan el fármaco. ¿Parece que la fatiga es una reacción adversa de Clarinex?
- 24. Tabaquismo y educación universitaria.** Una encuesta reveló que, de 785 sujetos seleccionados al azar y que completaron cuatro años de estudios universitarios, el 18.3% fuma y el 81.7% no fuma (según datos de la American Medical Association). Utilice un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de tabaquismo entre quienes tienen cuatro años de estudios universitarios es menor que el porcentaje del 27% registrado en la población general. ¿Por qué los graduados universitarios tienen una tasa menor de tabaquismo que el resto?

Uso de los conjuntos de datos del apéndice B. En los ejercicios 25 a 28, utilice los conjuntos de datos del apéndice B para probar la aseveración dada.

- 25. Uso de datos de M&M.** Remítase al conjunto de datos 13 del apéndice B y calcule la proporción muestral de dulces M&M que son azules. Utilice este resultado para probar la aseveración de Mars, Inc., de que el 24% de sus dulces M&M son azules.
- 26. Precipitación en Boston.** Remítase al conjunto de datos 10 del apéndice B, y observe que los días sin precipitación tienen valores distintos de 0. Utilice un nivel de significancia de 0.05 para probar la aseveración de que en Boston los domingos hay precipitación más del 25% de los días.
- 27. Exactitud de temperaturas pronosticadas.** Remítase al conjunto de datos 8 del apéndice B. Calcule la proporción de días con una temperatura máxima real que tenga una diferencia mayor de 2° con la temperatura máxima pronosticada un día antes. Permita que p sea la proporción de días con una temperatura máxima real que tenga una diferencia mayor de 2° con la temperatura máxima pronosticada un día antes. Utilice un nivel de significancia de 0.05 para probar la aseveración de que $p < 0.5$. ¿Qué indica el resultado sobre la exactitud del pronóstico?
- 28. Consumo de alcohol y tabaco en películas infantiles de dibujos animados.** Utilice los resultados del conjunto de datos 5 en el apéndice B para probar la aseveración de que la mayoría de las películas infantiles de dibujos animados muestran el consumo de alcohol o tabaco (o ambos). Utilice un nivel de significancia de 0.05.

8-3 MÁS ALLÁ DE LO BÁSICO

- 29. Uso de intervalos de confianza para probar hipótesis.** Al analizar los últimos dígitos de los números telefónicos de Port Jefferson, se encontró que, de 1000 dígitos seleccionados al azar, 119 son ceros. Si los números se seleccionan aleatoriamente, la proporción de ceros debería ser de 0.1.
- Utilice el método tradicional, con un nivel de significancia de 0.05, para probar la aseveración de que la proporción de ceros es igual a 0.1.
 - Utilice el método del valor P , con un nivel de significancia de 0.05, para probar la aseveración de que la proporción de ceros es igual a 0.1.

continúa

- c. Use los datos muestrales para construir un estimado del intervalo de confianza del 95% de la proporción de ceros. ¿Qué sugiere el intervalo de confianza sobre la aseveración de que la proporción de ceros es igual a 0.1?
- d. Compare los resultados obtenidos con el método tradicional con los del método del valor P y los del método del intervalo de confianza. ¿Conducen todos a la misma conclusión?
- 30. Uso de la corrección por continuidad.** Repita el ejercicio 28, pero incluya la corrección por continuidad que se estudió en la sección 6-6. ¿De qué manera se ven afectados los resultados al incluir la corrección por continuidad?
- 31. Método alternativo para probar una aseveración acerca de p .** En un estudio sobre percepción, se somete a prueba a 80 hombres y resulta que 7 de ellos tienen ceguera a los colores rojo y verde (según datos de *USA Today*). Deseamos emplear un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de hombres con ceguera a los colores rojo y verde es mayor que la tasa del 0.25% registrada entre las mujeres.
- ¿Por qué no podemos utilizar los métodos de esta sección?
 - Suponiendo que el porcentaje de ceguera a los colores rojo y verde de los hombres es igual al porcentaje de 0.25% de las mujeres, calcule la probabilidad de que, de 80 hombres seleccionados al azar, al menos 7 tengan este tipo de ceguera al color. Describa el método utilizado para calcular esa probabilidad.
 - Con base en los resultados del inciso b), ¿qué concluye?
- 32. Manejo de no éxitos.** En una muestra aleatoria simple de 50 dulces M&M sencillos, se encontró que ninguno de ellos era azul. Queremos emplear un nivel de significancia de 0.01 para probar la aseveración de Mars, Inc., de que la proporción de dulces M&M azules es igual a 0.10. ¿Podrían utilizarse los métodos de esta sección? Si es así, pruebe la aseveración; si no, explique por qué.
- 33. Potencia.** Para probar una hipótesis con un nivel de significancia α específico, la probabilidad de un error tipo I es α , mientras que la probabilidad β de un error tipo II depende del valor particular de p que se utiliza como alternativa a la hipótesis nula.
- Calcule la potencia de la prueba utilizando una hipótesis alternativa de $p < 0.4$, un tamaño muestral de $n = 50$ y suponga que el valor real de p es 0.25. Utilice el procedimiento dado en el inciso a) del ejercicio 47, en la sección 8-2. [Sugerencia: En el paso 3, utilice los valores $p = 0.25$ y $pq/n = (0.25)(0.75)/50$].
 - Calcule el valor de β , la probabilidad de cometer un error tipo II.
 - Dadas las condiciones citadas en el inciso a), ¿qué indican los resultados acerca de la eficacia de la prueba de hipótesis?

Prueba de una aseveración respecto 8-4 de una media: σ conocida

Concepto clave En esta sección se presentan métodos para poner a prueba una aseveración respecto de una media poblacional, cuando se conoce el valor de la desviación estándar poblacional. La siguiente sección presenta métodos para probar una aseveración respecto de una media cuando no se conoce el valor de σ . Aquí se usa la distribución normal con los mismos componentes de las pruebas de hipótesis que se presentaron en la sección 8-2.

Los requisitos, el estadístico de prueba, los valores críticos y el valor P se resumen de la siguiente manera:

Prueba de aseveraciones acerca de una media poblacional (σ desconocida)

Requisitos

1. La muestra es aleatoria simple.
2. Se conoce el valor de la desviación estándar poblacional σ .
3. Se satisface una o ambas de las siguientes condiciones: la población se distribuye normalmente o $n > 30$.

Estadístico de prueba para probar una aseveración sobre una media (σ conocida)

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Valores P : Utilice la distribución normal estándar (tabla A-2) y remítase a la figura 8-6.

Valores críticos: Utilice la distribución normal estándar (tabla A-2).

Antes de iniciar cualquier procedimiento de prueba de hipótesis, debemos *explorar* primero el conjunto de datos y verificar que los requisitos de la prueba específica se cumplan. Con los métodos descritos en los capítulos 2 y 3, investigue las medidas de tendencia central, la variación y la distribución dibujando una gráfica; calcule la media, la desviación estándar y el resumen de los cinco números; también identifique cualquier valor extremo.

EJEMPLO Método del valor P El conjunto de datos 13 del apéndice B incluye los pesos de 13 dulces M&M rojos elegidos al azar de una bolsa que contiene 465 dulces. La desviación estándar de los pesos de todos los dulces M&M que están en la bolsa es $\sigma = 0.0565$ g. A continuación se presentan los pesos muestrales (en gramos), que tienen una media de $\bar{x} = 0.8635$. En la bolsa se afirma que el peso neto del contenido es de 396.9 g, de manera que los dulces M&M deben tener un peso medio de al menos $396.9/465 = 0.8535$ g para dar la cantidad anunciada. Utilice los datos muestrales con un nivel de significancia de 0.05 para probar la aseveración de un gerente de producción de que los dulces M&M en realidad tienen una media mayor que 0.8535 g, por lo que los consumidores están recibiendo una cantidad mayor de la indicada en la etiqueta. Utilice el método del valor P siguiendo el procedimiento descrito en la figura 8-8.

0.751	0.841	0.856	0.799	0.966	0.859	0.857
0.942	0.873	0.809	0.890	0.878	0.905	

SOLUCIÓN

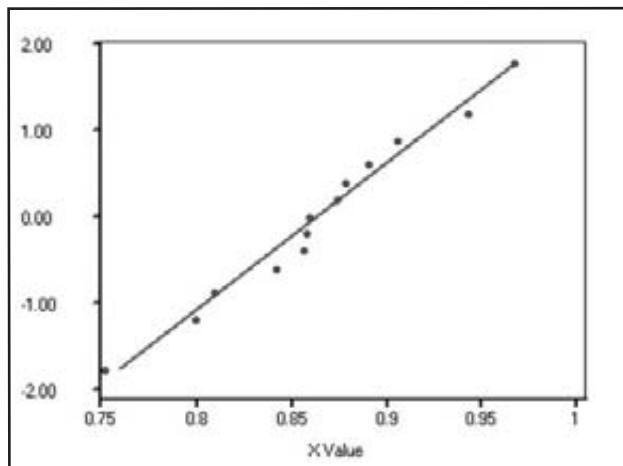
REQUISITOS ✓ Vea los requisitos descritos arriba. Se trata de una muestra aleatoria simple. Se conoce el valor de σ (0.0565 g). El tamaño muestral es $n = 13$, que no es mayor que 30, pero no existen valores extremos y la imagen de la gráfica cuantil normal de STATDISK sugiere que los pesos se distribuyen normalmente (ya que los puntos se acercan a la línea recta y no muestran un patrón sistemático). Un histograma también sugeriría que los pesos se distribuyen de manera normal. Los requisitos se satisfacen y podemos realizar la prueba de hipótesis. ✓



Comerciales

Las cadenas de televisión tienen sus propios departamentos para revisar los comerciales y verificar sus aseveraciones. La National Advertising Division, una rama del Council of Better Business Bureaus, investiga las afirmaciones publicitarias. La Federal Trade Commission y los fiscales locales de distrito también realizan este proceso. Hace algún tiempo, Firestone tuvo que eliminar la afirmación de que sus neumáticos frenaban un 25% más rápido, y Warner Lambert tuvo que gastar \$10 millones para informar a sus clientes que Listerine no previene ni cura el resfriado. Muchos anuncios engañosos son retirados de manera voluntaria, y muchos otros escapan al escrutinio simplemente porque los mecanismos regulatorios no pueden revisar una cantidad tan grande de comerciales.

STATDISK



Seguiremos el procedimiento del valor P que se resume en la figura 8-8.

- Paso 1: La aseveración de que la media es mayor que 0.8535 g se expresa en forma simbólica como $\mu > 0.8535$.
- Paso 2: La alternativa (en forma simbólica) a la aseveración original es $\mu \leq 0.8535$.
- Paso 3: Puesto que la afirmación $\mu > 0.8535$ no contiene la condición de igualdad, se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de que $\mu = 0.8535$.

$$H_0: \mu = 0.8535$$

$$H_1: \mu > 0.8535 \quad (\text{aseveración original})$$

- Paso 4: Tal como se especifica en el planteamiento del problema, el nivel de significancia es $\alpha = 0.05$.
- Paso 5: Puesto que la aseveración se refiere a la *media poblacional* μ , el estadístico muestral más relevante para esta prueba es la *media muestral* $\bar{x} = 0.8635$. Como se supone que conocemos σ (0.0565) y parece que la población se distribuye normalmente, el teorema del límite central indica que la distribución de medidas muestrales puede aproximarse por medio de una distribución *normal*.
- Paso 6: El estadístico de prueba se calcula de la siguiente manera:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{0.8635 - 0.8535}{\frac{0.0565}{\sqrt{13}}} = 0.64$$

Utilizando el estadístico de prueba $z = 0.64$, ahora procedemos al cálculo del valor P . Observe el diagrama de flujo de la figura 8.6 que resume el procedimiento para el cálculo de los valores P . Se trata de una prueba de cola derecha, de manera que el valor P es el

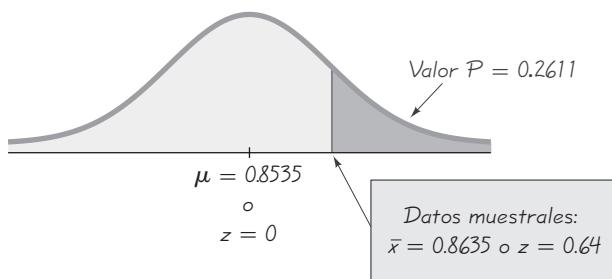


Figura 8-11 Prueba de la aseveración de que $\mu > 0.8535$

área a la derecha de $z = 0.64$. Ahora nos remitimos a la tabla A-2 para encontrar que el área a la izquierda de $z = 0.64$ es 0.7389, por lo que el valor P es $1 - 0.7389 = 0.2611$. (Si utilizamos un recurso tecnológico, un valor P más exacto es 0.2609). (Véase la figura 8-11).

Paso 7: Puesto que el valor P de 0.2611 es mayor que el nivel de significancia de $\alpha = 0.05$, no rechazamos la hipótesis nula.

INTERPRETACIÓN El valor P de 0.2611 nos indica que si $\mu = 0.8535$ g, existe una buena probabilidad (0.2611) de obtener una media muestral como la calculada en la muestra. Es decir, una media muestral como 0.8635 puede presentarse fácilmente por azar con una media poblacional de 0.8535. No existe evidencia suficiente para sustentar la conclusión de que la media poblacional sea mayor que 0.8535, como afirmó el gerente de producción.

Método tradicional Si se utiliza el método tradicional de prueba de hipótesis en el ejemplo anterior, los primeros cinco pasos serían los mismos. En el paso 6 calcularíamos el valor crítico de $z = 1.645$ en vez de calcular el valor P . Una vez más, no rechazaríamos la hipótesis nula, ya que el estadístico de prueba $z = 0.64$ no caería en la región crítica. La conclusión final sería la misma.

Método del intervalo de confianza Ahora usaremos un intervalo de confianza para probar una aseveración acerca de μ cuando conocemos σ . Para una prueba de hipótesis de una cola, con un nivel de significancia de 0.05, construimos un intervalo de confianza del 90%. Si utilizamos los datos muestrales del ejemplo anterior ($n = 13$ y $\bar{x} = 0.8635$) y suponemos que sabemos que $\sigma = 0.0565$, podemos probar la aseveración de que $\mu > 0.8535$ aplicando los métodos de la sección 7-3 para construir el siguiente intervalo de confianza del 90%: $0.8377 < \mu < 0.8893$. Puesto que el intervalo de confianza incluye el valor 0.8535, no podemos sustentar la afirmación de que μ es mayor que 0.8535. Con base en el intervalo de confianza, μ podría ser igual a 0.8535. Al igual que en la prueba de hipótesis, concluimos que no existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 0.8535 g.

En la sección 8-3 vimos que al probar una aseveración sobre una proporción poblacional, el método tradicional y el método del valor P son equivalentes, pero que el método del intervalo de confianza es un poco diferente. Cuando se prueba

una aseveración sobre una media poblacional, no existe tal diferencia, y los tres métodos son equivalentes.

En lo que resta del libro, aplicaremos métodos de prueba de hipótesis en otras circunstancias. Es fácil enredarse en una compleja red de pasos sin comprender el razonamiento que subyace en la prueba de hipótesis. La clave para comprenderlo radica en la regla del suceso infrecuente de la estadística inferencial: **Si, bajo un supuesto dado, existe una probabilidad excepcionalmente pequeña de obtener resultados muestrales que sean al menos tan extremos como los resultados que se obtuvieron, concluimos que probablemente el supuesto no sea correcto.** Al probar una aseveración, suponemos igualdad (hipótesis nula). Después comparamos el supuesto y los resultados muestrales para llegar a una de las siguientes conclusiones:

- Si los resultados muestrales (o resultados más extremos) pueden ocurrir con facilidad cuando el supuesto (hipótesis nula) es verdadero, atribuimos al azar la discrepancia relativamente pequeña entre el supuesto y los resultados muestrales.
- Si los resultados muestrales (o resultados más extremos) no pueden ocurrir con facilidad cuando el supuesto (hipótesis nula) es verdadero, explicamos la discrepancia relativamente grande entre el supuesto y los resultados muestrales mediante la conclusión de que el supuesto no es verdadero, por lo que rechazamos este último.

Uso de la tecnología

STATDISK Si trabaja con una lista de los valores muestrales originales, primero calcule el tamaño de muestra, la media muestral y la desviación estándar muestral por medio del procedimiento de STATDISK descrito en la sección 3-2. Después de obtener los valores de n , \bar{x} , y s , proceda a seleccionar **Analysis** de la barra del menú principal, después seleccione

Hypothesis Testing, seguido por **Mean-One Sample**.

MINITAB Minitab Release 14 le permite utilizar el resumen de estadísticos o una lista de los valores muestrales originales. (Las versiones anteriores no permiten el uso del resumen de estadísticos). Seleccione **Stat**, **Basic Statistics** y **1-Sample z** del menú e introduzca el resumen de estadísticos o **C1** en el cuadro localizado en la parte superior derecha. También ingrese el valor de s en el cuadro “Standard Deviation” o “Sigma”. Utilice el botón **Options** para cambiar la forma de la hipótesis alternativa.

EXCEL La función ZTEST creada de Excel es extremadamente confusa, ya que el valor P generado no siempre es el mismo valor P estándar utilizado en el resto del mundo. En su lugar, utilice Data Desk XL, que es un complemento de este libro. Primero introduzca los

datos muestrales en la columna A. Seleccione **DDXL**, después **Hypothesis Test**. En las opciones del tipo de función, seleccione **1 Var z Test**. Haga clic en el ícono del lápiz e introduzca el rango de valores de datos, como A1:A45, si tiene 45 valores listados en la columna A. Haga clic en **OK**. Siga los cuatro pasos del cuadro de diálogo. Después de hacer clic en **Compute** en el paso cuatro, obtendrá el valor P , el estadístico de prueba y la conclusión.

TI-83/84 PLUS Si utiliza la calculadora TI-83/84 Plus, presione **STAT**, luego seleccione **TESTS** y elija la primera opción **Z-Test**. Usted puede utilizar los datos originales o un resumen de los estadísticos (**Stats**) al dar las entradas indicadas en la pantalla. Los primeros tres elementos de los resultados de la TI-83/84 Plus incluirán la hipótesis alternativa, el estadístico de prueba y el valor P .

8-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Requisitos.** ¿Debe usted tener un tamaño muestral de $n > 30$ para poder utilizar los métodos de prueba de hipótesis presentados en esta sección? Si una muestra aleatoria simple tiene menos de 31 valores, ¿qué requisitos se deben cumplir para justificar el uso de los métodos de esta sección?



- 2. Verificación de requisitos.** Una muestra aleatoria simple consta de $n = 12$ valores, de manera que es necesario verificar el requisito de normalidad. ¿De qué manera puede verificar ese requisito de normalidad?
- 3. Intervalo de confianza.** Usted desea probar la aseveración de que $\mu < 100$ construyendo un intervalo de confianza. Si la prueba de hipótesis se realizara con un nivel de significancia de 0.01, ¿qué nivel de confianza debe emplear para el intervalo de confianza?
- 4. Muestreo sistemático.** Un decano obtiene una muestra en una universidad al seleccionar cada quincuagésimo nombre de la lista de los 6000 estudiantes actuales de tiempo completo. Luego, pone a prueba la aseveración de que la media de la calificación promedio es mayor que 2.50. ¿Se trata de una muestra aleatoria simple? ¿Es probable que la muestra sea representativa o tiene algún sesgo?

Verificación de supuestos. En los ejercicios 5 a 8, determine si las condiciones dadas justifican el uso de los métodos de esta sección cuando se prueba una aseveración acerca de la media poblacional μ .

5. El tamaño muestral es $n = 27$, $\sigma = 6.44$, y la población original se distribuye de manera aproximadamente normal.
6. El tamaño muestral es $n = 9$, $\sigma = 2.5$, y un histograma de los datos muestrales indica que la distribución dista mucho de tener forma de campana.
7. El tamaño muestral es $n = 24$, se desconoce σ , y un histograma de los datos muestrales revela una distribución aproximadamente normal.
8. El tamaño muestral es $n = 121$, $\sigma = 0.25$, y un histograma de los datos muestrales revela una distribución uniforme en vez de normal.

En los ejercicios 9 a 12, identifique los valores indicados o interprete la pantalla de resultados.

- 9. Dulces M&M.** El conjunto de datos 13 del apéndice B incluye una muestra de 27 dulces M&M, con un peso medio de 0.8560 g. Suponga que sabe que σ es 0.0565 g. Considere una prueba de hipótesis que utiliza un nivel de significancia de 0.05 para probar la aseveración de que el peso medio de todos los dulces M&M es igual a 0.8535 g (el peso necesario para que las bolsas de M&M tengan el peso que viene impreso en el empaque).
- a. ¿Cuál es el estadístico de prueba?
 - b. ¿Cuáles son los valores críticos?
 - c. ¿Cuál es el valor P ?
 - d. ¿Cuál es la conclusión sobre la hipótesis nula (se rechaza o no se rechaza)?
 - e. ¿Cuál es la conclusión final en términos sencillos y sin tecnicismos?
- 10. Temperaturas corporales humanas.** El conjunto de datos 2 del apéndice B incluye una muestra de 106 temperaturas corporales, con una media de 98.20°F. Suponga que se sabe que σ es 0.62°F. Considere una prueba de hipótesis que utiliza un nivel de significancia de 0.05 para probar la aseveración de que la temperatura media corporal de la población es menor que 98.6°F.
- a. ¿Cuál es el estadístico de prueba?
 - b. ¿Cuál es el valor crítico?
 - c. ¿Cuál es el valor P ?
 - d. ¿Cuál es la conclusión sobre la hipótesis nula (se rechaza o no se rechaza)?
 - e. ¿Cuál es la conclusión final en términos sencillos y sin tecnicismos?
- 11. Temperaturas de los Everglades.** Para verificar la salud ecológica de los Everglades de Florida, se efectúan varias mediciones en momentos diferentes. Las temperaturas más bajas se registran en la estación Garfield Bright, y se obtiene la media de 30.377°C para las 61 temperaturas registradas en 61 días diferentes. Suponiendo que $\sigma = 1.7^\circ\text{C}$, pruebe la aseveración de que la media poblacional es mayor que 30.0°C. Utilice un nivel de significancia de 0.05 y los resultados del complemento DDLX de Excel (que aparecen en la siguiente página).

Excel (DDXL) Resultados para el ejercicio 1

Test Summary	
H ₀ :	$\mu = 30$
H _a :	Upper tail: $\mu > 30$
z Statistic:	1.732
p-value:	0.0416

Test Results	
Conclusion	
Reject H ₀ at alpha = 0.05	

TI-83/84 Plus

```
Z-Test
μ≠200
z=-1.455441601
p=.1455471254
x=172.5
n=40
```

- 12. Niveles de cotinina de fumadores.** Cuando las personas fuman, la nicotina que absorben se convierte en cotinina, la cual puede medirse. Una muestra de 40 fumadores tiene un nivel medio de cotinina de 172.5 ng/ml. Suponga que se sabe que σ es 119.5 ng/ml y pruebe la aseveración de que el nivel medio de cotinina de todos los fumadores es igual a 200.0 ng/ml. Utilice un nivel de significancia de 0.05 y los resultados de una calculadora TI-83/84 Plus.

Prueba de hipótesis. En los ejercicios 13 a 20, pruebe la aseveración dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor (o valores) crítico(s) de P , la conclusión acerca de la hipótesis nula y la conclusión final que retoma la aseveración original. Utilice el método del valor P , a menos que su profesor especifique otra cosa.

- 13. Percepción del tiempo.** Alumnos del autor, seleccionados al azar, participaron en un experimento que ponía a prueba su habilidad para determinar el transcurso de 1 minuto (o 60 segundos). Cuarenta estudiantes produjeron una media muestral de 58.3 segundos. Suponiendo que $\sigma = 9.5$ segundos, utilice un nivel de significancia de 0.05 para probar la aseveración de que la media poblacional es igual a 60 segundos. Con base en el resultado, ¿parece haber una percepción general de 1 minuto razonablemente exacta?
- 14. Análisis de los últimos dígitos.** El análisis de los últimos dígitos de datos muestrales en ocasiones revela si éstos se han medido y reportado de forma exacta. Cuando se seleccionan al azar y con reemplazo dígitos únicos del 0 al 9, la media debe ser 4.50 y la desviación estándar debe ser 2.87. Los datos reportados (como pesos o estaturas) suelen redondearse, de manera que los últimos dígitos incluyen, de manera desproporcionada, más ceros y cinco. Utilice los últimos dígitos de las longitudes (en pies) de los 73 *home runs* anotados por Barry Bonds en 2001 para probar la aseveración de que provienen de una población con una media de 4.50. Utilice un nivel de significancia de 0.05. Los últimos dígitos de las 73 longitudes de *home runs* (listadas en el conjunto de datos 17 del apéndice B) tienen una media de 1.753. Con base en los resultados, ¿parece que las distancias se midieron con exactitud?
- 15. ¿La dieta funciona?** Cuando 40 personas pusieron en práctica la dieta Atkins durante un año, el cambio medio de su peso fue de -2.1 libras (según datos de "Comparison of the Atkins, Ornish, Weight Watchers, and Zone Diets for Weight Loss and Heart Disease Reduction", de Dansinger *et al.*, *Journal of the American Medical Association*, vol. 293, núm. 1). Suponga que la desviación estándar de todo este tipo de cambios de peso es $\sigma = 4.8$ libras, y utilice un nivel de significancia de 0.05 para probar la aseveración de que el cambio medio de peso es menor que 0. Con base en esos resultados, ¿parece que la dieta es eficaz? ¿Parece que el cambio medio de peso es lo suficientemente grande para justificar la dieta especial?

- 16. ¿Las latas de aluminio delgado son más endebles?** La carga axial de una lata de aluminio es el peso máximo que los costados pueden soportar antes de colapsar. La carga axial es una medida importante, ya que las cubiertas superiores ejercen presión sobre los costados con presiones que varían entre 158 y 165 libras. Pepsi experimentó con latas de aluminio más delgadas, y una muestra aleatoria de 175 de las latas más

delgadas tiene una carga axial media de 267.11 lb. Las latas estándar tienen una carga axial media de 281.81 lb y una desviación estándar de 27.77 lb. Utilice un nivel de significancia de 0.01 para probar la aseveración de que las latas más delgadas tienen una carga axial media menor que 281.81 lb. Suponga que $\sigma = 27.77$ lb. ¿Parece que las latas más delgadas tienen una carga axial media menor que 281.81 lb? ¿Parece que las latas más delgadas son lo suficientemente fuertes para no colapsar cuando las cubiertas superiores presionan los costados?

- 17. Niveles de presión sanguínea.** Cuando 14 estudiantes de segundo año de medicina del Bellevue Hospital midieron la presión sanguínea sistólica de la misma persona, obtuvieron los resultados que se listan abajo (en mmHg). Suponiendo que se sabe que la desviación estándar poblacional es de 10 mmHg, utilice un nivel de significancia de 0.05 para probar la aseveración de que el nivel medio de la presión sanguínea es menor que 140 mmHg. La hipertensión se define como un nivel de presión sanguínea de 140 mmHg o mayor. Con base en los resultados de la prueba de hipótesis, ¿es seguro concluir que la persona no tiene hipertensión?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

- 18. El mamífero más pequeño del mundo.** El mamífero más pequeño del mundo es el murciélagos abejorro, también conocido como murciélagos nariz de cochino (o *Craseonycteris thonglongyai*). Estos animales apenas alcanzan el tamaño de un abejorro grande. A continuación se incluyen los pesos (en gramos) de una muestra de estos murciélagos. Suponiendo que los pesos de todos estos murciélagos tienen una desviación estándar de 0.30 gramos, utilice un nivel de significancia de 0.05 para probar la aseveración de que estos murciélagos provienen de la misma población con una media conocida de 1.8 g. ¿Parece que los murciélagos provienen de la misma población?

1.7 1.6 1.5 2.0 2.3 1.6 1.6 1.8 1.5 1.7 2.2 1.4 1.6 1.6 1.6

- 19. Conjunto de datos del apéndice B: Pesos de monedas de 25 centavos.** Utilice los pesos de las monedas de 25 centavos acuñadas después de 1964, que se listan en el conjunto 14 de datos del apéndice B. Suponiendo que las monedas se acuñan para producir pesos con una desviación estándar poblacional de 0.068, utilice la muestra de pesos con un nivel de significancia de 0.01 para probar la aseveración de que las monedas provienen de una población con una media de 5.670 g. ¿Parece que las monedas se acuñaron de acuerdo con la especificación estadounidense de que la media es igual la 5.670 g?

- 20. Conjunto de datos del apéndice B: Errores de pronóstico.** Remítase al conjunto de datos 8 del apéndice B y reste cada temperatura máxima real de la temperatura máxima pronosticada un día antes. El resultado es una lista de errores. Suponiendo que todos los errores de este tipo tienen una desviación estándar de 2.5° , utilice un nivel de significancia de 0.05 para probar la aseveración de que todos los errores de este tipo tienen una media igual a 0. ¿Qué sugiere el resultado sobre la exactitud de las temperaturas pronosticadas?

8-4 MÁS ALLÁ DE LO BÁSICO

- 21. Potencia de una prueba.** En el ejercicio 47, de la sección 8-2, se presenta el procedimiento para calcular la potencia de una prueba de hipótesis que implica una proporción. Utilice el mismo procedimiento para calcular la potencia de una prueba de hipótesis sobre la aseveración de que $\mu > 100$, dada una muestra de tamaño 40, una desviación estándar poblacional conocida de 15 y una media poblacional de 108. (*Sugerencia:* En el paso 2 del procedimiento, utilice el estadístico de prueba presentado en esta sección, en vez del estadístico de prueba utilizado para una proporción, y en el paso 4 del procedimiento, utilice $\mu = 108$). ¿Qué indica el valor de la potencia acerca de la eficacia de la prueba para reconocer que la media es mayor que 100, cuando en realidad la media es 108? También calcule β , la probabilidad de un error tipo II.



Prueba de una aseveración respecto 8-5 de una media: σ desconocida

Concepto clave En la sección anterior se presentaron métodos para probar una aseveración sobre una media poblacional, cuando se conoce el valor de la desviación estándar poblacional σ . En esta sección se presentan métodos para probar una aseveración respecto de una media poblacional cuando se desconoce el valor de σ . Los métodos de esta sección son muy prácticos y realistas, porque generalmente se desconoce σ . Aquí se utiliza la distribución t de Student, que se presentó en la sección 7-4. Los requisitos, el estadístico de prueba, el valor P y los valores críticos se resumen a continuación.

Prueba de aseveraciones acerca de una media poblacional (σ desconocida)

Requisitos

1. La muestra es aleatoria simple.
2. Se *desconoce* el valor de la desviación estándar poblacional σ .
3. Se satisfacen una o ambas de las siguientes condiciones: la población se distribuye de manera normal o $n > 30$.

Estadístico de prueba para probar una aseveración acerca de una media (σ desconocida)

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$

Valores P y valores críticos: Utilice la tabla A-3 y utilice $gl = n - 1$ para el número de grados de libertad. (Véase la figura 8-6 para los procedimientos del cálculo del valor P).

Requisito de normalidad El requisito de una población distribuida normalmente no es estricto, y generalmente podemos tratar a la población como si tuviera una distribución normal después de utilizar los datos muestrales para confirmar que no existen valores extremos y que el histograma tenga una forma que no se aleje mucho de la normalidad. Se dice que esta prueba t es *robusta* con respecto a su alejamiento de la normalidad, lo que significa que la prueba funciona bastante bien si no se aleja demasiado de la forma normal.

Tamaño muestral Usamos el criterio simplificado de $n > 30$ como justificación para tratar la distribución de medias muestrales como una distribución normal, pero el tamaño muestral mínimo realmente depende de qué tanto la distribución poblacional se aleja de una distribución normal. Como no conocemos el valor de σ , lo estimamos con el valor de la desviación estándar muestral s , aunque esto introduce otra fuente de falta de confiabilidad, especialmente en el caso de muestras pequeñas. Para compensar esta falta de confiabilidad adicional, calculamos los valores P y los valores críticos utilizando la distribución t en vez de la distribución normal que se empleó en la sección 8-4, donde se conocía σ . Las siguientes son las propiedades importantes de la distribución t de Student:

Propiedades importantes de la distribución t de Student

1. La distribución t de Student difiere para tamaños de muestra distintos (véase la figura 7-5 en la sección 7-4).
2. La distribución t de Student tiene la misma forma general de campana que la distribución normal estándar; su forma más ancha refleja una mayor variabilidad, lo que se espera cuando se utiliza s para estimar σ .
3. La distribución t de Student tiene una media de $t = 0$ (del mismo modo que la distribución normal estándar tiene una media de $z = 0$).
4. La desviación estándar de la distribución t de Student varía de acuerdo con el tamaño muestral y es mayor que 1 (a diferencia de la distribución normal estándar, que tiene $\sigma = 1$).
5. Conforme aumenta el tamaño muestral n , la distribución t de Student se acerca más a la distribución normal estándar.

Elección de la distribución apropiada

Cuando se prueban aseveraciones acerca de medias poblacionales, en ocasiones se aplica la distribución normal, en otras la distribución t de Student y en algunas no se aplica ninguna de las dos, por lo que debemos utilizar métodos no paramétricos o técnicas *bootstrap* de muestreo. (Los métodos no paramétricos, que no requieren una distribución en particular, se estudian en el capítulo 13; la técnica *bootstrap* de muestreo se describe en el Proyecto tecnológico al final del capítulo 7). Consulte las páginas 354 y 355, donde la figura 7-6 y la tabla 7-1 resumen las decisiones que deben tomarse al elegir entre la distribución normal y la t de Student. En ellas se observa que cuando se prueban aseveraciones acerca de medias poblacionales, la distribución t de Student se aplica en estas condiciones:

Utilice la distribución t de Student cuando se desconozca σ y cuando cualquiera o ambas de las siguientes condiciones se satisfagan:

La población se distribuye normalmente o $n > 30$.

EJEMPLO Control de calidad de los dulces M&M El conjunto de datos 13 del apéndice B incluye los pesos de 13 dulces M&M rojos, elegidos al azar de una bolsa que contiene 465 M&M. A continuación se presentan los pesos (en gramos), los cuales tienen una media de $\bar{x} = 0.8635$ y una desviación estándar de $s = 0.0576$ g. En el empaque se afirma que el peso neto del contenido es 396.9 g, de manera que los M&M deben tener un peso medio de al menos $396.9/465 = 0.8535$ g para dar la cantidad anunciada. Utilice los datos muestrales con un nivel de significancia de 0.05, para probar la aseveración que hizo un gerente de producción de que los M&M tienen en realidad una media mayor que 0.8535 g, de manera que los consumidores están recibiendo más que la cantidad indicada en la etiqueta. Utilice el método tradicional, siguiendo el procedimiento descrito en la figura 8-9.

0.751	0.841	0.856	0.799	0.966	0.859	0.857
0.942	0.873	0.809	0.890	0.878	0.905	

SOLUCIÓN

REQUISITOS ✓ Vea los tres requisitos listados arriba. Se trata de una muestra aleatoria simple; no estamos empleando un valor conocido de σ . El tamaño



La ética en los experimentos

Con frecuencia se pueden obtener datos muestrales encuestando o simplemente observando a miembros seleccionados de la población. Muchas otras situaciones requieren que, de alguna manera, manipulemos circunstancias para obtener datos muestrales. En ambos casos surgen cuestiones éticas. Investigadores en Tuskegee, Alabama, negaron el tratamiento de penicilina eficaz a víctimas de sífilis para poder estudiar la enfermedad. ¡Este experimento continuó por un periodo de 27 años!



¿Los zurdos mueren antes?

Un estudio realizado por los psicólogos Diane Halpern y Stanley Coren recibió una gran atención de los medios de comunicación masiva y generó un gran interés cuando concluyó que las personas zurdas no viven tanto tiempo como las personas diestras. Con base en el estudio, parecía que las personas zurdas viven un promedio de nueve años menos que las diestras. El estudio de Halpern y Coren ha sido criticado por utilizar datos defectuosos, ya que se emplearon datos de segunda mano al encuestar a los parientes de individuos que habían muerto recientemente. El mito de que los zurdos mueren más jóvenes se convirtió en una idea generalizada que ha perdido durante muchos años. Sin embargo, estudios más recientes indican que los zurdos *no* viven menos que los diestros.

muestral es $n = 13$, que no es mayor que 30, pero no se encuentran valores extremos, y una gráfica cuantílica normal sugiere que los pesos se distribuyen normalmente porque los puntos se acercan a la línea recta y no muestran un patrón sistemático. (En la sección anterior se presenta la gráfica cuantílica normal). Un histograma también sugeriría que los pesos se distribuyen de forma normal. Los requisitos se satisfacen y podemos efectuar la prueba de hipótesis. ✓

Seguiremos el procedimiento tradicional que se resume en la figura 8-9.

- Paso 1: La aseveración de que la media es mayor que 0.8535 g se expresa simbólicamente como $\mu > 0.8535$.
- Paso 2: La alternativa (en forma simbólica) a la aseveración original es $\mu \leq 0.8535$.
- Paso 3: Puesto que la afirmación $\mu > 0.8535$ no contiene la condición de igualdad, se convierte en la hipótesis alternativa. La hipótesis nula es la aseveración de que $\mu = 0.8535$.

$$H_0: \mu = 0.8535$$

$$H_1: \mu > 0.8535 \quad (\text{aseveración original})$$

- Paso 4: Como se especifica en el planteamiento del problema, el nivel de significancia es $\alpha = 0.05$.
- Paso 5: Puesto que la aseveración se refiere a la *media poblacional* μ , y como los requisitos para utilizar el estadístico de prueba t se satisfacen, empleamos la distribución t . (Remítase a la figura 7-6 o a la tabla 7-1 para consultar los criterios de elección entre las distribuciones normal y t).
- Paso 6: El estadístico de prueba se calcula de la siguiente manera:

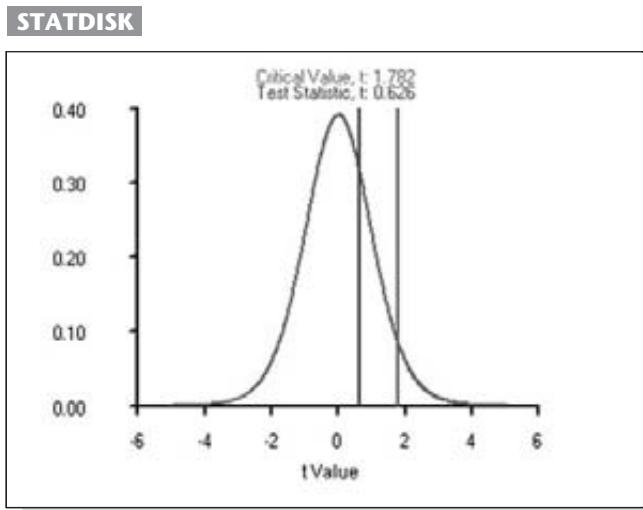
$$t = \frac{\bar{x} - \mu_x}{\frac{s}{\sqrt{n}}} = \frac{0.8635 - 0.8535}{\frac{0.0576}{\sqrt{13}}} = 0.626$$

El valor crítico de $t = 1.782$ se obtiene de la tabla A-3. Primero localice el número correcto de grados de libertad en la columna de la izquierda. En este ejemplo utilizamos $gl = n - 1 = 12$. Como se trata de una prueba de cola derecha, con $\alpha = 0.05$, nos remitimos a la columna que indica una área de 0.05 en una cola. El estadístico de prueba y el valor crítico se muestran en la pantalla de resultados de STATDISK que se presenta en la siguiente página.

- Paso 7: Como el estadístico de prueba de $t = 0.626$ no se localiza dentro de la región crítica, no rechazamos la hipótesis nula.

INTERPRETACIÓN (Remítase a la figura 8-7 para ayudarse a redactar la conclusión). No se rechaza la hipótesis nula. Con base en los datos muestrales disponibles, no tenemos evidencia suficiente para sustentar la aseveración de que el peso medio de los dulces M&M es mayor que 0.8535 g. Aunque la media muestral de 0.8635 g excede a 0.8535 g, no lo rebasa por una cantidad *significativa*.

El valor crítico en el ejemplo anterior fue $t = 1.782$, pero si se utiliza la distribución normal, el valor crítico habría sido $z = 1.645$. El valor crítico t de Student es más grande (más cargado a la derecha), lo que demuestra que con la distribución t de Student la evidencia muestral debe ser *más extrema*, antes de considerarla significativa.



Cálculo de valores P con la distribución t de Student

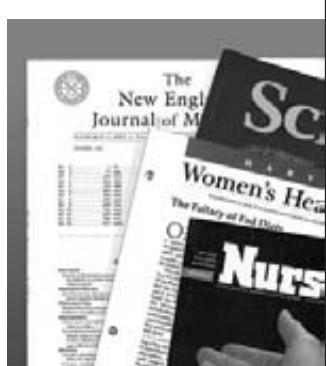
En el ejemplo anterior se empleó el método tradicional de prueba de hipótesis, pero STATDISK, Minitab, la calculadora TI-83/84 Plus y muchos artículos de revistas científicas presentan valores P . Para el ejemplo anterior, STATDISK y la calculadora TI-83/84 Plus dan un valor P de 0.2715, y Minitab y Excel dan un valor P de 0.271. Con un nivel de significancia de 0.05 y un valor P mayor que 0.05, no rechazamos la hipótesis nula, como hicimos al emplear el método tradicional en el ejemplo anterior. Si no dispone de un programa de cómputo o de una calculadora TI-83/84 Plus, puede utilizar la tabla A-3 para identificar un *rango de valores* que contenga el valor P . Recomendamos esta estrategia para el cálculo de valores P utilizando la distribución t :

1. Utilice un programa de cómputo o una calculadora TI-83/84 Plus. (STATDISK, Minitab, el complemento DDXL de Excel, la calculadora TI-83/84 Plus, SPSS y SAS dan valores P para las pruebas t).
2. Si no dispone de algún recurso tecnológico, consulte la tabla A-3 para identificar un rango de valores P de la siguiente manera: utilice el número de grados de libertad para localizar el renglón relevante de la tabla A-3, luego determine dónde cae el estadístico de prueba en relación con los valores t en esa fila. Con base en una comparación del estadístico de prueba t y los valores t en la fila de la tabla A-3, identifique un rango de valores al consultar el área de valores dada en la parte superior de la tabla.

EJEMPLO Cálculo del valor P Suponiendo que no disponemos de un programa de cómputo o de una calculadora TI-83/84 Plus, consultamos la tabla A-3 para obtener un rango de valores para el valor P , correspondientes al estadístico de prueba de $t = 0.626$ del ejemplo anterior. Observe que la prueba es de cola derecha, y que el número de grados de libertad es 12.

SOLUCIÓN Se trata de una prueba de cola derecha, de manera que el valor P es el área ubicada a la derecha del estadístico de prueba. Remítase a la tabla A-3 y localice el renglón correspondiente a 12 grados de libertad. El estadístico de prueba de $t = 0.626$ es menor que cualquiera de los valores en ese renglón de la tabla A-3, de manera que “el área en una cola” (a la derecha del estadístico

continúa



Meta-análisis

El término meta-análisis se refiere a una técnica para realizar un estudio que en esencia combina resultados de otros estudios. Tiene la ventaja de que se pueden combinar muestras separadas más pequeñas en una gran muestra, lo que hace más significativos los resultados colectivos. También tiene la ventaja de aprovechar trabajo que ya se realizó. El meta-análisis tiene la desventaja de que sólo es tan bueno como los estudios utilizados. Si los estudios anteriores tienen defectos, se presentará el fenómeno de “entra basura, sale basura”. El empleo de meta-análisis es actualmente de uso común en investigaciones médicas y psicológicas. Como un ejemplo, un estudio de tratamientos de jaqueca se basó en datos de otros 46 estudios (véase “Meta-Analysis of Migraine Headache Treatments: Combining Information from Heterogeneous Designs” de Dominici *et al.*, *Journal of the American Medical Association*, vol. 94, núm. 445).

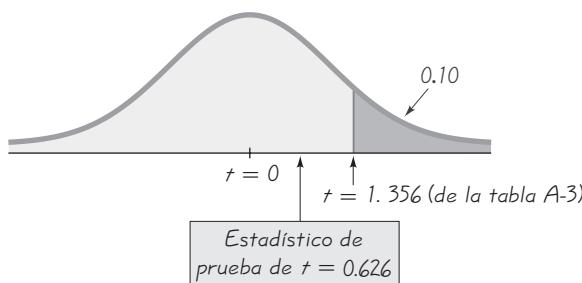


Figura 8-12 Rango de valores para el valor P .

de prueba) es mayor que 0.10. Observe la figura 8-12, que indica la ubicación del estadístico de prueba $t = 0.626$ con respecto al siguiente valor t menor en el decimosegundo renglón de la tabla A-3. En la figura 8-12 podemos ver con claridad que el área a la derecha de $t = 0.626$ es mayor que 0.10. Aunque no encontramos el valor P exacto, podemos concluir que el valor $P > 0.10$.

EJEMPLO Cálculo de un valor P Si no dispone de un programa de cómputo o de una calculadora TI-83/84 Plus, utilice la tabla A-3 para obtener un rango de valores para el valor P , correspondientes a un estadístico de prueba de t con los siguientes componentes: el estadístico de prueba es $t = 2.777$, el tamaño muestral es $n = 20$, el nivel de significancia es $\alpha = 0.05$ y la hipótesis alternativa es $H_1: \mu \neq 5$.

SOLUCIÓN Puesto que el tamaño muestral es 20, remítase a la tabla A-3 y localice el renglón correspondiente a 19 grados de libertad. Como el estadístico de prueba $t = 2.777$ se ubica entre los valores 2.861 y 2.539, el valor P se localiza entre las “áreas de dos colas” correspondientes de 0.01 y 0.02. (Asegúrese de utilizar el “área en dos colas” si la prueba es de dos colas). Aunque no encontramos el valor P exacto, podemos concluir que $0.01 < P < 0.02$. (Los programas de cómputo o la calculadora TI-83/84 Plus dan el valor P exacto de 0.0120).

Recuerde que es posible calcular los valores P con facilidad utilizando un programa de cómputo o una calculadora TI-83/84 Plus. Asimismo, el método tradicional de prueba de hipótesis se puede utilizar en vez del método del valor P .

Método del intervalo de confianza Podemos utilizar un intervalo de confianza para probar una aseveración acerca de μ , cuando desconocemos σ . Para una prueba de hipótesis de dos colas con un nivel de significancia de 0.05, construimos un intervalo de confianza del 95%, pero para una prueba de hipótesis de una cola con un nivel de significancia de 0.05, construimos un intervalo de confianza del 90% (véase la tabla 8-2). Utilizando los datos muestrales del primer ejemplo de esta sección ($n = 13$, $\bar{x} = 0.8635$ g, $s = 0.0576$), sin conocer σ y utilizando un nivel de significancia de 0.05, podemos probar la aseveración de que $\mu > 0.8535$ por medio del método del intervalo de confianza. Construya este intervalo de confianza del 90%: $0.8350 < \mu < 0.8920$ (véase la sección 7-4). Puesto que el valor supuesto de $\mu = 0.8535$ está contenido dentro del intervalo de confianza, no podemos rechazar ese supuesto. Con base en los 13 valores muestrales dados en el ejemplo, no tenemos evidencia suficiente para sustentar la aseveración de que el peso medio es mayor que 0.8535 g. Con base en el intervalo de confianza, es probable que el valor verdadero de μ esté entre 0.8350 g y 0.8920 g, incluyendo 0.8535 g.

Uso de la tecnología

STATDISK Si se trabaja con la lista de los valores muestrales originales, primero calcule el tamaño muestral, la media muestral y la desviación estándar muestral por medio del procedimiento de STATDISK descrito en la sección 3-2. Después de obtener los valores de n , \bar{x} , y s , proceda a seleccionar **Analysis**



de la barra del menú principal, después seleccione **Hypothesis Testing**, seguido por **Mean-One Sample**.

MINITAB Minitab Release 14 le permite usar el resumen de estadísticos o una lista de los valores muestrales originales. (Las versiones anteriores no permiten el uso del resumen de estadísticos). Seleccione **Stat**, **Basic Statistics** y **1-Sample t** del menú. Introduzca el resumen de estadísticos o ingrese CI en el cuadro que se ubica en la parte superior derecha. Utilice el botón **Options** para cambiar el formato de la hipótesis alternativa.

EXCEL Excel no posee una función para la prueba t , por lo tanto, utilice Data Desk XL, que es complemento de este libro. Primero introduzca los datos muestrales en la columna A. Seleccione **DDXL**, después **Hypothesis**

Tests. En las opciones del tipo de función, seleccione **1 Var t Test**. Haga clic en el icono del lápiz e introduzca el rango de valores de datos, como A1:A12, si tiene 12 valores listados en la columna A. Haga clic en **OK**. Siga los cuatro pasos del cuadro de diálogo. Después de hacer clic en **Compute** en el paso 4, obtendrá el valor P , el estadístico de prueba y la conclusión.

TI-83/84 PLUS Si utiliza la calculadora TI-83/84 Plus, presione **STAT**, luego seleccione **TESTS** y elija la segunda opción, **T-Test**. Puede utilizar los datos originales o un resumen de los estadísticos (**Stats**) al dar las entradas indicadas en la pantalla. Los primeros tres elementos de los resultados de la calculadora TI-83/84 Plus incluirán la hipótesis alternativa, el estadístico de prueba y el valor P .

En la sección 8-3 vimos que, cuando probamos una aseveración respecto de una proporción poblacional, el método tradicional y el método del valor P son equivalentes, pero el método del intervalo es un poco diferente. Cuando se prueba una aseveración sobre una media poblacional, no existe esta diferencia y los tres métodos son equivalentes.

8-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Grados de libertad.** Cuando se utiliza la tabla A-3 para obtener valores críticos, es necesario usar el número apropiado de grados de libertad. Si una muestra consiste en cinco valores, ¿cuál es el número apropiado de grados de libertad? Si usted no conoce ninguno de los cinco valores muestrales, pero sabe que su media es exactamente 20.0, ¿cuántos valores puede crear antes de que los valores restantes estén determinados por la restricción de que la media es 20.0?
- 2. Distribuciones normal y t .** Identifique dos características que la distribución normal estándar y la distribución t tengan en común, e identifique dos características que sean diferentes para estas mismas distribuciones.
- 3. Prueba innecesaria.** Un gerente de control de calidad afirma que latas de bebida de cola se llenan con cantidades que tienen una media menor que 12 onzas. Si los datos muestrales consisten en 24 latas con una media de 12.13 oz y una desviación estándar de 0.12 oz, ¿por qué no es necesario realizar una prueba formal de hipótesis para concluir que los datos muestrales no sustentan la aseveración del gerente?
- 4. Verificación de la realidad.** A diferencia de la sección anterior, esta sección no incluye el requisito de que se conozca el valor de la desviación estándar poblacional. ¿Qué sección es más probable que se aplique en situaciones reales: ésta o la anterior? ¿Por qué?

Uso de la distribución correcta. En los ejercicios 5 a 8, determine si la prueba de hipótesis incluye una distribución muestral de medias con distribución normal, distribución t de Student o ninguna de éstas (Sugerencia: Consulte la figura 7-6 y la tabla 7-1).

5. Aseveración: $\mu = 2.55$. Datos muestrales: $n = 7$, $\bar{x} = 2.41$, $s = 0.66$. Los datos muestrales parecen provenir de una población distribuida normalmente, con μ y σ desconocidas.
6. Aseveración: $\mu = 1002$. Datos muestrales: $n = 200$, $\bar{x} = 1045$, $s = 85$. Los datos muestrales parecen provenir de una población con una distribución que no es normal, con σ desconocida.
7. Aseveración: $\mu = 0.0105$. Datos muestrales: $n = 17$, $\bar{x} = 0.0134$, $s = 0.0022$. Los datos muestrales parecen provenir de una población que se aleja mucho de una distribución normal, con σ desconocida.
8. Aseveración: $\mu = 75$. Datos muestrales: $n = 15$, $\bar{x} = 66$, $s = 12$. Los datos muestrales parecen provenir de una población distribuida normalmente, con $\sigma = 14$.

Cálculo de valores P . En los ejercicios 9 a 12, utilice algún recurso tecnológico para calcular el valor P o la tabla A-3 para calcular un rango de valores para el valor P .

9. Prueba de cola derecha con $n = 7$ y estadístico de prueba $t = 3.500$.
10. Prueba de cola izquierda con $n = 27$ y estadístico de prueba $t = -1.500$.
11. Prueba de dos colas con $n = 21$ y estadístico de prueba $t = 9.883$.
12. Prueba de dos colas con $n = 11$ y estadístico de prueba $t = -2.000$.

Cálculo de los componentes de prueba. En los ejercicios 13 y 14, suponga que se seleccionó una muestra aleatoria simple de una población distribuida de manera normal. Obtenga la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor P (o rango de valores P), el valor (o valores) crítico(s) y establezca la conclusión final.

13. Aseveración: La puntuación media del CI de profesores de estadística es mayor que 120. Datos muestrales: $n = 12$, $\bar{x} = 132$, $s = 12$. El nivel de significancia es $\alpha = 0.05$.
14. Aseveración: La vida media de una computadora personal de escritorio es menor de 7 años. Datos muestrales: $n = 21$, $\bar{x} = 6.8$ años, $s = 2.4$ años. El nivel de significancia es $\alpha = 0.05$.

Interpretación de la pantalla de resultados. En los ejercicios 15 y 16, pruebe la aseveración dada interpretando la pantalla de los resultados de la prueba de hipótesis.

15. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los estudiantes de estadística tienen una puntuación media de CI mayor que 110. Los datos muestrales se resumen con los estadísticos $n = 25$, $\bar{x} = 118.0$, y $s = 10.7$. Consulte la siguiente pantalla de resultados de Minitab.

Test of mu = 110 vs > 110							
N	Mean	StDev	SE Mean	95%		T	P
				Lower	Bound		
25	118.000	10.700	2.140	114.339	114.339	3.74	0.001

16. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la vida media de teléfonos celulares es igual a 5 años. Los datos muestrales se resumen con los estadísticos $n = 27$, $\bar{x} = 4.6$ años y $s = 1.9$ años. Consulte la imagen de SPSS en la siguiente página.

SPSS

	Test Value = 5					
	t	df	Sig. (2-tailed)	Mean	95% Confidence Interval of the Difference	
				Difference	Lower	Upper
ADE	-1.094	26	.284	-4000.007

Prueba de hipótesis. En los ejercicios 17 a 32, suponga que se seleccionó una muestra aleatoria simple de una población distribuida de manera normal, y pruebe la aseveración dada. A menos que su profesor lo especifique, utilice el método tradicional o el método del valor P para probar las hipótesis.

17. **Temperaturas corporales.** El conjunto de datos 2 del apéndice B incluye 106 temperaturas corporales con una media de 98.20°F y una desviación estándar de 0.62°F. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la temperatura media corporal es menor que 98.6°F. Con base en esos resultados, ¿parece que la media de 98.6°F que suele utilizarse es errónea?
18. **Pelotas de béisbol.** En pruebas anteriores, se dejaron caer pelotas de béisbol desde una altura de 24 pies sobre una superficie de concreto; las pelotas rebotaron un promedio de 92.84 pulgadas. En una prueba realizada a una muestra de 40 pelotas nuevas, la altura del rebote tuvo una media de 92.67 in, con una desviación estándar de 1.79 in (según datos de Bookhaven National Laboratory y *USA Today*). Utilice un nivel de significancia de 0.05 para determinar si existe evidencia suficiente para sustentar la aseveración de que las nuevas pelotas tienen rebotes con una media distinta a 92.84 in. ¿Parecería que las pelotas son diferentes?
19. **Pesos al nacer.** En un estudio sobre los efectos del consumo de cocaína durante el embarazo sobre los bebés, se obtuvieron los siguientes datos muestrales de pesos al nacer: $n = 190$, $\bar{x} = 2700$ g, y $s = 645$ g (según datos de “Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure”, de Singer *et al.*, *Journal of the American Medical Association*, vol. 291, núm. 20). Utilice un nivel de significancia de 0.01 para probar la aseveración de que los pesos al nacer de los hijos de consumidoras de cocaína tienen una media menor que la media de 3103 g que los bebés de madres que no consumen esta droga. Con base en los resultados, ¿parecería que los pesos al nacer se ven afectados por el consumo de cocaína?
20. **Calificación de crédito.** Cuando los consumidores solicitan un crédito, éste se califica utilizando las puntuaciones FICO (Fair, Isaac and Company). Se obtiene una muestra aleatoria de calificaciones FICO de crédito, las cuales se resumen con estos estadísticos: $n = 18$, $\bar{x} = 660.3$, y $s = 95.9$. Utilice un nivel de significancia de 0.05 para probar la aseveración de que estas calificaciones de crédito provienen de una población con una media igual a 700. Si el Bank of Newport exige una calificación de crédito de 700 o mayor para otorgar un préstamo destinado a adquirir un automóvil, ¿los resultados indican que todos podrían ser elegidos para concederles un crédito de este tipo? ¿Por qué?
21. **Tratamiento del síndrome de fatiga crónica.** Se sometió a prueba a pacientes con síndrome de fatiga crónica, luego se les dio un tratamiento con fludrocortisona y después se sometieron a prueba nuevamente. Se utilizó una escala estándar de -7 a +7 para medir la fatiga antes y después del tratamiento. Los cambios se resumen con los siguientes estadísticos: $n = 21$, $\bar{x} = 4.00$, y $s = 2.17$ (según datos de “The Relationship Between Neurally Mediated Hypotension and the Chronic Fatigue Syndrome” de BouHalaigah, Rowe, Kan y Calkins, *Journal of the American Medical Association*, vol. 274, núm. 12). Los cambios se calcularon de tal forma que los valores positivos representan mejorías. Utilice un nivel de significancia de 0.01 para probar la aseveración de que el cambio medio es positivo. ¿Parece ser efectivo el tratamiento?
22. **Eficacia de una dieta.** Cuarenta individuos se sometieron a la dieta Weight Watchers durante un año. Los cambios en su peso se resumen con los siguientes estadísticos: $\bar{x} = -6.6$ lb, libras y $s = 10.8$ libras (según datos de “Comparison of the Atkins,

Ornish, Weight Watchers, and Zone Diets for Weight Loss and Heart Disease Reduction”, de Dansinger *et al.*, *Journal of the American Medical Association*, vol, 293, núm. 1). Utilice un nivel de significancia de 0.01 para probar la aseveración de que la dieta no tiene ningún efecto. Con base en los resultados, ¿parece que la dieta es eficaz?

- 23. Estaturas de supermodelos.** Se midió la estatura de las supermodelos Niki Taylor, Nadia Avermann, Claudia Schiffer, Elle MacPherson, Christy Turlington, Bridget Hall, Kate Moss, Valeria Mazza y Kristy Hume. Ellas tienen una media de 70.2 in y una desviación estándar de 1.5 in. Utilice un nivel de significancia de 0.01 para probar la aseveración de que las supermodelos tienen estaturas con una media que es mayor a la media de 63.6 in de las mujeres en la población general. Dado que sólo contamos con nueve estaturas, ¿realmente podemos concluir que las supermodelos son más altas que la mujer típica?
- 24. Periodo de vida de un director de orquesta.** Un artículo del *New York Times* señaló que la media del periodo de vida de 35 directores de orquesta varones era de 73.4 años, en contraste con la media de 69.5 años de la población general de hombres. Suponiendo que los 35 varones tienen periodos de vida con una desviación estándar de 8.7 años, utilice un nivel de significancia de 0.05 para probar la aseveración de que los directores de orquesta varones tienen un periodo medio de vida mayor que 69.5 años. ¿Parecería que los directores de orquesta varones viven más que los varones de la población general? ¿Por qué la experiencia de ser un director de orquesta varón no hace que los hombres vivan más tiempo? (*Sugerencia:* Preúntese si los directores de orquesta nacen, o se convierten en directores a una edad mucho más tardía).
- 25. Verificación de plomo en el aire.** Más adelante se listan cantidades medidas de plomo (en microgramos por metro cúbico o $\mu\text{g}/\text{m}^3$) en el aire. La Environmental Protection Agency estableció un estándar de calidad del aire para el plomo de $1.5 \mu\text{g}/\text{m}^3$. Las mediciones presentadas abajo se registraron en el edificio 5 del World Trade Center en diferentes días, inmediatamente después de la destrucción causada por los ataques terroristas del 11 de septiembre de 2001. Después del colapso de los dos edificios del World Trade Center, surgió una gran preocupación sobre la calidad del aire. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la muestra proviene de una población con una media mayor que el parámetro de la EPA de $1.5 \mu\text{g}/\text{m}^3$. ¿Existe algo en estos datos que sugiera que el supuesto de una población distribuida normalmente podría no ser válido?

5.40 1.10 0.42 0.73 0.48 1.10

- 26. Azúcar en el cereal.** Se seleccionaron al azar diferentes cereales y se obtuvo el contenido de azúcar (gramos de azúcar por gramo de cereal), con los siguientes resultados para Cheerios, Harmony, Smart Start, Cocoa Puffs, Lucky Charms, Corn Flakes, Fruit Loops, Wheaties, Cap'n Crunch, Frosted Flakes, Apple Jacks, Bran Flakes, Special K, Rice Krispies, Corn Pops y Trix. Utilice un nivel de significancia de 0.05 para probar la aseveración de un cabildero de que la media de todos los cereales es menor que 0.3 g.

0.03 0.24 0.30 0.47 0.43 0.07 0.47 0.13
0.44 0.39 0.48 0.17 0.13 0.09 0.45 0.43

- 27. El mamífero más pequeño del mundo.** El mamífero más pequeño del mundo es el murciélagos abejorro, también conocido como murciélagos nariz de cochino (o *Craseonycteris thonglongyai*). Estos animales apenas alcanzan el tamaño de un abejorro grande. A continuación se incluyen los pesos (en gramos) de una muestra de estos murciélagos. Ponga a prueba la aseveración de que estos murciélagos provienen de la misma población con un peso medio de 1.8 g.

1.7 1.6 1.5 2.0 2.3 1.6 1.6 1.8 1.5 1.7 2.2 1.4 1.6 1.6 1.6

- 28. Ganadores olímpicos.** A continuación se presentan los tiempos ganadores (en segundos) de hombres en la carrera de 100 metros, durante juegos olímpicos de verano consecutivos, listados en orden por renglón. Suponiendo que estos resultados son datos muestrales seleccionados al azar de la población de todos los juegos olímpicos pasados y futuros, pruebe la aseveración de que el tiempo medio es menor que 10.5 s. ¿Qué observa sobre la precisión de los números? ¿Qué característica sumamente importante del conjunto de datos no se toma en cuenta en esta prueba de hipótesis? ¿Sugieren los resultados de la prueba de hipótesis que los tiempos ganadores futuros estarán alrededor de los 10.5 s? ¿Es válida una conclusión como ésta?

12.0	11.0	11.0	11.2	10.8	10.8	10.8	10.6	10.8	10.3	10.3	10.3
10.4	10.5	10.2	10.0	9.95	10.14	10.06	10.25	9.99	9.92	9.96	

Uso de conjuntos de datos del apéndice B. En los ejercicios 29 a 32, utilice el conjunto de datos del apéndice B para probar la aseveración dada.

- 29. Conjunto de datos del apéndice B: Pesos de monedas de 25 centavos.** Utilice los pesos de las monedas de 25 centavos acuñadas después de 1964, incluidas en el conjunto de datos 14 del apéndice B. Pruebe la aseveración de que las monedas se fabricaron de acuerdo con la especificación estadounidense de acuñación, que establece una media igual a 5.670 g.
- 30. Conjunto de datos del apéndice B: Errores de pronóstico.** Remítase al conjunto de datos 8 del apéndice B y reste cada temperatura máxima real de la temperatura máxima pronosticada un día antes. El resultado es una lista de errores. Pruebe la aseveración de que todos esos errores tienen una media igual a 0. ¿Qué sugiere el resultado acerca de la exactitud de las temperaturas pronosticadas?
- 31. Conjunto de datos del apéndice B: Pulso.** En el momento más intenso de un programa de ejercicio, el autor aseveró que su pulso era menor que el pulso medio de un hombre típico. La medida del pulso del autor fue de 60 latidos por minuto. Utilice los pulsos de hombres del conjunto de datos 1 en el apéndice B para probar la aseveración de que tales pulsos provienen de la población con una media mayor que 60.
- 32. Conjunto de datos del apéndice B: consumo de tabaco en películas infantiles.** Remítase al conjunto de datos 5 del apéndice B y sólo considere las películas que indican algún consumo de tabaco. Ponga a prueba la aseveración de un crítico de cine de que “en las películas que muestran el consumo de tabaco, el tiempo medio de exposición es de 2 minutos”. De acuerdo con los datos muestrales, ¿la afirmación es engañosa?

8-5 MÁS ALLÁ DE LO BÁSICO

- 33. Método alternativo.** Cuando se prueba una aseveración respecto de una media poblacional μ utilizando una muestra aleatoria simple, obtenida de una población distribuida normalmente, con σ desconocida, un método alternativo (que no se utiliza en este libro) consiste en emplear los métodos de esta sección si la muestra es pequeña ($n \leq 30$), pero si la muestra es grande ($n > 30$) sustituya s por σ y proceda como si conociera σ (como en la sección 8-4). Una muestra de tamaño $n = 32$ tiene $\bar{x} = 105.3$ y $s = 15.0$. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la muestra proviene de una población con una media igual a 100. Utilice el método alternativo de comparar los resultados con los obtenidos por medio del método de esta sección.
- 34. Uso de la distribución incorrecta.** Cuando se prueba una aseveración acerca de una media poblacional, con una muestra aleatoria simple seleccionada de una población distribuida normalmente, con σ desconocida, se debe emplear la distribución t de Student para calcular los valores críticos y/o un valor P . Si en vez de ello usted utiliza de forma incorrecta una distribución normal estándar, ¿este error lo hace más o menos proclive a rechazar la hipótesis nula, o no hace ninguna diferencia? Explique.

35. Efecto de un valor extremo. Repita el ejercicio 25 después de cambiar el primer valor de 5.40 por 540. Con base en los resultados, describa el efecto de un valor extremo en una prueba t .

36. Cálculo de los valores críticos t . Cuando se calculan valores críticos, en ocasiones necesitamos niveles de significancia diferentes a los que están disponibles en la tabla A-3. Algunos programas de cómputo aproximan valores críticos t al calcular

$$t = \sqrt{df \cdot (e^{A^2/df} - 1)}$$

donde $gl = n - 1$, $e = 2.718$, $A = z(8 \cdot df + 3)/(8 \cdot df + 1)$, y z es la puntuación crítica z . Utilice esta aproximación para calcular la puntuación crítica t correspondiente a $n = 10$ y un nivel de significancia de 0.05 en un caso de cola derecha. Compare los resultados con el valor crítico t obtenido en la tabla A-3.

37. Potencia de una prueba. Remítase a los datos muestrales del ejercicio 27 y suponga que está utilizando nivel de significancia de 0.05 para probar la aseveración de que $\mu < 1.8$ g. Se utiliza Minitab para calcular que $\beta = 0.5873$, dado que la media real es 1.7 g. Calcule la potencia de la prueba y la probabilidad de un error tipo II. Al parecer, ¿la potencia es lo suficientemente alta para que la prueba sea muy eficaz al rechazar $\mu = 1.8$ g, cuando en realidad $\mu = 1.7$ g?

Prueba de una aseveración respecto de una desviación estándar o de una varianza

Concepto clave En esta sección se presentan métodos para probar una aseveración respecto de una desviación estándar poblacional σ o varianza poblacional σ^2 . Los métodos de esta sección utilizan la distribución chi cuadrada, que se explicó en la sección 7-5. A continuación se resumen los supuestos, el estadístico de prueba, el valor P y los valores críticos.

Prueba de aseveraciones acerca de σ o σ^2

Requisitos

1. La muestra es aleatoria simple.
2. La población tiene una distribución normal. (Éste es un requisito mucho más estricto que el de una distribución normal cuando se prueban aseveraciones acerca de medias, como en las secciones 8-4 y 8-5).

Estadístico de prueba para probar una aseveración acerca de σ o σ^2

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

Valores P y valores críticos: Utilice la tabla A-4, con $gl = n - 1$ para el número de grados de libertad. (La tabla A-4 está basada en *áreas acumulativas de la derecha*).

No utilice los métodos de esta sección con una población que tiene una distribución que se aleja mucho de la normalidad. En las secciones 8-4 y 8-5 vimos que los métodos de prueba de aseveraciones acerca de medias requieren de una población distribuida de forma normal, y que esos métodos funcionan razonablemente

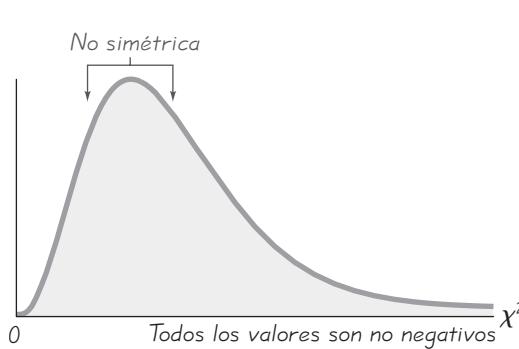


Figura 8-13 Propiedades de la distribución chi cuadrada

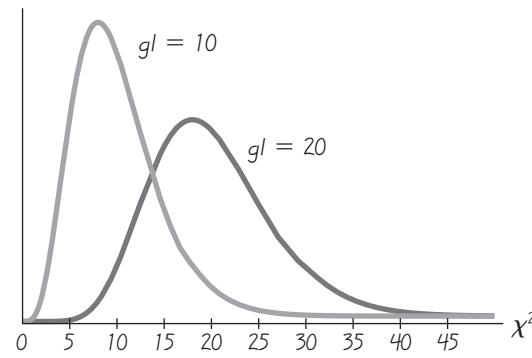


Figura 8-14 Distribución chi cuadrada para $gl = 10$ y $gl = 20$

bien siempre y cuando la distribución poblacional no se aleje mucho de la normalidad. Sin embargo, las pruebas de aseveraciones acerca de desviaciones estándar o varianzas no son tan *robustas* con respecto a formas alejadas de la normalidad, lo que quiere decir que no funcionan muy bien con distribuciones que no tienen una distribución normal. Por consiguiente, la condición de una población distribuida normalmente es un requisito mucho más estricto en esta sección.

La distribución chi cuadrada se explicó en la sección 7-5, donde señalamos las siguientes propiedades importantes.

Propiedades de la distribución chi cuadrada

1. Todos los valores de χ^2 son no negativos y la distribución no es simétrica (véase la figura 8-13).
2. Existe una distribución χ^2 diferente para cada número de grados de libertad (véase la figura 8-14).
3. Todos los valores críticos se encuentran en la tabla A-4, utilizando

$$\text{grados de libertad} = n - 1$$

La tabla A-4 está basada en áreas acumulativas de la zona *derecha* (a diferencia de los datos de la Tabla A-2 que representan áreas acumulativas de la zona izquierda). Para obtener los valores críticos en la tabla A-4, primero se localiza el renglón correspondiente al número apropiado de grados de libertad (donde $gl = n - 1$). Luego, se utiliza el nivel de significancia α para determinar la columna correcta. Los siguientes ejemplos se basan en un nivel de significancia de $\alpha = 0.05$, pero se puede emplear cualquier otro nivel de significancia de manera similar.

Prueba de cola derecha: Puesto que el área a la *derecha* del valor crítico es 0.05, localice 0.05 en la parte superior de la tabla A-4.

Prueba de cola izquierda: Con una área de cola izquierda de 0.05, el área a la *derecha* del valor crítico es 0.95, así que localice 0.95 en la parte superior de la tabla A-4.

Prueba de dos colas: Divida el nivel de significancia de 0.05 entre la cola de derecha y la cola izquierda, de manera que las áreas a la *derecha* de los dos valores críticos sean 0.975 y 0.025, respectivamente. Localice 0.975 y 0.025 en la parte superior de la tabla A-4 (véase la figura 7-10 y el ejemplo en la página 366).

EJEMPLO Control de calidad El mundo de la industria comparte esta meta común: mejorar la calidad reduciendo la variación. Los ingenieros de control de calidad desean asegurarse de que un producto tenga una media aceptable, pero también quieren producir artículos con una calidad *consistente*, eliminando los defectos. La Newport Bottling Company ha fabricado latas de bebidas de cola con cantidades que tienen una desviación estándar de 0.051 onzas. Se prueba una nueva máquina embotelladora, y una muestra aleatoria simple de 24 latas produce las cantidades (en onzas) que se listan a continuación. (Las 24 cantidades tienen una desviación estándar de $s = 0.039$ oz). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las latas de bebidas de cola de la nueva máquina tienen cantidades con una desviación estándar menor que 0.051 oz.

11.98	11.98	11.99	11.98	11.90	12.02	11.99	11.93
12.02	12.02	12.02	11.98	12.01	12.00	11.99	11.95
11.95	11.96	11.96	12.02	11.99	12.07	11.93	12.05

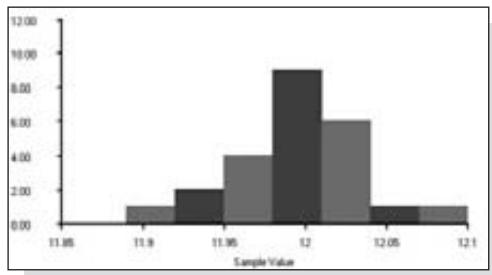
SOLUCIÓN

REQUISITOS

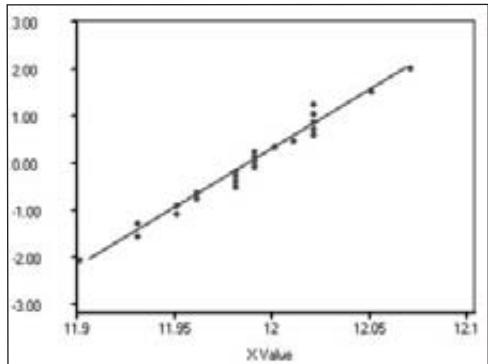
✓ Vea los dos requisitos que se nombran arriba.

1. La muestra es aleatoria simple.
2. Con base en el histograma y la gráfica cuantílar normal generados por STATDISK, parece que la muestra proviene de una población con una distribución normal. Al parecer, el histograma tiene forma de campana. Los puntos en la gráfica cuantílar normal se acercan mucho a un patrón de línea recta y no se observa ningún otro patrón. No hay valores extremos. La distribución se aleja muy poco de la normalidad. Con ambos requisitos satisfechos, procedemos a aplicar los métodos de esta sección y probar la aseveración de que las cantidades de bebidas de cola provienen de una población con una desviación estándar menor que 0.051 oz. ✓

STATDISK



STATDISK



Usaremos el método tradicional de prueba de hipótesis, como se describe en la figura 8-9.

- Paso 1: La expresión simbólica de la aseveración es $\sigma < 0.051$.
- Paso 2: Si la aseveración original es falsa, entonces $\sigma \geq 0.051$.
- Paso 3: La expresión $\sigma < 0.051$ no incluye igualdad, por lo que se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de que $\sigma = 0.051$.

$$H_0: \sigma = 0.051$$

$$H_a: \sigma < 0.051 \quad (\text{aseveración original})$$

- Paso 4: El nivel de significancia es $\alpha = 0.05$.
- Paso 5: Puesto de que la aseveración es respecto a σ , usamos la distribución chi cuadrada.
- Paso 6: De paso 6: El estadístico de prueba es

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2} = \frac{(24 - 1)(0.039)^2}{0.051^2} = 13.450$$

El valor crítico de 13.091 se encuentra en la tabla A-4, en el renglón 23 (grados de libertad = $n - 1 = 23$), en la columna correspondiente a 0.95. Observe el estadístico de prueba y los valores críticos que se muestran en la figura 8-15.

- Paso 7: Puesto que el estadístico de prueba no está en la región crítica, no rechazamos la hipótesis nula.

INTERPRETACIÓN No hay suficiente evidencia para sustentar la aseveración de que la desviación estándar de las cantidades con la nueva máquina sea menor que 0.051 onzas. Quizás la nueva máquina produce cantidades de bebida de cola que son más consistentes, con una desviación estándar menor que 0.051 oz, pero aún no tenemos evidencia suficiente para sustentar esa aseveración.

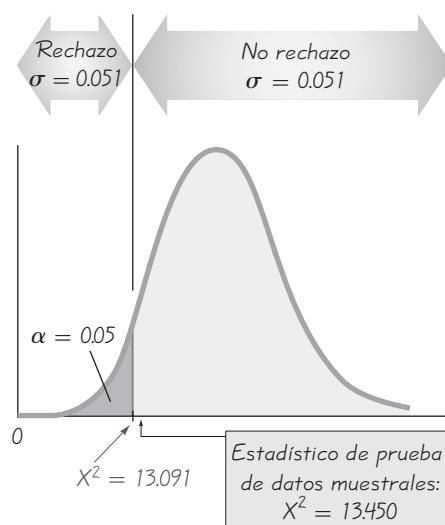
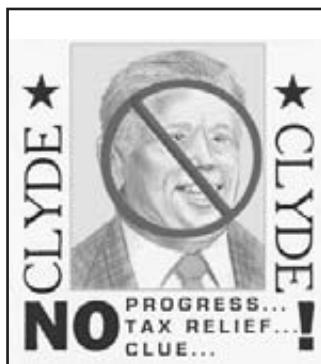


Figura 8-15 Prueba de la aseveración de que $\sigma < 0.051$



Sondeo de empuje

“Sondeo de empuje” es la práctica de efectuar campañas políticas fingiendo realizar un sondeo de opinión. Su nombre se deriva de su objetivo de empujar a los votantes para alejarlos de los candidatos de la oposición haciendo preguntas tendenciosas diseñadas para desacreditar a esos candidatos. He aquí un ejemplo de una pregunta de este tipo: “Dígame, por favor, si sería más o menos probable que usted votara por Roy Romer, si supiera que el gobernador Romer, desde que entró en funciones, nombró una junta de libertad bajo palabra que otorga la libertad, antes de cumplir la totalidad de su condena, a un promedio de cuatro delincuentes convictos al día”. El National Council on Public Polls considera que los sondeos de empuje son poco éticos, pero algunos encuestadores profesionales no censuran la práctica en tanto las preguntas no incluyan información falsa.

Método del valor P

En vez de utilizar el método tradicional de prueba de hipótesis para el ejemplo anterior, también podemos utilizar el método del valor P que se resume en las figuras 8-6 y 8-8. Si se usa STATDISK o una calculadora TI-83/84 Plus en el ejemplo anterior, se obtendrá el valor P de 0.0584. Puesto que el valor P es mayor que el nivel de significancia de 0.05, no rechazamos la hipótesis nula y llegamos a la misma conclusión. Si empleamos la tabla A-4, generalmente no podemos obtener valores P exactos, ya que la tabla de la distribución chi cuadrada únicamente incluye valores seleccionados de α . (A causa de esta limitación, es más fácil probar aseveraciones acerca de σ o σ^2 con la tabla A-4 utilizando el método tradicional, que si se emplea el método del valor P). Si utilizamos la tabla A-4 podemos identificar los límites que contienen al valor P . El estadístico de prueba del último ejemplo es $\chi^2 = 13.450$ y sabemos que la prueba es de cola izquierda con 23 grados de libertad. Remítase al renglón 23 de la tabla A-4 y observe que el estadístico de prueba de 13.450 se ubica entre los valores 13.091 y 14.848 del renglón, lo que significa que el área a la derecha del estadístico de prueba está entre 0.95 y 0.90. Se infiere que el área izquierda del estadístico de prueba está entre 0.05 y 0.10, de manera que concluimos que $0.05 < \text{valor } P < 0.10$. Puesto que el valor P es mayor que el nivel de significancia de $\alpha = 0.05$, no rechazamos la hipótesis nula. Una vez más, el método tradicional y el método del valor P son equivalentes en el sentido de que siempre conducen a la misma conclusión.

Método del intervalo de confianza

El ejemplo anterior también se resuelve con el método del intervalo de confianza de prueba de hipótesis. Aplicando los métodos descritos en la sección 7-5, podemos emplear los datos muestrales ($n = 24$, $s = 0.039$) para construir el siguiente intervalo de confianza del 90%: $0.032 < \sigma < 0.052$. (Recuerde que una prueba de hipótesis de una cola, con un nivel de significancia de 0.05, se efectúa utilizando un intervalo de confianza del 90%. Consulte la tabla 8-2). Con base en este intervalo de confianza, no podemos concluir que σ es menor que 0.051. Llegamos a la misma conclusión que el método tradicional y el del valor P .

Uso de la tecnología

STATDISK Seleccione **Analysis**, después **Hypothesis Testing** y luego **StDev-One Sample**. Proceda a introducir los datos requeridos en el cuadro de diálogo y después haga clic en **Evaluate**. STATDISK desplegará el

estadístico de prueba, los valores críticos, el valor P , la conclusión y el intervalo de confianza.

MINITAB Pruebe la hipótesis generando un intervalo de confianza. Con Minitab Release 14 ingrese los datos en la columna C1, haga clic en **Stat**, luego en **Basic Statistics** y seleccione **Graphical Summary** (o **Display Descriptive Statistics** en versiones anteriores). Ingrese C1 en el recuadro de Variables. (En las versiones anteriores de Minitab, haga clic en **Graphs** y luego en **Graphical Summary**). Indique el nivel de confianza y haga clic en **OK**. Los resultados incluirán un intervalo de confianza para la desviación estándar.

EXCEL Utilice DDXL, seleccione **Hypothesis Tests**, luego **ChiSquare for SD**.

Haga clic en el icono en forma de lápiz e indique el rango de los datos muestrales, tal como A1:A24. Haga clic en **OK** para continuar.

TI-83/84 PLUS La calculadora TI-83/84 Plus no hace pruebas de hipótesis respecto a σ o σ^2 directamente, pero se puede utilizar el programa **S2TEST**. Ese programa fue desarrollado por Michael Lloyd de Henderson State University, y se puede descargar de www.pearsoneducacion.net/Triola. El programa S2TEST utiliza el programa ZZINEWT, por lo que también se debe de instalar. Después de almacenar los programas en la calculadora, presione la tecla **PRGM**, seleccione **S2TEST** y proceda a ingresar la varianza aseverada σ^2 , la varianza muestral s^2 , y el tamaño muestral n . Seleccione el formato utilizado para la hipótesis alternativa y presione la tecla **ENTER**. La computadora mostrará el valor P .



8-6 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Robustez.** ¿A qué nos referimos cuando decimos que la prueba chi cuadrada de esta sección no es *robusta* para distribuciones que se alejan de la normalidad? ¿Cómo afecta esto las condiciones que deben satisfacerse para la prueba chi cuadrada de esta sección?
- Uso de un intervalo de confianza.** Suponga que debe utilizar un nivel de significancia de 0.01 para probar la aseveración de que $\sigma > 5.00$. Si usted planea poner a prueba esa afirmación construyendo un intervalo de confianza, ¿qué nivel de confianza debería usar para el intervalo de confianza? ¿La conclusión basada en el intervalo de confianza será igual a la conclusión basada en una prueba de hipótesis que utiliza el método tradicional o el método del valor P ?
- Requisitos.** Cuando se lanza un dado legal, los resultados tienen una media $\mu = 3.5$ y una desviación estándar $\sigma = \sqrt{35/12} \approx 1.7$. Se lanza un dado 100 veces para tratar de verificar si se comporta como un dado legal. Si usted calcula la desviación estándar de los 100 resultados, ¿podría usar ese valor con los métodos de esta sección para probar la aseveración de que $\sigma = 1.7$ para este dado? ¿Por qué?
- Desviación estándar y varianza.** ¿La prueba de la aseveración de que $\sigma = 2.00$ es equivalente a la prueba de la aseveración de que $\sigma^2 = 4.00$? Al probar la aseveración de que $\sigma = 2.00$, ¿utilizaría un estadístico de prueba diferente al que emplearía para probar la aseveración de que $\sigma^2 = 4.00$?

Cálculo de componentes de prueba. En los ejercicios 5 a 8, calcule el estadístico de prueba, luego utilice la tabla A-4 para obtener el valor (o los valores) crítico(s) de χ^2 , después consulte la tabla A-4 para encontrar los límites que contienen el valor P y determine si existe evidencia suficiente para sustentar la hipótesis alternativa dada.

- $H_1: \sigma \neq 2.00, \alpha = 0.05, n = 10, s = 3.00$.
- $H_1: \sigma > 15, \alpha = 0.05, n = 5, s = 30$.
- $H_1: \sigma < 15, \alpha = 0.01, n = 21, s = 10$.
- $H_1: \sigma \neq 75, \alpha = 0.01, n = 51, s = 70$.

Prueba de aseveraciones sobre variación. En los ejercicios 9 a 20 pruebe la aseveración dada. Suponga que se selecciona una muestra aleatoria simple de una población distribuida normalmente. Utilice el método tradicional de prueba de hipótesis, a menos que su profesor indique otra cosa.

- Pesos al nacer.** Se realizó un estudio de los hijos de madres que consumieron cocaína durante el embarazo y se obtuvieron los siguientes datos muestrales de pesos al nacer: $n = 190, \bar{x} = 2700$ g, y $s = 645$ g (según datos de “Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure”, de Singer *et al.*, *Journal of the American Medical Association*, vol. 291, núm. 20). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la desviación estándar de los pesos al nacer de hijos de consumidoras de cocaína difiere de la desviación estándar de 696 g de los pesos al nacer de hijos de mujeres que no consumieron cocaína durante el embarazo. (Como la tabla A-4 tiene un máximo de 100 grados de libertad, mientras que aquí se requieren 189 grados de libertad, utilice los siguientes valores críticos obtenidos por medio de STATDISK: $\chi_L^2 = 152.8222$ y $\chi_R^2 = 228.9638$.) Con base en el resultado, ¿parece que la cocaína consumida por las madres afecta la variación de los pesos de sus bebés?
- Acuñamiento de monedas de 25 centavos.** En la actualidad las monedas de 25 centavos se acuñan con un peso medio de 5.670 g y una desviación estándar de 0.062 g. Se prueba un nuevo equipo con la intención de mejorar la calidad reduciendo la variación. Se obtiene una muestra aleatoria simple de 24 monedas de 25 centavos acuñadas con el

nuevo equipo, y esta muestra tiene una desviación estándar de 0.049 g. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las monedas acuñadas con el nuevo equipo tienen pesos con una desviación estándar menor que 0.62 g. Al parecer, ¿el nuevo equipo es eficaz para reducir la variación de los pesos? ¿Cuál sería una consecuencia adversa del hecho de tener monedas con pesos muy variables?

11. **Variación en dulces M&M de cacahuate.** Utilice un nivel de significancia de 0.01 para probar la aseveración de que los dulces M&M de cacahuate tienen pesos que varían más que los pesos de los dulces M&M sencillos. La desviación estándar de los pesos de los dulces M&M sencillos es de 0.056 g. Una muestra de 41 dulces M&M de cacahuate tiene pesos con una desviación estándar de 0.31 g. ¿Por qué los dulces de cacahuate tendrán pesos que varían más que los pesos de los dulces sencillos?
12. **Fabricación de altímetros para aviones.** La Stewart Aviation Products Company utiliza un nuevo método de producción para fabricar altímetros para aviones. Se prueba una muestra aleatoria simple de 81 altímetros en una cámara de presión, y se registran los errores en la altitud como valores positivos (para las lecturas que son demasiado altas) o valores negativos (para las lecturas que son demasiado bajas). La muestra tiene una desviación estándar de $s = 52.3$ ft. Con un nivel 0.05 de significancia, pruebe la aseveración de que la nueva línea de producción tiene errores con una desviación estándar diferente de 43.7 ft, que era la desviación estándar del antiguo método de producción. Si parece que la desviación estándar ha cambiado, ¿el nuevo método es mejor o peor que el antiguo método de producción?
13. **Calificación de crédito.** Cuando los consumidores solicitan un crédito, éste se califica utilizando las puntuaciones FICO (Fair, Isaac and Company). Abajo se presentan puntuaciones de crédito para una muestra de solicitantes de préstamos para automóvil, y todos ellos provienen de una nueva sucursal del Bank of Newport. Utilice los datos muestrales para probar la aseveración de que esas calificaciones de crédito provienen de una población con una desviación estándar diferente de 83, que es la desviación estándar de los solicitantes del banco central. Utilice un nivel de significancia de 0.05. Con base en los resultados, ¿parece que los solicitantes de la sucursal tienen calificaciones de crédito que varían más que las de los solicitantes del banco central?

661 595 548 730 791 678 672 491 492 583 762 624 769 729 734 706

14. **El mamífero más pequeño del mundo.** El mamífero más pequeño del mundo es el murciélagos abejorro, también conocido como murciélagos nariz de cochino (*Craseonycteris thonglongyai*). Estos animales apenas alcanzan el tamaño de un abejorro grande. A continuación se incluyen los pesos (en gramos) de una muestra de estos murciélagos. Utilice un nivel de significancia de 0.05 para probar la aseveración de que esos pesos provienen de una población con una desviación estándar igual a 0.30 g, que es la desviación estándar de los pesos de los murciélagos abejorro de una región en Tailandia. ¿Parece que estos murciélagos tienen pesos con la misma variación que los murciélagos de esa región en Tailandia?

1.7 1.6 1.5 2.0 2.3 1.6 1.6 1.8 1.5 1.7 2.2 1.4 1.6 1.6 1.6

15. **Pesos de supermodelos.** Utilice un nivel de significancia de 0.01 para probar la aseveración de que los pesos de mujeres supermodelos varían menos que los pesos de las mujeres en general. La desviación estándar de los pesos de la población de mujeres es de 29 lb. A continuación se listan los pesos (en libras) de nueve supermodelos seleccionadas al azar.

125 (Taylor)	119 (Auermann)	128 (Schiffer)	128 (MacPherson)
119 (Turlington)	127 (Hall)	105 (Moss)	123 (Mazza)
115 (Hume)			

16. **Estaturas de supermodelos.** Utilice un nivel de significancia de 0.05 para probar la aseveración de que las estaturas de mujeres supermodelos varían menos que las estaturas de las mujeres en general. La desviación estándar de las estaturas de la población de

mujeres es de 2.5 in. A continuación se listan las estaturas (en pulgadas) de supermodelos seleccionadas al azar (Taylor, Harlow, Mudler, Goff, Evangelista, Auermann, Schiffer, MacPherson, Turlington, Hall, Crawford, Campbell, Herzogova, Seymour, Banks, Moss, Mazza, Hume).

71	71	70	69	69.5	70.5	71	72	70
70	69	69.5	69	70	70	66.5	70	71

- 17. Control del plomo en el aire.** A continuación se listan cantidades medidas de plomo (en microgramos por metro cúbico o $\mu\text{g}/\text{m}^3$) en el aire. La Environmental Protection Agency estableció un estándar de calidad del aire para el plomo de $1.5 \mu\text{g}/\text{m}^3$. Las mediciones presentadas abajo se registraron en el edificio 5 del World Trade Center en diferentes días, inmediatamente después de la destrucción causada por los ataques terroristas del 11 de septiembre de 2001. Ponga a prueba la aseveración de que estas cantidades provienen de una población con una desviación estándar mayor que $0.4 \mu\text{g}/\text{m}^3$. Utilice un nivel de significancia de 0.05. ¿Hay algo en los valores muestrales que podría violar uno de los requisitos para la aplicación de la prueba chi cuadrada?

5.40 1.10 0.42 0.73 0.48 1.10

- 18. Tiempos de espera de clientes bancarios.** Los valores listados son tiempos de espera (en minutos) de clientes del banco Jefferson Valley, donde los clientes se forman en una sola fila atendida por tres ventanillas. Ponga a prueba la aseveración de que la desviación estándar de los tiempos de espera es menor que 1.9 minutos, que es la desviación estándar de los tiempos de espera del mismo banco cuando se utilizan filas separadas para cada ventanilla. Utilice un nivel de significancia de 0.05. ¿Parece que el uso de una sola fila reduce la variación entre los tiempos de espera? ¿Cuál es una de las ventajas de reducir la variación entre los tiempos de espera?

6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7

- 19. Datos del apéndice B sobre pesos de monedas de 25 centavos.** Remítase al conjunto de datos 14 del apéndice B y utilice los datos muestrales que consisten en los pesos de monedas de 25 centavos acuñadas después de 1964. Ponga a prueba la aseveración de que estas monedas provienen de una población con una desviación estándar igual al valor especificado de 0.068 g. Utilice un nivel de significancia de 0.05. ¿Parece que la variación de los pesos es la deseada?

- 20. Datos del apéndice B sobre cantidades de precipitación.** Remítase al conjunto de datos 8 del apéndice B y utilice los datos muestrales que consisten en cantidades de precipitación. Ponga a prueba la aseveración de que estas cantidades provienen de una población con una desviación estándar menor que 1.00 in. Utilice un nivel de significancia de 0.05. ¿Al parecer se satisface el requisito de una distribución normal? ¿Cómo afecta esto a los resultados?

8-6 MÁS ALLÁ DE LO BÁSICO

- 21. Cálculo de valores críticos de χ^2** Para números grandes de grados de libertad podemos aproximar los valores críticos de χ^2 de la siguiente forma:

$$\chi^2 = \frac{1}{2} \left(z + \sqrt{2k - 1} \right)^2$$

Aquí k es el número de grados de libertad y z es el valor crítico, obtenido en la tabla A-2. Por ejemplo, si deseamos aproximar los dos valores críticos de χ^2 en una prueba de hipótesis de dos colas, con $\alpha = 0.05$ y un tamaño de muestra de 190, permitimos que $k = 189$ con $z = -1.96$, seguido por $k = 189$ y $z = 1.96$. Utilice esta aproximación para estimar los valores críticos de χ^2 en una prueba de hipótesis de dos colas con $n = 190$ y $\alpha = 0.05$. Compare los resultados con los valores críticos dados en el ejercicio 9.

- 22. Cálculo de los valores críticos de χ^2** Repita el ejercicio 21 aplicando esta aproximación (con k y z descritas en el ejercicio 21):

$$\chi^2 = k \left(1 - \frac{2}{9k} + z \sqrt{\frac{2}{9k}} \right)^3$$

- 23. Efectos de un valor extremo.** Al utilizar el procedimiento de prueba de hipótesis de esta sección, ¿se vería muy afectado el resultado con la presencia de un valor extremo? Describa como llegó a su respuesta.

Repaso

En este capítulo se presentaron métodos básicos para la prueba de aseveraciones acerca de una proporción poblacional, una media poblacional o una desviación estándar poblacional (o varianza). Los profesionales utilizan los métodos de este capítulo en una gran variedad de disciplinas, como se ilustra en muchas revistas científicas.

En la sección 8-2 presentamos los conceptos fundamentales de una prueba de hipótesis: la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, la región crítica, el nivel de significancia, el valor crítico, el valor P , el error tipo I y el error tipo II. También estudiamos las pruebas de dos colas, las pruebas de cola izquierda, las pruebas de cola derecha y el planteamiento de conclusiones. Empleamos estos componentes para identificar tres métodos diferentes de prueba de hipótesis:

1. El método del valor P (que se resume en la figura 8-8)
2. El método tradicional (que se resume en la figura 8-9)
3. Los intervalos de confianza (estudiados el capítulo 7)

En las secciones 8-3 a 8-6 estudiamos métodos específicos para manejar distintos parámetros. Puesto que es tan importante seleccionar correctamente la distribución y el estadístico de prueba, presentamos la tabla 8-3, que resume los procedimientos de este capítulo para la prueba de hipótesis.

Conocimientos estadísticos y pensamiento crítico

1. **Significancia práctica.** Para probar la aseveración de que la cantidad media de peso perdido en una dieta es mayor que 0 lb, utilizamos resultados muestrales que indican que 10,000 personas pierden una cantidad media de peso de 0.1 lb, con una desviación estándar de 0.8 lb. Los resultados de la prueba de hipótesis incluyen un estadístico de prueba $t = 12.500$ y un valor P de 0.0000 (redondeado a cuatro decimales). ¿La pérdida de peso es estadísticamente significativa? ¿La pérdida media de peso sugiere que la dieta es práctica? ¿La dieta debería recomendarse?
2. **Muestra.** Usted acaba de reunir una muestra muy grande ($n = 2575$) de respuestas, obtenidas de adultos estadounidenses que enviaron por correo sus respuestas a un cuestionario impreso en la revista *Fortune*. Si se usa esta muestra para hacer inferencias respecto de una población, ¿cuál es la población? Una prueba de hipótesis, realizada con un nivel de significancia de 0.01, lleva a la conclusión de que la mayoría de los adultos (más del 50%) se oponen a los impuestos estatales. ¿Podemos concluir que la mayoría de los adultos estadounidenses se oponen a los impuestos estatales? ¿Por qué?

Tabla 8-3 Pruebas de hipótesis			
Parámetro	Requisitos: Muestra aleatoria simple y . . .	Distribución y estadístico de prueba	Valores <i>P</i> y críticos
Proporción	$np \geq 5$ y $nq \geq 5$	Normal: $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$	Tabla A-2
Media	σ conocida y población distribuida normalmente o σ conocida y $n > 30$	Normal: $z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$	Tabla A-2
	σ conocida y población distribuida normalmente o σ conocida y $n > 30$	<i>t</i> de Student: $t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$	Tabla A-3
	Población no distribuida normalmente y $n \leq 30$	Usar método no paramétrico o <i>bootstrapping</i>	
Desviación estándar o varianza	Población distribuida normalmente	Chi cuadrada: $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$	Tabla A-4

3. **Valor *P*.** Usted acaba de crear una nueva cura para el resfriado común y planea realizar una prueba formal para justificar su eficacia. ¿Qué valor *P* preferiría: 0.99, 0.05, 0.5, 0.01 o 0.001? ¿Por qué?
4. **Valor *P*.** Para poner a prueba la aseveración de que la cantidad media de bebida de cola en latas es mayor que 12 oz, usted utiliza datos muestrales para obtener un valor *P* de 0.250. ¿Qué debe usted concluir? ¿Qué representa la probabilidad de 0.250?

Ejercicios de repaso

1. **Identificación de hipótesis, distribuciones y verificación de requisitos.** Con base en las condiciones dadas, identifique la hipótesis alternativa. Asimismo, identifique la distribución muestral (normal, *t*, chi cuadrada) del estadístico de prueba, o bien, indique que no deben emplearse los métodos de este capítulo.

- a. Aseveración: El ingreso anual medio de estudiantes universitarios de tiempo completo es menor que \$10,000. Datos muestrales: Para 750 estudiantes universitarios seleccionados al azar, la media es \$3662 y la desviación estándar es \$2996. Los datos muestrales sugieren que la población no tiene una distribución normal.
 - b. Aseveración: La mayoría de los estudiantes universitarios tienen al menos una tarjeta de crédito. Datos muestrales: De 500 estudiantes universitarios seleccionados al azar, el 82% tiene al menos una tarjeta de crédito.
 - c. Aseveración: La media del CI de adultos es igual a 100. Datos muestrales: $n = 150$ y $\bar{x} = 98.8$. Es razonable suponer que $\sigma = 15$.
 - d. Aseveración: Al ensamblar manualmente las partes de un teléfono, los tiempos de ensamble varían más que los tiempos del ensamble automático que, según se sabe, tienen una media de 27.6 segundos y una desviación estándar de 1.8 segundos. Datos muestrales: Para 24 tiempos elegidos al azar, la media es de 25.4 segundos, la desviación estándar es de 2.5 segundos, y los datos muestrales sugieren que la población tiene una distribución que se aleja mucho de lo normal.
 - e. Aseveración: Los profesores de estadística tienen puntuaciones de CI que varían menos que las puntuaciones de CI de la población adulta, la cual tiene una desviación estándar de 15. Datos muestrales: Para 24 profesores de estadística seleccionados al azar, la desviación estándar es 10.2 y los datos muestrales sugieren que la población tiene una distribución muy cercana a lo normal.
2. **Errores en las entrevistas de trabajo.** Una encuesta de Accountemps, realizada a 150 ejecutivos, reveló que el 44% de ellos dicen que el error más común de los aspirantes durante una entrevista de trabajo es decir que “no conocen o conocen poco a la empresa” (según datos de *USA Today*). Utilice un nivel de significancia de 0.05 para probar la aseveración de que menos de la mitad de todos los ejecutivos identifican ese error como el más común en una entrevista de trabajo.
 3. **Encuesta de la revista Glamour.** La revista *Glamour* patrocinó una encuesta de 2500 novias y encontró que el 60% de ellas gastan menos de \$750 en su vestido de novia. Utilice un nivel de significancia de 0.01 para probar la aseveración de que menos del 62% de las novias gastan menos de \$750 en su vestido de novia. ¿Cómo se ven afectados los resultados si se sabe que las respuestas se obtuvieron de lectoras de revistas que decidieron responder a la encuesta a través de una página de Internet?
 4. **Pesos de paquetes de azúcar.** Un alumno del autor seleccionó al azar 70 paquetes de azúcar y pesó el contenido de cada uno; obtuvo una media de 3.586 g y una desviación estándar de 0.074 g. Ponga a prueba la aseveración de que los pesos de los paquetes de azúcar tienen una media igual a 3.5 g, tal como se indica en la etiqueta. Si al parecer la media no es igual a 3.5 g, ¿la diferencia es lo suficientemente grande para causar problemas de salud a los consumidores?
 5. **¿Se engaña a los consumidores?** El Orange County Bureau of Weights and Measures recibió quejas de que la Windsor Bottling Company estaba engañando a los consumidores al incluir menos de 12 oz de cerveza de raíz en sus latas. Cuando se seleccionaron al azar 24 latas y se hicieron mediciones, se descubrió que la cantidad media era de 11.4 oz y la desviación estándar de 0.62 oz. El presidente de la compañía, Harry Windsor, asevera que la muestra es demasiado pequeña para ser significativa. Utilice los datos muestrales para probar la aseveración de que se está engañando a los consumidores. ¿Tiene alguna validez el argumento de Harry Windsor?
 6. **Prueba de los efectos adversos de Lipitor.** En pruebas clínicas del fármaco Lipitor (nombre genérico, atorvastatin), 863 pacientes fueron tratados con dosis de 10 mg de atorvastatin, y 19 de ellos experimentaron síntomas de resfriado (según datos de Parke-Davis). Utilice un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de pacientes tratados que presentaron síntomas de resfriado es mayor que la tasa del 1.9% de los pacientes que no recibieron el tratamiento. ¿Parece que los síntomas de resfriado son una reacción adversa del tratamiento?
 7. **Puntuaciones de CI.** Para una muestra aleatoria simple de adultos, las puntuaciones de CI se distribuyen normalmente, con una media de 100 y una desviación estándar de

15. Una muestra aleatoria simple de 13 profesores de estadística produce una desviación estándar de $s = 7.2$. Un psicólogo está muy seguro de que los profesores de estadística tienen puntuaciones de CI con una media mayor que 100. Él no comprende muy bien el concepto de la desviación estándar y no se da cuenta de que ésta debe ser menor que 15 (porque los profesores de estadística tienen una variación menor que la población general). En vez de ello, afirma que los profesores de estadística tienen puntuaciones de CI con una desviación estándar igual a 15, la misma desviación estándar de la población general. Suponga que las puntuaciones de CI de profesores de estadística se distribuyen normalmente y utilice un nivel de significancia de 0.05 para probar la aseveración de que $\sigma = 15$. Con base en el resultado, ¿qué concluye usted acerca de la desviación estándar de las puntuaciones de CI de los profesores de estadística?
- 8. Generación aleatoria de datos.** La calculadora TI-83/84 Plus puede generar datos aleatorios a partir de una población distribuida normalmente. El comando **randNorm(100,15,50)** genera 50 valores de una población distribuida normalmente, con $\mu = 100$ y $\sigma = 15$. Una muestra de 50 valores de este tipo tiene una media de 98.4 y una desviación estándar de 16.3. Utilice un nivel de significancia de 0.10 para probar la aseveración de que la muestra en realidad proviene de una población con una media igual a 100. Suponga que $\sigma = 15$. Con base en los resultados, ¿parecería que el generador de números aleatorios de la calculadora funciona correctamente?
- 9. Temperatura corporal.** Un estudiante del curso propedéutico para medicina, que toma un curso de estadística, necesita hacer un proyecto para su clase. Intrigado por las temperaturas corporales del conjunto de datos 2 en el apéndice B, planea reunir su propia muestra de datos para poner a prueba la aseveración de que la temperatura media corporal es menor que 98.6°F, como suele creerse. Como tiene el tiempo limitado por otros cursos y por el deseo de mantener una vida social que va más allá de hablar en sus sueños, se da cuenta de que sólo tiene tiempo para reunir datos de 12 personas. Después de planear cuidadosamente un procedimiento para obtener una muestra aleatoria simple de 12 adultos saludables, mide sus temperaturas corporales y obtiene los resultados que se muestran a continuación. Utilice un nivel de significancia de 0.05 para probar la aseveración de que estas temperaturas corporales provienen de una población con una media menor que 98.6°F.
- 98.0 97.5 98.6 98.8 98.0 98.5 98.6 99.4 98.4 98.7 98.6 97.6
- 10. Estaturas de mujeres.** Se utilizan datos antropométricos de encuesta para publicar valores que resultan útiles en el diseño de productos para adultos. Según Gordon, Churchill *et al.*, las mujeres tienen estaturas con una media de 64.1 in y una desviación estándar de 2.52 in. La muestra de las 40 estaturas de mujeres en el conjunto de datos 1 del apéndice B tienen una media de 63.20 in, y una desviación estándar de 2.74 in. Utilice un nivel de significancia de 0.05 para probar la aseveración de que esta muestra proviene de una población con una desviación estándar igual a 2.52 in. Al diseñar asientos de automóvil para mujeres, ¿cuál sería una de las consecuencias de creer que las estaturas de las mujeres varían menos de lo que sucede en la realidad?
- 11. Porcentaje de *strikes* declarados por árbitros.** En un año reciente, algunos jugadores profesionales de béisbol se quejaron de que los árbitros estaban declarando más *strikes* que la tasa promedio del 61.0% del año anterior. En cierto momento de la temporada, el árbitro Dan Morrison declaró *strikes* en 2231 de 3581 lanzamientos (según datos de *USA Today*). Utilice un nivel de significancia de 0.05 para probar la aseveración de que esta tasa de *strikes* es mayor al 61.0%.
- 12. Nicotina en cigarrillos.** La Carolina Tobacco Company anunció que sus cigarrillos sin filtro de mayor venta contenían a lo sumo 40 mg de nicotina, pero la revista *Consumer Advocate* realizó pruebas con 10 cigarrillos elegidos al azar y encontró las cantidades (en mg) que se muestran abajo. Es muy grave afirmar que el anuncio de la compañía es incorrecto, por lo que el editor de la revista eligió un nivel de significancia de $\alpha = 0.01$ para poner a prueba su creencia de que el contenido medio de nicotina es mayor que 40 mg. Utilice un nivel de significancia de 0.01 para probar la creencia del editor de que la media es mayor que 40 mg.
- 47.3 39.3 40.3 38.3 46.3 43.3 42.3 49.3 40.3 46.3

Ejercicios de repaso acumulativo

- 1. Estaturas de presidentes.** A continuación se presentan las estaturas (en pulgadas) de los presidentes estadounidenses del siglo xx. Suponga que estos valores son datos muestrales de alguna población más grande.

67	70	72	71	72	70	71	74	69
70.5	72	75	71.5	69.5	73	74	74.5	

- a. Calcule la media.
 - b. Calcule la mediana.
 - c. Calcule la desviación estándar.
 - d. Calcule la varianza.
 - e. Calcule el rango.
 - f. Suponiendo que los valores son datos muestrales, construya un estimado de un intervalo de confianza del 95% de la media poblacional.
 - g. La estatura media de los hombres es de 69.0 pulgadas. Utilice un nivel de significancia de 0.05 para probar la aseveración de que esta muestra proviene de una población con una media mayor que 69.0 in. ¿Parece que los presidentes son más altos que el hombre común?
- 2. Gráfica cuantílar normal.** Remítase a los datos del ejercicio 1 y construya una gráfica cuantílar normal. Con base en el resultado, ¿parece que las estaturas provienen de una población distribuida normalmente?
- 3. Calificaciones de mujeres en la prueba SAT de matemáticas.** Las calificaciones que obtienen mujeres en el área de matemáticas del SAT se distribuyen normalmente, con una media de 496 y una desviación estándar de 108.
- a. Si se selecciona al azar a una mujer que resuelve la parte de matemáticas del SAT, calcule la probabilidad de que su calificación esté por arriba de 500.
 - b. Si se seleccionan al azar cinco calificaciones de matemáticas del SAT de la población de mujeres que resolvieron el examen, calcule la probabilidad de que las cinco calificaciones estén por arriba de 500.
 - c. Si se seleccionan al azar cinco mujeres que resuelven la parte de matemáticas del SAT, calcule la probabilidad de que su media esté por encima de 500.
 - d. Calcule P_{90} , la calificación que separa al 90% inferior del 10% superior.

Actividades de cooperación en equipo

- 1. Actividad en clase** Sin utilizar ningún aparato de medición, cada estudiante debe dibujar dos líneas: una que, según sus cálculos, mida 3 pulgadas de longitud y otra que mida 3 centímetros de longitud. Después, utilice reglas para medir y registrar las longitudes de las líneas. Calcule las medias y las desviaciones estándar de los dos conjuntos de longitudes. Ponga a prueba la aseveración de que las líneas que se supone miden 3 in tienen una longitud media igual a 3 in. Ponga a prueba la aseveración de que las líneas que se supone miden 3 cm tienen una longitud media igual a 3 cm. Compare los resultados. ¿Las estimaciones de la línea de 3 in parecen ser más exactas que las de la línea de 3 cm?
- 2. Actividad en clase** Suponga que un método de selección del género puede afectar la probabilidad de que un bebé sea niña, de manera que la probabilidad es de 1/4. Cada estudiante debe simular 20 nacimientos al sacar 20

cartas de un mazo mezclado. Reemplace cada carta después de ser elegida y mezcle nuevamente. Considere que las cartas de corazones son niñas, y que el resto de las cartas representan niños. Después de hacer 20 selecciones y registrar los “géneros” de los bebés, utilice un nivel de significancia de 0.10 para probar la aseveración de que la proporción de niñas es igual a 1/4. ¿Cuántos estudiantes se espera que obtengan resultados que conduzcan a la conclusión errónea de que la proporción no es de 1/4? ¿Cómo se relaciona esto con la probabilidad de un error tipo I? ¿Al parecer este procedimiento sirve para identificar la eficacia del método de selección del género? (Si no dispone de un mazo de cartas, utilice algún otro recurso para simular los nacimientos, como el generador de números aleatorios de una calculadora o utilice los dígitos de números telefónicos o números de seguridad social).

- 3. Actividad fuera de clase** Grupos de tres o cuatro estudiantes deben acudir a la biblioteca y reunir una muestra de fechas de edición de libros (de acuerdo con las fechas de derecho de autor). Organice y describa el plan de muestreo, ejecute el procedimiento de muestreo y luego utilice los resultados para probar la aseveración de que la media de las antigüedades de todos los libros de la biblioteca es mayor que 20 años.
- 4. Actividad en clase** Cada estudiante debe anotar un estimado de la edad del presidente actual de Estados Unidos. Después, deben reunir todos los estimados y calcular la media muestral y la desviación estándar. Luego, los estudiantes deben poner a prueba la hipótesis de que la media de todos estos estimados es igual a la edad actual del presidente.
- 5. Actividad en clase** Se debe diseñar un proyecto de clase para realizar una prueba en la que cada estudiante recibe una prueba de Coca-Cola y una de Pepsi. A cada estudiante se le pide que identifique cuál de las muestras corresponde a la Coca-Cola. Después de reunir todos los resultados, pruebe la aseveración de que la tasa de éxitos es mejor que la tasa que se esperaría si se hicieran conjjeturas.
- 6. Actividad en clase** Cada estudiante debe estimar la longitud del salón de clases. Los valores deben basarse en estimaciones visuales, es decir, no deben tomarse mediciones reales. Una vez que se han reunido las estimaciones, midan la longitud de la habitación, después prueben la aseveración de que la media muestral es igual a la longitud real del salón de clases. ¿Existe una “sabiduría colectiva” por la que la media de la clase es aproximadamente igual a la longitud real del aula?
- 7. Actividad fuera de clase** Utilice un reloj de pulso que sea razonablemente preciso y póngalo a tiempo. Hágalo con una estación de radio o un reporte telefónico que diga “en el momento del tono, la hora es . . .”. Si no puede poner la hora exacta en segundos, registre el error del reloj que está utilizando. Ahora compare la hora de su reloj con la hora de los demás. Registre los errores con signo positivo para los relojes que están adelantados y con signo negativo para los que están atrasados. Utilice los datos para probar la aseveración de que el error medio de todos los relojes de pulso es igual a 0. ¿Están todos a tiempo, o están adelantados o atrasados? También pruebe la aseveración de que la desviación estándar de los errores es menor que un minuto. ¿Cuáles son las implicaciones prácticas de una desviación estándar excesivamente grande?
- 8. Actividad en clase** En un grupo de tres o cuatro estudiantes, realice un experimento de percepción extrasensorial (PES) seleccionando a uno de los miembros del grupo como sujeto. Dibuje un círculo en un trozo de papel y luego dibuje un cuadrado en otro papel del mismo tamaño. Repita este experimento 20 veces: seleccione al azar el círculo o el cuadrado y colóquelo en la mano del sujeto colocada detrás de él para que no pueda verlo; después pídale que identifique la forma (sin verla); registre si la respuesta es la correcta. Pruebe la aseveración de que el sujeto tiene capacidades de PES porque la proporción de respuestas correctas es mayor que 0.5.
- 9. Actividad en clase** Despues de formar grupos de entre 10 y 20 individuos, cada miembro debe registrar su número de latidos cardiacos por minuto. Despues de calcular \bar{x} y s , cada grupo probará la aseveración de que la media es mayor que 60, que es el resultado del autor. (Cuando las personas hacen ejercicio, tienden a tener pulsos más bajos y el autor corre 5 millas varias veces por semana. ¡Qué tipo!)
- 10. Actividad fuera de clase** Como parte de una encuesta de Gallup, se preguntó a unos sujetos: “¿Está usted a favor de la pena de muerte para las personas sentenciadas por homicidio?” El 65% de los individuos dijeron estar a favor, mientras que el 27% se manifestó en contra y el 8% no opinó. Utilice los métodos de la sección 7-2 para determinar el tamaño de muestra necesario para estimar la proporción de estudiantes de su universidad que están a favor. La clase debe determinar un intervalo de confianza y un margen de error. Despues divida el tamaño muestral entre el número de estudiantes en la clase; realice la encuesta, pidiéndole a cada miembro de la clase que pregunte al número apropiado de estudiantes de la universidad. Analice los resultados para determinar si los estudiantes difieren significativamente de los resultados de la encuesta de Gallup.
- 11. Actividad fuera de clase** Cada estudiante debe buscar un artículo en una revista científica, que incluya una prueba de hipótesis como las que analizamos en este capítulo. Redacte un breve informe que describa la prueba de hipótesis y su papel en el contexto del artículo.

Proyecto tecnológico

Este proyecto presenta las simulaciones como otra manera de probar hipótesis. Suponga que deseamos poner a prueba la aseveración de que los residentes de Maine tienen una media de CI mayor que 100, y obtenemos una muestra aleatoria de 50 puntuaciones de CI de residentes de Maine, con los siguientes resultados: $n = 50$, $\bar{x} = 106.0$, $s = 15.0$. Si utilizamos un nivel de significancia de 0.05, podemos probar la aseveración de que $\mu > 100$ utilizando el método del valor P , el método tradicional y construyendo un intervalo de confianza del 90%.

La idea que subyace en una prueba de hipótesis es la regla del suceso infrecuente para la estadística inferencial, que analizamos en el capítulo 4. Con el uso de esa regla, necesitamos determinar si la puntuación de 106.0 de CI es “infrecuente” o si puede ocurrir fácilmente por azar. Determine esto utilizando simulaciones. De manera repetida, genere 50 puntuaciones de CI de una población distribuida normalmente, con una media supuesta de 100 (como en la hipótesis nula) y una desviación estándar de 15. Luego, con base en las medias muestrales que obtenga, determine si una media como 106.0 o mayor es “infrecuente” o si puede ocurrir fácilmente por azar. Si la media de 106.0 (o mayor) se encuentra en menos del 5% de las muestras, concluya que un valor como 106.0 no puede ocurrir fácilmente por azar, de manera que hay evidencia suficiente para sustentar la aseveración de que la media es mayor que 100.

Utilice STATDISK, Minitab, Excel, la calculadora TI-83/84 Plus o cualquier otro recurso tecnológico que pueda generar datos aleatorios de una población distribuida normalmente, con una media y una desviación estándar dadas. Repita el proceso de generar una muestra aleatoria de 50 valores de una distribución normal con una media de 100 y una desviación estándar de 15. Calcule la media de cada muestra. Genere una cantidad de muestras suficiente para tener una confianza razonable al determinar que una media muestral como 106.0 o mayor es infrecuente o puede ocurrir fácilmente por azar. Registre todos sus resultados. ¿Qué sugieren estos resultados con respecto a la aseveración de que la media del CI es mayor que 100?

A continuación se presentan las instrucciones para generar valores aleatorios de una población distribuida normalmente.

STATDISK: Haga clic en **Data**, después en **Normal Generator**. Proceda a llenar el cuadro de diálogo y luego haga clic en **Generate**. Utilice las funciones copiar y pegar para copiar los datos en la ventana de Sample Editor, luego seleccione **Data** y **Descriptive Statistics** para calcular la media muestral.

Minitab:

Haga clic en **Calc** del menú principal, después en **Random Data** y luego en **Normal**. En el cuadro de diálogo introduzca el tamaño muestral para el número de renglones que se generarán, introduzca C1 para la columna en que se almacenarán los datos, introduzca la media y la desviación estándar y después haga clic en **OK**. Calcule la media muestral al seleccionar **Stat**, **Basic Statistics** y después **Display Descriptive Statistics**.

Excel:

Haga clic en **Tools**, seleccione **Data Analysis**, luego **Random Number Generation** y haga clic en **OK**. En el cuadro de diálogo, introduzca 1 para el número de variables, ingrese el tamaño muestral deseado para la cantidad de números aleatorios, seleccione la opción de distribución **Normal**, introduzca la media y la desviación estándar y después haga clic en **OK**. Obtenga la media muestral al seleccionar **Tools**, **Data Analysis** y luego **Descriptive Statistics**. En la ventana de Descriptive Statistics ingrese el rango de valores (como A1:A50) y asegúrese de hacer clic en el recuadro identificado como **Summary Statistics**.

TI-83/84 Plus:

Primero limpie la lista L1 presionando **STAT**, luego **4:ClrList** y después **L1**. Ahora presione **MATH**, después seleccione **PRB** y elija **6:randNorm(** del menú. Presione **ENTER** y luego proceda a introducir la media, la desviación estándar y el tamaño muestral. Por ejemplo, para generar 50 puntuaciones de CI de una población distribuida normalmente, con una media de 100 y una desviación estándar de 15, debe ingresar **randNorm(100, 15, 50)**. Presione **ENTER** y luego guarde los datos muestrales en la lista L1 presionando **STO L1**, seguido por la tecla **ENTER**. Ahora calcule la media de los valores muestrales en L1 presionando **STAT**, seleccione **CALC** y el primer elemento del menú **1-VAR Stats** e ingrese L1.

De los datos a la decisión

Pensamiento crítico: ¿Los perros pueden detectar enfermedades?

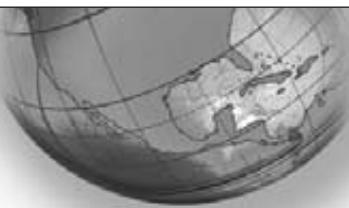
Un estudio publicado en el *British Medical Journal* describe la forma en que se utilizaron perros para identificar a pacientes con cáncer de vejiga. Un ensayo incluyó seis muestras diferentes de orina de personas saludables, además de otra muestra de orina de un individuo que tenía cáncer de vejiga. Se entrenó a los perros para que identificaran la muestra del paciente con cáncer de vejiga. El ensayo se repitió 54 veces, con 22 identificaciones correctas y 32 incorrectas (según datos del *New York Times*).

Análisis de los resultados

- a. Dado que cada ensayo incluyó seis muestras de personas saludables y una muestra de un paciente con cáncer de vejiga, ¿cuál es la probabilidad de que un perro seleccione la muestra del sujeto con cáncer si hizo una elección al azar?
- b. Dado que el ensayo se realizó 54 veces, ¿cuál es el número esperado de indicaciones correctas, suponiendo que se hicieron elecciones al azar?
- c. En los 54 ensayos hubo 22 identificaciones correctas. Ponga a prueba la hipótesis de que los perros tuvieron más éxitos de los que se esperaría por el azar. ¿Parece que los perros estaban adivinando, o

parece que tienen cierta habilidad para identificar la muestra con cáncer?

- d. Suponiendo que los perros tuvieron más éxitos de lo que se esperaría por el azar, ¿tuvieron una actuación tan buena que deberían usarse para hacer diagnósticos médicos reales? ¿Por qué?



Proyecto de Internet

Prueba de hipótesis

En este capítulo se explicó la metodología para la prueba de hipótesis, una técnica esencial para la estadística inferencial. Este proyecto de Internet requerirá que realice pruebas utilizando una variedad de conjuntos de datos en diferentes áreas de estudio. Para cada sujeto se le pedirá que

- reúna datos disponibles en Internet
- formule una hipótesis nula y una alternativa, con base en una pregunta dada

- realice una prueba de hipótesis con un nivel específico de significancia
- resuma sus conclusiones

Vaya a la página de Internet de *Estadística* en <http://www.pearsoneducacion.net/triola> y localice el proyecto de Internet para este capítulo. Ahí encontrará investigaciones guiadas en los campos de la educación, economía y deportes, así como un ejemplo clásico de las ciencias físicas.

La estadística en el trabajo

“Es extremadamente importante que cada uno de nosotros comprenda la estadística para poder procesar de forma efectiva las grandes cantidades de información que se nos presentan cada día en nuestras vidas profesionales y personales”.



Michael Saccucci

Director de estadística y gerencia de calidad para Consumers Union, encargado de probar productos y servicios, así como de dar calificaciones y recomendaciones a los consumidores en la revista Consumer Report.

¿Qué conceptos y procedimientos estadísticos utiliza en Consumers Union?

Durante cualquier día, los especialistas en estadística tienen que utilizar varios procedimientos estadísticos, muchos de los cuales se estudian en este libro de texto. Por ejemplo, en un estudio reciente, realizado para evaluar la calidad y seguridad del pollo, desarrollamos un esquema de muestreo complejo para que los distintos fabricantes estuvieran bien representados. En un estudio reciente de protectores solares, utilizamos la distribución normal para determinar el número adecuado de réplicas necesarias para evaluar correctamente los productos. Dependiendo del tipo de prueba, el especialista en estadística puede necesitar construir un diseño completamente aleatorizado, un diseño aleatorizado por bloques o algún otro tipo de diseño experimental para asegurarse de que los resultados sean exactos y sin sesgos. Durante la fase de análisis, el especialista utiliza diversas técnicas, como el análisis de varianza, el análisis de regresión, el análisis de series de tiempo, el análisis categórico y/o el análisis no paramétrico.

¿Qué hacen los especialistas en estadística en Consumers Union?

Realizan gran variedad de tareas. En las primeras etapas de un proyecto, el especialista en estadística trabaja con el equipo del proyecto para desarrollar el protocolo de prueba y ayudar a seleccionar los productos que van a probarse. Después, ayuda a crear un diseño experimental adecuado para la prueba. Una vez que se han obtenido los datos de prueba, analiza los resultados y presenta los hallazgos en un informe estadístico. El especialista también interviene en una variedad de proyectos especiales, dependiendo de las necesidades de la organización. Los consumidores confían en la información que ofrecemos, por lo que es importante que utilicemos

las técnicas estadísticas apropiadas para asegurarnos de que nuestras evaluaciones son correctas.

¿Qué pasos sigue para asegurar objetividad en sus procedimientos de prueba?

Es política de la Consumers Union que todas las pruebas se realicen de manera objetiva y científica y que se cuide la seguridad del personal de prueba. Hacemos grandes esfuerzos para respetar esta política. Por ejemplo, no aceptamos ningún tipo de publicidad externa en nuestras publicaciones. Empleamos compradores anónimos distribuidos en todo el territorio de Estados Unidos para adquirir nuestras muestras de prueba de las mismas formas disponibles a los consumidores. No aceptamos muestras gratuitas de nadie, incluyendo vendedores. No probamos muestras enviadas por un fabricante que no solicitamos. Además, los técnicos emplean diseños experimentales aleatorizados para asegurarse de que nuestras pruebas se realizan con integridad y objetividad científica. Cuando resulta práctico, los artículos que se prueban se codifican de forma ciega, de tal manera que los encargados de efectuar la prueba no saben qué marcas están evaluando.

¿Las calificaciones y recomendaciones de la revista Consumer Reports sólo se basan en la significancia estadística?

No. La información que ofrecemos debe ser útil para los consumidores. Nuestros técnicos realizan una variedad de pruebas para evaluar el desempeño de un producto. Estas pruebas están diseñadas para simular condiciones del uso predecible de los consumidores. Si resulta que existe una significancia estadística, pero no hay una diferencia importante en los resultados de la prueba, no consideramos una marca mejor que otra. Por ejemplo, al

Nota del autor: El autor se reunió con Mike Saccucci y los otros especialistas en estadística en Consumers Union: Keith Newsom-Stewart, Martin Romm y Eric Rosenberg. El autor visitó las instalaciones donde se prueban los productos y observó diversos experimentos en progreso. Quedó muy impresionado con la participación de los profesionales en las diferentes etapas de la prueba de los productos, con el cuidado extremo y detallado de los diseños de los experimentos y con el uso cuidadoso y eficaz de los análisis estadísticos en la prueba de resultados.

probar selladores de agua, podríamos encontrar que existe una diferencia estadísticamente significativa entre las cantidades de agua que se filtra en dos marcas diferentes de sellador. Sin embargo, si la diferencia consiste en unas cuantas gotas de agua, calificaríamos de forma similar a los productos respecto a esa característica.

¿Cree usted que se tiene una mejor percepción de los solicitantes de empleo cuando tienen algunos estudios de estadística?

Dado el nivel de oferta que existe ahora, considero que el conocimiento básico de la estadística sería considerado favorablemente en casi cualquier campo de estudio, sobre todo en las áreas cuantitativas, como las ciencias, la ingeniería y los negocios. Es extremadamente importante que cada uno de nosotros comprenda la estadística para poder procesar de forma efectiva las grandes cantidades de información que se nos presentan cada día en nuestras vidas profesionales y personales. Un enfoque en el pensamiento estadístico sería especialmente útil.

¿Qué tan esenciales considera que son sus antecedentes profesionales para llevar a cabo sus responsabilidades con excelencia?

La misión de Consumers Union es adelantarse a los intereses de los consumidores al brindar información y consejo acerca de productos y servicios, acerca de aspectos

que afectan su bienestar, defendiendo el punto de vista del consumidor. Para ser competitivos tuvimos que buscar formas más eficientes de ofrecer mayor información a los consumidores en menor tiempo. Mi historial, tanto en estadística como en gerencia de calidad, ha sido sumamente valioso para ayudar a que Consumers Union logre esta misión.

Cuando era estudiante universitario, ¿pensaba que utilizaría la estadística en su trabajo?

Inicié mi carrera en matemáticas y realmente no me interesé en la estadística si-no hasta el último año de la carrera. Fue durante el posgrado, mientras trabajaba bajo la dirección del profesor Hoerl en la Universidad de Delaware, cuando me di cuenta cuán interesante sería una carrera en estadística. A pesar de los sentimientos negativos que muchos estudiantes tienen por la estadística, creo tener uno de los trabajos más interesantes. Nunca sé qué esperar durante la jornada. Un día quizá esté sentado en una sesión de capacitación sobre cata de vinos para aprender acerca de los procedimientos de prueba; otro día tal vez tenga que discutir diversas formas para probar pinturas. Sin embargo, la mayoría de los días paso gran parte de tiempo utilizando una computadora para diseñar el próximo estudio o buscando entre grandes cantidades de datos aquellos que resultarán útiles como base de las evaluaciones de productos.



Inferencias a partir de dos muestras

9



- 9-1 Panorama general**
- 9-2 Inferencias acerca de dos proporciones**
- 9-3 Inferencias acerca de dos medias: Muestras independientes**
- 9-4 Inferencias a partir de datos apareados**
- 9-5 Comparación de la variación en dos muestras**

¿Qué es mejor para tratar el síndrome del túnel carpiano: el entablillado o la cirugía?

El síndrome del túnel del carpo es un padecimiento común de la muñeca, producido por la presión en un nervio. Con frecuencia es el resultado del uso constante de movimientos de muñeca repetitivos, como los que se asocian al uso de un teclado. De los distintos tratamientos disponibles, dos son los más comunes: aplicar un entablillado o practicar una cirugía. El tratamiento de entablillado tiene la ventaja de no ser invasivo, de ser más sencillo, más rápido y mucho menos costoso. Sin embargo, ¿estas ventajas justifican optar por el tratamiento del entablillado en vez del tratamiento quirúrgico? Un factor crucial es el éxito del tratamiento. En una prueba aleatoria controlada se identificó a 156 pacientes con síndrome del túnel carpiano, a los que se trató con entablillado o cirugía. Estos pacientes fueron evaluados un año después. El éxito se definió como “completamente recuperado” o “con una gran mejoría”, y se determinó utilizando puntuaciones de los pacientes y otros resultados que se midieron, como el número de noches que los pacientes permanecían despiertos por los síntomas. Después de un año, de los 73 pacientes que fueron tratados con cirugía, 67 resultaron exitosos. De los 83 pacientes tratados con entablillados, 60 resultaron exitosos. Estos resultados se resumen en la tabla 9-1 (según datos de “Splinting vs Surgery in the Treatment of Carpal Tunnel Syndrome”, de Gerritsen *et al.*, *Journal of the American Medical Association*, vol. 288, núm. 10).

Al examinar los resultados de las pruebas en la tabla 9-1, parece que la cirugía es un mejor tratamiento porque su porcentaje de éxito es del 92%, comparado con el 72% del tratamiento de entablillado. Sin embargo, no debemos extraer una conclusión únicamente con base en esos porcentajes de éxito. También debemos tomar en cuenta el tamaño de las muestras y la magnitud de las diferencias entre los dos porcentajes. Asimismo,

debemos considerar la distribución muestral que se aplica. Necesitamos un procedimiento que tome en cuenta todos estos factores importantes.

La tabla 9-1 incluye dos proporciones muestrales: 67/73 (para el grupo de tratamiento con cirugía) y 60/83 (para el grupo de tratamiento de entablillado). En un artículo científico acerca de la prueba, los autores afirmaron que “el tratamiento con cirugía abierta de liberación del túnel carpiano produjo mejores resultados que el tratamiento de entablillado de la muñeca en los pacientes con STC (síndrome del túnel carpiano)”. ¿Los resultados muestrales realmente sustentan la aseveración de que la cirugía es mejor? En este capítulo se estudiarán métodos para probar este tipo de aseveraciones. Después, podremos determinar si el tratamiento con cirugía es *significativamente* mejor que el tratamiento de entablillado.

El problema de este capítulo incluye dos proporciones poblacionales, pero los métodos que analiza también nos permitirán comparar dos medias independientes, las diferencias de datos apareados y dos desviaciones estándar o varianzas independientes.

Tabla 9-1 Tratamientos del síndrome del túnel carpiano

	Tratamiento	
	Cirugía	Entablillado
Éxito un año después del tratamiento	67	60
Número total de sujetos tratados	73	83
Porcentaje de éxito	92%	72%

9-1 Panorama general

En los capítulos 7 y 8 se estudiaron conceptos importantes y fundamentales de la estadística inferencial: métodos para *estimar* valores de parámetros poblacionales y métodos para *probar hipótesis* (o aseveraciones) acerca de parámetros poblacionales. Los capítulos 7 y 8 incluyen métodos que se aplican a una sola muestra, utilizada para hacer una inferencia acerca de un solo parámetro poblacional. Sin embargo, en realidad existen muchas situaciones importantes y significativas en las que es necesario comparar *dos* conjuntos de datos muestrales. Los siguientes son ejemplos típicos incluidos en este capítulo, el cual presenta métodos para utilizar datos muestrales de dos poblaciones de manera que se puedan hacer inferencias acerca de tales poblaciones.

- Poner a prueba la aseveración de que, cuando se trata el síndrome del túnel carpiano, la cirugía es más exitosa que la aplicación de un entabillado.
- Cuando se prueba la eficacia de la vacuna de Salk en la prevención de la poliomielitis paralítica, determinar si el grupo de tratamiento tiene una menor incidencia de poliomielitis que el grupo al que se administró un placebo.
- Cuando se prueba la eficacia del Lipitor, determinar si los sujetos tienen niveles más bajos de colesterol después de tomar el fármaco.
- Dados dos grupos similares de sujetos con depresión bipolar, determinar si el grupo tratado con paroxetina obtiene puntuaciones más bajas en la escala de depresión Hamilton que el grupo que recibió un placebo.

Los capítulos 7 y 8 incluyeron métodos que se aplicaron a proporciones, medias y medidas de variación (desviación estándar y varianza), y en el presente capítulo nos referiremos a esos mismos parámetros. Este capítulo aplica los mismos métodos utilizados en los capítulos 7 y 8 a situaciones que implican comparaciones de dos muestras en vez de una sola muestra.

9-2 Inferencias acerca de dos proporciones

Concepto clave En esta sección se presentan métodos en los que se utilizan dos proporciones muestrales para construir un estimado de un intervalo de confianza de la diferencia entre las proporciones poblacionales correspondientes o para probar una aseveración acerca de dos proporciones poblacionales. Si bien esta sección se basa en proporciones, podemos utilizar los mismos métodos para tratar con probabilidades o podemos tratar con porcentajes utilizando los equivalentes decimales correspondientes. Por ejemplo, tal vez queramos determinar si existe una diferencia entre el porcentaje de reacciones adversas en un grupo placebo y el porcentaje de reacciones adversas en un grupo de tratamiento con un fármaco. Podemos convertir los porcentajes a sus valores decimales correspondientes y proceder a utilizar los métodos de esta sección.

Cuando se prueba una hipótesis acerca de dos proporciones poblacionales o cuando se construye un estimado del intervalo de confianza de la diferencia entre dos proporciones poblacionales, los requisitos y la notación son los siguientes. Observe que cuando se prueba la hipótesis nula de $p_1 = p_2$, no hay necesidad de estimar los parámetros individuales p_1 y p_2 , sino que estimamos su valor común con la proporción muestral agrupada que se describe en el cuadro de resumen.

Requisitos

1. Tenemos proporciones de dos muestras aleatorias simples que son *independientes*. (Las muestras son independientes si los valores muestrales seleccionados de una población no están relacionados ni apareados de alguna forma con los valores muestrales seleccionados de la otra población).
2. Para ambas muestras, el número de éxitos es de al menos 5 y el número de fracasos es de al menos 5.

Notación para dos proporciones

Para la población 1 permitimos que

p_1 = proporción *poblacional*

n_1 = tamaño muestral

x_1 = número de éxitos en la muestra

$$\hat{p} = \frac{x_1}{n_1} \quad (\text{la proporción muestral})$$

$$\hat{q}_1 = 1 - \hat{p}_1$$

Se adjuntan los significados correspondientes a p_2 , n_2 , x_2 , \hat{p}_2 y \hat{q}_2 , que provienen de la población 2.

Proporción muestral agrupada

La proporción de la muestra agrupada se denota por \bar{p} y está dada por:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Denotamos el complemento de \bar{p} como \bar{q} , de manera que $\bar{q} = 1 - \bar{p}$.

Estadístico de prueba para dos proporciones (con $H_0: p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

donde $p_1 - p_2 = 0$ (**supuesto en la hipótesis nula**)

$$\hat{p}_1 = \frac{x_1}{n_1} \quad \text{y} \quad \hat{p}_2 = \frac{x_2}{n_2}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\bar{q} = 1 - \bar{p}$$

Valor P: Utilice la tabla A-2. (Use el valor calculado del estadístico de prueba z y obtenga el valor P siguiendo el procedimiento que se resume en la figura 8-6).

Valores críticos: Utilice la tabla A-2. (Con base en el nivel de significancia α , obtenga valores críticos utilizando los procedimientos de la sección 8-2).

continúa



El margen de error líder

Los autores Stephen Ansola-behere y Thomas Belin escribieron lo siguiente en su artículo “Poll Faulting” (revista *Chance*): “Nuestra mayor crítica al reporte de los resultados de la encuesta es para el margen de error de una proporción individual ($\pm 3\%$), tomando en cuenta que la atención de los medios de comunicación está claramente dirigida al *liderazgo* o delantera de un candidato”. Señalan que el liderazgo es realmente la *diferencia* entre dos proporciones ($p_1 - p_2$) y explican cómo desarrollaron la siguiente regla práctica: el liderazgo es aproximadamente $\sqrt{3}$ veces más grande que el margen de error para cualquier proporción individual. Para una encuesta típica de preelección, un margen de error reportado de $\pm 3\%$ se convierte en alrededor de $\pm 5\%$ para el liderazgo de un candidato sobre el otro. Los autores señalan que debe reportarse el margen de error para el liderazgo.

Estimado del intervalo de confianza de $p_1 - p_2$

El estimado del intervalo de confianza de $p_1 - p_2$ es:

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

donde el margen de error E está dado por $E = z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$

Cálculo del número de éxitos x_1 y x_2 Los cálculos para pruebas de hipótesis e intervalos de confianza requieren que tengamos valores específicos de x_1 , n_1 , x_2 y n_2 . Algunas veces los datos muestrales disponibles incluyen estos números específicos, pero algunas veces es necesario calcular los valores de x_1 y x_2 . Por ejemplo, considere la afirmación de que “cuando se hizo una encuesta a 1125 personas, el 47% de ellas dijeron que nunca viajan por avión o que lo hacen pocas veces”. A partir de esta afirmación podemos ver que $n_1 = 1125$ y $\hat{p}_1 = 0.47$, pero no tenemos el número real de éxitos x_1 . Sin embargo, a partir de $\hat{p}_1 = x_1/n_1$, sabemos que

$$x_1 = n_1 \cdot \hat{p}_1$$

de manera que $x_1 = 1125 \cdot 0.47 = 528.75$. Pero usted no puede tener 528.75 personas que nunca viajan por avión o que lo hacen pocas veces, de manera que el número de éxitos x_1 debe ser el número entero 529. Ahora podemos utilizar $x_1 = 529$ en los cálculos que requieren de este valor. En realidad es bastante sencillo: el 47% de 1125 es $0.47 \cdot 1125$, que da por resultado 528.75, cifra que se redondea a 529.

Pruebas de hipótesis

Ahora consideraremos pruebas de hipótesis acerca de dos proporciones poblacionales, pero sólo nos ocuparemos de pruebas que tienen la hipótesis nula $p_1 = p_2$. (Para aseveraciones de que la diferencia entre p_1 y p_2 es igual a una constante que no sea cero, véase el ejercicio 35 en esta sección). El siguiente ejemplo ayudará a aclarar los papeles que desempeñan x_1 , n_1 , \hat{p}_1 , \bar{p} , etcétera. En específico, usted debe reconocer que bajo el supuesto de igualdad de proporciones, el mejor estimado de la proporción común se obtiene al combinar ambas muestras en una muestra grande, de manera que \bar{p} es el estimador de la proporción poblacional común.



EJEMPLO ¿La cirugía es mejor que el entablillado? El problema del capítulo incluye los resultados de una prueba clínica en la que se dio tratamiento a pacientes con síndrome de túnel carpiano, y los resultados se resumen en la tabla 9-1. Utilice los datos muestrales de la tabla 9-1, con un nivel de significancia de 0.05 para probar la aseveración de que la tasa de éxito de la cirugía es mejor que la tasa de éxito del entablillado.

SOLUCIÓN

REQUISITOS Primero debemos verificar que se satisfagan los requisitos necesarios. Dado el diseño del experimento, es razonable suponer que se trata de una muestra aleatoria simple. Además, el grupo de tratamiento con cirugía es independiente del grupo de tratamiento con entablillado. Para el segundo requisito, observe que el grupo de tratamiento con cirugía tiene 67 éxitos en 73 pacientes, de manera que existen 6 fracasos. Por lo tanto, el grupo de tratamiento

con cirugía tiene al menos 5 éxitos y al menos 5 fracasos. Por otra parte, el grupo de tratamiento con entablillado tiene 60 éxitos en 83 pacientes, de manera que el número de fracasos es 23. Por lo tanto, el grupo de tratamiento con entablillado tiene al menos 5 éxitos y al menos 5 fracasos. Para cada una de las dos muestras revisamos que tengan al menos 5 éxitos y al menos 5 fracasos. La verificación de los requisitos se completó con éxito y procedemos con la prueba de hipótesis.

Ahora utilizaremos el método del valor P para la prueba de hipótesis, tal como se resume en la figura 8-8. En los siguientes pasos estipulamos que los pacientes sometidos a cirugía constituyen la muestra 1, y que los pacientes tratados con entablillado constituyen la muestra 2.

- Paso 1: La aseveración de una mayor proporción de éxitos en el grupo de tratamiento con cirugía se expresa como $p_1 > p_2$.
- Paso 2: Si $p_1 > p_2$ es falso, entonces $p_1 \leq p_2$.
- Paso 3: Puesto que nuestra aseveración de $p_1 > p_2$ no contiene igualdad, se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de igualdad, entonces tenemos

$$H_0: p_1 = p_2 \quad H_1: p_1 > p_2 \quad (\text{aseveración original})$$

- Paso 4: El nivel de significancia es $\alpha = 0.05$.
- Paso 5: Utilizaremos la distribución normal (con el estadístico de prueba dado con anterioridad) como una aproximación de la distribución binomial. Estimamos el valor común de p_1 y p_2 con el estimado de la muestra agrupada \bar{p} , calculado como se indica a continuación, con espacios decimales adicionales para minimizar los errores de redondeo en cálculos posteriores.

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{67 + 60}{73 + 83} = 0.81410256$$

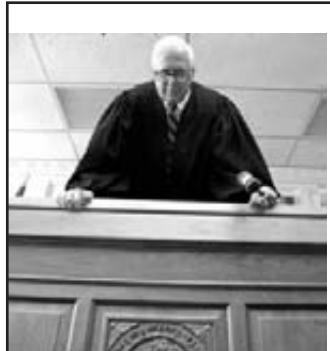
Con $\bar{p} = 0.81410256$, se deduce que $\bar{q} = 1 - 0.81410256 = 0.18589744$.

- Paso 6: Ahora podemos calcular el valor del estadístico de prueba.

$$\begin{aligned} z &= \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} \\ &= \frac{\left(\frac{67}{73} - \frac{60}{83}\right) - 0}{\sqrt{\frac{(0.81410256)(0.18589744)}{73} + \frac{(0.81410256)(0.18589744)}{83}}} \\ &= 3.12 \end{aligned}$$

El valor P de 0.0009 se calcula como sigue: se trata de una prueba de cola derecha, por lo que el valor P es el área ubicada a la derecha del estadístico de prueba $z = 3.12$ (como se indica en la figura 8-6). Remítase a la tabla A-2 y encuentre que el área a la izquierda del estadístico de prueba $z = 3.12$ es 0.9991; por lo tanto, el valor P es

continúa



El autor como testigo

El autor fue requerido para testificar en la Suprema Corte del estado de Nueva York en el caso de un ex alumno que impugnaba una reelección perdida ante la oficina del ayuntamiento del condado Dutchess. El autor testificó utilizando la estadística para demostrar que el comportamiento de votación en un distrito impugnado fue significativamente diferente del comportamiento en todos los demás distritos. Cuando el abogado de la oposición se refirió a los resultados de un intervalo de confianza, preguntó si el 5% de error (de un intervalo de confianza del 95%) podría añadirse a los tres puntos porcentuales del margen de error para obtener un error total de 8%, indicando de esa forma que no entendía el concepto básico de un intervalo de confianza. El juez citó el testimonio del autor, apoyó la demanda del ex alumno y ordenó una nueva elección en el distrito impugnado. Este juicio después fue anulado por la corte de apelación con base en que las irregularidades de la votación debían haberse impugnado antes de la elección, no después.



¿Ayuda la aspirina a prevenir ataques cardíacos?

En un estudio reciente realizado a 22,000 médicos, la mitad recibió dosis normales de aspirina mientras que a la otra mitad se le dieron placebos. El estudio duró seis años, con un costo de \$4.4 millones. Entre quienes tomaron la aspirina, 104 sufrieron ataques cardíacos. De los que tomaron los placebos, 189 sufrieron ataques cardíacos. (Estas cifras están basadas en datos de *Time* y el *New England Journal of Medicine*, vol. 318, núm. 4). Éste es un experimento clásico con un grupo de tratamiento (quienes tomaron la aspirina) y un grupo placebo (los que tomaron tabletas que tenían la apariencia y el sabor de tabletas de aspirina, pero que no contenían aspirina). Podemos utilizar los métodos que se presentan en este capítulo para señalar el hecho de que los resultados muestran una tasa menor estadísticamente significativa de ataques cardíacos en el grupo muestral que tomó aspirina.

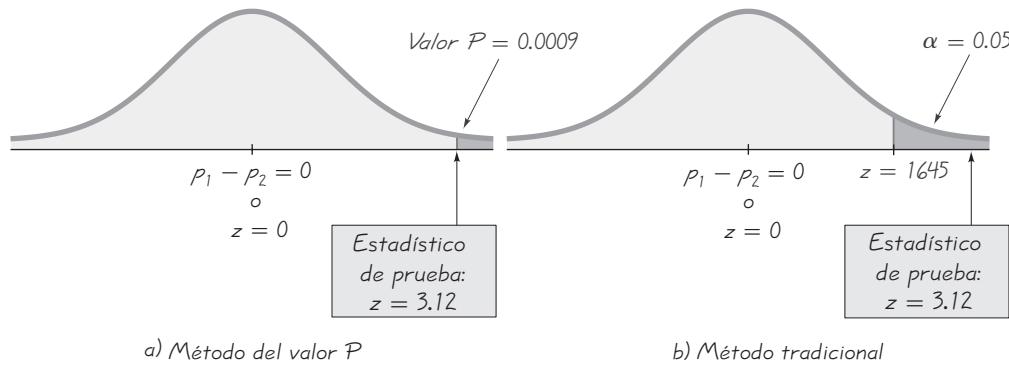


Figura 9-1 Prueba de la aseveración de que la cirugía es mejor que el entablillado

$1 - 0.9991 = 0.0009$. El estadístico de prueba y el valor P se presentan en la figura 9-1a).

Paso 7: Puesto que el valor P de 0.0009 es menor que el nivel de significancia de $\alpha = 0.05$, rechazamos la hipótesis nula de $p_1 = p_2$.

INTERPRETACIÓN Debemos retomar la afirmación original de que el porcentaje de éxito con la cirugía es mayor que el porcentaje de éxito con el entablillado. Puesto que rechazamos la hipótesis nula, concluimos que existe evidencia suficiente para sustentar la aseveración de que la proporción de éxitos con la cirugía es mayor que con el entablillado. (Véase la figura 8-7 para la redacción de la conclusión final). Los investigadores tenían una justificación al escribir en un artículo científico que “el tratamiento con cirugía abierta de liberación del túnel carpiano produjo mejores resultados que el tratamiento de entablillado de la muñeca en los pacientes con STC (síndrome del túnel carpiano)”. Con base en el ensayo clínico y el análisis estadístico posterior de los resultados, los médicos tienen mayores fundamentos cuando recomiendan un tratamiento para el síndrome del túnel carpiano, una afección que a menudo es muy dolorosa.

Método tradicional de prueba de hipótesis

El ejemplo anterior ilustra el método del valor P para prueba de hipótesis, pero sería bastante fácil utilizar en su lugar el método tradicional. En el paso 6, en vez de calcular el valor P , podríamos calcular el valor crítico. Con un nivel de significancia de $\alpha = 0.05$ en una prueba de cola derecha, basada en una distribución normal, remítase a la tabla A-2 para encontrar que una área de $\alpha = 0.05$ en la cola derecha corresponde al valor crítico de $z = 1.645$. Consulte la figura 9-1b), donde podemos observar que el estadístico de prueba se localiza en la región crítica limitada por el valor crítico de $z = 1.645$. Una vez más, rechazamos la hipótesis nula y concluimos que hay suficiente evidencia para sustentar la aseveración de que la cirugía es más exitosa que el entablillado.

Intervalos de confianza

Podemos construir un estimado del intervalo de confianza de la diferencia entre las proporciones poblacionales ($p_1 - p_2$) utilizando el formato anterior. Si un estimado del intervalo de confianza de $p_1 - p_2$ no incluye a 0, tenemos evidencia que sugiere que p_1 y p_2 tienen valores diferentes. Sin embargo, la desviación estándar utilizada

para un intervalo de confianza difiere de la que se utiliza para una prueba de hipótesis. Cuando se prueban las aseveraciones acerca de diferencias entre dos proporciones poblacionales, el método tradicional y el método del valor P son equivalentes en el sentido de que siempre producen los mismos resultados, pero el estimado del intervalo de confianza de la diferencia podría sugerir otra conclusión (véase el ejercicio 34). Si se obtienen diferentes conclusiones, se debe a que los métodos tradicional y del valor P utilizan una desviación estándar *exacta* con base en la suposición de que no existe diferencia entre las proporciones poblacionales (como se estableció en la hipótesis nula). Sin embargo, el intervalo de confianza se construye utilizando una desviación estándar que se basa en valores *estimados* de las dos proporciones poblacionales. Como consecuencia, si usted desea estimar la diferencia entre dos proporciones poblacionales, hágalo construyendo un intervalo de confianza, pero si usted desea probar alguna aseveración acerca de dos proporciones poblacionales, utilice el método del valor P o el método tradicional.

Además, *no pruebe la igualdad de dos proporciones poblacionales determinando si existe un traslape entre dos estimados individuales del intervalo de confianza de las dos proporciones poblacionales individuales*. Cuando se compara con el estimado del intervalo de confianza de $p_1 - p_2$, el análisis del traslape de dos intervalos de confianza individuales es más conservador (rechazando la igualdad con menos frecuencia) y tiene menos potencia (porque es menos probable rechazar $p_1 = p_2$ cuando en realidad $p_1 \neq p_2$). (Véase “On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals”, de Schenker y Gentleman, *American Statistician*, vol. 55, núm. 3). Consulte el ejercicio 33.

PROBLEMA 9 DEL CAPÍTULO

EJEMPLO **¿La cirugía es mejor que el entablillado?** Utilice los datos muestrales que se presentan en la tabla 9-1 para construir un estimado del intervalo de confianza del 90% de la diferencia entre las dos proporciones poblacionales. (El nivel de confianza del 90% es comparable al nivel de significancia de $\alpha = 0.05$ que se utilizó en la anterior prueba de hipótesis de cola derecha. Véase la tabla 8-2 en la sección 8-2).

SOLUCIÓN

REQUISITO La solución del ejemplo anterior comienza con la misma verificación de requisitos que se necesita aquí. Se cumple la verificación de requisitos, así que podemos continuar con la construcción del intervalo de confianza.

Con un nivel de confianza del 90%, $z_{\alpha/2} = 1.645$ (de la tabla A-2). Primero calculamos el valor del margen de error E como se indica.

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} = 1.645 \sqrt{\frac{\left(\frac{67}{73}\right)\left(\frac{6}{73}\right)}{73} + \frac{\left(\frac{60}{83}\right)\left(\frac{23}{83}\right)}{83}} = 0.0966$$

Con $\hat{p}_1 = 67/73 = 0.9178$, $\hat{p}_2 = 60/83 = 0.7299$ y $E = 0.0966$, el intervalo de confianza se evalúa como sigue.

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

$$(0.9178 - 0.7299) - 0.0966 < (p_1 - p_2) < (0.9178 - 0.7299) + 0.0966$$

$$0.0983 < (p_1 - p_2) < 0.292$$

continúa



Experimento en relación con la poliomielitis

En 1954 se realizó un experimento para probar la efectividad de la vacuna de Salk como protección contra los devastadores efectos de la poliomielitis. Aproximadamente 200,000 niños recibieron una inyección de solución salina inocua, y otros 200,000 una inyección de la vacuna. El experimento fue “doble ciego” porque los niños inyectados no sabían si estaban recibiendo la vacuna real o el placebo, en tanto que los doctores que aplicaban las inyecciones y evaluaban los resultados tampoco lo sabían. Sólo 33 de los 200,000 niños vacunados padecieron posteriormente poliomielitis paralítica, mientras que 115 de los 200,000 inyectados con la solución salina padecieron posteriormente poliomielitis paralítica. Un análisis estadístico de éstos y otros resultados llevó a la conclusión de que la vacuna de Salk realmente era efectiva contra la poliomielitis paralítica.



La pena de muerte como correctivo

Un argumento utilizado comúnmente para apoyar la pena de muerte es que ésta desanima a otros individuos de cometer asesinatos. Jeffrey Grogger, de la Universidad de California, analizó los datos sobre los homicidios diarios en ese estado durante cuatro años, en una época en que las ejecuciones eran frecuentes. Entre sus conclusiones, publicadas en el *Journal of the American Statistical Association* (vol. 85, núm. 410), está la siguiente: "El análisis realizado de forma consistente indica que estos datos no sustentan la hipótesis de que la ejecución desanima el asesinato en el corto plazo". Éste es uno de los temas más importantes de política social y los esfuerzos de personas como el profesor Grogger ayudan a disipar las ideas erróneas, de manera que nos permiten tener información precisa para analizar temas como éste.

INTERPRETACIÓN Los límites del intervalo de confianza no contienen a 0, lo que sugiere que existe una diferencia significativa entre las dos proporciones. Además, con base en esos resultados, tenemos una confianza del 90% de que el porcentaje de éxitos con la cirugía es mayor que el porcentaje de éxitos con el entablillado por una cantidad que está entre el 9.83 y el 29.2%.

Fundamentos: ¿Por qué funcionan los procedimientos de esta sección? El estadístico de prueba dado para la prueba de hipótesis se justifica por lo siguiente:

1. Con $n_1 p_1 \geq 5$ y $n_1 q_1 \geq 5$, la distribución de \hat{p}_1 puede aproximarse con una distribución normal con media p_1 , una desviación estándar $\sqrt{p_1 q_1 / n_1}$ y una varianza $p_1 q_1 / n_1$. Estas conclusiones se basan en las secciones 6-6 y 7-2, y también se aplican a la segunda muestra.
2. Puesto que \hat{p}_1 y \hat{p}_2 se aproximan por medio de una distribución normal, $\hat{p}_1 - \hat{p}_2$ también se aproximarán por medio de una distribución normal con media $p_1 - p_2$ y varianza

$$\sigma_{(\hat{p}_1 - \hat{p}_2)}^2 = \sigma_{\hat{p}_1}^2 + \sigma_{\hat{p}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

(El resultado anterior se basa en esta propiedad: la varianza de las *diferencias* entre dos variables aleatorias independientes es la *suma* de sus varianzas individuales).

3. Puesto que generalmente no se conocen los valores de p_1 , q_1 , p_2 y q_2 , y a partir de la hipótesis nula, suponemos que $p_1 = p_2$ y podemos agrupar (o combinar) los datos muestrales. El estimado combinado del valor común de p_1 y p_2 es $\bar{p} = (x_1 + x_2) / (n_1 + n_2)$. Si reemplazamos p_1 y p_2 por \bar{p} y reemplazamos q_1 y q_2 por $\bar{q} = 1 - \bar{p}$, la varianza del paso 2 nos lleva a la siguiente desviación estándar:

$$\sigma_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{\bar{p} \bar{q}}{n_1} + \frac{\bar{p} \bar{q}}{n_2}}$$

4. Sabemos ahora que la distribución de $p_1 - p_2$ es aproximadamente normal, con media $p_1 - p_2$ y desviación estándar como se indica en el paso 3, por lo tanto, el estadístico de prueba z tiene la forma dada antes.

La forma del intervalo de confianza requiere una expresión para la varianza diferente de la que se dio en el paso 3. En el paso 3 suponemos que $p_1 = p_2$, pero si no hacemos esta suposición (como en la construcción del intervalo de confianza), estimamos la varianza de $p_1 - p_2$ como

$$\sigma_{(\hat{p}_1 - \hat{p}_2)}^2 = \sigma_{\hat{p}_1}^2 + \sigma_{\hat{p}_2}^2 = \frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}$$

y la desviación estándar se convierte en

$$\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Uso de la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego seleccione **Hypothesis Testing** o **Confidence Intervals**. Seleccione **Proportion-Two Samples** del menú principal. Ingrese los elementos requeridos en el cuadro de diálogo y luego haga clic en el botón **Evaluate**. La siguiente imagen es del primer ejemplo de esta sección.

STATDISK

```

Claim: p1 > p2
Pooled proportion: 0.8141036
Test Statistic, z: 3.1226
Critical z: 1.6449
P-Value: 0.0009
90% Confidence interval:
0.0983474 < p1-p2 < 0.2914859
Reject the Null Hypothesis
Sample provides evidence to support the
claim

```

MINITAB

Ahora Minitab puede manejar un resumen de estadísticos para dos muestras. Seleccione **Stat** de la barra del menú principal, luego seleccione **Basic Statistics**, luego **2 Proportions**. Haga clic en el botón **Summarize Data** e ingrese los valores muestrales. Haga clic en la barra de **Options**. Introduzca el nivel de confianza deseado. Si realiza una prueba de hipótesis, ingrese el valor aseverado de $p_1 - p_2$, seleccione el formato para la hipótesis alternativa, y haga clic en el cuadro para utilizar el estimado agrupado de p para la prueba. Haga clic en **OK** dos veces.

EXCEL

Usted debe utilizar el complemento Data Desk XL, que complementa a este libro. Primero haga lo siguiente: en la celda A1 ingrese el número de éxitos para la muestra 1, en la celda B1 introduzca el número de ensayos para la muestra 1, en la celda C1 ingrese el número de éxitos para la muestra 2 y en la celda D1 introduzca el número de

ensayos para la muestra 2. Haga clic en **DDXL**. Seleccione **Hypothesis Tests** y **Summ 2 Var Prop Test** o seleccione **Confidence Intervals** y **Summ 2 Var Prop Interval**. En el cuadro de diálogo, haga clic en los cuatro iconos con forma de lápiz e introduzca A1, B1, C1 y D1 en los cuatro campos de entrada. Haga clic en **OK**. Proceda a completar el nuevo cuadro de diálogo.

TI-83/84 PLUS

La calculadora TI-83/84 Plus puede utilizarse para pruebas de hipótesis e intervalos de confianza. Oprima **STAT** y seleccione **TEST**. Luego elija la opción de **2-PropZTest** (para una prueba de hipótesis) o **2-PropZInt** (para un intervalo de confianza). Cuando se prueban hipótesis, la calculadora TI-83/84 Plus mostrará en la pantalla un valor P en vez de valores críticos, por lo que utiliza el método del valor P para prueba de hipótesis.

En el estadístico de prueba

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}$$

utilice los valores positivo y negativo de z (para dos colas) y despeje $p_1 - p_2$. Los resultados son los límites del intervalo de confianza que se dieron antes.

9-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Proporción muestral agrupada.** ¿Qué es una proporción muestral agrupada y qué símbolo la representa? ¿Se utiliza para prueba de hipótesis e intervalos de confianza?
- 2. \hat{p}, \bar{p}, p .** ¿Qué representan los símbolos \hat{p}, \bar{p}, p ?
- 3. Interpretación del intervalo de confianza.** Se utilizan los datos muestrales de dos poblaciones diferentes para construir el siguiente intervalo de confianza del 95%: $0.200 < p_1 - p_2 < 0.300$. Redacte una aseveración que interprete ese intervalo de confianza.
- 4. Equivalencia de prueba de hipótesis e intervalos de confianza.** Dadas las proporciones muestrales de dos poblaciones diferentes, deseamos utilizar un nivel de significancia de 0.05 para probar la aseveración de que $p_1 = p_2$. Una forma de hacerlo es con el método del valor P para prueba de hipótesis; otra opción consiste en utilizar el método tradicional de prueba de hipótesis; también podemos basar la conclusión en

el estimado del intervalo de confianza del 95% de $p_1 - p_2$. ¿Los tres métodos siempre conducen a la misma conclusión? Explique.

Cálculo del número de éxitos. En los ejercicios 5 a 8, calcule el número de éxitos x sugeridos por la afirmación dada.

5. Tomado de un artículo del *New York Times*: de 3250 aparatos que existen en la ciudad de Nueva York para permitir el paso a los peatones mediante la activación de un botón, el 23% funciona.
6. De una encuesta Gallup: de 1018 sujetos encuestados, el 22% fumó cigarrillos la semana anterior.
7. De una encuesta Gallup: de 976 sujetos encuestados, el 7% consumió una bebida alcohólica cada día.
8. De una encuesta de CNN/USA Today/Gallup: de 1003 sujetos encuestados, el 11% dice que la opinión de una celebridad influiría en su propia opinión.

Cálculos para probar aseveraciones. En los ejercicios 9 y 10, suponga que usted planea utilizar un nivel de significancia de $\alpha = 0.05$ para probar la aseveración de que $p_1 = p_2$. Utilice los tamaños muestrales y los números de éxitos dados para calcular a) el estimado agrupado \bar{p} , b) el estadístico de prueba z , c) los valores críticos z y d) el valor P .

	Grupo de tratamiento	Grupo placebo
9.	$n_1 = 500$ $x_1 = 100$	$n_2 = 400$ $x_2 = 50$
10.	Hombres	Mujeres
	$n_1 = 1068$ $x_1 = 332$	$n_2 = 1220$ $x_2 = 420$

11. **Ventaja del equipo local.** Cuando se muestrearon los partidos de una temporada, se encontró que el equipo local ganó 127 de 198 partidos de *básquetbol* profesional, y que el equipo local ganó 57 de 99 partidos de *fútbol* profesional (según datos de “Predicting Professional Sports Game Outcomes from Intermediate Game Scores”, de Cooper *et al.*, *Chance*, vol. 5, núm. 3-4). Observe la pantalla de la computadora TI-83/84 Plus que aparece al margen con los resultados de la prueba de la aseveración de igualdad de proporciones. Con base en los resultados, ¿parece existir una diferencia significativa entre las proporciones de triunfos de los equipos locales? ¿Qué concluye usted acerca de la ventaja de ser el equipo local?

TI-83/84 Plus

```
2-PropZTest
P1≠P2
z=1.098647687
P=.2719218487
P1=.6414141414
P2=.5757575758
↓P=.6195286195
```

12. **Prueba de guantes de laboratorio.** El *New York Times* publicó un artículo acerca de un estudio en el que el profesor Denise Korniewicz y otros investigadores de Johns Hopkins sometieron a tensión guantes de laboratorio. De 240 guantes de vinilo, el 63% presentó filtración de virus. De 240 guantes de látex, el 7% presentó filtración de virus. Observe la siguiente pantalla de resultados de Minitab. Con un nivel de significancia de 0.005, pruebe la aseveración de que los guantes de vinilo tienen una tasa de filtración de virus mayor que los guantes de látex.

Minitab

```
Difference = p (1) - p (2)
Estimate for difference: 0.558333
95% lower bound for difference: 0.500263
Test for difference = 0 (vs > 0): Z = 12.82 P-Value = 0.000
```

- 13. Selección del género.** El Genetics and IVF Institute llevó a cabo un ensayo clínico de sus métodos de selección del género. Mientras se escribía este libro, los resultados incluían 325 bebés nacidos de padres que utilizaron el método XSORT para incrementar la probabilidad de concebir una niña, y 295 de esos bebés fueron niñas. Asimismo, había 51 bebés nacidos de padres que utilizaron el método YSORT para incrementar la probabilidad de concebir un niño, y 39 de ellos fueron niños.
- Construya un estimado del intervalo de confianza del 95% de la diferencia entre la proporción de niñas y la proporción de niños.
 - ¿Al parecer hay una diferencia? ¿Parece que el método XSORT es efectivo? ¿Parece que el método YSORT es efectivo?
- 14. ¿Los mosquiteros reducen la malaria?** En un ensayo aleatorizado y controlado en Kenia, se probaron los mosquiteros tratados con insecticida como una forma de reducir la malaria. De 343 bebés en cuyas cunas se instalaron mosquiteros, 15 desarrollaron malaria. De los 294 bebés que no recibieron la protección de mosquiteros, 27 desarrollaron la enfermedad (según datos de “Sustainability of Reductions in Malaria Transmission and Infant Mortality in Western Kenya with Use of Insecticide-Treated Bednets”, de Lindblade *et al.*, *Journal of the American Medical Association*, vol. 291, núm. 21). Utilice un nivel de significancia de 0.01 para probar la aseveración de que la incidencia de malaria es menor en los bebés que estuvieron protegidos con mosquiteros. ¿Al parecer los mosquiteros son efectivos?
- 15. Métodos de encuesta telefónica.** En un estudio sobre la exactitud de las encuestas telefónicas, 720 personas se rehusaron a responder cuando formaron parte de los 1720 individuos incluidos en una encuesta “estándar” de 5 días. En el mismo estudio, 429 personas se rehusaron a responder cuando formaron parte de los 1640 individuos incluidos en una encuesta “rigurosa” de 8 semanas. (Los datos se basan en resultados de “Consequences of Reducing Nonresponse in a National Telephone Survey”, de Keeter *et al.*, *Public Opinion Quarterly*, vol. 64, núm. 2). Utilice un nivel de significancia de 0.01 para poner a prueba la aseveración de que el porcentaje de negativas es menor en la encuesta rigurosa. ¿Al parecer la encuesta rigurosa produce resultados más exactos?
- 16. Métodos de encuesta telefónica.** El ejercicio anterior incluye una hipótesis unilateral con un nivel de significancia de 0.01. Si usted planeara probar la aseveración utilizando un intervalo de confianza, ¿qué nivel de confianza debe usar? (*Sugerencia:* Consulte la tabla 8-2 en la sección 8-2). Utilice un nivel de confianza apropiado y construya un intervalo de confianza de la diferencia entre las dos tasas de negativas. Con base en el intervalo de confianza, ¿parece que la tasa de negativas es menor en la encuesta rigurosa? ¿Por qué?
- 17. Correo electrónico y privacidad.** En una encuesta con 436 trabajadores, 192 de ellos dijeron que era muy poco ético revisar el correo electrónico de los empleados. Cuando se entrevistó a 121 jefes de alto nivel, 40 dijeron que era muy poco ético revisar el correo electrónico de los empleados (según datos de una encuesta Gallup). Utilice un nivel de significancia de 0.05 para probar la aseveración de que, de aquellos que dicen que es muy poco ético revisar el correo electrónico de los empleados, la proporción de empleados es mayor que la proporción de jefes.
- 18. Correo electrónico y privacidad.** Remítase a los datos muestrales del ejercicio 17 y construya un estimado del intervalo de confianza del 90% de la diferencia entre las dos proporciones poblacionales. ¿Existe una brecha sustancial entre los empleados y los jefes?
- 19. Eficacia de las prohibiciones de fumar.** En 1994, la Joint Commission on Accreditation of Health-care Organizations exigió a los hospitales que prohibieran fumar. En un estudio sobre los efectos de esta prohibición, se seleccionaron al azar a sujetos fumadores de dos poblaciones diferentes. De 843 empleados fumadores que trabajaban en hospitales que aplicaban la prohibición, 56 dejaron de fumar un año después de establecida la prohibición. De 703 empleados fumadores que trabajaban en lugares donde no se aplicaba la prohibición, 27 dejaron de fumar un año después de establecida la prohibición (según datos de “Hospital Smoking Bans and Employee Smoking Behavior”, de Longo, Brownson

et al., Journal of the American Medical Association, vol. 275, núm. 16). ¿Existe una diferencia significativa entre las dos proporciones a un nivel de significancia de 0.05? ¿Existe una diferencia significativa entre las dos proporciones a un nivel de significancia de 0.01? ¿Parece que la prohibición tuvo un efecto en la tasa de abandono del hábito de fumar?

20. **Prueba de la eficacia de una vacuna.** En un artículo de *USA Today* sobre una vacuna infantil experimental en forma de atomizador nasal se presentó la siguiente aseveración: “En una prueba que incluyó a 1602 niños, sólo 14 (el 1%) de los 1070 que recibieron la vacuna, desarrollaron gripe, en comparación con 95 (el 18%) de los 532 que recibieron un placebo”. El artículo también se refería a un estudio que afirmaba que el atomizador nasal experimental “elimina las probabilidades de que los niños se contagien de gripe”. ¿Existe evidencia muestral suficiente para sustentar esa aseveración?
21. **Efectos adversos de Clarinex.** El fármaco Clarinex se utiliza para tratar los síntomas alérgicos. En una prueba clínica de este fármaco, el 2.1% de los 1655 sujetos tratados experimentaron fatiga. De los 1652 sujetos que recibieron placebos, el 1.2% experimentó fatiga (según datos de Schering Corporation). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la incidencia de fatiga es mayor entre los individuos que usaron Clarinex. Al parecer, ¿la fatiga debe preocupar a quienes toman Clarinex?
22. **Efectos adversos de Clarinex.** Utilice los mismos datos muestrales del ejercicio anterior. Si usted planeara probar la aseveración dada construyendo un intervalo de confianza, ¿qué nivel de confianza debería usar? (*Sugerencia:* Consulte la tabla 8-2 en la sección 8-2). Utilice el nivel apropiado de confianza y construya un estimado del intervalo de confianza de la diferencia entre las tasas de fatiga de los usuarios de Clarinex y las de los sujetos que recibieron placebos. Al parecer, ¿la fatiga debe preocupar a quienes toman Clarinex?
23. **Conducir al trabajo.** En una encuesta sobre los hábitos de transporte al trabajo, se descubrió que de 1068 propietarios de su casa, el 82.4% conduce a su trabajo. De las 1064 personas que rentan una casa, el 68.1% conduce a su trabajo (según datos del U.S. Census American Housing Survey). Construya un estimado de un intervalo de confianza del 95% para las diferencias entre la proporción de propietarios de su casa que conducen al trabajo y la proporción de individuos que rentan una casa y que conducen al trabajo. Con base en el resultado, ¿parece existir una diferencia significativa entre las dos proporciones? Identifique al menos un factor importante que podría explicar cualquier diferencia.
24. **Equipaje perdido.** De 5000 piezas de equipaje elegidas al azar, manejadas por American Airlines, se perdieron 22. De 4000 piezas de equipaje elegidas al azar, manejadas por Delta Airlines, se perdieron 15 (según datos del Departamento de Transporte de Estados Unidos). Utilice los datos muestrales para construir un estimado del intervalo de confianza del 95% para la diferencia entre las dos proporciones de equipaje perdido. Con base en el resultado, ¿parece existir una diferencia entre las dos proporciones de equipaje perdido? ¿Por qué?
25. **¿Diferencia de género en el uso del cinturón de seguridad?** De 2200 hombres, pasajeros de automóvil, mayores de 8 años de edad, elegidos al azar, el 72% utiliza el cinturón de seguridad. De 2380 mujeres, pasajeras de automóvil, mayores de 8 años de edad, elegidas al azar, el 84% utiliza el cinturón de seguridad (según datos del Departamento de Transporte de Estados Unidos). Utilice un nivel de significancia de 0.05 para probar la aseveración de que ambos géneros tienen el mismo porcentaje de uso del cinturón de seguridad. Con base en el resultado, ¿parece que existe una diferencia entre géneros?
26. **Incidencia de radón.** El radón es un gas que se produce cuando el radio se degrada, y puede entrar a los hogares y convertirse en una amenaza para la salud. Se descubrió que, de 186 hogares en Hyde Park, Nueva York (lugar de origen de Franklin D. Roosevelt), el 16% tenía niveles inseguros de radón (por arriba de 4 picocurios por litro). Se encontró que, de 237 hogares en LaGrange, Nueva York (lugar de origen del autor), el 19% tenía niveles inseguros de radón (según datos del Departamento de Salud del estado de Nueva

York). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las dos regiones del condado Dutchess tienen distintos porcentajes de niveles inseguros de radón. ¿Parece que la presencia de niveles inseguros de radón varía según la región?

27. **Actitudes hacia el matrimonio.** En una encuesta del Time y CNN, el 24% de 205 mujeres solteras dijeron que “definitivamente deseaban casarse”. En la misma encuesta, el 27% de 260 hombres solteros dieron esta misma respuesta. Construya un estimado del intervalo de confianza del 99% para la diferencia entre la proporción de mujeres solteras y hombres solteros que definitivamente desean casarse. ¿Existe una diferencia de género al respecto?
28. **Actitudes hacia el matrimonio.** Remítase a los mismos datos muestrales del ejercicio anterior y utilice un nivel de significancia de 0.01 para probar la aseveración de que existe una diferencia entre la proporción de hombres y la proporción de mujeres que definitivamente desean casarse. Al parecer, ¿hay alguna diferencia?
29. **Conjunto de datos del apéndice B: Precipitación.** El conjunto de datos 10 del apéndice B lista cantidades de precipitación en Boston. Considere que un día tiene precipitación si la cantidad es cualquier valor positivo. Ponga a prueba la aseveración de que el porcentaje de días entre semana (de lunes a viernes) con precipitación es igual al porcentaje de días del fin de semana (sábado y domingo) con precipitación. Varios artículos de periódicos han reportado que llueve más durante los fines de semana. ¿Los datos de Boston sustentan esa aseveración?
30. **Conjunto de datos del apéndice B: Precipitación.** El conjunto de datos 8 del apéndice B lista cantidades de precipitación del lugar donde vive el autor en el condado Dutchess, Nueva York. El conjunto de datos 10 del apéndice B lista cantidades de precipitación en Boston. Considere que un día tiene precipitación si la cantidad es cualquier valor positivo. Pruebe la aseveración de que el porcentaje de días con precipitación en el condado Dutchess es igual al porcentaje de días con precipitación en Boston.
31. **Conjunto de datos del apéndice B: Alcohol y tabaco en películas infantiles.** Ponga a prueba la aseveración de que la proporción de 25, de un total de 50 películas infantiles elegidas al azar, que presentan algún consumo de alcohol es significativamente menor que la proporción muestral de 28, de otras 50 películas de este tipo que presentan algún consumo de tabaco. ¿Los resultados son válidos si se toman del conjunto de datos 5 del apéndice B?
32. **Conjunto de datos del apéndice B: Encuesta de salud.** Remítase al conjunto de datos 1 del apéndice B y utilice los datos muestrales para probar la aseveración de que la proporción de hombres mayores de 30 años es igual a la proporción de mujeres mayores de 30 años.

9-2 MÁS ALLÁ DE LO BÁSICO

33. **Interpretación del traslape de intervalos de confianza.** En el artículo “On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals”, de Schenkrt Juntoer y Gentleman (*American Statistician*, vol. 55, núm. 3), los autores consideran datos muestrales en esta afirmación: “Se han seleccionado muestras aleatorias simples independientes, cada una de tamaño 200; en la primera muestra 112 personas tienen el atributo, mientras que 88 personas en la segunda muestra tienen el atributo”.
 - a. Utilice los métodos de esta sección para construir un estimado del intervalo de confianza del 95% para la diferencia $p_1 - p_2$. ¿Qué sugiere el resultado acerca de la igualdad de p_1 y p_2 ?
 - b. Utilice los métodos de la sección 7-2 para construir estimados individuales del intervalo de confianza del 95% para cada una de las dos proporciones poblacionales. Después de comparar el traslape entre los dos intervalos de confianza, ¿qué concluye acerca de la igualdad de p_1 y p_2 ?

- c. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las dos proporciones poblacionales son iguales. ¿Qué concluye?
- d. Con base en los resultados anteriores, ¿qué concluye acerca de la igualdad de p_1 y p_2 ? ¿Cuál de los tres métodos anteriores es el menos efectivo para poner a prueba la igualdad de p_1 y p_2 ?
- 34. Equivalencia de prueba de hipótesis e intervalo de confianza.** Se obtienen dos muestras aleatorias simples a partir de dos poblaciones diferentes. La primera muestra consiste en 20 personas, 10 de las cuales tienen un atributo en común. La segunda muestra consiste en 2000 personas, 1404 de las cuales tienen el mismo atributo en común. Compare los resultados a partir de una prueba de hipótesis de $p_1 = p_2$ (con un nivel de significancia de 0.05) y un estimado del intervalo de confianza del 95% de $p_1 - p_2$.
- 35. Prueba para diferencia constante.** Para probar la hipótesis nula de que la diferencia entre dos proporciones poblacionales es igual a una constante c diferente de 0, utilice el estadístico de prueba

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - c}{\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}}$$

Siempre y cuando n_1 y n_2 sean grandes, la distribución muestral del estadístico de prueba z será aproximadamente la distribución normal estándar. Remítase al ejercicio 12 y utilice un nivel de significancia de 0.05 para probar la aseveración de que la tasa de filtración de virus de los guantes de vinilo es mayor por 50 puntos porcentuales que la filtración de virus de los guantes de látex.

- 36. Determinación de tamaño muestral.** El tamaño muestral necesario para estimar la diferencia entre dos proporciones poblacionales dentro de un margen de error E , con un nivel de confianza de $1 - \alpha$, se calcula como sigue. En la expresión

$$E = z_{\alpha/2} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

sustituya n_1 y n_2 por n (suponiendo que ambas muestras tienen el mismo tamaño) y sustituya p_1, q_1, p_2 y q_2 por 0.5 (puesto que sus valores no se conocen). Luego, despeje n .

Utilice este método para calcular el tamaño de cada muestra si usted quiere estimar la diferencia entre las proporciones de hombres y mujeres que tienen automóvil. Suponga que usted quiere tener una confianza del 95% de que su error no sea mayor de 0.03.

- 37. Verificación de la propiedad de varianzas.** Cuando se analizaron los fundamentos de los métodos de esta sección, se estableció que puesto que \hat{p}_1 y \hat{p}_2 se aproximan cada una a una distribución normal, $\hat{p}_1 - \hat{p}_2$ también se aproximarán a una distribución normal con media $p_1 - p_2$ y varianza $\sigma_{(\hat{p}_1 - \hat{p}_2)}^2 = \sigma_{\hat{p}_1}^2 + \sigma_{\hat{p}_2}^2$. Haga lo siguiente para verificar que la varianza de la *diferencia* entre dos variables aleatorias independientes es la *suma* de sus varianzas individuales.
- a. Suponiendo que se lanzan al aire dos monedas de 10 centavos de dólar, haga una lista del espacio muestral de 4 sucesos simples, luego calcule la proporción de caras en cada uno de los cuatro casos. Utilice la fórmula $\sigma^2 = \sum(x - \mu)^2/N$ para calcular la varianza para la población de las cuatro proporciones.
- b. Suponiendo que se lanzan dos monedas de un cuarto de dólar, el espacio muestral y la varianza serán las mismas que en el inciso a). Haga una lista de 16 diferencias en las proporciones ($\hat{p}_D - \hat{p}_Q$) que son posibles cuando cada resultado de las dos monedas de 10 centavos de dólar se aparea con cada posible resultado de las dos monedas de un cuarto de dólar. Calcule la varianza de σ^2 de la población de las 16 diferencias en las proporciones.
- c. Utilice los resultados anteriores para verificar que la *diferencia* entre dos variables aleatorias independientes es la *suma* de sus varianzas individuales.

9-3 Inferencias acerca de dos medias: Muestras independientes

Concepto clave En esta sección se presentan métodos para utilizar datos muestrales de dos muestras independientes para someter a prueba hipótesis acerca de dos medias poblacionales o para construir estimados de intervalos de confianza de la diferencia entre dos medias poblacionales. En la parte 1 de esta sección se incluyen situaciones en las que se desconocen las desviaciones estándar de las dos poblaciones, entre las cuales no se supone igualdad. La parte 2 incluye otras dos situaciones: 1. se conocen las dos desviaciones estándar poblacionales; 2. se desconocen las dos desviaciones estándar poblacionales, pero se supone igualdad entre ellas. Por razones que se explicarán más adelante, es necesario poner especial atención a los métodos que se describen en la parte 1, ya que son los más realistas y tienen un mejor desempeño general.

Parte 1: Muestras independientes con σ_1 y σ_2 desconocidas y sin suposición de igualdad

Comencemos con definiciones formales que distinguen entre muestras *independientes* y *dependientes*.

Definiciones

Dos muestras son **independientes** si los valores muestrales seleccionados de una población no están relacionados, apareados o asociados de alguna manera con los valores muestrales seleccionados de la otra población.

Dos muestras son **dependientes** (o consisten en **datos apareados**) si los miembros de una muestra se pueden usar para determinar los miembros de la otra muestra. [Las muestras que consisten en datos apareados (por ejemplo, datos de esposo/esposa) son dependientes. Además de los datos muestrales apareados, la dependencia también puede ocurrir con muestras relacionadas a través de asociaciones tales como miembros de la familia]. (En este libro usaremos el término *datos apareados*, ya que describe mejor la naturaleza de los datos).

EJEMPLO Prueba de fármaco

Muestras independientes: Se trata a un grupo de sujetos con el fármaco reductor del colesterol Lipitor, mientras que a un segundo grupo separado de sujetos se les da un placebo. Estos dos grupos muestrales son independientes, puesto que los individuos en el grupo de tratamiento no están en ninguna forma apareados o asociados con miembros correspondientes en el grupo placebo.

Datos apareados (o muestras dependientes): Se prueba la eficacia de una dieta utilizando los pesos de los sujetos medidos antes y después de someterse a la dieta. Cada valor “antes” se aparea con el valor “después”, puesto que cada par de mediciones antes/después proviene de la misma persona.

Esta sección considera dos muestras independientes, y la siguiente sección se enfoca en datos apareados. Cuando se utilizan dos muestras independientes para probar una aseveración acerca de la diferencia $\mu_1 - \mu_2$, o para construir un estimado del intervalo de confianza de $\mu_1 - \mu_2$, utilice el siguiente cuadro.

Requisitos

1. σ_1 y σ_2 se desconocen y no se hace una suposición sobre la igualdad de σ_1 y σ_2 .
2. Las dos muestras son *independientes*.
3. Ambas muestras son *aleatorias simples*.
4. Cualquiera o ambas de estas condiciones se satisfacen: los dos tamaños muestrales son grandes (con $n_1 > 30$ y $n_2 > 30$) o ambas muestras provienen de poblaciones que tienen distribuciones normales. (En muestras pequeñas, el requisito de normalidad es menos estricto, en el sentido de que los procedimientos se comportan bien siempre y cuando no existan valores extremos ni grandes sesgos).

Estadístico de prueba de hipótesis para dos medias: muestras independientes

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Grados de libertad: Cuando calcule valores críticos o valores P , utilice lo siguiente para determinar el número de grados de libertad, denotados por gl. (Si bien estos dos métodos generalmente dan por resultado números diferentes de grados de libertad, la conclusión de una prueba de hipótesis rara vez se ve afectada por la elección del método).

1. En este libro utilizamos el estimado sencillo y conservador: gl = el más pequeño de $n_1 - 1$ y $n_2 - 1$.
2. Los programas de cómputo de estadística por lo regular utilizan el estimado más exacto, pero más difícil, dado en la fórmula 9-1. (Nosotros no utilizaremos la fórmula 9-1 en los ejemplos y ejercicios de este libro).

Fórmula 9-1

$$gl = \frac{(A + B)^2}{\frac{A^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}}$$

donde $A = \frac{s_1^2}{n_1}$ y $B = \frac{s_2^2}{n_2}$

Valores P : Remítase a la tabla A-3. Utilice el procedimiento resumido en la figura 8-6.

Valores críticos: Remítase a la tabla A-3.

Estimado del intervalo de confianza de $\mu_1 - \mu_2$: muestras independientes

El estimado del intervalo de confianza de la diferencia $\mu_1 - \mu_2$ es

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

donde $E = t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

y el número de grados de libertad gl se obtiene como se describió antes para las pruebas de hipótesis. (En este libro utilizamos gl = el menor de $n_1 - 1$ y $n_2 - 1$).

Puesto que la prueba de hipótesis y el intervalo de confianza utilizan la misma distribución y el mismo error estándar, son equivalentes en el sentido de que dan como resultado las mismas conclusiones. En consecuencia, la hipótesis nula de $\mu_1 = \mu_2$ (o $\mu_1 - \mu_2 = 0$) puede probarse determinando si el intervalo de confianza incluye a 0. Para pruebas de hipótesis de dos colas con un nivel de significancia de 0.05, utilice un intervalo de confianza del 95%. Para una prueba de una cola, con un nivel de significancia de 0.05, utilice un intervalo de confianza del 90%. (Véase la tabla 8-2 para casos comunes).

Exploración de los conjuntos de datos

Debemos verificar los supuestos requeridos cuando utilizamos dos muestras independientes para hacer inferencias acerca de dos medias poblacionales. En vez de realizar de inmediato una prueba de hipótesis o construir un intervalo de confianza, primero debemos *explorar* las dos muestras utilizando los métodos descritos en los capítulos 2 y 3. Para cada una de las dos muestras, debemos investigar las medidas de tendencia central, la variación, la distribución, los valores extremos y si la población parece cambiar con el tiempo (CVDVT). Podría ser muy útil hacer lo siguiente:

- Calcule estadísticos descriptivos para ambos conjuntos de datos, incluyendo n , \bar{x} y s .
- Genere gráficas de cuadro de ambos conjuntos de datos, hechas en la misma escala para que puedan compararse.
- Genere histogramas de ambos conjuntos de datos, de manera que las distribuciones puedan compararse.
- Identifique cualquier valor extremo.

EJEMPLO Discriminación por edad Los Revenue Commissioners de Irlanda realizaron un concurso de promoción. A continuación se muestran las edades de los solicitantes que tuvieron éxito y de los que no tuvieron éxito (según datos de “Debating the Use of Statistical Evidence in Allegations of Age Discrimination”, de Barry y Boland, *American Statistician*, vol. 58, núm. 2). Algunos de los solicitantes que no tuvieron éxito para obtener la promoción se quejaron de que hubo discriminación por edad en la competencia. Maneje los datos como muestras de poblaciones más grandes y utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que los solicitantes sin éxito provienen de una población con una edad media mayor que la de los solicitantes exitosos. Con base en el resultado, ¿parece haber discriminación por la edad?

Edades de solicitantes sin éxito	Edades de solicitantes con éxito
34 37 37 38 41 42 43 44 44 45	27 33 36 37 38 38 39 42 42 43
45 45 46 48 49 53 53 54 54 55	43 44 44 44 45 45 45 45 46 46
56 57 60	47 47 48 48 49 49 51 51 52 54

SOLUCIÓN

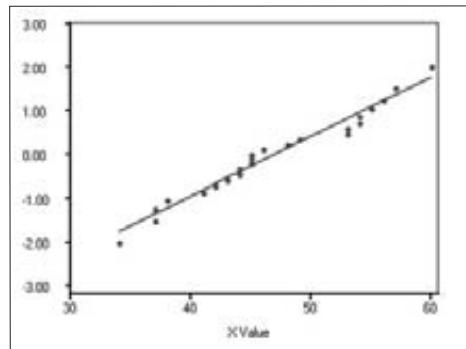
REQUISITO ✓ Los valores de las dos desviaciones estándar poblacionales se desconocen, y no estamos haciendo una suposición de igualdad entre ellas. Las dos muestras son independientes porque los valores de una muestra no están apareados con valores de la otra muestra. Podemos suponer que las muestras son aleatorias simples. Ambas muestras son pequeñas, por lo que

continúa

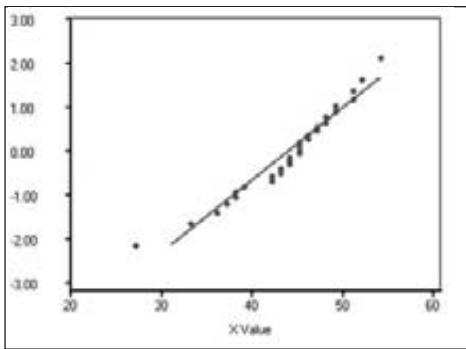
debemos verificar que cada muestra provenga de una población con una distribución normal. Al parecer no existen valores extremos. Las gráficas de cuadro que se presentan abajo sugieren que los datos se distribuyen normalmente. Asimismo, las gráficas cuantilares normales muestran patrones de puntos que se acercan razonablemente a una línea recta, sin presentar ningún otro patrón sistemático. Con base en estas imágenes, podemos concluir con seguridad que las distribuciones no se alejan mucho de una distribución normal. Los requisitos se satisfacen y podemos continuar con la prueba de hipótesis. ✓

Sin éxito	Con éxito	Sin éxito	Con éxito
$n = 23$	$n = 30$		2 7
$\bar{x} = 47.0$	$\bar{x} = 43.9$	4	3 3
$s = 7.2$	$s = 5.9$	877	3 67889
		44321	4 2233444
		986555	4 555566778899
		4433	5 1124
		765	5
		0	6

STATDISK



STATDISK



Una vez que verificamos que los requisitos se satisfacen, procedemos con la prueba de hipótesis; utilizaremos el método tradicional que se resume en la figura 8-9.

Paso 1: La aseveración de que en los solicitantes sin éxito tienen una edad media mayor que la edad media de los solicitantes con éxito se expresa simbólicamente como $\mu_1 > \mu_2$.

Paso 2: Si la aseveración original es falsa, entonces $\mu_1 \leq \mu_2$.

Paso 3: La hipótesis alternativa es la expresión que no contiene igualdad, y la hipótesis nula es una expresión de igualdad, de manera que tenemos

$$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 > \mu_2 \quad (\text{aseveración original})$$

Ahora procedemos con la suposición de que $\mu_1 = \mu_2$ o $\mu_1 - \mu_2 = 0$.

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Puesto que tenemos dos muestras independientes y estamos probando una aseveración acerca de dos medias poblacionales, utilizamos una distribución t con el estadístico de prueba dado antes en esta sección.

Paso 6: El estadístico de prueba se calcula como sigue:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(47.0 - 43.9) - 0}{\sqrt{\frac{7.2^2}{23} + \frac{5.9^2}{30}}} = 1.678$$

Como estamos utilizando una distribución t , los valores críticos de $t = 1.717$ se encuentran en la tabla A-3. (Con una área de 0.05 en la cola derecha, buscamos el valor t correspondiente a 22 grados de libertad, que es el más pequeño de $n_1 - 1$ y $n_2 - 1$ [o el más pequeño de 22 y 29]). El estadístico de prueba, el valor crítico y la región crítica se muestran en la figura 9-2.

Si usamos STATDISK, Minitab, Excel o una calculadora TI-83/84 Plus, también encontramos que el valor P es 0.0548 (con base en $gl = 41.868$).

Paso 7: Puesto que el estadístico de prueba no se ubica dentro de la región crítica, no se rechaza la hipótesis nula $\mu_1 = \mu_2$ (o $\mu_1 - \mu_2 = 0$).

INTERPRETACIÓN No existe evidencia suficiente para sustentar la aseveración de que la edad media de los solicitantes sin éxito es mayor que la edad media de los solicitantes con éxito. Con base en esta prueba de hipótesis, no parece existir discriminación por edad. Sin embargo, la prueba de hipótesis toma en cuenta una

continúa

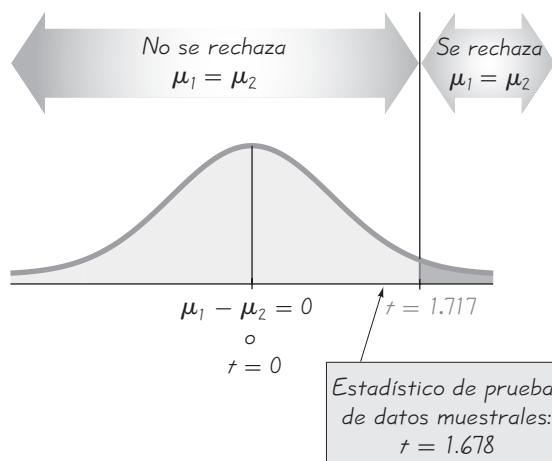
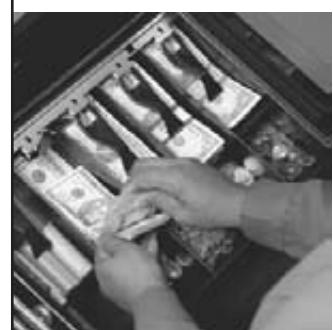


Figura 9-2 Prueba de la aseveración de discriminación por edad



Uso de la estadística para identificar ladrones

Los métodos de la estadística resultan útiles para determinar si un empleado está robando y también para estimar la cantidad robada. Los siguientes son algunos de los indicadores que se han utilizado. Para períodos de tiempo comparables, las muestras de ventas tienen medias que son significativamente diferentes. La cantidad media de ventas decrece de manera significativa. Existe un incremento significativo en la proporción de registros de "no venta" de las aperturas de caja. Existe una disminución significativa en la proporción de la recepción de efectivo y la de cheques. Se pueden aplicar los métodos de prueba de hipótesis para identificar indicadores como éstos. (Véase "How to Catch a Thief", de Manly y Thomson, *Chance*, vol. 11, núm. 4).



Costosa píldora de dieta

Existen muchos ejemplos pasados en los que se comercializaron tratamientos sin eficacia para obtener ganancias sustanciales. Las cápsulas de "Fat Trapper" y "Exercise in a Bottle", fabricadas por la compañía Enforma Natural Products, se anunciaron como si fueran tratamientos efectivos para la reducción de peso. Los anuncios afirmaban que después de tomar las cápsulas, la grasa sería bloqueada y las calorías serían quemadas, aun sin realizar ejercicio. Puesto que la Federal Trade Commission identificó aseveraciones que parecían no tener fundamento, se multó a la compañía con \$10 millones por publicidad engañosa.

La eficacia de tratamientos como éstos puede determinarse con experimentos en los cuales a un grupo de sujetos seleccionados al azar se les da el tratamiento, mientras que a otro grupo de sujetos seleccionados al azar se les da un placebo. Las pérdidas de peso resultantes se comparan utilizando métodos estadísticos, como los descritos en esta sección.

faceta de la situación general, y existen otros factores que podríamos considerar. Por ejemplo, podríamos comparar la proporción de solicitantes mayores de 50 años que tuvieron éxito con la proporción de solicitantes menores de 50 años que tuvieron éxito. Dos especialistas en estadística presentaron evidencias en el juicio y llegaron a conclusiones diferentes. La corte determinó que había "un vínculo entre la edad de los candidatos y su éxito o fracaso en la competencia".

EJEMPLO Intervalo de confianza para las edades de solicitantes de un ascenso Con los datos muestrales del ejemplo anterior, construya un estimado del intervalo de confianza del 90% de la diferencia entre la edad media de los solicitantes sin éxito y la edad media de los solicitantes con éxito. (Recuerde, una prueba de hipótesis unilateral, con un nivel de significancia de 0.05, se puede realizar utilizando un intervalo de confianza del 90%).

SOLUCIÓN

REQUISITO ✓ Puesto que estamos utilizando los mismos datos del ejemplo anterior, se aplica la misma verificación de requisitos. Los requisitos se satisfacen y procedemos a construir el intervalo de confianza. ✓

Primero calculamos el valor del margen de error E . Utilizamos $t_{\alpha/2} = 1.717$, que se encuentra en la tabla A-3 como la puntuación t que corresponde a una área de 0.10 en dos colas y $gl = 22$. (Como en el ejemplo anterior, buscamos la puntuación t correspondiente a 22 grados de libertad, que es la más pequeña de $n_1 - 1$ y $n_2 - 1$ [o el más pequeño de 22 y 29]).

$$E = t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1.717 \sqrt{\frac{7.2^2}{23} + \frac{5.9^2}{30}} = 3.2$$

Ahora calculamos el intervalo de confianza deseado como sigue:

$$\begin{aligned} (\bar{x}_1 - \bar{x}_2) - E &< (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E \\ (47.0 - 43.9) - 3.2 &< (\mu_1 - \mu_2) < (47.0 - 43.9) + 3.2 \\ -0.1 &< (\mu_1 - \mu_2) < 6.3 \end{aligned}$$

Si utilizamos programas de cómputo o la calculadora TI-83/84 Plus para obtener resultados más precisos, obtenemos el intervalo de confianza de $-0.1 < (\mu_1 - \mu_2) < 6.3$, y vemos que el intervalo de confianza de arriba es bastante adecuado.

INTERPRETACIÓN Tenemos una confianza del 90% de que los límites de -0.1 años y 6.3 años realmente contienen la diferencia entre las dos medias poblacionales. Puesto que esos límites contienen a 0, este intervalo de confianza sugiere que es muy posible que las medias de las dos poblaciones sean iguales. No existe una diferencia significativa entre las dos medias.

Fundamentos: ¿Por qué el estadístico de prueba y el intervalo de confianza tienen las formas particulares que hemos presentado? Si los supuestos dados se satisfacen, la distribución muestral de $\bar{x}_1 - \bar{x}_2$ puede aproximarse por medio de una distribución t , con media igual a $\mu_1 - \mu_2$ y desviación estándar igual a $\sqrt{s_1^2/n_1 + s_2^2/n_2}$. Esta última expresión para la desviación estándar se basa en la propiedad de que la varianza de las diferencias entre dos variables

aleatorias independientes es igual a la varianza de la primera variable aleatoria más la varianza de la segunda variable aleatoria. Es decir, la varianza de los valores muestrales $\bar{x}_1 - \bar{x}_2$ tiende a igualar $s_1^2/n_1 + s_2^2/n_2$, siempre y cuando \bar{x}_1 y \bar{x}_2 sean independientes. (Véase el ejercicio 35).

Parte 2: Métodos alternativos

La parte 1 de esta sección se refirió a situaciones en las que no se conocen las dos desviaciones estándar poblacionales, entre las cuales no se supone igualdad. En la parte 2, consideramos otras dos situaciones: 1. se conocen las dos desviaciones estándar poblacionales; 2. se desconocen las dos desviaciones estándar poblacionales, pero se supone igualdad entre ellas. Ahora describiremos los procedimientos para estos casos alternativos.

Método alternativo: σ_1 y σ_2 conocidas

En realidad, las desviaciones estándar poblacionales σ_1 y σ_2 casi nunca se conocen, pero si son conocidas, el estadístico de prueba y el intervalo de confianza están basados en una distribución normal y no en una distribución *t*. Veamos el siguiente resumen.

Requisitos

1. Se conocen las dos desviaciones estándar poblacionales.
2. Las dos muestras son *independientes*.
3. Ambas muestras son *aleatorias simples*.
4. Cualquiera de estas condiciones (o ambas) se satisfacen: los dos tamaños muestrales son *grandes* (con $n_1 > 30$ y $n_2 > 30$) o las dos muestras provienen de poblaciones que tienen distribuciones normales. (Para muestras pequeñas, el requisito de normalidad es menos estricto en el sentido de que los procedimientos funcionan bien siempre y cuando no existan valores extremos ni desviaciones de la normalidad demasiado pronunciadas).

Prueba de hipótesis para dos medias: muestras independientes y $\sigma_1 = \sigma_2$ conocidas

$$\text{Estadístico de prueba: } z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Valor *P* y valores críticos: Remítase a la tabla A-2.

Estimado del intervalo de confianza de $\mu_1 - \mu_2$: muestras independientes y $\sigma_1 = \sigma_2$ conocidas

Intervalo de confianza: $(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$

$$\text{donde } E = z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$



Súper Bowl

Se invitó a un grupo de estudiantes a un juego del Súper Bowl, y a la mitad de ellos se les dieron tazones grandes, con capacidad de 4 litros, con botanas, mientras que la otra mitad recibió tazones más pequeños, de 2 litros. Los que recibieron los tazones grandes consumieron un 56% más que los que utilizaron los tazones pequeños. (Véase “Super Bowls: Serving Bowl Size and Food Consumption”, de Wansink y Cheney, *Journal of the American Medical Association*, vol. 293, núm. 14).

Otro estudio demostró que existe “un incremento significativo de los accidentes automovilísticos fatales las horas posteriores a la transmisión del Súper Bowl en Estados Unidos”. Los investigadores analizaron 20,377 muertes en 27 domingos de Súper Bowl y en otros 54 domingos que se utilizaron como controles. Encotraron un incremento del 41% en las muertes después de los juegos del Súper Bowl. (Véase “Do Fatal Crashes Increase Following A Super Bowl Telecast?”, de Redelmeier y Stewart, *Chance*, vol. 18, núm. 1).

Inferencias acerca de dos medias independientes

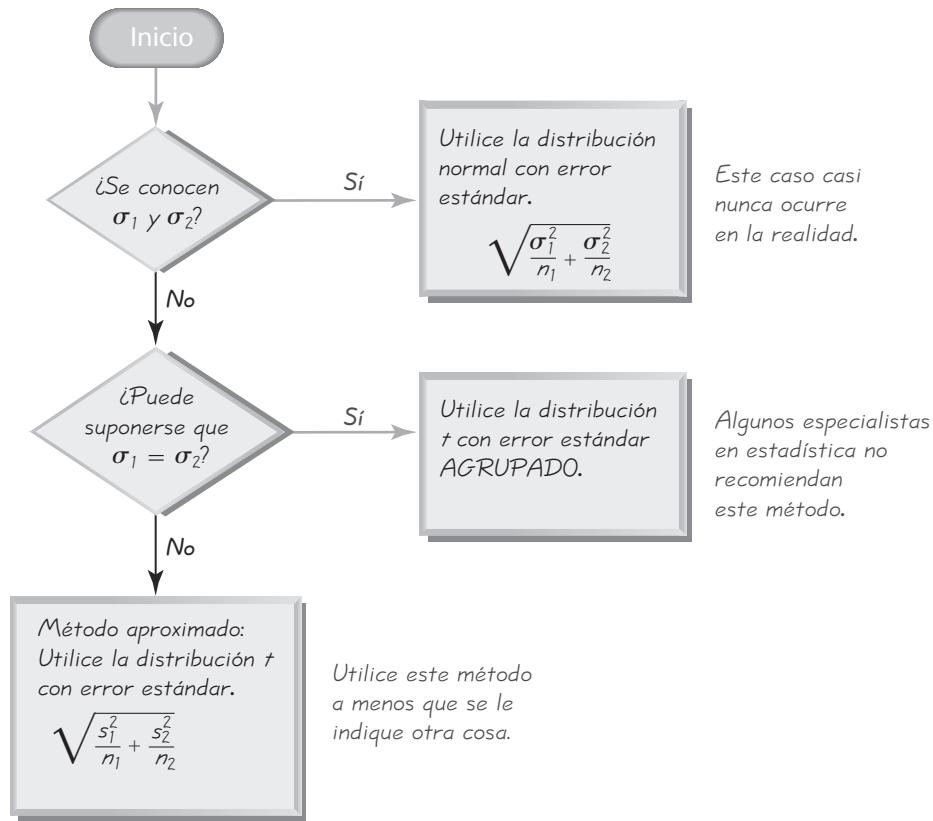


Figura 9-3 Métodos para inferencias acerca de dos medias independientes

Un método alternativo (que no se utiliza en este libro) consiste en usar las expresiones de arriba si se desconocen σ_1 y σ_2 , pero si ambas muestras son grandes (con $n_1 > 30$ y $n_2 > 30$). Este método alternativo se usa con σ_1 reemplazada por s_1 y σ_2 reemplazada por s_2 . Puesto que σ_1 y σ_2 en la realidad rara vez se conocen, este libro no utilizará tal método alternativo. Vea la figura 9-3, donde se describe procedimiento que se utiliza en este texto.

Método alternativo: se supone que $\sigma_1 = \sigma_2$ y se agrupan las varianzas muestrales

Aun cuando los valores específicos de σ_1 y σ_2 no se conozcan, si se puede suponer que tienen el *mismo* valor, las varianzas muestrales s_1^2 y s_2^2 pueden *agruparse* para obtener un estimado de la varianza poblacional σ^2 común. El **estimado agrupado de σ^2** se denota por s_p^2 y es un promedio ponderado de s_1^2 y s_2^2 , que se incluye en el siguiente cuadro.

Requisitos

1. Se desconocen las dos desviaciones estándar poblacionales, pero se supone que son iguales. Es decir, $\sigma_1 = \sigma_2$.
2. Las dos muestras son *independientes*.
3. Ambas muestras son *aleatorias simples*.
4. Cualquiera de estas condiciones (o ambas) se satisfacen: los dos tamaños muestrales son *grandes* (con $n_1 > 30$ y $n_2 > 30$) o las dos muestras provienen de poblaciones que tienen distribuciones normales. (Para muestras pequeñas, el requisito de normalidad es menos estricto en el sentido de que los procedimientos funcionan bien siempre y cuando no existan valores extremos ni desviaciones de la normalidad demasiado pronunciadas).

Estadístico de prueba de hipótesis para dos medias: muestras independientes y $\sigma_1 = \sigma_2$

Estadístico de prueba: $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$

donde $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$ (Varianza agrupada)

y el número de grados de libertad está dado por $gl = n_1 + n_2 - 2$.

Estimado del intervalo de confianza de $\mu_1 - \mu_2$: muestras independientes y $\sigma_1 = \sigma_2$

Intervalo de confianza: $(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$

donde $E = t_{\alpha/2} \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$ y s_p^2 es como se expresó en el estadístico de prueba anterior; el número de grados de libertad está dado por $gl = n_1 + n_2 - 2$.

Si queremos utilizar este método, ¿cómo determinamos que $\sigma_1 = \sigma_2$? Un enfoque consiste en utilizar una prueba de la hipótesis nula $\sigma_1 = \sigma_2$, como en la sección 9-5, pero este enfoque no se recomienda y, en este libro, no utilizaremos la prueba preliminar de $\sigma_1 = \sigma_2$. En el artículo “Homogeneity of Variance in the Two-Sample Means Test” (de Moser y Stevens, *American Statistician*, vol. 46, núm. 1), los autores señalan que rara vez se sabe que $\sigma_1 = \sigma_2$. Ellos analizan el desempeño de las diferentes pruebas considerando los tamaños muestrales y la potencia de las pruebas; concluyen que debe dedicarse más esfuerzo al aprendizaje del método presentado en la parte 1 y que debe ponerse menos énfasis en el método basado en el supuesto de que $\sigma_1 = \sigma_2$. Si no se indica otra cosa, utilizamos la siguiente estrategia, que es consistente con las recomendaciones del artículo de Moser y Stevens:

Suponga que se desconocen σ_1 y σ_2 , no suponga que $\sigma_1 = \sigma_2$, y utilice el estadístico de prueba y el intervalo de confianza dados en la parte 1 de esta sección (véase la figura 9-3).



Mejores resultados con grupos más pequeños

Un experimento realizado en la Universidad Estatal de Nueva York, en Stony Brook, reveló que los estudiantes tenían mejores resultados en clases limitadas a 35 estudiantes, que en clases grandes donde hay entre 150 y 200 alumnos. En un curso de cálculo, los porcentajes de reprobación fueron del 19% en las clases pequeñas, en comparación con un 50% en las clases grandes. Los porcentajes de calificación A (equivalente a 10) fueron del 24% para las clases pequeñas y del 3% para las clases grandes. Estos resultados sugieren que los estudiantes se benefician de los grupos más pequeños, los cuales permiten una interacción más directa entre los alumnos y los maestros.

Uso de la tecnología

STATDISK Seleccione el elemento **Analysis** del menú, luego elija **Hypothesis Testing o Confidence Intervals**; luego seleccione **Mean-Two Independent Samples**. Ingrese los valores requeridos en el cuadro de diálogo. Usted tiene la opción de “Not Eq vars: NO POOL”, “Eq vars: POOL” o “Prelim F Test”. Se recomienda la opción de **Not Eq vars: NO POOL**. (La prueba *F* se describe en la sección 9-5).

MINITAB Comenzando con Release 14, Minitab ahora permite el uso del resumen de estadísticos o de listas de datos muestrales originales. Si se conocen los valores muestrales originales, ingréselos en las columnas C1 y C2. Seleccione las opciones **Stat, Basic Statistics y 2-Sample t**. Ingrese

la información requerida en la ventana que se abre. Utilice el botón **Options** para elegir un nivel de confianza, ingrese un valor aseverado de la diferencia o seleccione un formato para la hipótesis alternativa. El resultado de Minitab también incluye los límites del intervalo de confianza.

Si las dos varianzas poblacionales parecen ser iguales, Minitab permite utilizar un estimado agrupado de la varianza común. Aparecerá un cuadro junto a **Assume equal variances**; haga clic en este cuadro sólo si usted desea suponer que las dos poblaciones tienen varianzas iguales. Este método no se recomienda.

EXCEL Ingrese los datos para las dos muestras en las columnas A y B.

Para utilizar el complemento del programa Data Desk XL, haga clic en **DDXL**. Seleccione **Hypothesis Tests y 2 Var t Test** o seleccione **Confidence Intervals y 2 Var t Interval**. En el cuadro de diálogo, haga clic en el ícono del lápiz para la primera columna cuantitativa e ingrese el rango de valores para la primera muestra, como por ejemplo A1:A23. Haga clic en el ícono del lápiz para la segunda columna cuantitativa e ingrese el rango de valores para la segunda muestra. Haga clic en **OK**. Ahora complete el nuevo cuadro de diálogo siguiendo los pasos indicados. En el paso 1, seleccione **2-Sample** para la suposición de varianzas poblacionales desiguales. (También

puede seleccionar **Pooled** para la suposición de varianzas poblacionales iguales, pero no se recomienda este método).

Para utilizar el complemento del programa, Data Analysis de Excel, haga clic en **Tools** y seleccione **Data Analysis**. Seleccione uno de los siguientes dos elementos (recomendamos la suposición de varianzas *desiguales*):

prueba *t*: dos muestras suponiendo variables iguales

prueba *t*: dos muestras suponiendo varianzas desiguales

Proceda a ingresar el rango de valores de la primera muestra (como A1:A23) y después el rango de valores para la segunda muestra. Introduzca un valor para la diferencia aseverada entre las dos medias poblacionales, que con frecuencia será de 0. Ingrese el nivel de significancia en el cuadro Alpha y haga clic en **OK**. (Excel no proporciona un intervalo de confianza).

TI-83/84 PLUS Para realizar pruebas como las que se encuentran en esta sección, oprima **STAT**, luego seleccione **TESTS** y elija **2-SampTTest** (para una prueba de hipótesis) o **2-SampTInt** (para un intervalo de confianza). La calculadora TI-83/84 Plus le da la opción de utilizar varianzas “agrupadas” (si usted cree que $\sigma_1^2 = \sigma_2^2$) o de no agruparlas, pero nosotros recomendamos que no se agrupen.

9-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Requisitos.** Consulte los datos de las edades utilizadas en los ejemplos de la prueba de hipótesis y el intervalo de confianza de esta sección. Si incluimos otra edad de 80 años para un solicitante sin éxito, ¿se satisfacen los requisitos de la parte 1? ¿Por qué?
- 2. Intervalo de confianza para prueba de hipótesis.** Usted planea construir un intervalo de confianza que utilizará para probar la aseveración de que una población tiene una media mayor que la otra. Si la prueba se hará con un nivel de significancia de 0.01, ¿qué nivel de confianza debe usar para el intervalo de confianza? ¿Qué nivel de confianza debe utilizar para poner a prueba la aseveración de que las dos poblaciones tienen medias diferentes (una vez más con un nivel de significancia de 0.01)?
- 3. Grados de libertad.** En el ejemplo de la prueba de hipótesis de esta sección, el valor crítico $t = 1.717$ se obtuvo utilizando $gl =$ más pequeño de $n_1 - 1$ y $n_2 - 1$. Con tamaños muestrales de 23 y 30, usamos $gl = 22$. Si calculamos gl con la fórmula 9-1, obtenemos $gl = 41.868$, y el valor crítico correspondiente es 1.682. ¿En qué sentido el uso de un valor crítico de $t = 1.717$ es “más conservador” que el uso de valor crítico de 1.682?



- 4. Datos antes y después.** Una muestra aleatoria simple de sujetos es tratada con un fármaco que se supone debe disminuir sus niveles de colesterol. Para cada sujeto, el colesterol se mide una ocasión antes del tratamiento y una ocasión después del tratamiento. ¿Podemos utilizar los métodos de esta sección para probar la aseveración de que los niveles de colesterol “antes” tienen una media mayor que los niveles “después”? ¿Por qué?

Muestras independientes y datos apareados. En los ejercicios 5 a 8, determine si las muestras son independientes o si consisten en datos apareados.

5. La eficacia del Prilosec para tratar la acidez estomacal se prueba midiendo la secreción de ácido gástrico en un grupo de pacientes tratados con Prilosec y otro grupo de pacientes a quienes se da un placebo.
6. La eficacia de Prilosec para tratar la acidez estomacal se prueba midiendo la secreción de ácido gástrico en pacientes antes y después del tratamiento con el fármaco. Los datos consisten en mediciones antes y después para cada paciente.
7. En un experimento se prueba la eficacia de la dieta de Weight Watchers y se registra el peso de cada sujeto antes y después de la dieta.
8. La eficacia de una vacuna para la gripe se pone a prueba al tratar a un grupo de sujetos con la vacuna y a otro grupo con un placebo.

En los ejercicios 9 a 28, suponga que las dos muestras son aleatorias simples independientes, seleccionadas de poblaciones distribuidas normalmente. No suponga que las desviaciones estándar poblacionales son iguales, a menos que su profesor indique otra cosa.

9. **Prueba de hipótesis sobre la eficacia de la equinácea.** En una prueba aleatorizada, doble ciego y controlada con placebo se puso a prueba la equinácea como tratamiento para las infecciones de vías respiratorias superiores en niños. Los “días con fiebre” fue uno de los criterios utilizados para medir los efectos. En 337 niños tratados con equinácea, el número medio de días con fiebre fue de 0.81, con una desviación estándar de 1.50 días. En 370 niños que recibieron un placebo, el número medio de días con fiebre fue de 0.64, con una desviación estándar de 1.16 días (según datos de “Efficacy and Safety of Echinacea in Treating Upper Respiratory Tract Infections in Children”, de Taylor *et al.*, *Journal of the American Medical Association*, vol. 290, núm. 21). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la equinácea afecta el número de días con fiebre. Con base en esos resultados, ¿parece que la equinacea es efectiva?
10. **Prueba de hipótesis sobre los efectos de la cocaína en niños.** Se realizó un estudio para evaluar los efectos de la exposición a la cocaína antes del nacimiento. Cuando los niños tenían 4 años de edad, se evaluó su habilidad para ensamblar objetos, la cual fue descrita como “una tarea que requiere de destrezas visoespaciales relacionadas con las habilidades matemáticas”. Los 190 hijos de consumidoras de cocaína tuvieron una media de 7.3 y una desviación estándar de 3.0. Los 186 niños que no estuvieron expuestos a la cocaína tuvieron una puntuación media de 8.2, con una desviación estándar de 3.0. (Los datos están basados en “Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure”, de Singer *et al.*, *Journal of the American Medical Association*, vol. 291, núm. 20). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la exposición prenatal a la cocaína está asociada con puntuaciones más bajas en niños de cuatro años en la prueba de ensamblaje de objetos.
11. **Intervalo de confianza para el efecto del peso al nacer sobre la puntuación de CI.** Al investigar la relación entre el peso al nacer y el CI, los investigadores encontraron que 258 sujetos con pesos sumamente bajos al nacer (menos de 1000 g), a los ocho años tuvieron puntuaciones de CI en la prueba Wechsler con una media de 95.5 y una desviación estándar de 16.0. En el caso de los 220 sujetos con pesos normales al nacer, la media a los ocho años fue de 104.9, con una desviación estándar de 14.1 (según datos de “Neurobehavioral Outcomes of School-age Children Born Extremely Low Birth Weight or Very Preterm in the 1900s”, de Anderson *et al.*, *Journal of the American Medical Association*,

vol. 289, núm. 24). Construya un estimado de un intervalo de confianza del 95% de la diferencia entre la puntuación media de CI de los niños de ocho años nacidos con bajo peso y la media de los niños de ocho años nacidos con peso normal. ¿Parece que la puntuación de CI puede verse afectada por el peso al nacer?

- 12. Intervalo de confianza para comparar dietas.** Un ensayo aleatorizado puso a prueba la eficacia de dietas en adultos. En 40 sujetos que utilizaron la dieta Weight Watchers, la pérdida media de peso después de un año fue de 3.0 lb, con una desviación estándar de 4.9 lb. En los 40 sujetos que utilizaron la dieta Atkins, la pérdida media de peso después de un año fue de 2.1 lb, con una desviación estándar de 4.8 lb. Construya un estimado de un intervalo de confianza del 95% de la diferencia entre las medias poblacionales. Al parecer, ¿hay una diferencia en la eficacia de las dos dietas? ¿La cantidad de peso perdida justifica cada dieta?

- 13. Prueba de hipótesis del efecto del consumo de marihuana en estudiantes universitarios.** Se han realizado muchos estudios para probar los efectos del consumo de marihuana en las capacidades mentales. En uno de esos estudios se probó la capacidad de recuperación de memoria en grupos de consumidores de marihuana ocasionales y frecuentes en la universidad, con los resultados que se dan abajo (según datos de “The Residual Cognitive Effects of Heavy Marijuana Use in College Students”, de Pope y Yurgelun-Todd, *Journal of the American Medical Association*, vol. 275, núm. 7). Utilice un nivel de significancia de 0.01 para probar la aseveración de que la población de consumidores frecuentes de marihuana tiene una media más baja que la de los consumidores ocasionales. ¿Debería preocupar el consumo de marihuana a los estudiantes universitarios?

Artículos ordenados correctamente por consumidores ocasionales de marihuana:

$$n = 64, \bar{x} = 53.3, s = 3.6$$

Artículos ordenados correctamente por consumidores frecuentes de marihuana:

$$n = 65, \bar{x} = 51.3, s = 4.5$$

- 14. Intervalo de confianza del efecto del consumo de marihuana en estudiantes universitarios.** Remítase a los datos muestrales utilizados en el ejercicio 13 y construya un intervalo de confianza del 98% para la diferencia entre las dos medias poblacionales. ¿El intervalo de confianza incluye a 0? ¿Qué sugiere el intervalo de confianza acerca de la igualdad de las dos medias poblacionales?

- 15. Intervalo de confianza para tratamiento de depresión bipolar.** En experimentos clínicos que implican diferentes grupos de muestras independientes, es importante que los grupos sean similares en los aspectos importantes que afectan el experimento. En un experimento diseñado para probar la eficacia de la paroxetina en el tratamiento de la depresión bipolar, se midió la depresión de los sujetos utilizando la escala de Hamilton, con los resultados que se presentan abajo (según datos de “Double-Blind, Placebo-Controlled Comparison of Imipramine and Paroxetine in the Treatment of Bipolar Depression”, de Nemeroff *et al.*, *American Journal of Psychiatry*, vol. 158, núm. 6). Construya un intervalo de confianza del 95% para la diferencia entre las dos medias poblacionales. Con base en los resultados, ¿parece que las dos poblaciones tienen medias diferentes? ¿Debería recomendarse la paroxetina como un tratamiento para la depresión bipolar?

Grupo placebo: $n = 43, \bar{x} = 21.57, s = 3.87$

Grupo de tratamiento con paroxetina: $n = 33, \bar{x} = 20.38, s = 3.91$

- 16. Prueba de hipótesis para tratamiento de depresión bipolar.** Remítase a los datos muestrales del ejercicio 15 y utilice un nivel de significancia de 0.05 para probar la aseveración de que el grupo de tratamiento y el grupo placebo provienen de poblaciones con la misma media. ¿Qué sugiere el resultado de la prueba de hipótesis acerca de la paroxetina como tratamiento para la depresión bipolar?

- 17. Prueba de hipótesis para tratamiento magnético del dolor.** La gente gasta enormes sumas de dinero (actualmente alrededor de \$5,000 millones al año) en la compra de magnetos que se utilizan para tratar una amplia variedad de dolores. Investigadores

realizaron un estudio para determinar si los magnetos son efectivos en el tratamiento del dolor de espalda. El dolor se midió utilizando la escala análoga visual y los resultados que se presentan abajo son algunos de los obtenidos en el estudio (según datos de “Bipolar Permanent Magnets for the Treatment of Chronic Lower Back Pain: A Pilot Study”, de Collacott, Zimmerman, White y Rindone, *Journal of the American Medical Association*, vol. 283, núm. 10). Utilice un nivel de significancia de 0.05 para probar la aseveración de que los sujetos tratados con magnetos tienen una mayor reducción del dolor que aquellos a quienes se dio un tratamiento simulado (similar a un placebo). ¿Parece que los magnetos sirven para tratar el dolor de espalda? ¿Es válido argumentar que los magnetos podrían parecer efectivos si los tamaños muestrales fueran mayores?

Reducción en el nivel del dolor después del tratamiento magnético:

$$n = 20, \bar{x} = 0.49, s = 0.96$$

Reducción en el nivel del dolor después del tratamiento simulado:

$$n = 20, \bar{x} = 0.44, s = 1.4$$

- 18. Intervalo de confianza para tratamiento magnético del dolor.** Remítase a los datos muestrales del ejercicio 17 y construya un estimado de un intervalo de confianza del 90% de la diferencia entre la reducción media del dolor para los sujetos tratados con magnetos y la reducción media del dolor para quienes recibieron un tratamiento simulado. Con base en el resultado, ¿parece que los magnetos sirven para reducir el dolor?

- 19. Prueba de hipótesis para identificar trastornos psiquiátricos.** ¿Están relacionados los trastornos psiquiátricos graves con factores biológicos observables médicalemente? Un estudio utilizó tomografía computarizada por rayos X (TC) para reunir datos de los volúmenes cerebrales de un grupo de pacientes con trastorno obsesivo-compulsivo y de un grupo control de personas sanas. Los resultados muestrales de los volúmenes (en mL) se presentan abajo (según datos de “Neuroanatomical Abnormalities in Obsessive-Compulsive Disorder Detected with Quantitative X-Ray Computed Tomography”, de Luxenberg *et al.*, *American Journal of Psychiatry*, vol. 145, núm. 9). Construya un estimado de un intervalo de confianza del 99% de la diferencia entre el volumen cerebral medio para el grupo control saludable y el volumen cerebral medio para el grupo obsesivo compulsivo. ¿Qué sugiere el intervalo de confianza acerca de la diferencia entre las dos medias poblacionales? Con base en este resultado, ¿parece que el trastorno obsesivo compulsivo tiene una base biológica?

$$\text{Grupo control: } n = 10, \bar{x} = 0.45, s = 0.08$$

$$\text{Pacientes obsesivo compulsivos: } n = 10, \bar{x} = 0.34, s = 0.08$$

- 20. Intervalo de confianza para identificar trastornos psiquiátricos.** Remítase a los datos muestrales del ejercicio 19 y utilice un nivel de significancia de 0.01 para probar la aseveración de que existe una diferencia entre las dos medias poblacionales. Con base en el resultado, ¿parece que el trastorno obsesivo-compulsivo tiene una base biológica?

- 21. Intervalo de confianza para efectos del alcohol.** Se realizó un experimento para probar los efectos del alcohol. Se registraron los *errores* en una prueba de destrezas visuales y motrices de un grupo de tratamiento de personas que bebieron etanol y de otro grupo al que se administró un placebo. Los resultados se muestran en la tabla adjunta (según datos de “Effects of Alcohol Intoxication on Risk Taking, Strategy, and Error Rate in Visuomotor Performance”, de Streufert *et al.*, *Journal of Applied Psychology*, vol. 77, núm. 4). Construya un estimado de un intervalo de confianza del 95% de la diferencia entre las dos medias poblacionales. ¿Sustentan los resultados la creencia común de que beber es peligroso para conductores, pilotos, capitanes de navíos, etcétera? ¿Por qué?

- 22. Prueba de hipótesis para efectos del alcohol.** Remítase a los datos muestrales del ejercicio 21 y utilice un nivel de significancia de 0.05 para probar la aseveración de que existe una diferencia entre el grupo de tratamiento y el grupo de control. Si existe una diferencia significativa, ¿podemos concluir que el tratamiento causa una disminución de las destrezas visuales y motrices?

Grupo de tratamiento	Grupo placebo
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 0.049$	$\bar{x}_2 = 0.000$
$s_1 = 0.015$	$s_2 = 0.000$

- 23. Prueba de hipótesis para la diferencia de valores de casas.** A continuación se muestran valores de mercado (en miles de dólares) de casas elegidas al azar en Long Beach Island y Nueva Jersey. Utilice un nivel de significancia de 0.05 para probar la aseveración de un agente inmobiliario de que las casas con vista al mar (ubicadas en la playa) valen más que las casas que están a un lado del océano, no directamente sobre la playa. Puesto que sólo hay cinco valores en cada muestra, ¿realmente podemos concluir que las casas con vista al mar valen más?

Vista al mar:	2199	3750	1725	2398	2799	
Junto al mar:	700	1355	795	1575	759	

- 24. Prueba de hipótesis para la diferencia de la antigüedad de automóviles y taxis.** Cuando el autor visitó Dublín, en Irlanda (lugar de origen de William Gosset, empleado de Guinness Brewery, creador de la distribución t), registró la antigüedad de automóviles y taxis elegidos al azar. (Cuando se viaja con el autor la diversión nunca termina). A continuación se listan las antigüedades (en años). Utilice un nivel de significancia de 0.05 para probar la aseveración de que hay una diferencia entre la antigüedad media de un automóvil de Dublín y la antigüedad media de un taxi de Dublín. Esperaríamos que los taxis fueran más nuevos, pero, ¿que sugieren los resultados?

Automóviles										Taxis									
4	0	8	11	14	3	4	4	3	5	8	8	0	3	8	4	3	3	6	11
8	3	3	7	4	6	6	1	8	2	7	7	6	9	5	10	8	4	3	4
15	11	4	1	6	1	8													

- 25. Alquitrán y cigarrillos.** Utilice los datos muestrales que se listan a continuación y con un nivel de significancia de 0.05 ponga a prueba la aseveración de que la cantidad media de alquitrán en cigarrillos largos con filtro es *menor que* la cantidad media de alquitrán en cigarrillos largos sin filtro. Todas las mediciones son en miligramos, y los datos provienen de la Federal Trade Commission.

Con filtro	16	15	16	14	16	1	16	18	10	14	12
	11	14	13	13	13	16	16	8	16	11	
Sin filtro	23	23	24	26	25	26	21	24			

- 26. Bloqueo en exámenes.** Muchos estudiantes han tenido la experiencia poco placentera de sentir pánico en los exámenes porque la primera pregunta era excepcionalmente difícil. Se estudió el efecto que tiene el orden de las preguntas de exámenes sobre la ansiedad. Las siguientes puntuaciones son mediciones de la “ansiedad debilitante por exámenes”, que la mayoría de nosotros llamamos pánico o bloqueo (según datos de “Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance”, de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4). ¿Existe suficiente evidencia para sustentar la aseveración de que las dos poblaciones de puntuaciones tienen la misma media? ¿Existe evidencia suficiente para sustentar la aseveración de que el orden de las preguntas de examen tiene un efecto en la calificación?

Preguntas ordenadas de fácil a difícil					Preguntas ordenadas de difícil a fácil			
24.64	39.29	16.32	32.83	28.02	33.62	34.02	26.63	30.26
33.31	20.60	21.13	26.69	28.90	35.91	26.68	29.49	35.32
26.43	24.23	7.10	32.86	21.06	27.24	32.34	29.34	33.53
28.89	28.71	31.73	30.02	21.96	27.62	42.91	30.20	32.54
25.49	38.81	27.85	30.29	30.72				

- 27. Conjunto de datos del apéndice B: pesos de monedas de 25 centavos.** Las máquinas expendedoras utilizan los pesos de las monedas de 25 centavos para detectar monedas

falsas. Remítase al conjunto de datos 14 del apéndice B y ponga a prueba la aseveración de que el peso medio de las monedas de plata de 25 centavos acuñadas antes de 1964 es igual al peso medio de las monedas de 25 centavos acuñadas después de esa fecha. Dados los tamaños muestrales relativamente pequeños, obtenidos de poblaciones grandes de millones de monedas, ¿realmente podemos concluir que los pesos medios son diferentes?

- 28. Conjunto de datos del apéndice B: pesos de Coca-Cola.** Remítase al conjunto de datos 12 del apéndice B para poner a prueba la aseveración de que, debido a que contienen la misma cantidad de bebida de cola, los pesos de latas de Coca-Cola regular tienen la misma media que los pesos de latas de Coca-Cola dietética. Si existe una diferencia en los pesos medios, identifique la explicación más viable de esta diferencia.

Agrupamiento. En los ejercicios 29 a 32, suponga que dos muestras son aleatorias simples independientes, seleccionadas de poblaciones distribuidas normalmente. También suponga que las desviaciones estándar poblacionales son iguales ($\sigma_1 = \sigma_2$), de manera que el error estándar de la diferencia entre las medias se obtiene agrupando las varianzas muestrales.

- 29. Intervalo de confianza con agrupamiento.** Resuelva el ejercicio 12 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?
- 30. Prueba de hipótesis con agrupamiento.** Resuelva el ejercicio 13 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?
- 31. Prueba de hipótesis con agrupamiento.** Resuelva el ejercicio 17 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?
- 32. Intervalo de confianza con agrupamiento.** Resuelva el ejercicio 18 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?

9-3 MÁS ALLÁ DE LO BÁSICO

- 33. Efectos de un valor extremo.** Remítase al ejercicio 24 e incluya un valor extremo consistente en un automóvil de 50 años de antigüedad. ¿La prueba de hipótesis se ve muy afectada por la presencia del valor extremo?
- 34. Efectos de las unidades de medida.** ¿De qué manera se ven afectados los resultados del ejercicio 24, si todas las edades se convierten de años a meses? En general, ¿afecta la elección de la escala las conclusiones acerca de la igualdad de dos medias poblacionales? ¿Afecta esa elección al intervalo de confianza?
- 35. Verificación de una propiedad de las varianzas**
- Calcule la varianza para esta *población* de valores x : 5, 10, 15. (Vea la sección 3-3 para la varianza σ^2 de una población).
 - Calcule la varianza para esta *población* de valores y : 1, 2, 3.
 - Haga una lista de la *población* de todas las diferencias posibles $x - y$, y calcule la varianza de esta población.
 - Utilice los resultados de los incisos a), b) y c) para verificar que la varianza de las *diferencias* entre dos variables aleatorias independientes es la *suma* de sus varianzas individuales ($\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2$). (Este principio se utiliza para obtener el estadístico de prueba y el intervalo de confianza dados en esta sección).
 - ¿Cómo se relaciona el *rango* de las diferencias $x - y$ con el rango de los valores x y con el rango de los valores y ?
- 36. Efecto de no variación en una muestra.** Se realizó un experimento para probar los efectos del alcohol. Los niveles de alcohol exhalado se midieron en un grupo de tratamiento de personas que bebieron etanol y en otro grupo al que se administró un placebo. Los resultados se presentan en la tabla adjunta. Utilice un nivel de significancia

Grupo de tratamiento	Grupo placebo
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 0.049$	$\bar{x}_2 = 0.000$
$s_1 = 0.015$	$s_2 = 0.000$

de 0.05 para probar la aseveración de que los dos grupos muestrales provienen de poblaciones con la misma media. Los resultados están basados en datos de “Effects of Alcohol Intoxication on Risk Taking, Strategy, and Error Rate in Visuomotor Performance”, de Streufert *et al.*, *Journal of Applied Psychology*, vol. 77, núm. 4.

37. **Cálculo de grados de libertad.** ¿De qué manera se ve afectado el número de grados de libertad en los ejercicios 19 y 20 si se utiliza la fórmula 9-1 en vez de seleccionar el más pequeño entre $n_1 - 1$ y $n_2 - 1$? Si se utiliza la fórmula 9-1 para el número de grados de libertad en vez del más pequeño entre $n_1 - 1$ y $n_2 - 1$, ¿de qué manera se ven afectados el valor P y el ancho del intervalo de confianza? ¿En qué sentido “gl = el más pequeño entre $n_1 - 1$ y $n_2 - 1$ ” es un estimado más conservador del número de grados de libertad que el estimado que se obtiene con la fórmula 9-1?

Inferencias a partir de 9-4 datos apareados

Concepto clave Dos muestras son **dependientes** (o consisten en **datos apareados**) si los miembros de una muestra pueden utilizarse para determinar los miembros de la otra muestra. En esta sección estudiamos métodos para poner a prueba aseveraciones acerca de la diferencia media de datos apareados. Para cada par de datos de valores muestrales, calculamos la diferencia entre los dos valores, y luego utilizamos esas diferencias muestrales para probar aseveraciones acerca de la diferencia poblacional o para construir estimados de intervalos de confianza para la diferencia poblacional. En el capítulo 10 también hablaremos sobre datos apareados, sólo que entonces nos interesaremos en la *asociación* entre dos variables y no en la *media* de las diferencias.

En los datos apareados, existe alguna relación para que cada valor en una muestra se aparezca con un valor correspondiente en la otra muestra. A continuación se presentan algunos ejemplos típicos de datos apareados:

- Los datos muestrales son datos apareados de mediciones de colesterol de lipoproteínas de baja densidad (LDL), tomadas antes y después del tratamiento con Lipitor. Ejemplo: LDL antes de Lipitor = 182; LDL después de Lipitor = 155.
- Los datos muestrales son datos apareados de los índices de masa corporal (IMC) del esposo y de la esposa. Ejemplo: IMC del esposo = 25.1; IMC de la esposa = 19.7.
- Los datos muestrales son las estaturas de candidatos ganadores de la presidencia apareados con las estaturas de los candidatos que recibieron el segundo número más alto de votos. Ejemplo: estatura de Truman = 69 pulgadas; estatura de Dewey = 68 pulgadas.

Para tratar con inferencias acerca de medias y datos apareados, abajo se incluyen resúmenes de los requisitos, la notación, el estadístico de prueba de hipótesis y el intervalo de confianza. Puesto que la prueba de hipótesis y el intervalo de confianza utilizan la misma distribución y el mismo error estándar, son equivalentes en el sentido de que arrojan las mismas conclusiones. En consecuencia, la hipótesis nula de que la diferencia de la media es igual a 0 puede probarse determinando si el intervalo de confianza incluye a 0. Una prueba de hipótesis de dos colas con un nivel de significancia de 0.05 puede probarse con un 95% de confianza, mientras que una prueba de hipótesis de una cola con nivel de significancia de 0.05 puede probarse con un intervalo de confianza del 90%. [Véase la tabla 7.2 para casos comunes].

Requisitos

1. Los datos muestrales consisten en datos apareados.
2. Las muestras son aleatorias simples.
3. Cualquiera o ambas de estas condiciones se satisfacen: el número de datos apareados de datos muestrales es grande ($n > 30$) o los pares de valores tienen diferencias que se toman de una población con una distribución aproximadamente normal. (Si existe una desviación radical de la distribución normal, no debemos utilizar los métodos que se estudian en esta sección, pero quizás podemos utilizar los métodos no paramétricos que se analizan en el capítulo 13).

Notación para datos apareados

d = diferencia individual entre los dos valores en un solo dato apareado

μ_d = valor medio de las diferencias d para la *población* de todos los datos apareados

\bar{d} = valor medio de las diferencias d para los datos *muestrales* apareados (igual a la media de los valores $x - y$)

s_d = desviación estándar de las diferencias d para la *muestra* de datos apareados

n = número de *pares* de datos

Estadístico de prueba de hipótesis para datos apareados

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

donde los grados de libertad = $n - 1$.

Valores P y valores críticos: tabla A-3 (distribución t)

Intervalos de confianza para datos apareados

$$\bar{d} - E < \mu_d < \bar{d} + E$$

donde

$$E = t_{\alpha/2} \frac{s_d}{\sqrt{n}}$$

Valores críticos de $t_{\alpha/2}$: Utilice la tabla A-3 con $n - 1$ grados de libertad.

EJEMPLO Prueba de hipótesis con temperaturas reales y pronosticadas. La tabla 9-2 incluye cinco temperaturas mínimas reales y las correspondientes temperaturas mínimas que se pronosticaron cinco días antes (según datos registrados cerca de la casa del autor). Se trata de datos apareados, puesto que cada par de valores representa al mismo día. Las temperaturas pronosticadas parecen ser muy diferentes de las temperaturas reales, pero ¿existe suficiente evidencia para concluir que la diferencia media no es de cero? Utilice un nivel de significancia de 0.05 para probar la aseveración de que existe una diferencia entre las temperaturas mínimas reales y las temperaturas mínimas pronosticadas cinco días antes.

continúa



Crest y muestras dependientes

A fines de la década de 1950, Procter & Gamble lanzó al mercado la pasta dental Crest como el primer dentífrico con fluoruro. Con el fin de probar la eficacia de Crest en la reducción de las caries, los investigadores realizaron experimentos con varios pares de gemelos. Uno de los gemelos de cada par usó Crest con fluoruro, mientras que el otro continuó con el uso de una pasta dental ordinaria sin fluoruro. Se creía que cada par de gemelos tendría hábitos semejantes de alimentación y de cepillado, así como características genéticas similares. Los resultados mostraron que los gemelos que usaron Crest tenían un número significativamente menor de caries que los que no lo usaron. Este empleo de gemelos como muestras dependientes permitió a los investigadores controlar muchas de las diferentes variables que afectan las caries.

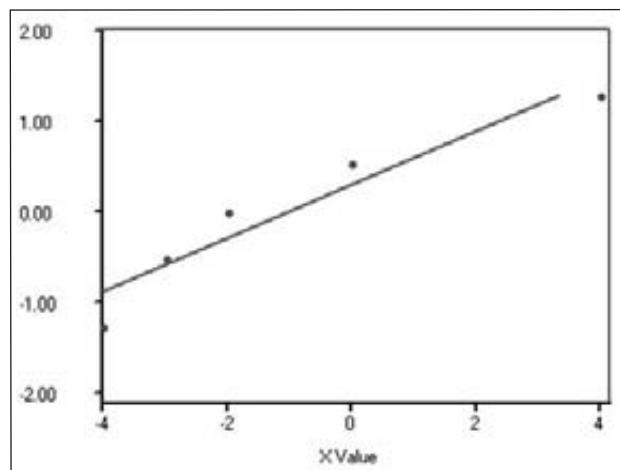
Tabla 9-2 Temperatura real y pronosticada

Mínima real	54	54	55	60	64
Mínima pronosticada cinco días antes	56	57	59	56	64
Diferencia $d = \text{real} - \text{pronosticada}$	-2	-3	-4	4	0

SOLUCIÓN

REQUISITOS ✓ Consideramos cada uno de los tres requisitos. **1.** Los datos muestrales consisten en datos apareados referentes al mismo día. La temperatura real mínima es la temperatura que se registró el día en cuestión, y el otro valor es la temperatura mínima pronosticada para ese día cinco días antes. **2.** En vez de tratarse de una muestra aleatoria simple, tenemos resultados de los primeros cinco días consecutivos del conjunto de datos 8 en el apéndice B. Éste podría ser un problema que surge de factores como un pronosticador sumamente malo (o bueno) que solamente hizo el pronóstico para esos cinco días. Supondremos que existe un sistema más general de pronóstico, y que los días son típicos de los que resultarían de una muestra aleatoria simple. **3.** El número de datos apareados no es grande, por lo que debemos verificar la normalidad de las diferencias. En la siguiente imagen de resultados de STATDISK se observa la gráfica cuantil normal de las diferencias, y concluimos que provienen de una población distribuida normalmente, porque los puntos se acercan razonablemente al patrón de una línea recta, sin presentar otro patrón sistemático. Los requisitos se satisfacen, así que procedemos con la prueba de hipótesis. ✓

STATDISK



Seguiremos el mismo método básico de prueba de hipótesis que se empleó en el capítulo 8, pero utilizaremos el estadístico de prueba para datos apareados que se presentó antes en esta sección.

Paso 1: La aseveración de que existe una diferencia entre las temperaturas mínimas reales y las temperaturas mínimas pronosticadas para cinco días puede expresarse como $\mu_d \neq 0$.

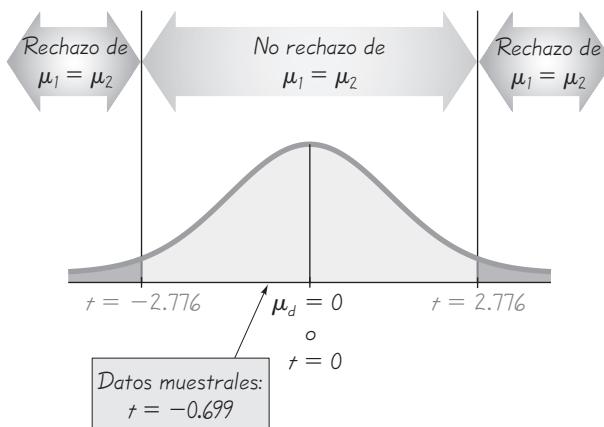


Figura 9-4 Distribución de diferencias d entre valores de datos apareados

- Paso 2: Si la aseveración original no es verdadera, tenemos $\mu_d = 0$.
 Paso 3: La hipótesis nula debe expresar igualdad y la hipótesis alternativa no puede incluir igualdad, por lo tanto, tenemos

$$H_0: \mu_d = 0 \quad H_1: \mu_d \neq 0 \quad (\text{aseveración original})$$

- Paso 4: El nivel de significancia es $\alpha = 0.05$.
 Paso 5: Utilizamos la distribución t de Student.
 Paso 6: Antes de calcular el valor del estadístico de prueba, primero debemos calcular los valores de \bar{d} y s_d . Remítase a la tabla 9-2 y utilice las diferencias de $-2, -3, -4, 4$ y 0 para calcular estos estadísticos muestrales: $\bar{d} = -1.0$ y $s_d = 3.2$. Utilizando estos estadísticos muestrales y la suposición de la prueba de hipótesis de que $\mu_d = 0$, ahora podemos calcular el valor del estadístico de prueba.

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-1.0 - 0}{\frac{3.2}{\sqrt{5}}} = -0.699$$

Los valores críticos de $t = \pm 2.776$ se encuentran en la tabla A-3 como sigue: utilice la columna para 0.05 (área en dos colas) y utilice el renglón con grados de libertad $n - 1 = 4$. La figura 9-4 nos indica el estadístico de prueba, los valores críticos y la región crítica.

- Paso 7: Puesto que el estadístico de prueba no cae en la región crítica, no rechazamos la hipótesis nula.

INTERPRETACIÓN Los datos muestrales de la tabla 9-2 no ofrecen suficiente evidencia para sustentar la aseveración de que las temperaturas mínimas reales y pronosticadas para cinco días son diferentes. Esto *no* significa que las temperaturas reales y pronosticadas sean iguales. Quizá datos muestrales adicionales darían la evidencia necesaria para concluir que las temperaturas mínimas reales y pronosticadas son diferentes. (Véase el ejercicio 21 donde se utilizan resultados para 35 días).



¿Salvan vidas las bolsas de aire?

La National Highway Transportation Safety Administration reportó que en un año reciente se salvaron 3,448 vidas gracias a las bolsas de aire. Se reportó que para los conductores de automóviles involucrados en choques frontales, la tasa de mortalidad se redujo en un 31%, mientras que para los pasajeros se redujo en un 27%. Se señaló que “el cálculo de vidas salvadas se realiza con un análisis matemático de casos reales de accidentes mortales en vehículos con bolsas de aire en comparación con vehículos sin bolsas de aire. Los análisis de este tipo se llaman estudios de comparación de doble apareo, y son métodos de análisis estadístico ampliamente aceptados”.

Método del valor P . En el ejemplo anterior se utilizó el método tradicional, pero se puede usar el método del valor P . Con STATDISK, Excel, Minitab o la calculadora TI-83/84 Plus, se calcula el valor P , que es 0.5185. (Si se emplea la tabla A-3 con el estadístico de prueba $t = -0.699$, podemos determinar que el valor P es mayor que 0.20). Una vez más, no rechazamos la hipótesis nula, puesto que el valor P es mayor que el nivel de significancia de $\alpha = 0.05$.

EJEMPLO Intervalo de confianza con temperaturas reales y pronosticadas Utilice los mismos datos apareados de la tabla 9-2 y construya un estimado de un intervalo de confianza del 95% de μ_d , que es la media de las diferencias entre las temperaturas mínimas reales y las temperaturas mínimas pronosticadas de cinco días. Interprete el resultado.

SOLUCIÓN

REQUISITO ✓ La solución del ejemplo anterior incluye la verificación de que los requisitos se cumplan. Procedemos con la construcción del intervalo de confianza. ✓

Utilizamos los valores $\bar{d} = -1.0$, $s_d = 3.2$, $n = 5$ y $t_{\alpha/2} = 2.776$ (encontrado en la tabla A-3 con $n - 1 = 4$ grados de libertad y área de 0.05 en dos colas). Primero calculamos el valor del margen de error E .

$$E = t_{\alpha/2} \frac{s_d}{\sqrt{n}} = 2.776 \cdot \frac{3.2}{\sqrt{5}} = 4.0$$

Ahora puede calcularse el intervalo de confianza.

$$\begin{aligned}\bar{d} - E &< \mu_d < \bar{d} + E \\ -1.0 - 4.0 &< \mu_d < -1.0 + 4.0 \\ -5.0 &< \mu_d < 3.0\end{aligned}$$

INTERPRETACIÓN Algunas veces el resultado se expresa como -1.0 ± 4.0 o como $(-5.0, 3.0)$. (Si se utiliza un programa de cómputo o una calculadora TI-83/84 Plus, obtenemos el resultado de $-4.9 < \mu_d < 2.9$. La discrepancia se debe al uso de la desviación estándar redondeada de 3.2, en vez del valor sin redondeo de 3.16227766). A la larga, el 95% de las muestras de este tipo conducirán a límites del intervalo de confianza que realmente no contienen la media poblacional real de las diferencias. Note que los límites del intervalo de confianza contienen a 0, lo que indica que el valor real de μ_d no es significativamente diferente de 0. No podemos concluir que existe una diferencia significativa entre las temperaturas mínimas reales y pronosticadas.

Diseño experimental Suponga que deseamos realizar un experimento con dos fertilizantes diferentes (uno nuevo y uno viejo), y que contamos con 20 parcelas de terreno con la misma área. En vez de utilizar 10 parcelas para el fertilizante nuevo y 10 parcelas para el fertilizante viejo, sería mejor dividir cada parcela a la mitad, y tratar una mitad con el fertilizante nuevo y la otra mitad con el fertilizante viejo. Después, se podrían aparear las cosechas por las parcelas que comparten. La gran ventaja de utilizar datos apareados es que se reduce la variación extraña, y disminuye la probabilidad de confundirse por cosechas más productivas en parcelas con un suelo más fértil. Si resulta que una parcela tiene un suelo sumamente fértil, ésta sería compartida por ambos fertilizantes; si el experimento se hubiera

Uso de la tecnología

STATDISK Primero ingrese los datos apareados en columnas de la ventana de datos de STATDISK, luego seleccione **Analysis** del menú principal. Elija **Hypothesis Testing** o **Confidence Intervals** y luego seleccione **Mean-Matched Pairs**. Complete las entradas, indique las selecciones en el cuadro de diálogo y luego dé clic en **Evaluate**.

MINITAB Ingrese los datos apareados en las columnas C1 y C2. Haga clic en **Stat**, seleccione **Basic Statistics**, luego seleccione

Paired t. Ingrese C1 para la primera muestra, ingrese C2 para la segunda muestra y luego haga clic en el cuadro de **Options** para cambiar el nivel de confianza o el formato de la hipótesis alternativa.

EXCEL Ingrese los datos muestrales apareados en las columnas A y B.

Haga clic en **DDXL** para utilizar el complemento Data Desk XL. Seleccione **Hypotheses Tests** y **Paired t Test** o seleccione **Confidence Intervals** y **Paired t Interval**. En el cuadro de diálogo, haga clic en el ícono del lápiz para la primera columna cuantitativa e ingrese el rango de valores para la primera muestra, por ejemplo, A1:A25. Haga clic en el ícono del lápiz para la segunda columna cuantitativa e ingrese el rango de valores para la segunda muestra. Haga clic en **OK**. Ahora complete el nuevo cuadro de diálogo siguiendo los pasos indicados.

Para utilizar el complemento de Data Analysis de Excel, haga clic en **Tools**, que se encuentra en la barra del menú principal, luego seleccione **Data Analysis** y proceda a seleccionar **t-test Paired Two Sample for Means**.

En el cuadro de diálogo, ingrese el rango de valores para cada una de las dos muestras, ingrese el valor supuesto de la diferencia media poblacional (generalmente 0) e ingrese el nivel de significancia. Los resultados en la pantalla incluirán el estadístico de prueba, los valores *P* para una prueba de una cola y para una prueba de dos colas, así como los valores críticos para una prueba de una cola y para una prueba de dos colas.

TI-83/84 PLUS *Advertencia:* No utilice el elemento del menú **2-SampTTest**, porque éste se aplica a muestras *independientes*. En vez de ello, ingrese los datos de la primera variable en la lista L1, ingrese los datos de la segunda variable en la lista L2, luego despeje la pantalla e ingrese **L1 – L2 → L3** de manera que la lista L3 contenga las diferencias individuales *d*. Después oprima **STAT**, luego seleccione **TESTS** y elija la opción de **T-Test** (para una prueba de hipótesis) o **TInterval** (para un intervalo de confianza). Utilice la opción de alimentación de **Data**. Para la lista, ingrese L3. (Si utiliza **T-Test**, también ingrese el valor 0 para μ_0). Oprima **ENTER** cuando termine.



diseñado de otra forma, se pensaría que uno de los fertilizantes es muy eficaz, cuando en realidad es el suelo el responsable de la mayor producción y no el fertilizante. Al diseñar un experimento o planear un estudio observacional, generalmente es mejor utilizar datos apareados que dos muestras independientes.

9-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Estaturas y pesos de osos.** Los datos muestrales consisten en las estaturas y los pesos de 12 osos seleccionados al azar. Cuando se analiza la relación entre estatura y peso, ¿se aplican los métodos de esta sección? ¿Por qué?
- 2. Apareamiento con lanzamientos de monedas.** Un investigador está estudiando unas mediciones de inteligencia y obtiene puntuaciones de CI de 24 sujetos configuradas como 12 datos apareados; el apareamiento se efectuó con base en lanzamientos de monedas. ¿Es posible aplicar los métodos de esta sección?
- 3. Notación.** ¿Qué denotan los símbolos \bar{d} y s_d ?
- 4. Gráficas de cuadro.** Los ejemplos en esta sección están basados en datos apareados de temperaturas reales y pronosticadas. Para esos ejemplos, ¿tendría sentido construir una gráfica de cuadro para las temperaturas reales y otra para las temperaturas pronosticadas?

Cálculos para datos apareados. En los ejercicios 5 y 6, suponga que usted quiere utilizar un nivel de significancia de 0.05 para probar la aseveración de que los datos muestrales apareados provienen de una población en la que la diferencia media es $\mu_d = 0$. Calcule a) d, b) s_d c) el estadístico de prueba t y d) los valores críticos.

5.	x	2	5	7	1	1
	y	2	3	9	4	0
6.	x	8	8	5	2	7
	y	9	6	5	5	3

7. **Intervalo de confianza.** Utilice los datos muestrales apareados del ejercicio 5 y construya un intervalo de confianza del 95% para la media poblacional de todas las diferencias $x - y$.
8. **Intervalo de confianza.** Utilice los datos muestrales apareados del ejercicio 6 y construya un intervalo de confianza del 99% para la media poblacional de todas las diferencias $x - y$.

9. **Interpretación de resultados de prueba de tratamiento del malestar por movimiento.** La siguiente pantalla de Minitab es el resultado de un experimento en el que se hizo una prueba del malestar por movimiento a 10 sujetos, antes y después de tomar el fármaco astemizole. Los resultados de Minitab se basan en las diferencias del número de movimientos de cabeza que los sujetos podían soportar sin sufrir náuseas. (Las diferencias se obtuvieron restando los valores “después” de los valores “antes”).

- a. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el astemizole tiene un efecto (para bien o para mal) en la vulnerabilidad al malestar por movimiento. Con base en el resultado, ¿utilizaría el astemizole si se preocupara por el malestar por movimiento cuando estuviera a bordo de un crucero?
- b. Suponga que, en vez de probar si existe algún efecto (para bien o para mal), queremos probar la aseveración de que el astemizole es eficaz en la *prevención* del malestar por movimiento, ¿cuál es el valor P y qué concluye usted?

95% CI for mean difference: (-48.8, 33.8)
T-Test of mean difference = 0 (vs not = 0):
T-value = -0.41 P-value = 0.691

10. **Interpretación de resultados de prueba antes y después del tratamiento.** El Captopril es un fármaco diseñado para reducir la presión sanguínea sistólica. Cuando se probaron sujetos con este fármaco, sus lecturas de presión sanguínea sistólica (en mmHg) se midieron antes y después de tomar el fármaco. Se utilizó Excel para aplicar una prueba t con los datos apareados de las lecturas de antes y después, y abajo se presentan los resultados (según datos de “Essential Hypertension: Effect of an Oral Inhibitor of Angiotensin-Converting Enzyme”, de MacGregor *et al.*, *British Medical Journal*, vol. 2). Utilice los resultados de Excel para probar la aseveración de que Captopril sirve para reducir la presión sanguínea sistólica. ¿El fármaco es eficaz?

Excel

t-Test: Paired Two Sample for Means		
	Variable 1	Variable 2
Mean	185.3333333	166.75
Variance	291.3333333	220.9318182
Observations	12	12
Pearson Correlation	0.808392337	
Hypothesized Mean Difference	0	
df	11	
t Stat	6.371428571	
P(T<=t) one-tail	2.64274E-05	
t Critical one-tail	1.795884814	
P(T<=t) two-tail	5.28547E-05	
t Critical two-tail	2.200985159	

11. Temperaturas pronosticadas. Algunos ejemplos de esta sección incluyen temperaturas mínimas reales y temperaturas mínimas que se pronosticaron cinco días antes. Abajo se presentan las temperaturas máximas reales y las temperaturas máximas que se pronosticaron un día antes (según datos registrados cerca de la casa del autor).

- Utilice un nivel de significancia de 0.05 para probar la aseveración de que hay una diferencia media de cero entre las temperaturas máximas reales y las temperaturas máximas pronosticadas un día antes. ¿Qué sugieren los resultados sobre la exactitud de las temperaturas pronosticadas?
- Construya un estimado de un intervalo de confianza del 95% de la diferencia media entre las temperaturas máximas reales y las temperaturas máximas pronosticadas un día antes. Interprete el intervalo de confianza resultante y analice las implicaciones si los límites del intervalo de confianza contuvieran a 0.

Máxima real	80	77	81	85	73
Máxima pronosticada un día antes	78	75	81	85	76

12. ¿El viernes 13 es de mala suerte? Investigadores reunieron datos del número de admisiones hospitalarias por accidentes automovilísticos; abajo se incluyen los resultados de los días viernes 6 del mes y del siguiente viernes 13 del mismo mes (según datos de “Is Friday 13th Bad for your Health?”, por Scanlon *et al.*, *British Medical Journal*, vol. 307, tal como aparece en *Data and Story Line*, el recurso *online* de conjuntos de datos). Utilice un nivel de significancia de 0.05 para probar la aseveración de que cuando el día 13 de un mes cae en viernes, el número de admisiones hospitalarias por accidentes automovilísticos no se ve afectado.

viernes 6:	9	6	11	11	3	5
viernes 13:	13	12	14	10	4	12

13. Medición de la presión sanguínea. Catorce estudiantes de medicina midieron la presión sanguínea del mismo paciente y repitieron la medición al día siguiente. A continuación se listan las lecturas sistólicas en mmHg (según datos del Bellevue Hospital en la ciudad de Nueva York). Construya un estimado de un intervalo de confianza del 95% de la diferencia media entre las lecturas sistólicas de un día y las tomadas al día siguiente. ¿Qué sugiere el resultado?

Día 1:	138	130	135	140	120	125	120	130	130	144	143	140	130	150
Día 2:	116	120	125	110	120	135	124	118	120	130	140	140	130	138

14. Estaturas de ganadores y segundos lugares. Abajo se presentan las estaturas de candidatos que ganaron las elecciones presidenciales y las estaturas de candidatos con el siguiente número más alto de votos populares. Los datos se encuentran en orden cronológico, de manera que las estaturas correspondientes de las dos listas están apareadas. Para los candidatos que ganaron más de una vez, sólo se incluyen las estaturas de la primera elección y no se incluyen elecciones previas a 1900.

- Una hipótesis conocida plantea que los candidatos ganadores tienden a ser más altos que los candidatos perdedores correspondientes. Utilice un nivel de significancia de 0.05 para poner a prueba esa hipótesis. ¿Parece que la estatura es un factor importante para ganar la presidencia?
- Si usted planeara poner a prueba la afirmación del inciso a) utilizando un intervalo de confianza, ¿qué nivel de confianza debería usar? Construya un intervalo de confianza con ese nivel de confianza y luego interprete el resultado.

Ganador de la presidencia	Segundo lugar
71 74.5 74 73 69.5 71.5 75 72	73 74 68 69.5 72 71 72 71.5
70.5 69 74 70 71 72 70 67	70 68 71 72 70 72 72 72

- 15. Curso de preparación para la prueba SAT.** El artículo “An SAT Coaching Program That Works”, de Kaplan (*Chance*, vol. 15, núm. 1) incluyó una gráfica que describe las puntuaciones de la prueba SAT de 50 sujetos en un grupo de control. Se eligieron al azar nueve de las 50 puntuaciones, donde cada punto representa la puntuación de la prueba SAT aplicada la primera vez y la puntuación de la prueba SAT aplicada la segunda vez. Entre las dos pruebas no se tomó un curso de preparación. La gráfica se utilizó para identificar las puntuaciones que se listan a continuación. Ponga a prueba la aseveración de que las diferencias tienen una media de 0. ¿Qué sugieren los resultados?

Estudiante	A	B	C	D	E	F	G	H	I
Puntuación en la primera aplicación del SAT	480	510	530	540	550	560	600	620	660
Puntuación en la segunda aplicación del SAT	460	500	530	520	580	580	560	640	690

- 16. Estaturas de hombres reportadas y medidas.** Como parte de la National Health and Nutrition Examination Survey realizada por el Departamento de Salud y Servicios Humanos de EUA, se obtuvieron estaturas reportadas y medidas de hombres de 12 a 16 años de edad. Abajo se listan resultados muestrales.
- ¿Existe evidencia suficiente para sustentar la aseveración de que existe una diferencia entre las estaturas reportadas y las estaturas medidas de hombres de 12 a 16 años de edad? Utilice un nivel de significancia de 0.05.
 - Construya un estimado del intervalo de confianza del 95% de la diferencia media entre las estaturas reportadas y las estaturas medidas. Interprete el intervalo de confianza resultante y comente las implicaciones sobre si los límites del intervalo de confianza contienen a 0.

Estatura reportada	68	71	63	70	71	60	65	64	54	63	66	72
Estatura medida	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67.0	55.6	74.2	65.0	70.8

- 17. Eficacia de la hipnosis en la reducción del dolor.** Se realizó un estudio para investigar la eficacia de la hipnosis en la reducción del dolor. Los resultados de sujetos seleccionados al azar se incluyen en la tabla adjunta (según “An Analysis of Factors That Contribute to the Efficacy of Hipnotic Analgesia”, de Price y Barber, *Journal of Abnormal Psychology*, vol. 96, núm. 1). Los valores se tomaron antes y después de la hipnosis; la unidad de medición son centímetros en una escala de dolor.
- Construya un intervalo de confianza del 95% para la media de las diferencias “antes-después”.
 - Utilice un nivel de significancia 0.05 para probar la afirmación de que las mediciones sensoriales son más bajas después de la hipnosis.
 - ¿Parece ser eficaz la hipnosis en la reducción del dolor?

Sujeto	A	B	C	D	E	F	G	H
Antes	6.6	6.5	9.0	10.3	11.3	8.1	6.3	11.6
Después	6.8	2.4	7.4	8.5	8.1	6.1	3.4	2.0

- 18. Medición de inteligencia en niños.** A menudo se miden las capacidades mentales de niños pequeños dándoles cubos y pidiéndoles que construyan una torre tan alta como sea posible. Un experimento de construcción con cubos se repitió un mes después, con los tiempos (en segundos) que se listan en la tabla adjunta (según datos de “Tower Building”, de Johnson y Courtney, *Child Development*, vol. 3).
- ¿Existe suficiente evidencia para sustentar la aseveración de que hay una diferencia entre los dos tiempos? Utilice un nivel de significancia de 0.01.

- b. Construya un intervalo de confianza del 99% para la media de las diferencias.

¿Los límites del intervalo de confianza contienen a 0, indicando que no existe una diferencia significativa entre los tiempos del primer y segundo ensayo?

Niño	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Primer ensayo	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
Segundo ensayo	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15

19. **Prueba de semillas de maíz.** En 1908 William Gosset publicó el artículo “The Probable Error of a Mean” bajo el seudónimo de “Student” (*Biometrika*, vol. 6, núm. 1). El artículo incluyó los datos listados abajo para dos tipos diferentes de semillas de maíz (comunes y secadas al horno) que se utilizaron en parcelas adyacentes. Los valores listados son las cosechas de cabezas de maíz o mazorcas en libras por acre.

- a. Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que no existe diferencia entre las cosechas de los dos tipos de semillas.
 b. Construya un estimado del intervalo de confianza del 95% de la diferencia media entre las cosechas de los dos tipos de semillas.
 c. ¿Parece que algún tipo de semilla es mejor?

Comunes	1903	1935	1910	2496	2108	1961	2060	1444	1612	1316	1511			
Secadas al horno	2009	1915	2011	2463	2180	1925	2122	1482	1542	1443	1535			

20. **Intervalos de confianza para comparar teclados.** La configuración tradicional del teclado se conoce como QWERTY debido a la posición de esas letras en su fila superior. Se supone que el teclado Dvorak, creado en 1936, ofrece un arreglo más eficiente porque las teclas más utilizadas se ubican en la fila intermedia, donde son más accesibles. Un artículo de la revista *Discover* sugirió que se puede medir la facilidad para teclear utilizando este sistema de calificación: contar cada letra en la fila intermedia como 0, contar cada letra en la fila superior como 1 y contar cada letra en la fila inferior como 2. Utilizando este sistema de calificación en cada una de las 52 palabras en el preámbulo de la Constitución estadounidense, obtenemos los valores que se presentan abajo. Construya un estimado de un intervalo de confianza del 95% de la diferencia entre las medias de las puntuaciones de las palabras con ambos sistemas. ¿Parece que el resultado sustenta la aseveración de que la configuración del teclado Dvorak es más fácil?

Teclado QWERTY

2	2	5	1	2	6	3	3	4	2	4	0	5	7	7	5	6	6	8	10
7	2	2	10	5	8	2	4	4	2	6	2	1	7	2	7	2	3	8	
1	5	2	5	2	14	2	2	6	3	1	7								

Teclado Dvorak

2	0	3	1	0	0	0	2	0	4	0	3	4	0	3	3	1	3	5
2	0	4	1	5	0	4	0	1	3	0	1	0	3	0	1	2	0	0
0	1	0	3	0	1	2	0	0	0	1								

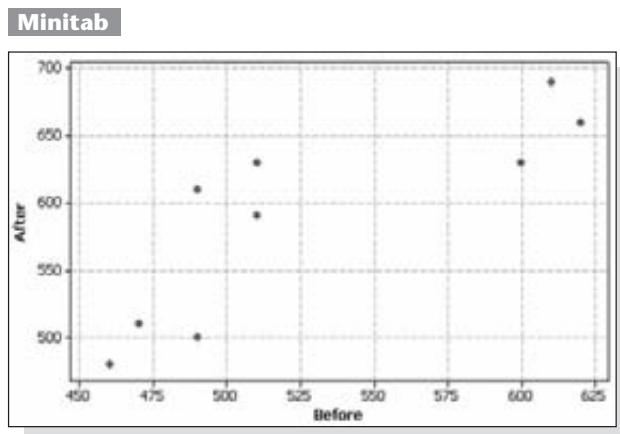
21. **Conjunto de datos del apéndice B: temperaturas pronosticadas.** Para los ejemplos de esta sección se utilizaron sólo cinco pares de datos muestrales para que los cálculos fueran sencillos. Remítase al conjunto de datos 8 del apéndice B y utilice todas las temperaturas mínimas reales y las temperaturas mínimas que se pronosticaron cinco días antes.

- a. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que no existe diferencia entre las temperaturas mínimas reales y las temperaturas mínimas pronosticadas cinco días antes.
 b. Construya un estimado del intervalo de confianza del 95% de la diferencia media entre las temperaturas mínimas reales y las temperaturas mínimas pronosticadas cinco días antes.
 c. Compare los resultados con los obtenidos en los ejemplos de esta sección. ¿Parece que las temperaturas mínimas pronosticadas son exactas?

- 22. Conjunto de datos del apéndice B: alcohol y tabaco en películas infantiles.** Remítase al conjunto de datos 5 en el apéndice B. Utilice los datos apareados que consisten en las ocasiones en que las películas mostraron consumo de tabaco y las ocasiones en que mostraron consumo de alcohol.
- ¿Existe evidencia suficiente para concluir que difiere el número de ocasiones?
 - Construya un estimado del intervalo de confianza del 99% de la media de las diferencias entre las ocasiones en que hubo consumo de tabaco y consumo de alcohol. Con base en el resultado, ¿existe una diferencia significativa en el número de veces que los niños se vieron expuestos a consumo de tabaco y el número de veces que se vieron expuestos a consumo de alcohol?
- 23. Conjunto de datos del apéndice B: precios de casas.** Remítase al conjunto de datos 18 del apéndice B y utilice los precios de venta y los precios de lista de las casas vendidas.
- Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de un agente inmobiliario de que el mercado de bienes raíces tiene tanto auge que no existe la diferencia entre los precios de venta y los precios de lista.
 - Construya un estimado de un intervalo de confianza del 95% de la diferencia media entre los precios de venta y los precios de lista.
- 24. Conjunto de datos del apéndice B: Old Faithful.** Remítase al conjunto de datos 11 del apéndice B y utilice los intervalos antes y después de cada erupción.
- Utilice un nivel de significancia de 0.05 para probar la aseveración de que las diferencias antes/después tienen una media igual a 0. ¿Qué sugiere el resultado acerca de los intervalos entre erupciones?
 - Construya un estimado de un intervalo de confianza del 95% de la diferencia media entre los intervalos antes de las erupciones y los intervalos después de las erupciones.

9-4 MÁS ALLÁ DE LO BÁSICO

- 25. Curso de preparación para el examen SAT.** El artículo “An SAT Coaching Program That Works”, de Kaplan (*Chance*, vol. 15, núm. 1) incluyó una gráfica similar a la que se muestra a continuación. Los puntos representan las puntuaciones de la prueba SAT de 9 sujetos antes y después de tomar un curso de preparación para la prueba. Para cada sujeto, identifique las puntuaciones que obtuvo en la prueba SAT antes y después del curso, y después ponga a prueba la aseveración de que el curso de preparación sirve para aumentar las puntuaciones de la prueba.



- 26. Efectos de un valor extremo y unidades de medida**
- Al utilizar los métodos de esta sección, ¿puede un valor extremo tener un efecto drástico en la prueba de hipótesis y en el intervalo de confianza?

- b.** Para los ejemplos de esta sección se utilizaron temperaturas medidas en grados Fahrenheit. Si convertimos todas las temperaturas muestrales en grados Fahrenheit a grados Celsius, ¿se ve afectada la prueba de hipótesis por un cambio de este tipo en las unidades? ¿Cómo?
- 27. Uso del procedimiento correcto**
- Considere que los datos muestrales que se presentan abajo son datos apareados y utilice un nivel de significancia de 0.05 para probar la aseveración de que $\mu_d > 0$.
 - Considere que los datos muestrales que se presentan a continuación son dos muestras independientes. Utilice un nivel de significancia de 0.05 para probar la aseveración de que $\mu_1 > \mu_2$.
 - Compare los resultados de los incisos *a*) y *b*). ¿Es crucial utilizar el método correcto? ¿Por qué?

x	1	3	2	2	1	2	3	3	2	1
y	1	2	1	2	1	2	1	2	1	2

9-5 Comparación de la variación en dos muestras

Concepto clave Puesto que la característica de variación entre los datos es extremadamente importante, esta sección presenta la prueba *F* para comparar dos varianzas (o desviaciones estándar) poblacionales utilizando dos muestras. En esta sección estudiaremos la distribución *F*, que se utiliza para la prueba *F*. Es sumamente importante estar conscientes de una grave desventaja de este procedimiento: la prueba *F* para comparar dos varianzas (o desviaciones estándar) poblacionales es *muy* sensible a las desviaciones que se alejan de la distribución normal. Al final de esta sección analizaremos brevemente alternativas a la prueba *F* que no son tan sensibles a esta característica.

Prueba *F* para comparar varianzas

En esta sección utilizamos varianzas muestrales (o desviaciones estándar) para comparar dos varianzas (o desviaciones estándar) poblacionales. Debemos saber que la varianza es el cuadrado de la desviación estándar y también debemos conocer la siguiente notación.

Medidas de variación

$$s = \text{desviación estándar muestral} \quad s^2 = \text{varianza muestral} \text{ (desviación estándar muestral al cuadrado)}$$

$$\sigma = \text{desviación estándar poblacional} \quad \sigma^2 = \text{varianza poblacional} \text{ (desviación estándar poblacional al cuadrado)}$$

Los cálculos de esta sección se simplificarán considerablemente si designamos las dos muestras de manera que s_1^2 represente a *la más grande* de las dos varianzas muestrales. Matemáticamente no importa cuál muestra se designe como la muestra 1, así que la vida será más fácil si permitimos que s_1^2 represente a la mayor de las dos varianzas muestrales, como en el estadístico de prueba incluido en el cuadro de resumen.

Requisitos

1. Las dos poblaciones son *independientes* una de la otra. (En la sección 9-2 aprendimos que dos muestras son independientes si la muestra seleccionada de una población no está relacionada con la muestra seleccionada de la otra población. Las muestras no están apareadas o asociadas).
2. Las dos poblaciones están *distribuidas normalmente*. (Este supuesto es importante, ya que los métodos esta sección son sumamente sensibles a las desviaciones de la normalidad).

Notación para pruebas de hipótesis con dos varianzas o desviaciones estándar

s_1^2 = la más grande de dos varianzas muestrales

n_1 = tamaño de la muestra que tiene la varianza más grande

σ_1^2 = varianza de la población de donde se obtiene la muestra con la varianza más grande

Los símbolos s_2^2 , n_2 y σ_2^2 se utilizan para la otra muestra y la otra población.

Estadístico de prueba para pruebas de hipótesis con dos varianzas

$$F = \frac{s_1^2}{s_2^2} \quad (\text{donde } s_1^2 \text{ es la más grande de las dos varianzas muestrales})$$

Valores críticos: Utilice la tabla A-5 para obtener valores críticos F que se determinan por lo siguiente:

1. El nivel de significancia α (la tabla A-5 tiene cuatro páginas de valores críticos para $\alpha = 0.025$ y 0.05).
2. **Grados de libertad del numerador** = $n_1 - 1$
3. **Grados de libertad del denominador** = $n_2 - 1$

Para dos poblaciones distribuidas normalmente con varianzas iguales (es decir, $\sigma_1^2 = \sigma_2^2$), la distribución muestral del estadístico de prueba $F = s_1^2/s_2^2$ es la **distribución F** que se muestra en la figura 9-5 con los valores críticos que se listan en la tabla de A-5. Si usted continúa repitiendo el experimento de seleccionar muestras al azar a partir de dos poblaciones distribuidas normalmente con varianzas iguales, la distribución de la proporción s_1^2/s_2^2 de las varianzas muestrales es la distribución F .

En la figura 9-5, observe las siguientes propiedades de la distribución F :

- La distribución F no es simétrica.
- Los valores de la distribución F no pueden ser negativos.
- La forma exacta de la distribución F depende de dos diferentes grados de libertad.

Valores críticos: Para calcular un valor crítico, primero remítase a la parte de la tabla A-5 correspondiente a α (para una prueba de una cola) o $\alpha/2$ (para una prueba de dos colas), luego intercepte la columna que representa los grados de libertad para s_1^2 con el renglón que representa los grados de libertad para s_2^2 . Puesto que estamos estipulando que la varianza muestral más grande es s_1^2 , todas

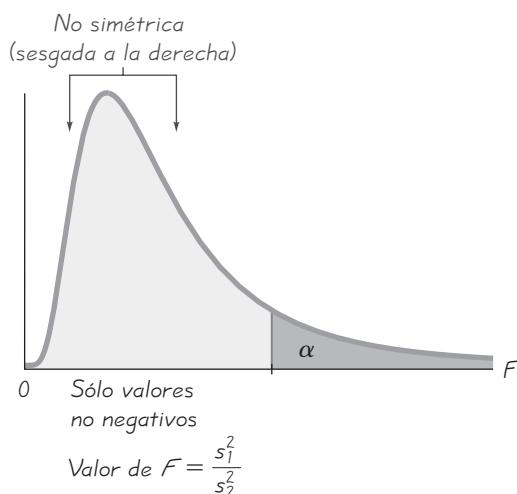


Figura 9-5 Distribución F

Existe una distribución F diferente para cada par diferente de grados de libertad para el numerador y el denominador.

las pruebas de una cola serán de cola derecha y todas las pruebas de dos colas requerirán que encontremos sólo el valor crítico localizado a la derecha. Buenas noticias: No tenemos necesidad de calcular un valor crítico separando una región crítica de cola izquierda. (Puesto que la distribución F no es simétrica y sólo tiene valores no negativos, un valor crítico de cola izquierda no puede encontrarse utilizando el negativo del valor crítico de cola derecha; en vez de ello, el valor crítico de cola izquierda se calcula utilizando el recíproco del valor de cola derecha con los números de grados de libertad invertidos. Véase el ejercicio 23).

Con frecuencia tenemos números de grados de libertad que no están incluidos en la tabla a A-5. Podríamos utilizar interpolación lineal para aproximar los valores que no aparecen, pero en la mayoría de los casos esto no es necesario puesto que el estadístico de prueba F es menor que el valor crítico más bajo posible o mayor que el valor crítico más alto posible. Por ejemplo, la tabla de A-5 indica que para $\alpha = 0.025$ en la cola derecha, 20 grados de libertad para el numerador y 34 grados de libertad para el denominador, el valor crítico F está entre 2.0677 y 2.1952. Cualquier estadístico de prueba F menor que 2.0677 determina que no se rechace la hipótesis nula, y cualquier estadístico de prueba F mayor que 2.1952 dará por resultado el rechazo de la hipótesis nula; la interpolación sólo es necesaria si el estadístico de prueba F cae entre 2.0677 y 2.1952. El uso de la tecnología adecuada elimina por completo este problema al dar valores P o valores críticos para cualquier número de grados de libertad.

Interpretación del estadístico de prueba F Si en realidad las dos poblaciones tienen varianzas iguales, entonces la proporción s_1^2/s_2^2 tiende a 1, puesto que los valores de s_1^2 y s_2^2 tienden a acercarse. Pero si las dos poblaciones tienen varianzas radicalmente diferentes, s_1^2 y s_2^2 tienden a ser números muy distintos. Si denotamos la más grande de las varianzas muestrales como s_1^2 , vemos que la proporción s_1^2/s_2^2 será un número grande siempre que s_1^2 y s_2^2 tengan valores lejanos entre sí. En consecuencia, un valor de F cercano a 1 será evidencia a favor de la



Menor variación, mayor calidad

Ford y Mazda estaban produciendo transmisiones similares que, al parecer, se fabricaban con las mismas especificaciones. Pero las transmisiones hechas en Estados Unidos requerían más reparaciones por garantía que las transmisiones hechas en Japón. Cuando los investigadores inspeccionaron muestras de las cajas de engranajes de las transmisiones japonesas, al principio pensaron que sus instrumentos de medición estaban defectuosos porque no detectaban ninguna variabilidad entre las cajas de engranajes de las transmisiones Mazda. Se dieron cuenta de que, aunque las transmisiones estadounidenses cumplían con las especificaciones, las transmisiones Mazda no sólo estaban dentro de las especificaciones, sino también uniformemente cercanas al valor deseado. Al reducir la variabilidad entre las cajas de engranajes de las transmisiones, Mazda redujo los costos de inspección, de desecho, de corrección y de las reparaciones por garantía.

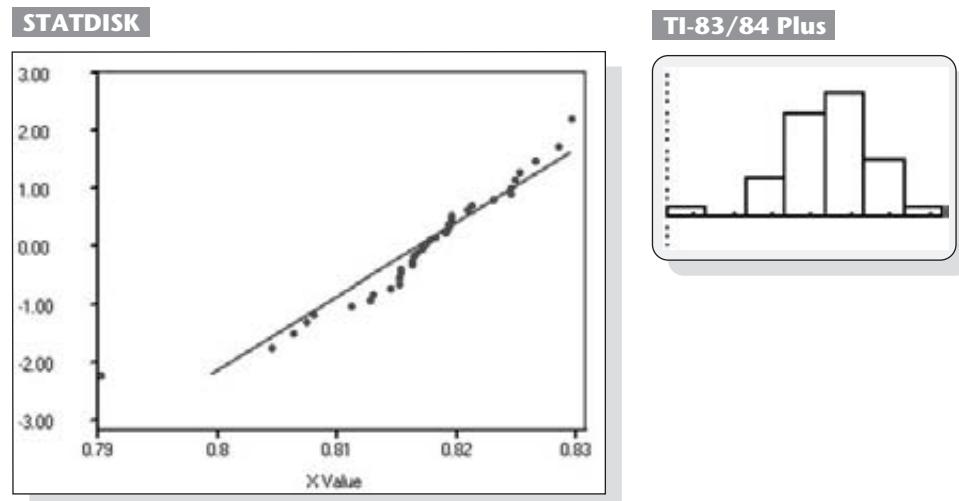
conclusión de que $\sigma_1^2 = \sigma_2^2$, y un valor grande de F será evidencia en contra de la conclusión de igualdad de las varianzas poblacionales.

Los valores de F grandes son evidencia en contra de $\sigma_1^2 = \sigma_2^2$.

Aseveraciones acerca de desviaciones estándar: El estadístico de prueba F se aplica a una aseveración acerca de dos varianzas, pero también podemos utilizarlo para aseveraciones acerca de dos desviaciones estándar poblacionales. Cualquier aseveración acerca de dos desviaciones estándar poblacionales puede replantearse en términos de las varianzas correspondientes.

Exploración de los datos

Puesto que el requisito de distribuciones normales es muy importante y muy estricto, debemos comenzar comparando los dos conjuntos de datos muestrales con la ayuda de herramientas como los histogramas, las gráficas de cuadro y las gráficas cuantilares normales (véase la sección 6-7); luego, debemos buscar valores extremos. Es necesario calcular los valores de los estadísticos muestrales, en especial las desviaciones estándar. Por ejemplo, considere los 36 pesos de Coca-Cola clásica en 36 latas diferentes. (Los pesos se listan en el conjunto de datos 12 del apéndice B). Aquí se presenta un histograma de una calculadora TI-83/84 Plus y una gráfica cuantilar normal de STATDISK. El histograma indica que los datos tienen una distribución aproximadamente normal y que existe un dato que es un valor extremo potencial. La gráfica cuantilar normal señala que los puntos se aproximan razonablemente a una línea recta, aunque no se ajustan a la línea recta perfectamente. Este conjunto de datos, a todas luces, satisface el requisito de una distribución que es *aproximadamente* normal, pero no es muy claro que este conjunto de datos satisfaga los requisitos más estrictos de normalidad que se aplican a los métodos de esta sección.



EJEMPLO Coca contra Pepsi El conjunto de datos 12 del apéndice B incluye los pesos (en libras) de muestras de Coca clásica y Pepsi clásica. Los estadísticos muestrales se resumen en la tabla adjunta. Utilice un nivel de significancia 0.05 para probar la aseveración de que los pesos de Coca clásica y los pesos de Pepsi clásica tienen la misma desviación estándar.

	Coca clásica	Pepsi clásica
<i>n</i>	36	36
\bar{x}	0.81682	0.82410
<i>s</i>	0.007507	0.005701

SOLUCIÓN

REQUISITO ✓ En primer lugar, es evidente que las poblaciones son independientes entre sí. Los valores muestrales no están apareados o asociados de ninguna forma. En segundo lugar, las muestras sugieren que provienen de una población con una distribución aproximadamente normal. Consulte el párrafo anterior a este ejemplo y observe la gráfica cuantil normal y el histograma de los 36 pesos de Coca-Cola clásica. Los 36 pesos de Pepsi clásica se podrían explorar con una gráfica cuantil normal y un histograma, y los resultados sugerirían que estos pesos provienen de una población con distribución normal. Los requisitos se cumplen y podemos continuar con la prueba. ✓

En vez de utilizar las desviaciones estándar muestrales para probar la aseveración de desviaciones estándar poblacionales iguales, utilizamos las varianzas muestrales para probar la aseveración de varianzas poblacionales iguales, pero podemos plantear conclusiones en términos de desviaciones estándar. Puesto que estipulamos en esta sección que la varianza mayor se denota por s_1^2 , permitimos que $s_1^2 = 0.007507^2$, $n_1 = 36$, $s_2^2 = 0.005701^2$ y $n_2 = 36$. Ahora procedemos a utilizar el método tradicional de prueba de hipótesis como se describe en la figura 8-9.

- Paso 1: La aseveración de desviaciones estándar iguales es equivalente a una aseveración de varianzas iguales, lo cual se expresa simbólicamente como $\sigma_1^2 = \sigma_2^2$.
- Paso 2: Si la aseveración original es falsa, entonces $\sigma_1^2 \neq \sigma_2^2$.
- Paso 3: Puesto que la hipótesis nula es la afirmación de igualdad y dado que la hipótesis alternativa no puede contener igualdad, tenemos

$$H_0: \sigma_1^2 = \sigma_2^2 \quad (\text{aseveración original}) \quad H_1: \sigma_1^2 \neq \sigma_2^2$$

- Paso 4: El nivel de significancia es $\alpha = 0.05$.
- Paso 5: Puesto que esta prueba implica dos varianzas poblacionales, utilizamos la distribución *F*.
- Paso 6: El estadístico de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{0.007507^2}{0.005701^2} = 1.7339$$

En cuanto a los valores críticos, primero observe que se trata de una prueba de dos colas con 0.025 en cada cola. En tanto que estamos estipulando que la varianza más grande se coloca en el numerador para el estadístico de prueba *F*, necesitamos encontrar sólo el valor crítico de cola derecha. En la tabla de A-5 vemos que el valor crítico de *F* está entre 1.8752 y 2.0739, que obtenemos al remitirnos a 0.025 en la cola derecha, con 35 grados de libertad para el numerador y 35 grados de libertad para el denominador (STATDISK y Excel dan un valor crítico de 1.9611).

continúa



Gemelos en Twinsburg

Durante el primer fin de semana de agosto de cada año, Twinsburg, Ohio celebra su festival anual "Día de los gemelos en Twinsburg". Miles de gemelos de todo el mundo han acudido a este festival en el pasado. Los científicos consideraron el festival como una oportunidad para estudiar a gemelos idénticos. Puesto que tienen la misma estructura genética básica, los gemelos idénticos son ideales para estudiar los distintos efectos de la herencia y del ambiente en una gran variedad de rasgos, como la calvicie masculina, las enfermedades cardíacas y la sordera (rasgos que se estudiaron recientemente en uno de los festivales de Twinsburg). Un estudio de gemelos demostró que la miopía se ve muy afectada por factores hereditarios y no por factores ambientales como ver televisión, navegar en Internet o participar en juegos de video o de computadora.

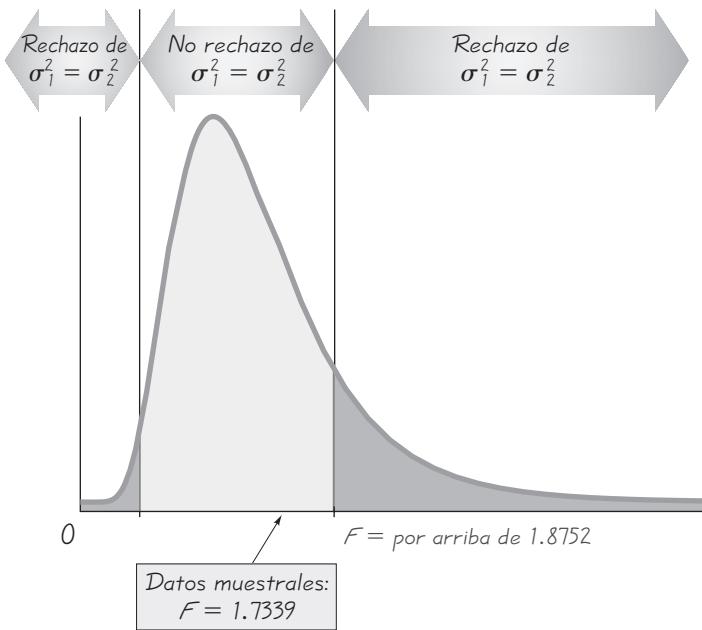


Figura 9-6 Distribución de s_1^2/s_2^2 para pesos de Coca-Cola clásica y Pepsi clásica

Paso 7: La figura 9-6 indica que el estadístico de prueba $F = 1.7339$ no se localiza dentro de la región crítica, por lo tanto, no rechazamos la hipótesis nula de varianzas iguales. Se deduce que no existe evidencia suficiente para sustentar el rechazo de la aseveración de desviaciones estándar iguales.

INTERPRETACIÓN No existe suficiente evidencia para justificar el rechazo de la aseveración de que las dos desviaciones estándar son iguales. Sin embargo, debemos reconocer que la prueba F es sumamente sensible a distribuciones que no son normales, de manera que podría parecer que esta conclusión indica que no existe una diferencia significativa entre las desviaciones estándar poblacionales, cuando realmente existe una diferencia que quedó oculta debido a distribuciones no normales.

El valor de 0.7901 de Coca es un valor extremo potencial, especialmente porque su puntuación z es -3.56 . Si repitiéramos la prueba sin este valor cuestionable, concluiríamos una vez más que no existe una diferencia significativa entre las dos desviaciones estándar poblacionales. Los métodos alternativos que se analizan más adelante en esta sección también nos conducen a la conclusión de que no existe una diferencia significativa entre las dos desviaciones estándar poblacionales.

Ahora utilicemos un poco de sentido común básico. Sabemos que las latas de Coca y Pepsi provienen de dos procesos de fabricación completamente separados e independientes, de manera que es poco probable que las dos varianzas poblacionales sean exactamente iguales. No obstante, con base en nuestro análisis, podemos concluir que cualquier diferencia entre las dos desviaciones estándar poblacionales no es significativa.

En el ejemplo anterior utilizamos pruebas de dos colas para la aseveración de varianzas iguales. Una prueba de cola derecha produciría el mismo estadístico de prueba $F = 1.7339$, pero un valor crítico de F diferente.

Prueba del valor P e intervalos de confianza Ya describimos el método tradicional para utilizar la prueba F con aseveraciones acerca de dos varianzas poblacionales. Es fácil utilizar el método del valor P con programas de cómputo capaces de dar valores P . En el ejemplo anterior, STATDISK, Excel, Minitab y la calculadora TI-83/84 Plus dan un valor P de 0.1081. El ejercicio 24 se refiere a la construcción de intervalos de confianza.

Métodos alternativos

Ya presentamos la prueba F para comparar varianzas, pero esa prueba es muy sensible a distribuciones alejadas de la normalidad. A continuación se describen brevemente algunas alternativas que son más robustas.

Conteo de cinco El método del *conteo de cinco* es una alternativa relativamente sencilla a la prueba F , y no requiere de poblaciones con distribución normal. (Véase “A Quick, Compact, Two-Sample Dispersion Test: Count Five”, de McGrath y Yeh, *American Statistician*, vol. 59, núm. 1). Si los dos tamaños muestrales son iguales, y una de las muestras tiene al menos cinco de las desviaciones medias absolutas (DMA) más grandes, entonces concluimos que su población tiene una varianza mayor. Véase el ejercicio 21, donde se describe el procedimiento específico.

Prueba Levene-Brown-Forsythe La *prueba Levene-Brown-Forsythe* (o prueba de Levene modificada) es otra alternativa a la prueba F y es mucho más robusta. Esta prueba comienza con una transformación de cada conjunto de valores muestrales. Dentro de la primera muestra, reemplace cada valor x por $|x - \text{mediana}|$ y haga lo mismo en la segunda muestra. Con los valores transformados, realice una prueba t de igualdad de medias para muestras independientes, como se describió en la parte 1 de la sección 9-3. Puesto que los valores transformados ahora son desviaciones, la prueba t para igualdad de medias es en realidad una prueba que compara la variación en las dos muestras. Consulte el ejercicio 22.

Además de la prueba del conteo de cinco y de la prueba Levene-Brown-Forsythe, existen otras alternativas a la prueba F , así como ajustes que mejoran el desempeño de la prueba F . Véase “Fixing the F Test for Equal Variances”, de Shoemaker, *American Statistician*, vol. 57, núm. 2.

Uso de la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, después seleccione **Hypothesis Testing** o **Confidence Intervals** y luego **StDev-Two Samples**. Ingrese los elementos requeridos en el cuadro de diálogo y luego haga clic en el botón de **Evaluate**.

MINITAB Obtenga el resumen de estadísticos de ambas muestras o ingrese todos los datos muestrales individuales en dos columnas. (Si está utilizando Minitab Release 13 u otra versión anterior, debe usar la lista original de datos en bruto). Seleccione **Stat**, luego **Basic Statistics** y después **2 Variances**; aparecerá un cuadro de diálogo: seleccione la opción de “Samples in different columns”, proceda a ingresar los nombres de las columnas o seleccione “Summarized data”; ingrese el resumen de estadísticos (si está utilizando Minitab Release 14 una versión posterior). Haga clic en el botón de **Options** e ingrese a nivel de confianza. (Ingrese 0.95 para una prueba de hipótesis con un nivel de significancia de 0.05). Haga clic en **OK** y luego clic en **OK** en el cuadro de diálogo principal.

EXCEL Primero ingrese los datos de la primera muestra en la primera columna A,

luego ingrese los valores de la segunda muestra en la columna B. Seleccione **Tools**, **Data Analysis** y luego **F-Test Two-Sample for Variances**. En el cuadro de diálogo, ingrese el rango de valores para la primera muestra (por ejemplo A1:A36) y el rango de valores para la segunda muestra. Ingresa el valor del nivel de significancia en el cuadro “Alfa”. Excel dará el estadístico de prueba F , el valor P para el caso de una cola y el valor crítico F para el caso de una cola. Para una prueba de dos colas haga dos ajustes: ingrese $\alpha/2$ en vez de α y duplique el valor P dado por Excel.

TI-83/84 PLUS Oprima la tecla **STAT**, luego seleccione **TESTS** y después **2-SampFTEST**. Puede utilizar el resumen de estadísticos o los datos que se han ingresado como listas.



9-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Debilidad de la prueba F .** ¿Cuál es la principal debilidad de la prueba F que describimos en esta sección? Nombre dos alternativas que sirven para superar esa debilidad.
- Efectos de distribuciones no normales.** ¿Cuál es la consecuencia de utilizar la prueba F con muestras de poblaciones que tienen distribuciones no normales? Es decir, ¿en qué falla la prueba?
- Distribución F .** En el contexto de esta sección, ¿qué es la distribución F ?
- Distribución F .** Identifique dos propiedades diferentes de la distribución F .

Prueba de hipótesis de varianzas iguales. En los ejercicios 5 y 6, pruebe la aseveración dada. Utilice un nivel de significancia de $\alpha = 0.05$ y suponga que todas las poblaciones están distribuidas normalmente.

- Aseveración:** La población de tratamiento y la población placebo tienen varianzas diferentes.

Grupo de tratamiento: $n = 16, \bar{x} = 21.33, s = 0.80$
 Grupo placebo: $n = 41, \bar{x} = 25.34, s = 0.40$

- Aseveración:** Las puntuaciones de CI de estudiantes de estadística varían menos que las puntuaciones de CI de otros estudiantes.

Estudiantes de estadística: $n = 28, \bar{x} = 118, s = 10.0$
 Otros estudiantes: $n = 25, \bar{x} = 112, s = 12.0$

- Interpretación de resultados de pesos de Coca-Cola clásica y Coca-Cola dietética.**

En esta sección se incluyó un ejemplo de una prueba de hipótesis para la aseveración de que los pesos de Coca clásica y Pepsi clásica tienen la misma desviación estándar. Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que la Coca clásica y la Coca dietética tienen pesos con desviaciones estándar diferentes. Los pesos muestrales se encuentran en el conjunto de datos 12 del apéndice B, pero consulte la imagen que aparece al margen con los resultados de la prueba F obtenidos con una calculadora TI-83/84 Plus. Si los resultados mostraran que las desviaciones estándar son significativamente diferentes, ¿cuál sería un factor importante que explicaría la diferencia?

- Interpretación de resultados de una prueba de equinácea.** En una prueba aleatorizada, doble ciego y controlada con placebo se puso a prueba la equinácea como tratamiento para las infecciones de vías respiratorias superiores en niños. Los “días con fiebre” fue uno de los criterios utilizados para medir los efectos. En 337 niños tratados con equinácea, el número medio de días con fiebre fue de 0.81, con una desviación estándar de 1.50 días. En 370 niños que recibieron un placebo, el número medio de días con fiebre fue de 0.64, con una desviación estándar de 1.16 días (según datos de “Efficacy and Safety of Echinacea in Treating Upper Respiratory Tract Infections in Children”, de Taylor *et al.*, *Journal of the American Medical Association*, vol. 290, núm. 21). Utilice un nivel de significancia de 0.05 para probar la aseveración de que los niños tratados con equinácea tienen una desviación estándar más grande que los que recibieron el placebo. Consulte la imagen adjunta de Excel.

TI-83/84 Plus

2-SampFTest
 $\sigma_1 \neq \sigma_2$
 $F=2.923274259$
 $P=.0020697599$
 $Sx_1=.007507372$
 $Sx_2=.004390896$
 $\bar{x}_1=.8168222222$

Excel

F.Test Two-Sample for Variances		
	Variable 1	Variable 2
4 Mean	0.043962632	-0.066075532
5 Variance	2.250000402	1.345600694
6 Observations	337	370
7 df	336	369
8 F	1.672115965	
9 P(F<=f) one-tail	7.47504E-07	
10 F Critical one-tail	1.191519202	

9. Pesos de Coca-Cola dietética y Pepsi dietética. En esta sección se incluyó un ejemplo de una prueba de hipótesis para la aseveración de que los pesos de Coca clásica y Pepsi clásica tienen la misma desviación estándar. Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que la Coca dietética y la Pepsi dietética tienen pesos con desviaciones estándar diferentes. Los pesos muestrales se encuentran en el conjunto de datos 12 del apéndice B, pero aquí tenemos el resumen de estadísticos: Coca de dieta: $n = 36$, $\bar{x} = 0.784794$ lb, $s = 0.004391$ lb; Pepsi de dieta: $n = 36$, $\bar{x} = 0.783858$ lb, $s = 0.004362$ lb.

- 10. Tratamiento de depresión bipolar.** En experimentos clínicos con diferentes grupos de muestras independientes, es importante que los grupos sean similares en los aspectos relevantes que afectan el experimento. En un experimento diseñado para probar la eficacia de la paroxetina en el tratamiento de la depresión bipolar, a los sujetos se les midió la depresión utilizando la escala de Hamilton, con los resultados que se presentan abajo (según datos de “Double-Blind, Placebo-Controlled Comparison of Imipramine and Paroxetine in the Treatment of Bipolar Depression”, de Nemeroff *et al.*, *American Journal of Psychiatry*, vol. 158, núm. 6). Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que ambas poblaciones tienen la misma desviación estándar. Con base en los resultados, ¿parece que las dos poblaciones tienen desviaciones estándar diferentes?

Grupo placebo: $n = 43$, $\bar{x} = 21.57$, $s = 3.87$

Grupo de tratamiento con paroxetina: $n = 33$, $\bar{x} = 20.38$, $s = 3.91$

- 11. Prueba de hipótesis para el tratamiento magnético del dolor.** Investigadores realizaron un estudio para determinar si los magnetos son eficaces en el tratamiento del dolor de espalda, con los resultados que se dan a continuación (según datos de “Bipolar Permanent Magnets for the Treatment of Chronic Lower Back Pain: A Pilot Study”, de Collacott, Zimmerman, White y Rindone, *Journal of the American Medical Association*, vol. 283, núm. 10). Los valores representan mediciones del dolor con la escala análoga visual. Utilice un nivel de significancia de 0.05 para probar la aseveración de que quienes recibieron un tratamiento simulado (similar a un placebo) presentan reducciones del dolor que varían más que las reducciones del dolor en quienes recibieron el tratamiento con magnetos.

Reducción en el nivel de dolor después del tratamiento simulado: $n = 20$, $\bar{x} = 0.44$, $s = 1.4$

Reducción en el nivel de dolor después del tratamiento magnético: $n = 20$, $\bar{x} = 0.49$, $s = 0.96$

- 12. Prueba de hipótesis para el efecto del consumo de marihuana en estudiantes universitarios.** En un estudio sobre los efectos del consumo de marihuana en una universidad se probó la capacidad de memoria de consumidores ocasionales y frecuentes de marihuana, con los resultados que se dan a continuación (según datos de “The Residual Cognitive Effects of Heavy Marijuana Use in College Students” de Pope y Yurgelun-Todd, *Journal of the American Medical Association*, vol. 275, núm. 7). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la población de consumidores frecuentes de marihuana tiene una desviación estándar diferente de la de los consumidores ocasionales.

Artículos acomodados correctamente por consumidores ocasionales de marihuana: $n = 64$, $\bar{x} = 53.3$, $s = 3.6$

Artículos acomodados correctamente por consumidores frecuentes de marihuana: $n = 65$, $\bar{x} = 51.3$, $s = 4.5$

- 13. Efectos del alcohol.** Se realizó un experimento para probar los efectos del alcohol. Se registraron los errores en una prueba de destrezas visuales y motrices de un grupo de tratamiento de personas que bebieron etanol y de otro grupo que recibió un placebo. Los resultados se muestran en la tabla adjunta (según datos de “Effects of Alcohol Intoxication on Risk Taking, Strategy, and Error Rate in Visuomotor Performance”, de Streufert *et al.*, *Journal of Applied Psychology*, vol. 77, núm. 4). Utilice un nivel de significancia de 0.05 para probar la aseveración de que el grupo de tratamiento tiene puntuaciones que varían más que las puntuaciones de grupo placebo.

Grupo de tratamiento	Grupo placebo
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 4.20$	$\bar{x}_2 = 1.71$
$s_1 = 2.20$	$s_2 = 0.72$

- 14. Antigüedad de automóviles de profesores y estudiantes.** Los estudiantes en la universidad del autor seleccionaron al azar 217 automóviles de estudiantes y encontraron que tienen antigüedades con una media de 7.80 años y una desviación estándar de 3.67 años. También seleccionaron al azar 152 automóviles de profesores y encontraron que tenían antigüedades con una media de 5.99 años y una desviación estándar de 3.65 años. ¿Existe evidencia suficiente para sustentar la aseveración de que las antigüedades de los automóviles de los profesores varían menos que las antigüedades de los automóviles de los estudiantes?
- 15. Pesos de monedas de 25 centavos.** Las máquinas expendededoras utilizan los pesos de las monedas de 25 centavos para detectar monedas falsas. El conjunto de datos 14 del apéndice B incluye los pesos de monedas de plata de 25 centavos acuñadas antes y después de 1964. El resumen de estadísticos es el siguiente: antes de 1964: $n = 40$, $\bar{x} = 6.19267$ g, $s = 0.08700$ g; después de 1964: $n = 40$, $\bar{x} = 5.63930$ g, $s = 0.06194$. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las dos poblaciones de monedas de 25 centavos tienen la misma desviación estándar. Si las cantidades de variación difieren, quizás las máquinas expendededoras necesiten ajustes más complicados. ¿Parece que este tipo de ajustes son necesarios?
- 16. Pesos de monedas de un centavo y de 25 centavos.** El conjunto de datos 14 del apéndice B incluye pesos de monedas de un centavo acuñadas después de 1983 y de monedas de 25 centavos acuñadas después de 1964. El resumen de estadísticos es el siguiente: monedas de un centavo después de 1983: $n = 37$, $\bar{x} = 2.49910$ g, $s = 0.01648$ g; monedas de 25 centavos después de 1964: $n = 40$, $\bar{x} = 5.63930$, $s = 0.06194$. Ponga a prueba la aseveración de que las monedas de un centavo acuñadas después de 1983 y las monedas de 25 centavos acuñadas después de 1964 tienen la misma cantidad de variación. ¿Deberían tener la misma cantidad de variación?
- 17. Bloqueo en exámenes.** Muchos estudiantes han tenido la experiencia poco placentera de sentir pánico en exámenes porque la primera pregunta era excepcionalmente difícil. Se estudió el orden de las preguntas de exámenes y sus efectos en la ansiedad. Se obtuvieron valores muestrales que consisten en mediciones de “ansiedad debilitante por exámenes” (que la mayoría de nosotros llamamos pánico o bloqueo) de un grupo de sujetos a quienes se les presentaron preguntas de examen ordenadas de fácil a difícil, y de otro grupo con preguntas de examen ordenadas de difícil a fácil. El resumen de estadísticos es el siguiente: grupo de fácil a difícil: $n = 25$, $\bar{x} = 27.115$, $s = 47.020$; grupo de difícil a fácil: $n = 16$, $\bar{x} = 31.728$, $s = 18.150$ (según datos de “Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance”, de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4). Ponga a prueba la aseveración de que las dos muestras provienen de poblaciones con la misma varianza.
- 18. Efecto del peso al nacer sobre puntuaciones de CI.** Al investigar la relación entre el peso al nacer y el CI, los investigadores encontraron que 258 sujetos con pesos sumamente bajos al nacer (menos de 1000 g), a los ocho años tuvieron puntuaciones de CI en la prueba Wechsler con una media de 95.5 y una varianza de 256.0. En el caso de los 220 sujetos con pesos normales al nacer, la media a los ocho años fue de 104.9, con una varianza de 198.8 (según datos de “Neurobehavioral Outcomes of School-age Children Born Extremely Low Birth Weight or Very Preterm in the 19002”, de Anderson *et al.*, *Journal of the American Medical Association*, vol. 289, no. 24). Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que los bebés con pesos extremadamente bajos al nacer y los bebés con pesos normales al nacer tienen cantidades diferentes de variación. (*Sugerencia:* Tome en cuenta que la conclusión no se aclara con el uso de la tabla A-5; utilice el siguiente valor crítico superior: $F = 1.2928$).
- 19. Conjunto de datos del apéndice B: precipitación pluvial en fines de semana.** El *USA Today* y otros periódicos reportaron un estudio que pretendía demostrar que llueve más durante los fines de semana. El estudio se refería a áreas en la región oriental de Estados Unidos, cerca del océano. El conjunto de datos 10 del apéndice B lista las cantidades de lluvia en Boston en un año.

- a. Suponiendo que queremos utilizar los métodos de esta sección para probar la aseveración de que las cantidades de precipitación pluvial del miércoles y del domingo tienen la misma desviación estándar, calcule el estadístico de prueba F , el valor crítico y plantee la conclusión. Utilice un nivel de significancia de 0.05.
 - b. Considere el prerrequisito de poblaciones distribuidas normalmente. En vez de construir histogramas o gráficas cuantilares normales, examine simplemente el número de días sin lluvia. ¿Las cantidades de lluvia del miércoles se distribuyen normalmente? ¿Las cantidades de lluvia del domingo se distribuyen normalmente?
 - c. ¿Qué se concluye de los resultados de los incisos a) y b)?
- 20. Conjunto de datos del apéndice B: consumo de tabaco y alcohol en películas infantiles de dibujos animados.** El conjunto de datos 5 del apéndice B lista tiempos (en segundos) en los que las películas de dibujos animados para niños presentan consumo de tabaco y consumo de alcohol.
- a. Suponiendo que queremos utilizar los métodos de esta sección para probar la aseveración de que los tiempos de consumo de tabaco y los tiempos de consumo de alcohol tienen desviaciones estándar diferentes, calcule el estadístico de prueba F , el valor crítico y plantee la conclusión. Utilice un nivel de significancia de 0.05.
 - b. Al parecer, ¿los datos satisfacen los requisitos de poblaciones independientes? Explique.
 - c. ¿Los datos parecen satisfacer los requisitos de poblaciones con distribución normal? En vez de construir histogramas o gráficas cuantilares normales, examine simplemente el número de películas que no presentan consumo de tabaco o alcohol. ¿Están distribuidos normalmente los tiempos de consumo de tabaco? ¿Están distribuidos normalmente los tiempos de consumo de alcohol?
 - d. ¿Qué se puede concluir de los resultados anteriores?

9-5 MÁS ALLÁ DE LO BÁSICO

- 21. Prueba del conteo de cinco para comparar la variación en dos poblaciones.** Consulte el ejemplo en esta sección y, en vez de utilizar la prueba F , utilice el siguiente procedimiento para una prueba de “conteo de cinco” de una igualdad de variación. ¿Qué concluye usted?

- a. Para la primera muestra, calcule la desviación media absoluta (DMA) de cada valor. (Recuerde que en la sección 3-3 aprendimos que la DMA de un valor muestral x es $|x - \bar{x}|$. Ordene los valores de la DMA. Haga lo mismo con la segunda muestra.)
- b. Permita que c_1 sea el conteo del número de valores DMA en la primera muestra que son mayores que el valor DMA más grande en la otra muestra. También permita que c_2 sea el conteo del número de valores DMA en la segunda muestra que son mayores que el valor DMA más grande en la otra muestra. (Uno de estos conteos siempre será igual a cero).
- c. Si los tamaños muestrales son iguales ($n_1 = n_2$), utilice el valor crítico de 5. Si $n_1 \neq n_2$, calcule el valor crítico que se muestra a continuación

$$\frac{\log(\alpha/2)}{\log\left(\frac{n_1}{n_1 + n_2}\right)}$$

- d. Si $c_1 \geq$ el valor crítico, entonces concluya que $\sigma_1^2 > \sigma_2^2$. Si $c_2 \geq$ el valor crítico, entonces concluya que $\sigma_2^2 > \sigma_1^2$. De cualquier otra forma, no rechace la hipótesis nula de $\sigma_1^2 = \sigma_2^2$.
- 22. Prueba Levene-Brown-Forsythe para comparar la variación en dos poblaciones.** Consulte el ejemplo de esta sección y, en vez de utilizar la prueba F , utilice la prueba Levene-Brown-Forsythe descrita en la parte final de esta sección. ¿Qué concluye usted?
- 23. Cálculo de valores críticos F inferiores.** En esta sección calculamos sólo el valor crítico superior para pruebas de hipótesis de dos colas. Denotemos este valor por F_D , donde el subíndice indica el valor crítico para la cola derecha. El valor crítico inferior F_I (para la cola izquierda) puede calcularse como sigue: primero intercambie los

grados de libertad y después tome el recíproco de valor F resultante encontrado en la tabla A-5. Calcule los valores críticos F_I y F_D para pruebas de hipótesis de dos colas basadas en los siguientes valores.

- a. $n_1 = 10, n_2 = 10, \alpha = 0.05$
- b. $n_1 = 10, n_2 = 7, \alpha = 0.05$
- c. $n_1 = 7, n_2 = 10, \alpha = 0.05$

- 24. Construcción de intervalos de confianza.** Además de probar aseveraciones que implican a σ_1^2 y σ_2^2 , también podemos construir estimados del intervalo de confianza de la proporción σ_1^2/σ_2^2 , utilizando la siguiente expresión.

$$\left(\frac{s_1^2}{s_2^2} \cdot \frac{1}{F_D} \right) < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2} \cdot \frac{1}{F_I} \right)$$

Aquí F_D y F_I son tal como se describe en el ejercicio 23. Remítase a los datos muestrales del ejercicio 5 y construya un estimado del intervalo de confianza del 95% para la proporción de la varianza del grupo de tratamiento con respecto a la varianza del grupo placebo.

Repaso

Cuando utilizamos datos muestrales para obtener conclusiones acerca de poblaciones, empleamos métodos de estadística inferencial. Dos de las principales actividades de la estadística inferencial son: **1.** la construcción de estimados de intervalos de confianza de parámetros poblacionales (como p , μ , σ) y **2.** la prueba de hipótesis o aseveraciones acerca de parámetros poblacionales. En los capítulos 7 y 8 analizamos la estimación de parámetros poblacionales y los métodos de prueba de hipótesis acerca de parámetros poblacionales; sin embargo, en los capítulos 7 y 8 sólo se consideraron casos con una sola población. En este capítulo consideramos dos muestras obtenidas de dos poblaciones.

- En la sección 9-2 se consideraron inferencias acerca de dos proporciones poblacionales. Dadas las condiciones en las que los requisitos listados se satisfacen, utilizamos la distribución normal para construir estimados del intervalos de confianza de la diferencia $p_1 - p_2$ y para probar aseveraciones, como la de que $p_1 = p_2$.
- En la sección 9-3 se consideraron inferencias acerca de las medias de dos poblaciones independientes. La sección 9-3 incluyó tres métodos diferentes, pero un método se utiliza en raras ocasiones puesto que requiere que las dos desviaciones estándar poblacionales se conozcan. Otro método implicó el agrupamiento de las dos desviaciones estándar muestrales para desarrollar un estimado del error estándar; sin embargo, este método se basa en el supuesto de que las dos desviaciones estándar poblacionales son iguales, y esta suposición suele ser riesgosa. Vea la figura 9-3 para determinar cuál método aplicar. El procedimiento que generalmente se recomienda es la prueba t , que no supone igualdad de las varianzas poblacionales.
- La sección 9-4 consideró inferencias acerca de la diferencia media para una población consistente en datos apareados.
- La sección 9-5 presentó la prueba F para poner a prueba aseveraciones acerca de la igualdad de dos desviaciones estándar o varianzas poblacionales. Es importante saber que la prueba F no es robusta, es decir, muestra un mal desempeño con poblaciones que no tienen una distribución normal. Se describieron brevemente alternativas a la prueba F .

Conocimientos estadísticos y pensamiento crítico

1. **¿Cuál método?** Un candidato de un puesto político está preocupado por los reportes de una “brecha de género”, que afirma que lo prefieren más los hombres que las mujeres. Usted ha sido contratado para investigar la brecha de género. ¿Qué métodos de este capítulo usaría?
2. **Muestra aleatoria simple.** Usted ha sido contratado para comparar la deuda media de crédito de los hombres en su estado con la deuda media de crédito de las mujeres

en su estado. Usted recibió muestras obtenidas mediante el siguiente proceso: primero se obtuvo una lista completa de los acreedores, después se utilizó una computadora para ordenar la lista de manera aleatoria, y luego se seleccionó una muestra aleatoria de 200 acreedores. ¿Este proceso de selección satisface el requisito de una muestra aleatoria simple? Explique.

- 3. Comparación de ingresos.** Con el uso de datos del Bureau of Labor Statistics, un investigador obtiene el ingreso medio de hombres y el ingreso medio de mujeres en cada una de las 50 entidades de Estados Unidos. Después realiza una prueba t de la hipótesis nula de que los hombres y las mujeres en Estados Unidos tienen ingresos medios iguales. ¿Es adecuado este procedimiento? ¿Por qué?
- 4. Muestras independientes.** ¿Qué diferencia hay entre dos muestras que son independientes y dos muestras que no lo son?

Ejercicios de repaso

- 1. Prácticas policiales discriminatorias.** Las prácticas policiales discriminatorias se refieren a la práctica polémica de acusar a alguien de posible conducta criminal con base en la raza, el país de origen o el origen étnico. En la tabla siguiente se incluyen datos de conductores elegidos al azar, detenidos por la policía en un año reciente (según datos del Departamento de Justicia de Estados Unidos, *Bureau of Justice Statistics*).
 - a. Utilice un nivel de significancia de 0.05 para poner a prueba la aseveración de que la proporción de afroestadounidenses detenidos por la policía es significativamente mayor que la proporción de caucásicos.
 - b. Construya un intervalo de confianza que se pueda utilizar para poner a prueba la aseveración planteada en el inciso a). Asegúrese de utilizar el nivel de significancia correcto. ¿Qué concluye usted con base en el intervalo de confianza?

Raza y origen étnico		
	Afroestadounidense y no hispano	Caucásico y no hispano
Conductores detenidos por la policía	24	147
Número total de conductores estudiados	200	1400

- 2. Estaturas reportadas y medidas de estudiantes de estadística varones.** Se aplicó una encuesta a 11 estudiantes de estadística varones, la cual incluía una pregunta en la que se les pedía reportar su estatura en pulgadas. No se les dijo que después se mediría su estatura, pero una vez terminada la encuesta se tomaron medidas exactas de su estatura. Se mantuvo el anonimato por medio del uso de códigos numéricos en vez de nombres, y los resultados se presentan a continuación. ¿Existe evidencia suficiente para sustentar la aseveración de que los estudiantes de estadística varones exageran su estatura?

Estatura reportada	68	74	66.5	69	68	71	70	70	67	68	70
Estatura medida	66.8	73.9	66.1	67.2	67.9	69.4	69.9	68.6	67.9	67.6	68.8

- 3. Comparación de facilidad de lectura de J. K. Rowling y León Tolstoi.** A continuación se presentan las puntuaciones Flesch de facilidad de lectura, tomadas de páginas elegidas al azar del libro *Harry Potter y la piedra filosofal* de J. K. Rowling y del libro *La guerra y la paz* de León Tolstoi. (Las puntuaciones Flesch más altas de facilidad de lectura indican que un escrito es más fácil de leer). Utilice un nivel de significancia de 0.05 para probar la aseveración de que es más fácil leer *Harry Potter y la piedra filosofal* de J. K. Rowling que *La guerra y la paz* de León Tolstoi.

Rowling: 85.3 84.3 79.5 82.5 80.2 84.6 79.2 70.9 78.6 86.2 74.0 83.7

Tolstoi: 69.4 64.2 71.4 71.6 68.5 51.9 72.2 74.4 52.8 58.4 65.4 73.6

- 4. Variación de J. K. Rowling y León Tolstoi.** Remítase a los datos del ejercicio 3 y utilice un nivel de significancia de 0.05 para probar la aseveración de que las páginas

de *Harry Potter y la piedra filosofal* de J. K. Rowling y de *La guerra y la paz* de León Tolstoi tienen puntuaciones Flesch de facilidad de lectura con la misma desviación estándar.

- 5. ¿Se recuperan mejor de una cirugía los pacientes cuyo cuerpo está tibio?** Un artículo publicado en *USA Today* afirmó que “en un estudio de 200 pacientes de cirugía rectal, a 104 se les mantuvieron tibios con mantas y líquidos intravenosos, y a 96 se les mantuvieron frescos. Los resultados indican que sólo 6 de quienes se mantuvieron tibios presentaron infecciones en la herida, a diferencia de 18 de los que se mantuvieron frescos”.

- Utilice un nivel de significancia de 0.05 para probar la aseveración del encabezado del artículo: “Los pacientes quirúrgicos se recuperan mejor si su cuerpo está tibio”. Si estos resultados se confirman, ¿deberá entibiararse por rutina a los pacientes que se sometieron a cirugía?
- Si se utilizara un intervalo de confianza para probar la aseveración del inciso a), ¿qué nivel de confianza debería utilizarse?
- Utilice el nivel de confianza del inciso b) y construya un estimado del intervalo de confianza de la diferencia entre las dos proporciones poblacionales.
- En general, si se utiliza un estimado del intervalo de confianza de la diferencia entre las dos proporciones poblacionales para probar alguna aseveración acerca de las proporciones, ¿la conclusión basada en el intervalo de confianza será siempre la misma que la conclusión de una prueba de hipótesis estándar?

- 6. Volumen cerebral y trastornos psiquiátricos.** Un estudio utilizó tomografía computarizada por rayos X (TC) para reunir datos de volúmenes cerebrales de un grupo de pacientes con trastorno obsesivo-compulsivo y de un grupo de control de personas saludables. Abajo se presentan los resultados muestrales (en mL) de los volúmenes cerebrales totales (según datos de “Neuroanatomical Abnormalities in Obsessive-Compulsive Disorder Detected with Quantitative X-Ray Computed Tomography”, de Luxemberg, *et al.*, *American Journal of Psychiatry*, vol. 145, núm. 9).

- Construya un intervalo de confianza del 95% para la diferencia entre el volumen cerebral medio de los pacientes obsesivo-compulsivos y el volumen cerebral medio de las personas saludables. No suponga que las dos poblaciones tienen varianzas iguales.
- Utilice un nivel de significancia de 0.05 para probar la aseveración de que no existe diferencia entre la media de los pacientes obsesivo-compulsivos y la media de las personas saludables. No suponga que las dos poblaciones tienen varianzas iguales.
- Con base en los resultados de los incisos a) y b), ¿parece que el volumen cerebral total puede utilizarse como un indicador del trastorno obsesivo compulsivo?

Pacientes obsesivo-compulsivos: $n = 10, \bar{x} = 1390.03, s = 156.84$

Grupo de control: $n = 10, \bar{x} = 1268.41, s = 137.97$

- 7. Variación de volúmenes cerebrales.** Utilice los datos muestrales del ejercicio 6, con un nivel de significancia de 0.05, para probar la aseveración de que las poblaciones de volúmenes cerebrales totales de los pacientes obsesivo-compulsivos y del grupo de control tienen diferentes cantidades de variación.

- 8. Conjunto de datos históricos.** En 1908 “Student” (William Gosset) publicó el artículo “The Probable Error of a Mean” (*Biometrika*, vol. 6, núm. 1). Él incluyó los datos listados abajo para dos tipos diferentes de semillas de paja (común y secada al horno) que se utilizaron en terrenos adyacentes. Los valores listados son las cosechas de paja en cwt por acre. (cwt es una unidad de masa británica, que significa *c-weight*, y equivale a 100 libras).

- Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que no existe diferencia entre las cosechas de los dos tipos de semilla.

- b.** Construya un estimado del intervalo de confianza de 95% de la diferencia media entre las cosechas de los dos tipos de semillas.
c. ¿Parece que algún tipo de semilla es mejor?

Común	19.25	22.75	23	23	22.5	19.75	24.5	15.5	18	14.25	17
Secada al horno	25	24	24	28	22.5	19.5	22.25	16	17.25	15.75	17.25

Ejercicios de repaso acumulativo

- 1. Multas por exceso de velocidad en carretera.** Una sección de la carretera 405 en Los Ángeles tiene un límite de velocidad de 65 mi/h; a continuación se presentan las velocidades registradas de automóviles elegidos al azar que viajaban en los carriles hacia el norte y hacia el sur (según datos de Sigalert.com).
- Utilice las velocidades de los carriles que van hacia el norte, calcule la media, la mediana, la desviación estándar, la varianza y el rango.
 - Utilice todas las velocidades combinadas para probar la aseveración de que la media es mayor que la velocidad límite de 65 mi/h.
 - Al parecer, ¿las velocidades de los carriles que van hacia el norte provienen de una población distribuida normalmente? Explique.
 - Suponiendo que las velocidades provienen de poblaciones distribuidas normalmente, ponga a prueba la aseveración de que la velocidad media de los carriles que van hacia el norte es igual a la velocidad media de los carriles que van hacia el sur. Con base en el resultado del inciso c), ¿es probable que esta hipótesis sea válida?

Carretera 405 hacia el norte: 68 68 72 73 65 74 73 72 68 65 65 73 66 71 68 74 66
71 65 73

Carretera 405 hacia el sur: 59 75 70 56 66 75 68 75 62 72 60 73 61 75 58 74 60 73
58 75

- 2. Lanzamiento de monedas.** Un ilusionista asegura tener la capacidad de lanzar una moneda y que resulte una cara. A continuación se presentan los resultados de una prueba de sus habilidades.
- Considere únicamente los resultados de los lanzamientos de una moneda de 25 centavos. ¿Cuál es la probabilidad de obtener nueve caras en nueve lanzamientos, si los resultados están determinados sólo por el azar? ¿Qué sugiere el resultado acerca de la aseveración de que es posible lanzar una moneda de forma que siempre resulte cara? Explique.
 - ¿Los resultados de la moneda de 25 centavos son independientes de los resultados de la moneda de un centavo, o se trata de datos apareados? Explique.
 - Utilice todos los resultados combinados, con un nivel de significancia de 0.01, para probar la aseveración de que una moneda puede lanzarse de tal manera que resulten caras con mayor frecuencia de lo esperado por el azar.

Moneda de 25 centavos: C C C C C C C C C

Moneda de 1 centavo: C C X C X C C X X

- 3. Teléfonos celulares y choques: análisis de reporte de periódico.** En un artículo de la Associated Press, se reportó que unos investigadores “seleccionaron al azar a 100 conductores que habían estado involucrados en un accidente y a 100 que no. De aquellos que estuvieron involucrados en accidentes, el 13.7% tenían un teléfono celular, mientras que sólo el 10.6% de los conductores que no sufrieron accidentes tenían un teléfono en el automóvil”. Analice estos resultados.

Actividades de cooperación en equipo

- 1. Actividad fuera de clase** ¿Los estimados se ven afectados por números anclados? Remítase a la actividad de cooperación en grupos relacionada en el capítulo 3. En el capítulo 3 señalamos que, según el autor John Rubin, cuando las personas tienen que estimar un valor, su estimado suele estar “anclado” a (o influido por) un número anterior. En esa actividad del capítulo 3, se pidió a algunos sujetos que estimaran rápidamente el valor de $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$, y a otros se les pidió que estimaran rápidamente el valor de $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$. En el capítulo 3 pudimos comparar los dos conjuntos de resultados utilizando estadísticos (como la media) y gráficas (como las gráficas de cuadro). Ahora los métodos del capítulo 9 nos permiten comparar los resultados con una prueba formal de hipótesis. En específico, reúna sus propios datos muestrales y pruebe la aseveración de que cuando comenzamos con números mayores (como en $8 \times 7 \times 6$), nuestros estimados tienden a ser más grandes.
- 2. Actividad en clase** Dividan la clase en grupos de acuerdo al género, con alrededor de 10 o 12 estudiantes en cada grupo. Cada miembro del grupo deberá registrar su pulso contando el número de latidos en un minuto; luego se calcularán los estadísticos del grupo (n , \bar{x} y s). Los grupos deben poner a prueba la hipótesis nula de no diferencia entre sus pulsos medios y la media del pulso de la población de la que se seleccionaron los sujetos del mismo género para el conjunto de datos 1 en el apéndice B.
- 3. Actividad fuera de clase** Seleccione al azar una muestra de estudiantes varones y una muestra de estudiantes mujeres y pregunte a cada persona seleccionada si apoya la pena de muerte para personas sentenciadas por homicidio. Utilice una prueba formal de hipótesis para determinar si existe una diferencia de género en este tema. Además, haga un registro de las respuestas de acuerdo al género de la persona que realiza las preguntas. ¿Parece que la respuesta se ve influida por el género del entrevistador?
- 4. Actividad fuera de clase** Utilice un reloj para registrar los tiempos de espera de una muestra de clientes de McDonald's y los tiempos de espera de una muestra de clientes de Burger King. Realice una prueba de hipótesis para determinar si existe una diferencia significativa.
- 5. Actividad fuera de clase** Elabore una encuesta breve, de sólo unas cuantas preguntas, incluyendo una pregunta que pida al sujeto que reporte su estatura. Cuando el sujeto termine de responder la encuesta, mida su estatura (sin zapatos) por medio de un sistema de medición exacto. Registre el género, la estatura reportada y la estatura medida de cada sujeto. (Véase el ejercicio de repaso 2). ¿Parece que los hombres exageran su estatura? ¿Parece que las mujeres exageran su estatura? ¿Parece que los errores de los hombres tienen la misma media que los errores de las mujeres?
- 6. Actividad en clase** Sin utilizar un instrumento de medición, cada estudiante debe dibujar una línea que, según sus cálculos, mida 3 pulgadas de largo. Luego, utilice reglas para medir y registrar las longitudes de las líneas dibujadas. Ponga a prueba una diferencia entre la longitud media de las líneas dibujadas por los hombres y la longitud media de las líneas dibujadas por las mujeres.
- 7. Actividad en clase** Utilice una regla para medir el tiempo de reacción. Una persona debe suspender la regla sosteniéndola de un extremo, mientras el sujeto coloca sus dedos pulgar e índice en el extremo inferior, preparado para atrapar la regla cuando sea soltada. Registre la distancia que cae la regla antes de ser atrapada. Convierta esa distancia en el tiempo (en segundos) que tardó el sujeto en reaccionar y atrapar la regla. (Si la distancia se mide en pulgadas, utilice $t = \sqrt{d/192}$. Si la distancia se mide en centímetros, utilice $t = \sqrt{d/487.68}$). Pruebe a cada sujeto una vez con la mano dominante y una vez con la otra mano, y luego registre los datos apareados. Al parecer, ¿existe una diferencia entre la media de los tiempos de reacción cuando se usa la mano dominante y la media de la otra mano? ¿Los hombres y las mujeres parecen tener diferentes tiempos de reacción?

Proyecto tecnológico

Los programas STATDISK, Minitab y Excel, así como la calculadora TI-83/84 Plus pueden generar datos distribuidos normalmente, obtenidos de una población con una media y una desviación estándar específicas. Genere dos conjuntos de datos muestrales que representen puntuaciones de CI simuladas, como se muestra abajo.

Puntuaciones de CI del grupo de tratamiento: Genere 10 valores muestrales a partir de una población distribuida normalmente con media 100 y desviación estándar 15.

Puntuaciones de CI del grupo placebo: Genere 12 valores muestrales de una población distribuida normalmente, con media 100 y desviación estándar 15.

STATDISK: Seleccione **Data**, luego seleccione **Normal Generator**.

Minitab: Seleccione **Calc, Random Data, Normal**.

Excel: Seleccione **Tools, Data Analysis, Random Number Generator**, y asegúrese de seleccionar **Normal** para la distribución.

TI-83/84 Plus: Oprima **MATH**, seleccione **PRB** y luego **6:randNorm(** y proceda a ingresar la media, la desviación estándar y el número de puntuaciones (como 100, 15, 10).

Usted puede ver, por la forma en que los datos se generan, que ambos conjuntos de datos provienen de la misma población, por lo que no debería existir ninguna diferencia entre las dos medias muestrales.

- a. Después de generar los dos conjuntos de datos, utilice un nivel de significancia de 0.10 para probar la aseveración de que las dos muestras provienen de poblaciones con la misma media.
- b. Si este experimento se repitiera muchas veces, ¿cuál sería el porcentaje esperado de ensayos que nos lleven a la conclusión de que las dos medias poblacionales son diferentes? ¿Cómo se relaciona esto con un error tipo I?
- c. Si sus datos generados lo llevaran a la conclusión de que las dos medias poblacionales son diferentes, ¿sería en realidad correcta o incorrecta esta conclusión? ¿Cómo lo sabe?
- d. Si el inciso a) se repitiera 20 veces, ¿cuál es la probabilidad de que ninguna de las pruebas de hipótesis nos lleve al rechazo de la hipótesis nula?
- e. Repita el inciso a) 20 veces. ¿Con qué frecuencia se rechazó la hipótesis nula de medias iguales? ¿Es éste el resultado que usted esperaba?

De los datos a la decisión

Pensamiento crítico: El temor de viajar en avión

Las vidas de muchas personas se ven afectadas por un temor que les impide viajar en avión. El comentarista de deportes John Madden ganó popularidad cuando cruzaba el país en ferrocarril o en un remolque vienda, viajando de un estadio de fútbol a otro. El Marist Institute for Public Opinion realizó una encuesta a 1014 adultos, 48% de los cuales eran varones. Los resultados reportados en *USA Today* indican que el 12% de los hombres y el 33% de las mujeres temen viajar en avión.

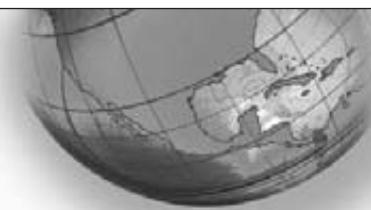
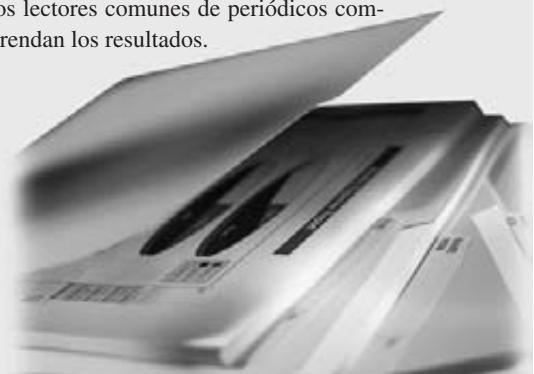
Análisis de los resultados

1. ¿A cuántos hombres se encuestó? ¿A cuántas mujeres se encuestó? ¿Cuántos de los hombres encuestados temen viajar en avión? ¿Cuántas de las mujeres encuestadas temen viajar en avión?
2. ¿Existe evidencia suficiente para concluir que hay una diferencia significativa entre el porcentaje de hombres y el porcentaje de mujeres que temen viajar en avión?

3. Construya un estimado del intervalo de confianza del 95% de la diferencia entre el porcentaje de hombres y el porcentaje de mujeres que temen viajar en avión. ¿Los límites del intervalo de confianza contienen a 0, y cuál es la significancia de si lo contienen o no?
4. Construya un intervalo de confianza del 95% para el porcentaje de hombres que temen viajar en avión.
5. Con base en el resultado del intervalo de confianza obtenido en el ejercicio 4, complete la siguiente afirmación, que es típica de las declaraciones reportadas en un periódico o en una revista: "Con base en la encuesta del Marist Institute for Public Opinion, el porcentaje de hombres que temen viajar en avión es del 22%, con un margen de error de ____".
6. Examine la aseveración completa del ejercicio 5. ¿Qué parte importante de la información debería incluirse, pero no está incluida?
7. En una encuesta separada de Gallup, se pidió a 1001 adultos seleccionados al azar que respondieran esta pregunta: "Si usted tuviera que viajar mañana en avión, ¿cómo

describiría sus sensaciones acerca del viaje? ¿Estaría usted muy temeroso, un poco temeroso, no muy temeroso o sin temor alguno?" He aquí las respuestas: muy temeroso (18%), un poco temeroso (26%), no muy temeroso (17%), sin temor alguno (38%) y sin opinión (1%). ¿Son consistentes estos resultados de la encuesta de Gallup con los que se obtuvieron de la encuesta realizada por el Marist Institute for Public Opinion? Explique. ¿Se pueden explicar las discrepancias por el hecho de que la encuesta de Gallup se realizó después de los ataques terroristas del 11 de septiembre de 2001, mientras que la otra encuesta se realizó antes de esa fecha?

8. Construya una gráfica que permita que los lectores comunes de periódicos comprendan los resultados.



Proyecto de Internet

Comparación de poblaciones

El capítulo anterior le presentó métodos para realizar pruebas de hipótesis acerca de una sola población. Este capítulo amplía esas ideas, permitiéndole probar hipótesis acerca de las relaciones entre dos poblaciones. De forma similar, el Proyecto de Internet para este capítulo difiere del proyecto del capítulo anterior en que usted necesitará datos de dos poblaciones o grupos para realizar las investigaciones. Consulte el proyecto de Internet para este capítulo en

Ahí encontrará varios problemas de prueba de hipótesis que implican múltiples poblaciones. En esos problemas usted analizará la imparcialidad de los salarios, la demografía y la superstición tradicional. En cada caso formule el problema como una prueba de hipótesis, reúna datos relevantes y después realice y resuma la prueba apropiada.

La estadística en el trabajo

“Sería imposible realizar investigación arqueológica sin tener al menos un conocimiento funcional de estadística básica”.



Mark T. Lycett



Kathleen Morrison

Mark T. Lycett y Kathleen Morrison forman parte del personal docente del Departamento de Antropología en la Universidad de Chicago.

La investigación del doctor Lycett se relaciona con temas de transformación económica, social y política asociados con el colonialismo español en el suroeste de Estados Unidos, y la investigación de la doctora Morrison en el sur de India trata con problemas de cambios en la agricultura, el imperialismo y la organización económica regional.

¿Qué tan importante es la estadística en la arqueología?

Sería imposible realizar investigación arqueológica sin tener al menos un conocimiento funcional de estadística básica.

¿Qué conceptos de estadística utiliza usted?

Los arqueólogos hacemos un uso extensivo de la estadística descriptiva e inferencial a diario. El análisis exploratorio de datos mediante una variedad de resúmenes gráficos y numéricos cada vez es más común en la arqueología moderna. Los problemas arqueológicos incluyen estudios de asociación para variables categóricas, pruebas de hipótesis para datos de 2 muestras y k muestras, problemas de correlación y regresión, así como una serie de métodos no paramétricos.

Por favor, exponga un ejemplo específico que ilustre el uso de la estadística en su trabajo.

Hemos explorado el tamaño de la distribución de granos de polen de hierba antiguos para investigar los cambios en la agricultura en el Viejo y el Nuevo Mundo durante los primeros siglos de la expansión colonial europea. Aunque casi todos los cultivos importantes son de hierba con polen morfológicamente similar, los cultivos alimenticios básicos del Nuevo Mundo (maíz) tienen granos de polen mucho más grandes que la hierba silvestre, en tanto que los cultivos del Viejo Mundo (principalmente trigo, cebada y arroz) son de tamaño intermedio. Estudiando la distribución de tamaño de muestras de referencia de estos cultivos alimenticios, así como el polen de hierba fosilizada a partir de contextos arqueológicos, hemos lo-

grado especificar el rango de los cultivos que se introdujeron y crecieron en lugares del periodo colonial en Nuevo México y la India.

Nuestros datos se han utilizado para hacer inferencias acerca del número y tipo de lugares arqueológicos que existieron en nuestras áreas de estudio, para reconstruir patrones de vegetación, de agricultura y economía antiguos, así como para estudiar los efectos del colonialismo y el imperialismo en las prácticas sociales, económicas y religiosas locales.

¿El uso que usted hace de la probabilidad y de la estadística está aumentando, disminuyendo o se mantiene estable?

El número y la variedad de aplicaciones estadísticas en la arqueología va en aumento, particularmente a medida que las bases de datos espaciales más complejas se vuelven disponibles a través del uso muy difundido de la tecnología de los Sistemas de Información Geográfica.

En términos de la estadística, ¿qué recomendaría usted a futuros empleados en el campo de la arqueología?

Cuando éramos estudiantes universitarios, comprendimos que la estadística sería parte de nuestras vidas profesionales, pero nunca imaginamos cuánto la utilizaríamos diariamente. Los estudiantes interesados en la arqueología deben comenzar con un curso de introducción a la probabilidad y estadística. Aquellos con metas profesionales o académicas deben considerar un curso más avanzado en la licenciatura o a nivel de posgrado en análisis de datos cuantitativos.



Correlación y regresión

10



- 10-1 Panorama general**
- 10-2 Correlación**
- 10-3 Regresión**
- 10-4 Variación e intervalos de predicción**
- 10-5 Regresión múltiple**
- 10-6 Elaboración de modelos**

¿Podemos predecir el momento de la siguiente erupción del géiser Old Faithful?

El géiser Old Faithful es la atracción más visitada del Parque Nacional Yellowstone. Está ubicado cerca del hotel Old Faithful Inn, que tal vez sea la segunda atracción más visitada de Yellowstone. Los turistas disfrutan la comida, las bebidas, el alojamiento y las tiendas del hotel, pero quieren asegurarse de ver al menos una erupción del famoso géiser Old Faithful. Los guardabosques del parque ayudan a los turistas publicando el momento predicho de la siguiente erupción. ¿Cómo hacen esas predicciones?

Cuando el Old Faithful hace erupción, se registran las siguientes mediciones: duración (en segundos) de la erupción, el intervalo de tiempo (en minutos) entre la erupción anterior y la erupción actual, el intervalo de tiempo (en minutos) entre la erupción actual y la siguiente, y la altura (en pies) de la erupción. En la tabla 10-1 se incluyen mediciones de ocho erupciones. (Las mediciones de la tabla 10-1 son ocho de las 40 erupciones incluidas en el conjunto de datos 11 del apéndice B. La tabla incluye una muestra pequeña con el fin de que los cálculos sean más fáciles cuando los datos se utilicen para analizar los métodos de las siguientes secciones).

Una vez que ocurre una erupción, queremos predecir el momento de la siguiente, que es el “intervalo de tiempo posterior” a la erupción. Para ver cuáles variables afectan los “intervalos de tiempo posteriores”, podríamos

comenzar construyendo diagramas de dispersión como los que genera Minitab, que se muestran en la figura 10-1. (Los diagramas de dispersión se presentaron en la sección 2-4). Al examinar los patrones de los puntos en los tres diagramas de dispersión, podemos plantear las siguientes conclusiones subjetivas:

1. Al parecer hay una relación entre el intervalo de tiempo posterior a una erupción y la duración de la erupción. [Véase la figura 10-1a)].
2. Parece que no existe una relación entre el intervalo de tiempo posterior a una erupción y la altura de la erupción. [Véase la figura 10-1b)].
3. Parece que no existe una relación entre el intervalo de tiempo posterior a una erupción y el intervalo de tiempo previo a la erupción. [Véase la figura 10-1c)].

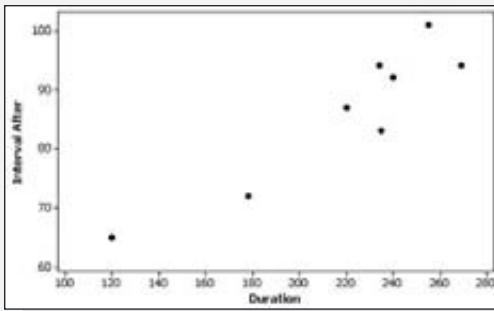
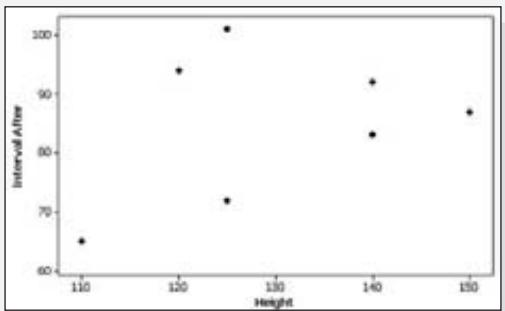
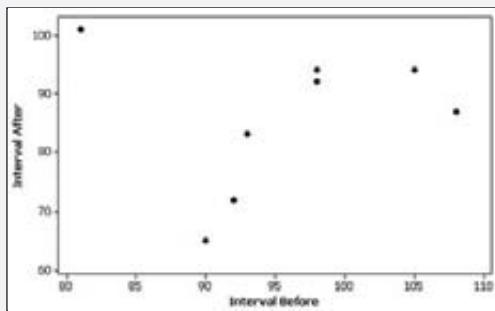
Este tipo de conclusiones basadas en diagramas de dispersión son subjetivas, y en este capítulo presentamos herramientas para analizar temas como éstos:

- ¿De qué manera se pueden utilizar métodos estadísticos para determinar objetivamente si hay una relación entre dos variables, como los intervalos de tiempo posteriores a las erupciones y la duración de éstas?

Tabla 10-1 Erupciones del géiser Old Faithful

Duración	240	120	178	234	235	269	255	220
Intervalo previo	98	90	92	98	93	105	81	108
Intervalo posterior	92	65	72	94	83	94	101	87
Altura	140	110	125	120	140	120	125	150

- Si existe una relación entre dos variables, ¿cómo podemos describirla? ¿Hay alguna ecuación que se pueda usar para predecir el momento de la siguiente erupción del géiser, dada la duración de la erupción actual?
- Si podemos predecir el momento de la siguiente erupción del Old Faithful, ¿qué tan exacta resultará esa predicción?

Minitab**a)****Minitab****b)****Minitab****c)****Figura 10-1** Gráficas de dispersión de mediciones de las erupciones del Old Faithful

10-1 Panorama general

Este capítulo presenta métodos importantes para hacer inferencias sobre una *correlación* (o relación) entre dos variables. También se explicará cómo describir este tipo de relación con una ecuación que pueda emplearse para predecir el valor de una variable, dado el valor de la otra variable. Consideramos datos muestrales ordenados en *pares*. La sección 9-4 también utilizó muestras apareadas, pero el objetivo era diferente en esencia. Las inferencias en la sección 9-4 se referían a la *media de las diferencias* entre pares de valores; en cambio, este capítulo tiene el objetivo de hacer inferencias sobre *relaciones* entre dos variables.

En la sección 10-2 se analiza la relación entre dos variables. La sección 10-3 presenta el análisis de regresión, el cual nos permite describir la relación por medio de una ecuación. En la sección 10-4 analizamos las diferencias entre valores predichos y valores reales observados. Las secciones 10-2 a la 10-4 se ocupan de relaciones entre *dos* variables, en tanto que la sección 10-5 presenta la regresión múltiple para relaciones entre tres o más variables. Por último, en la sección 10-6 se describen algunos métodos básicos para desarrollar un modelo matemático que puede emplearse para describir la relación entre dos variables. Aunque la sección 10-3 está limitada a las relaciones lineales, la sección 10-6 incluye algunas relaciones comunes no lineales.



10-2 Correlación

Concepto clave En esta sección se explica el *coeficiente de correlación lineal* r que es una medida numérica de la fuerza de la relación entre dos variables que representan datos cuantitativos. Utilizando datos muestrales apareados (que en ocasiones se llaman **datos bivariados**), calculamos el valor de r (generalmente con la ayuda de recursos tecnológicos) y luego utilizamos este valor para concluir que existe (o no) una relación entre las dos variables. En esta sección sólo consideramos las relaciones *lineales*, lo que quiere decir que cuando se grafican, los puntos se aproximan al patrón de una línea recta. Puesto que los programas de cómputo o las calculadoras suelen emplearse para calcular el valor de r , es importante enfocarse en los conceptos de esta sección, sin entretenerte demasiado con cálculos aritméticos tediosos.

Parte 1: Conceptos básicos de correlación

Iniciamos con la definición básica de *correlación*, un término que se utiliza comúnmente en el contexto de una relación entre dos variables.



Definición

Una **correlación** existe entre dos variables cuando una de ellas está relacionada con la otra de alguna manera.

La tabla 10-1, por ejemplo, consiste en datos apareados de las duraciones y los intervalos posteriores a las erupciones del géiser Old Faithful. (El “intervalo de tiempo posterior” a una erupción es el tiempo que transcurre hasta la siguiente erupción). Determinaremos si existe una correlación entre la variable x (duración) y la variable y (intervalo de tiempo posterior).

Exploración de los datos

Antes de trabajar con los métodos más formales de cálculo de esta sección, primero debemos explorar el conjunto de datos para ver qué podemos aprender. Con frecuencia podemos encontrar una relación entre dos variables al construir un diagrama de dispersión. Cuando examinamos un diagrama de dispersión, debemos estudiar el patrón general de los puntos graficados. Si existe un patrón, es necesario observar su dirección. Si los datos van hacia arriba, esto sugiere que cuando una variable aumenta, la otra también lo hace. Si los datos van hacia abajo, esto sugiere que cuando una variable aumenta, la otra disminuye. Debemos buscar valores extremos, que son puntos que se ubican muy lejos del resto de los puntos.

La figura 10-2 muestra diagramas de dispersión con características diferentes. Los diagramas de dispersión de la figura 10-2a), b) y c) describen un patrón en el que los valores crecientes de y corresponden a valores crecientes de x . Si observa las figuras de la a) a la c), se dará cuenta de que el patrón de puntos se acerca cada vez más a una línea recta, lo que sugiere que la relación entre x y y se vuelve más fuerte. Los diagramas de dispersión de la figura 10-2d), e) y f) describen patrones en los que los valores de y disminuyen cuando los valores de x aumentan. Nuevamente, si observa las figuras de la d) a la f), verá que la relación se vuelve más fuerte. En contraste con las primeras seis gráficas, el diagrama de dispersión de la figura 10-2g) no muestra patrón alguno y sugiere que no existe una correlación (o relación) entre x y y . Por último, el diagrama de dispersión de la figura 10-2h) muestra un patrón muy diferente, que sugiere una relación entre x y y , pero no es un patrón lineal (línea recta). [Las figuras 10-2a) a la g) son de ActivStats, y la figura 10-2h) es de Minitab].

Coeficiente de correlación lineal

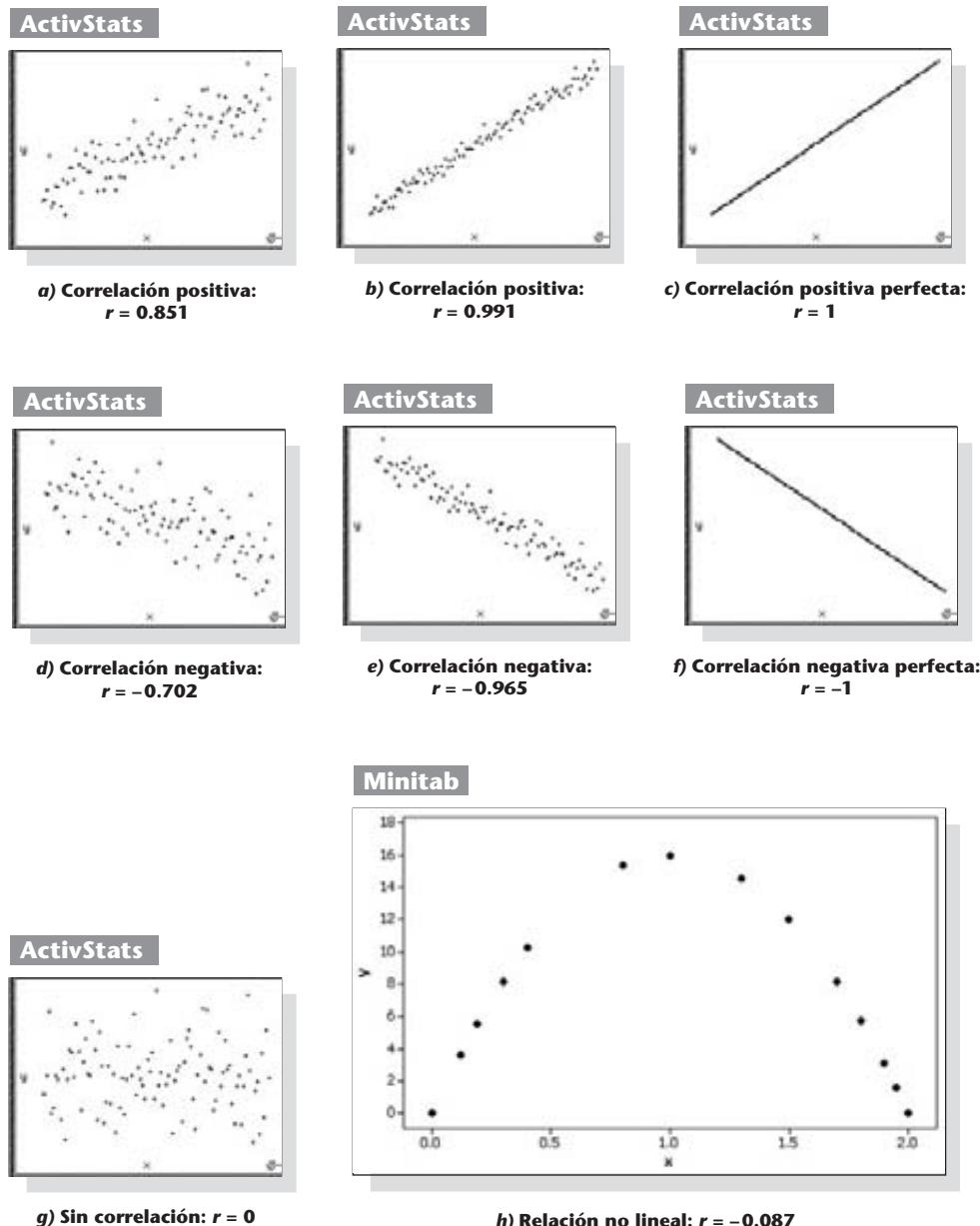
Puesto que el examen visual de los diagramas de dispersión es muy subjetivo, necesitamos medidas más precisas y objetivas. Empleamos el coeficiente de correlación lineal r , que sirve para detectar patrones lineales.



Definición

El **coeficiente de correlación lineal r** mide la fuerza de la relación lineal entre los valores cuantitativos apareados x y y en una *muestra*. Su valor se calcula con la fórmula 10-1, incluida en el recuadro de la página 520. [El coeficiente de correlación lineal también se conoce como **coeficiente de correlación producto momento de Pearson**, en honor de Karl Pearson (1857-1936), quien lo desarrolló originalmente].

Puesto que el coeficiente de correlación lineal r se calcula utilizando datos muestrales, se trata de un estadístico muestral empleado para medir la fuerza de la correlación lineal entre x y y . Si tuviéramos cada par de los valores poblacionales de x y y , el resultado de la fórmula 10-1 sería un parámetro poblacional, representado por ρ (rho griega). El recuadro en la página 520 incluye los supuestos requeridos, la notación y la fórmula 10-1.



Lectura de las manos

Algunas personas piensan que la longitud de la línea de la vida de la palma de las manos puede utilizarse para predecir la longevidad. En una carta publicada en el *Journal of the American Medical Association*, los autores M. E. Wilson y L. E. Mather refutaron esta creencia con un estudio de cadáveres. Se registraron las edades de los sujetos al morir, junto con las longitudes de la línea de la vida de sus palmas. Los autores concluyeron que no existe una correlación significativa entre la edad al morir y la longitud de la línea de la vida. La quiromancia pierde: ¡Abajo las manos!

Figura 10-2 Diagramas de dispersión

Requisitos

Dado cualquier conjunto de datos muestrales apareados, siempre se puede calcular el coeficiente de correlación lineal r , pero se deben satisfacer los siguientes requisitos cuando se prueban hipótesis o cuando se hacen inferencias acerca de r .

1. La muestra de datos apareados (x, y) es una muestra *aleatoria* de datos cuantitativos. (Es importante que los datos muestrales no se hayan reunido por medio de algún método inapropiado, como una muestra de respuesta voluntaria).
2. El examen visual del diagrama de dispersión debe confirmar que los puntos se acercan al patrón de una línea recta.
3. Es necesario eliminar cualquier valor extremo, si se sabe que se trata de un error. Los efectos de cualquier otro valor extremo deben tomarse en cuenta calculando r con y sin el valor extremo incluido.

Nota: Los requisitos 2 y 3 se simplifican al verificar el siguiente requisito formal:

Los pares de datos (x, y) tienen una **distribución normal bivariada**. (Las distribuciones normales se estudiaron en el capítulo 6, pero este supuesto requiere que, para cualquier valor fijo de x , los valores correspondientes de y tengan una distribución con forma de campana, y que para cualquier valor fijo de y , los valores de x tengan también una distribución con forma de campana). Suele ser difícil verificar este supuesto, así que, por ahora, usaremos los requisitos 2 y 3 descritos arriba.

Notación para el coeficiente de correlación lineal

n	representa el número de pares de datos presentes.
Σ	denota la suma de los elementos indicados.
Σx	denota la suma de todos los valores de x .
Σx^2	indica que cada valor de x debe elevarse al cuadrado y después deben sumarse esos cuadrados.
$(\Sigma x)^2$	indica que los valores de x deben sumarse y el total elevarse al cuadrado. Es sumamente importante evitar confundirse entre Σx^2 y $(\Sigma x)^2$.
Σxy	indica que cada valor de x debe multiplicarse primero por su valor y correspondiente. Después de obtener todos estos productos, se calcula su suma.
r	representa el coeficiente de correlación lineal de una <i>muestra</i> .
ρ	la letra griega rho se usa para representar el coeficiente de correlación lineal de una <i>población</i> .

Fórmula 10-1
$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2} \sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

Esta fórmula abreviada simplifica los cálculos manuales. Su formato la hace fácil de usar en una hoja de cálculo o en un programa de cómputo. Consulte otras fórmulas equivalentes que se presentan más adelante en esta sección; también consulte los fundamentos del cálculo de r .

Interpretación de r por medio de la tabla A-6: Si el valor absoluto del valor calculado de r excede el valor de la tabla de A-6, concluya que existe una correlación lineal significativa. De lo contrario, no existe evidencia suficiente para sustentar la conclusión de una correlación lineal.

Interpretación de r por medio de un programa de cómputo: Si el valor P calculado es menor o igual que el nivel de significancia, concluya que existe una correlación lineal. De lo contrario, no existe evidencia suficiente para sustentar la conclusión de una correlación lineal.

Redondeo del coeficiente de correlación lineal

Redondee el coeficiente de correlación lineal r a tres decimales (de manera que su valor pueda compararse directamente con los valores críticos de la tabla de A-6). Al calcular a mano r y otros estadísticos de este capítulo, hacer un redondeo a la mitad de un cálculo suele generar errores importantes, así que, trate de utilizar la memoria de su calculadora para almacenar los resultados inmediatos y redondee sólo al final.

EJEMPLO Cálculo de r Utilice la muestra aleatoria simple de datos que aparece al margen para calcular el valor del coeficiente de correlación lineal r .

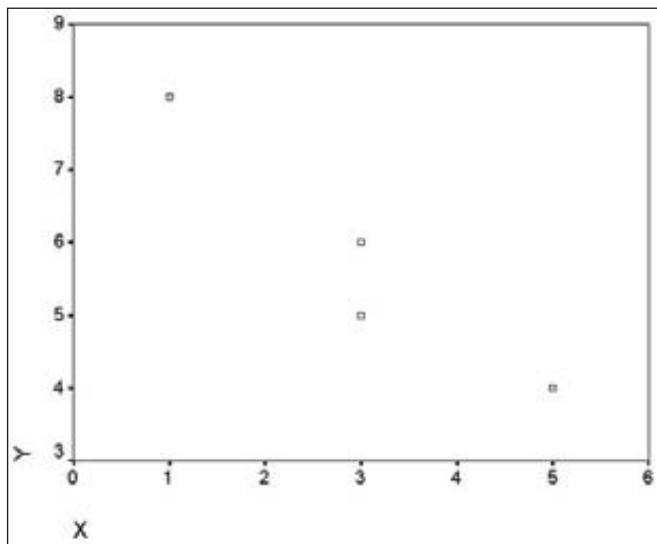
SOLUCIÓN

REQUISITO ✓ Los datos constituyen una muestra aleatoria simple. El diagrama de dispersión generado por SPSS que aparece abajo sugiere un patrón de puntos que se acerca al patrón de una línea recta. No hay valores extremos. Los requisitos se satisfacen y podemos continuar con los cálculos del coeficiente de correlación lineal r . ✓

continúa

x	3	1	3	5
y	5	8	6	4

SPSS





Las evaluaciones de profesores se correlacionan con las calificaciones

Con frecuencia se utilizan las evaluaciones que hacen los estudiantes de los profesores para medir su efectividad. Muchos estudios revelan una correlación de altas calificaciones de los estudiantes con evaluaciones positivas de los profesores. Un estudio realizado en la Universidad Duke incluyó evaluaciones de los estudiantes, recabadas antes y después de la entrega de las calificaciones finales. El estudio reveló que “las expectativas de las calificaciones o las calificaciones recibidas causaron un cambio en la forma en que los estudiantes percibían a los maestros y la calidad de su enseñanza”. Se señaló que con las evaluaciones de los estudiantes “aumentan los incentivos de los profesores para manipular sus políticas de calificación con el fin de mejorar sus evaluaciones”. Se concluyó que “la consecuencia final de este tipo de manipulaciones es la degradación de la calidad de la educación en Estados Unidos”. (Véase “Teacher Course Evaluation and Student Grades: An Academic Tango”, de Valen Johnson, *Chance*, vol. 15, núm. 3).

Tabla 10-2 Cálculo de estadísticos empleados para calcular r

	x	y	$x \cdot y$	x^2	y^2
	3	5	15	9	25
	1	8	8	1	64
	3	6	18	9	36
	5	4	20	25	16
Total	12	23	61	44	141
	↑	↑	↑	↑	↑
	Σx	Σy	Σxy	Σx^2	Σy^2

Para la muestra de datos apareados, $n = 4$, ya que existen cuatro pares de datos. Los otros componentes requeridos para la fórmula 10-1 se obtienen de los cálculos en la tabla 10-2. Observe que este formato vertical facilita los cálculos.

Con los valores calculados y la fórmula 10-1, ahora podemos evaluar r de la siguiente forma:

$$\begin{aligned} r &= \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2} \sqrt{n(\Sigma y^2) - (\Sigma y)^2}} \\ &= \frac{4(61) - (12)(23)}{\sqrt{4(44) - (12)^2} \sqrt{4(141) - (23)^2}} \\ &= \frac{-32}{\sqrt{32} \sqrt{35}} = -0.956 \end{aligned}$$

Estos cálculos se vuelven confusos con conjuntos grandes de datos, por lo que es una suerte que el coeficiente de correlación lineal pueda calcularse automáticamente por medio de muchas calculadoras y programas de cómputo distintos. Consulte el recuadro “Uso de la tecnología” al final de esta sección, donde encontrará comentarios acerca de STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus.

Interpretación del coeficiente de correlación lineal

Necesitamos interpretar un valor calculado de r , tal como el valor de -0.956 obtenido en el ejemplo anterior. Dada la manera en que la fórmula 10-1 está construida, el valor de r siempre debe estar entre -1 y $+1$, inclusive. Si r se acerca a 0 , concluimos que no existe una correlación lineal entre x y y , pero si r se acerca -1 o $+1$, concluimos que hay una correlación lineal entre x y y . Interpretaciones tales como “cerca a” 0 , a 1 o a -1 son vagas, por lo que utilizamos el siguiente criterio específico de decisión:

Uso de la tabla A-6: Si el valor absoluto del valor calculado de r excede el valor de la tabla a A-6, se concluye que existe una correlación lineal. De lo contrario, no existe evidencia suficiente para sustentar la conclusión de una correlación lineal.

Uso de un programa de cómputo: Si el valor P calculado es menor o igual que el nivel de significancia, se concluye que existe una correlación lineal. De lo contrario, no existe evidencia suficiente para sustentar la conclusión de una correlación lineal.

Cuando en realidad no existe una correlación lineal entre x y y , la tabla A-6 lista valores que son “críticos” en este sentido: separan valores *comunes* de r de aquellos que son *poco comunes*. Por ejemplo, la tabla de A-6 nos indica que con $n = 4$ pares de datos muestrales, los valores críticos son 0.950 (para $\alpha = 0.05$) y 0.999 (para $\alpha = 0.01$). Los valores críticos y el papel de α se describen detalladamente en los capítulos 7 y 8. He aquí cómo interpretamos estos números: con 14 pares de datos y ninguna correlación lineal entre x y y , existe una probabilidad del 5% de que el valor absoluto del coeficiente de correlación lineal r calculado excede 0.950. Con $n = 4$ y sin correlación lineal, existe una probabilidad del 1% de que $|r|$ excede 0.999. Dado $r = -0.956$ calculado en el ejemplo anterior, si usamos un nivel de significancia de 0.05, concluimos que existe una correlación lineal entre x y y (porque $|-0.956|$ excede el valor crítico de 0.950). Sin embargo, si usamos un nivel de significancia de 0.01, no concluimos que existe una correlación lineal (porque $|-0.956|$ no excede el valor crítico de 0.999).



EJEMPLO Old Faithful Utilice los datos apareados de la duración y el intervalo de tiempo posterior a la erupción que aparecen en la tabla 10-1 para calcular el valor del coeficiente de correlación lineal r . Despues remítase a la tabla A-6 para determinar si existe una correlación lineal significativa entre la duración y los intervalos de tiempo posteriores a las erupciones. En la tabla A-6, utilice el valor crítico para $\alpha = 0.05$. (Con $\alpha = 0.05$ concluimos que existe una correlación lineal significativa sólo si la muestra es improbable en este sentido: si no existe una correlación lineal entre dos variables, un valor de r como éste ocurre el 5% de las veces o menos).

SOLUCIÓN

REQUISITO ✓ Los datos provienen de erupciones elegidas al azar. El diagrama de dispersión que aparece en la figura 10-1a) muestra un patrón de puntos que se aproxima a una línea recta. Al parecer, no hay valores extremos. Los requisitos se satisfacen, por lo que procedemos a calcular r y a determinar si existe una correlación lineal entre la duración de los intervalos posteriores a las erupciones. ✓

Al utilizar el mismo procedimiento ilustrado en el ejemplo anterior o empleando algún recurso tecnológico, obtenemos que los ocho pares de datos duración/intervalo después de la erupción que aparecen en la tabla 10-1 dan como resultado $r = 0.926$. Éstos son los resultados de Minitab:

Minitab

Pearson correlation of Duration and Interval After = 0.926
P-Value = 0.001

Si nos remitimos a la tabla A-6, localizamos el renglón en el que $n = 8$ (porque hay 8 pares de datos). Este renglón contiene los valores críticos de 0.707 (para $\alpha = 0.05$) y 0.8345 (para $\alpha = 0.01$). Con el valor crítico para $\alpha = 0.05$, observamos que existe una probabilidad menor al 5% de que, sin correlación

continúa

lineal, el valor absoluto de r calculado excede a 0.707. Puesto que $r = 0.926$, su valor absoluto excede a 0.707, por lo que concluimos que existe una correlación lineal entre la duración y los intervalos de tiempo posteriores a la erupción.

Ya hemos señalado que la fórmula 10-1 requiere que el valor calculado de r caiga siempre entre -1 y $+1$, inclusive. Consideramos esa propiedad, junto con otras igualmente importantes, en la siguiente lista.

Propiedades del coeficiente de correlación lineal r

1. El valor de r está siempre entre -1 y $+1$, inclusive. Es decir,

$$-1 \leq r \leq +1$$

2. *El valor de r no cambia si todos los valores de cualquiera de las variables se convierten a una escala diferente.*
3. *El valor de r no se ve afectado por la elección de x o y .* Intercambie todos los valores de x y y , y el valor de r no sufrirá cambios.
4. *r mide la fuerza de una relación lineal.* No está diseñada para medir la fuerza de una relación que no sea lineal.

Interpretación de r : Variación explicada

Si concluimos que existe una correlación lineal entre x y y , podemos obtener una ecuación lineal que exprese y en términos de x , y la ecuación puede emplearse para predecir valores de y a partir de valores dados de x . En la sección 10-3 describiremos un procedimiento para el cálculo de tales ecuaciones y mostraremos cómo predecir valores de y cuando se tienen valores dados de x . Pero un valor predicho de y no será necesariamente el resultado exacto porque, además de x , existen otros factores que afectan a y , como la variación aleatoria y otras características que no están incluidas en el estudio. En la sección 10-4 presentaremos los fundamentos y más detalles acerca de este principio importante:

El valor de r^2 es la proporción de la variación de y que está explicada por la relación lineal entre x y y .



EJEMPLO Old Faithful Con los datos de la duración y los intervalos posteriores a la erupción que se presentan en la tabla 10-1, hemos encontrado que el coeficiente de correlación lineal es $r = 0.926$. ¿Qué proporción de la variación en los intervalos posteriores a la erupción puede explicarse por la variación en las duraciones?

SOLUCIÓN Con $r = 0.926$, obtenemos $r^2 = 0.857$.

INTERPRETACIÓN Concluimos que 0.857 (o aproximadamente el 86%) de la variación de los intervalos posteriores a la erupción puede explicarse por la relación lineal entre los tiempos de duración y los intervalos posteriores a la erupción. Esto implica que cerca del 14% de la variación de los intervalos posteriores a la erupción no pueden explicarse por los tiempos de duración.

Errores comunes en las correlaciones

Ahora identificamos tres de las fuentes más comunes de errores que se cometan al interpretar los resultados de correlaciones:

1. *Un error común es concluir que la correlación implica causalidad.* Con los datos muestrales de la tabla 10-1, podemos concluir que existe una correlación entre la duración y los intervalos de tiempo posteriores a la erupción, pero no podemos concluir que una mayor duración *cause* intervalos más prolongados posteriores a la erupción. Los intervalos de tiempo posteriores a la erupción pueden verse afectados por alguna otra variable intervventora en los antecedentes. (Una **variable intervventora** es aquella que afecta las variables que se estudian, pero que no está incluida en la investigación). Por ejemplo, la temperatura del aire exterior podría afectar la duración de una erupción y los intervalos posteriores a una erupción. Por lo tanto, la temperatura del aire exterior sería una variable intervventora.
2. *Otro error proviene de los datos basados en promedios.* Los promedios eliminan la variación individual y pueden inflar el coeficiente de correlación. Un estudio produjo un coeficiente de correlación lineal de 0.4 para datos apareados que relacionaban el ingreso y la educación de individuos, pero el coeficiente de correlación lineal se convirtió en 0.7 cuando se utilizaron promedios regionales.
3. *Un tercer error implica la propiedad de linealidad.* Puede existir una relación entre x y y , aun cuando no haya una correlación lineal. Los datos presentados en la figura 10-2h) dan por resultado un valor de $r = -0.087$, lo que indica que no existe una correlación *lineal* entre las dos variables. Sin embargo, al observar la figura, con facilidad podemos ver que existe un patrón que refleja una relación *no lineal* muy fuerte. [La figura 10-2h) es un diagrama de dispersión que representa la relación entre la distancia del suelo hacia arriba y el tiempo transcurrido para un objeto lanzado hacia arriba].

Parte 2: Prueba formal de hipótesis (requiere el estudio del capítulo 8)

Presentamos dos métodos (resumidos en el siguiente recuadro y en la figura 10-3) para utilizar una prueba formal de hipótesis con el fin de determinar si existe una correlación lineal significativa entre dos variables. Algunos profesores prefieren el método 1 porque refuerza conceptos estudiados en capítulos anteriores. Otros prefieren el método 2 porque supone cálculos más fáciles. El método 1 emplea la distribución t de Student, con un estadístico de prueba con la forma $t = r/s_r$, donde s_r denota la desviación estándar muestral de valores de r . El estadístico de prueba incluido en el recuadro (para el método 1) refleja el hecho de que la desviación estándar de los valores de r puede expresarse como $\sqrt{(1 - r^2)/(n - 2)}$.

La figura 10-3 indica que el criterio de decisión es el rechazo de la hipótesis nula de $\rho = 0$, si el valor absoluto del estadístico de prueba excede los valores críticos; el rechazo de $\rho = 0$ significa que existe evidencia suficiente para sustentar una aseveración de una correlación lineal entre las dos variables. Si el valor absoluto del estadístico de prueba no excede los valores críticos, entonces no rechazamos $\rho = 0$; es decir, no existe suficiente evidencia para concluir que existe una correlación lineal entre las dos variables.

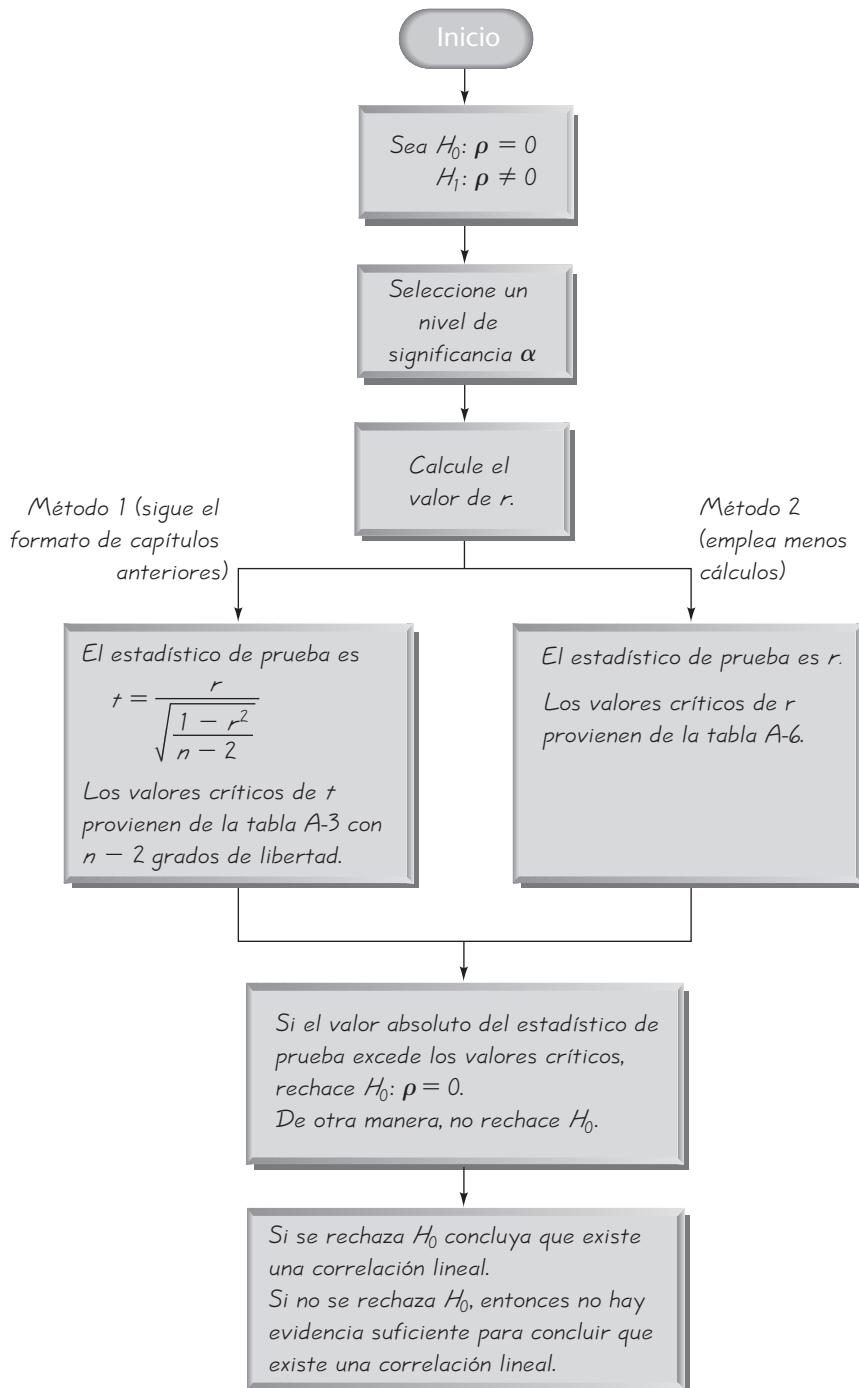


Figura 10-3 Prueba de hipótesis para una correlación lineal

Prueba de hipótesis de correlación (véase la figura 10-3).

$H_0: \rho = 0$ (No existe una correlación lineal).

$H_1: \rho \neq 0$ (Existe una correlación lineal).

Método 1: El estadístico de prueba es t

Estadístico de prueba:
$$t = \frac{r}{\sqrt{\frac{1 - r^2}{n - 2}}}$$

Valores críticos: Utilice la tabla A-3 con $n - 2$ grados de libertad.

Valor P : Utilice la tabla A-3 con $n - 2$ grados de libertad.

Conclusión: Si $|t| >$ el valor crítico de la tabla A-3, rechace H_0 y concluya que existe una correlación lineal. Si $|t| \leq$ valor crítico, no rechace H_0 ; no hay evidencia suficiente para concluir que existe una correlación lineal.

Método 2: El estadístico de prueba es r

Estadístico de prueba: r

Valores críticos: Remítase a la tabla A-6.

Conclusión: Si $|r| >$ el valor crítico de la tabla A-6, rechace H_0 y concluya que existe una correlación lineal. Si $|r| \leq$ valor crítico, no rechace H_0 ; no hay evidencia suficiente para concluir que existe una correlación lineal.



EJEMPLO Old Faithful Con los datos muestrales de la tabla 10-1 sobre la duración y los intervalos posteriores a las erupciones, pruebe la aseveración de que existe una correlación lineal entre la duración de una erupción y el intervalo de tiempo posterior a la erupción. Para obtener el estadístico de prueba utilice *a) el método 1 y b) el método 2*.

SOLUCIÓN

REQUISITO ✓ Un ejemplo previo incluye la verificación de los requisitos. [Los datos muestrales se seleccionaron al azar; el diagrama de dispersión en la figura 10-1a) indica un patrón de puntos de apariencia lineal, sin valores extremos]. Una vez completada esta verificación simplificada de requisitos, procedemos con nuestro análisis. ✓

Esta solución seguirá el procedimiento que se resume en la figura 10-3. Afirmar que existe una correlación lineal equivale a aseverar que el coeficiente de correlación lineal poblacional ρ es diferente de 0. Por lo tanto, tenemos las siguientes hipótesis:

$H_0: \rho = 0$ (No existe una correlación lineal).

$H_1: \rho \neq 0$ (Existe una correlación lineal).

Puesto que no se especificó un nivel de significancia, utilice $\alpha = 0.05$.

continúa

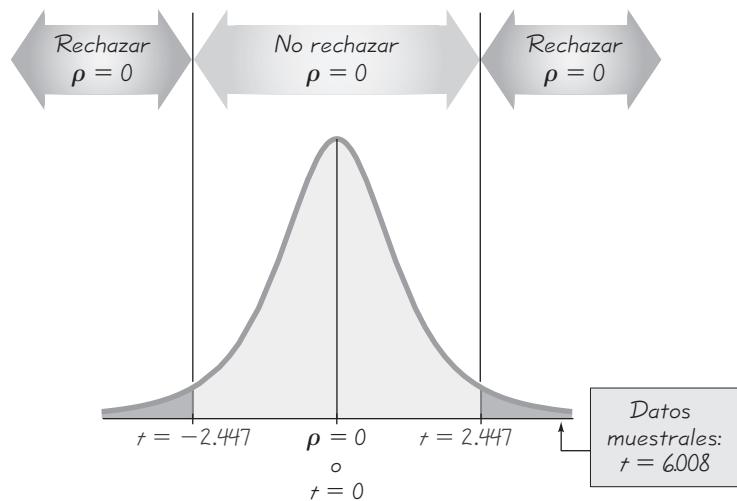


Figura 10-4 Prueba de $H_0: \rho = 0$ con el método 1

En un ejemplo previo ya calculamos que $r = 0.926$. Con ese valor ahora calculamos el estadístico de prueba y el valor crítico por medio de los dos métodos descritos.

- a. *Método 1:* El estadístico de prueba es

$$t = \frac{r}{\sqrt{\frac{1 - r^2}{n - 2}}} = \frac{0.926}{\sqrt{\frac{1 - 0.926^2}{8 - 2}}} = 6.008$$

Los valores críticos de $t = \pm 2.447$ se encuentran en la tabla A-3, donde 2.447 corresponde a una área de 0.05, dividida entre dos colas, y el número de grados de libertad es $n - 2 = 6$. Observe en la figura 10-4 la gráfica que incluye el estadístico de prueba y los valores críticos.

- b. *Método 2:* El estadístico de prueba es $r = 0.926$. Los valores críticos de $r = \pm 0.707$ se encuentran en la tabla A-6, con $n = 8$ y $\alpha = 0.05$. Observe en la figura 10-5 la gráfica que incluye el estadístico de prueba y los valores críticos.

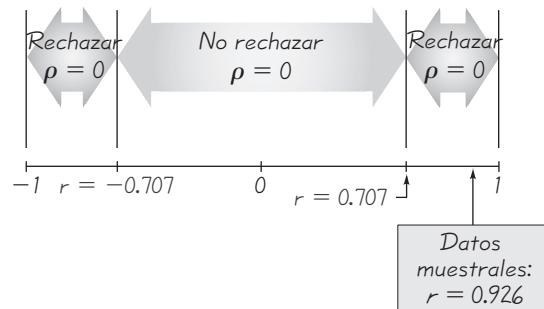


Figura 10-5 Prueba de $H_0: \rho = 0$ con el método 2

Al utilizar de cualquiera de los dos métodos, encontramos que el valor absoluto del estadístico de prueba excede el valor crítico (Método 1: $6.008 > 2.447$. Método 2: $0.926 > 0.707$); es decir, el estadístico de prueba cae en la región crítica. Por lo tanto, rechazamos $H_0: \rho = 0$. Existe evidencia suficiente para sustentar la aseveración de una correlación lineal entre las duraciones de las erupciones y los intervalos posteriores a éstas.

Pruebas de una cola: El ejemplo anterior y las figuras 10-4 y 10-5 ilustran una prueba de hipótesis de dos colas. En general, los ejemplos y ejercicios de esta sección implicarán únicamente pruebas de dos colas, pero puede presentarse una prueba de una cola en una aseveración de una correlación lineal positiva o una aseveración de una correlación lineal negativa. En estos casos, las hipótesis serán como las que se presentan a continuación.

Aseveración de correlación negativa (prueba de cola izquierda)	Aseveración de correlación positiva (prueba de cola derecha)
---	---

$$H_0: \rho = 0 \quad H_0: \rho = 0$$

$$H_1: \rho < 0 \quad H_1: \rho > 0$$

Con estas pruebas de una cola, el método 1 puede realizarse como se hizo en capítulos anteriores. En el caso del método 2, calcule el valor crítico como se describió en el ejercicio 38, o bien, modifique la tabla A-6 reemplazando los encabezados de columna de $\alpha = 0.05$ y $\alpha = 0.01$ por los niveles unilaterales de significancia de $\alpha = 0.025$ y $\alpha = 0.005$, respectivamente.

Fundamentos: Ya presentamos la fórmula 10-1 para el cálculo de r e ilustramos su uso. A continuación se presenta la fórmula 10-1 junto con algunas otras fórmulas que son “equivalentes” en el sentido de que todas producen los mismos valores. Los distintos autores de los libros prefieren expresiones diferentes por diversas razones, y las fórmulas que se presentan abajo son las expresiones de r más utilizadas. Estas fórmulas simplemente son versiones diferentes de la misma expresión. (Invitamos a los amantes del álgebra a divertirse probando la equivalencia de las fórmulas; los demás pueden confiar en la afirmación del autor de que las diversas fórmulas para r producen los mismos resultados). El formato de la fórmula 10-1 simplifica tanto los cálculos manuales como los que se realizan con hojas de cálculo, y también es fácil de incluir en un programa de cómputo original.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \quad r = \frac{\sum \left[\frac{(x - \bar{x})(y - \bar{y})}{s_x s_y} \right]}{n - 1}$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n - 1)s_x s_y} \quad r = \frac{s_{xy}}{\sqrt{s_{xx}} \sqrt{s_{yy}}}$$

Aun cuando la fórmula 10-1 simplifica los cálculos manuales, otras fórmulas para r son mejores para *entender* la manera en que funciona r . Para tratar de comprender los fundamentos que subyacen en el desarrollo del coeficiente de correlación lineal, usaremos la siguiente versión del coeficiente de correlación lineal:

$$r = \frac{\sum \left[\frac{(x - \bar{x})}{s_x} \frac{(y - \bar{y})}{s_y} \right]}{n - 1}$$

De manera temporal utilizaremos esta última versión de la fórmula 10-1, ya que se relaciona de manera más directa con la teoría subyacente. Ahora considere los siguientes datos apareados, que están representados en el diagrama de dispersión de la figura 10-6.

x	1	1	2	4	7
y	4	5	8	15	23

La figura 10-6 incluye el punto $(\bar{x}, \bar{y}) = (3, 11)$, denominado el *centroide* de los puntos muestrales.



Definición

Dado un conjunto de datos apareados (x, y) , el punto (\bar{x}, \bar{y}) se denomina **centroide**.

El estadístico r , que en ocasiones se llama *producto momento de Pearson*, fue creado por Karl Pearson. Se basa en la sumatoria de los productos $(x - \bar{x})(y - \bar{y})$. El estadístico $\sum(x - \bar{x})(y - \bar{y})$ es la expresión básica del fundamento de r , y ahora explicaremos por qué el estadístico es la clave.

En cualquier diagrama de dispersión, las líneas vertical y horizontal que pasan a través del centroide (\bar{x}, \bar{y}) dividen el diagrama en cuatro cuadrantes, como se ob-

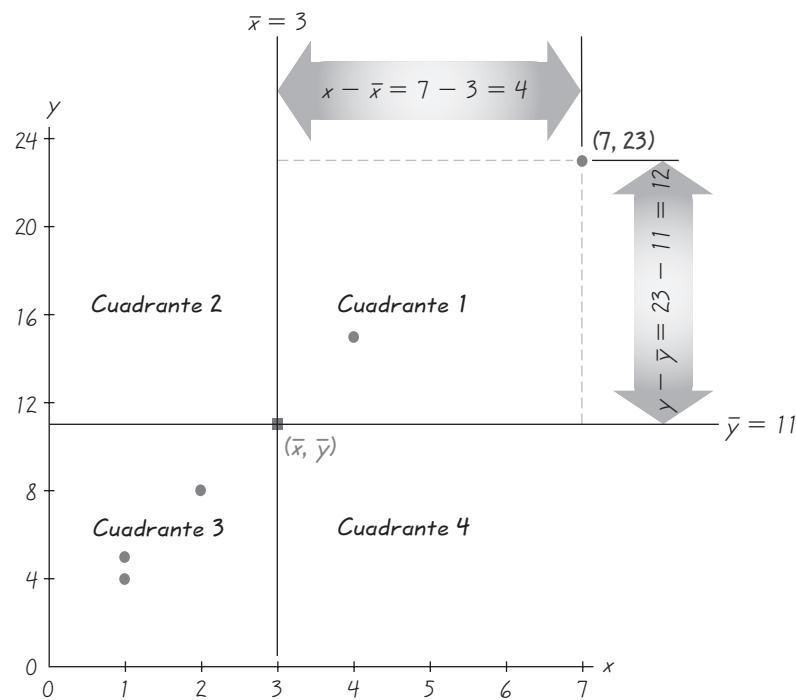


Figura 10-6 Diagrama de dispersión dividido en cuadrantes

serva en la figura 10-6. Si los puntos del diagrama de dispersión tienden a aproximarse a una línea ascendente (como en la figura), los valores individuales del producto $(x - \bar{x})(y - \bar{y})$ tienden a ser positivos ya que la mayoría de los puntos se encuentran en el primer y tercer cuadrantes, donde los productos de $(x - \bar{x})$ y $(y - \bar{y})$ son positivos. Si los puntos del diagrama de dispersión se aproximan a una línea descendente, la mayoría de los puntos se encuentran en el segundo y cuarto cuadrantes, donde $(x - \bar{x})$ y $(y - \bar{y})$ tienen el signo opuesto, de manera que $\sum(x - \bar{x})(y - \bar{y})$ es negativo. Los puntos que no siguen un patrón lineal tienden a dispersarse en los cuatro cuadrantes, de manera que el valor de $\sum(x - \bar{x})(y - \bar{y})$ tiende a 0. Por lo tanto, podemos usar $\sum(x - \bar{x})(y - \bar{y})$ para medir el arreglo de los puntos. Una suma positiva grande sugiere que los puntos están predominantemente en el primer y en el tercer cuadrantes (lo que corresponde a una correlación lineal positiva); una suma negativa grande sugiere que los puntos están predominantemente en el segundo y en el cuarto cuadrantes (lo que corresponde a una correlación lineal negativa), y una suma cercana a cero sugiere que los puntos se dispersan en los cuatro cuadrantes (y no existe una correlación lineal).

Por desgracia, la suma $\sum(x - \bar{x})(y - \bar{y})$ depende de la magnitud de los números utilizados. Por ejemplo, si se convierte x de pulgadas a pies, la suma cambiará. Para lograr que r sea independiente de la escala utilizada, incluimos la desviación estándar muestral. En la sección 3-4 aprendimos que podemos “estandarizar” los valores al convertirlos en puntuaciones z , como en $z = (x - \bar{x})/s$. (Aquí utilizamos s_x para denotar la desviación estándar de los valores muestrales x , y utilizamos s_y para denotar la desviación estándar de los valores muestrales y). Empleamos una técnica similar al estandarizar cada desviación $(x - \bar{x})$ dividiéndola entre s_x . También logramos que las desviaciones $(y - \bar{y})$ sean independientes de las magnitudes de los números al dividirlas entre s_y . Ahora tenemos el estadístico

$$\sum \left[\frac{(x - \bar{x})}{s_x} \frac{(y - \bar{y})}{s_y} \right]$$

que modificamos aún más al introducir el divisor de $n - 1$, que nos da un tipo de promedio en vez de una suma que aumenta simplemente porque tenemos más datos. (Las razones para dividir entre $n - 1$ y no entre n son básicamente las mismas relacionadas con la desviación estándar). El resultado final es la expresión

$$r = \frac{\sum \left[\frac{(x - \bar{x})}{s_x} \frac{(y - \bar{y})}{s_y} \right]}{n - 1}$$

Esta expresión puede manipularse algebraicamente en la forma equivalente de la fórmula 10-1 o cualquiera de las otras expresiones para r .

Intervalos de confianza En capítulos anteriores estudiamos métodos de estadística inferencial al utilizar métodos de prueba de hipótesis y métodos para construir estimados de intervalos de confianza. Se puede emplear un procedimiento similar para calcular intervalos de confianza para ρ . Sin embargo, como la construcción de tales intervalos de confianza implica transformaciones que son hasta cierto punto complicadas, ese proceso se presenta en el ejercicio 39.

Podemos utilizar el coeficiente de correlación lineal para determinar si existe una relación lineal entre dos variables. Con los datos de la tabla 10-1 sobre la duración y los intervalos posteriores a las erupciones, concluimos que existe una correlación

Uso de la tecnología

STATDISK Ingrese los datos apareados en columnas de la ventana de datos de Statdisk. Seleccione **Analysis** de la barra del menú principal, después utilice la opción **Correlation and Regression**. Introduzca un valor para el nivel de significancia. Haga clic en el botón **Evaluate**. Los resultados de STATDISK incluirán el valor del coeficiente de correlación lineal, junto con el valor crítico de r , la conclusión y otros resultados que se estudiarán en secciones posteriores. También es posible obtener un diagrama de dispersión al hacer clic en el botón de **Scatterplot**. Observe la siguiente pantalla de resultados de Statdisk sobre el ejemplo de esta sección.

STATDISK

Sample size, n:	8
Degrees of freedom:	6
<i>Correlation Results:</i>	
Correlation coeff, r:	0.9255912
Critical r:	±0.706734
P-value (two-tailed):	0.00097
Reject the Null Hypothesis	
Sample provides evidence to support linear correlation	

MINITAB Introduzca los datos apareados en las columnas C1 y C2, después seleccione **Stat** de la barra del menú principal, elija **Basic Statistics** y después **Correlation**; proceda a introducir C1 y C2, las columnas que serán utilizadas. Minitab dará el valor del coeficiente de correlación lineal r , así como también un valor P . Para obtener un diagrama de dispersión seleccione **Graph**, luego **Plot** y después introduzca C1 y C2 para X y Y ; finalmente haga clic en **OK**.

EXCEL Excel tiene una función que calcula el valor del coeficiente de correlación lineal. Primero introduzca los datos muestrales apareados en las columnas A y B. Haga clic en la tecla de función **fx** localizada en la barra del menú principal. Seleccione la categoría **Statistical** y el nombre de la función **CORREL**, después haga clic en **OK**. En el cuadro de diálogo ingrese el rango de la celda de los valores para x , como A1:A8. También ingrese el rango de la celda de los valores para y , como B1:B8. Para obtener un diagrama de dispersión, haga clic en el Chart Wizard del menú principal, después seleccione la gráfica identificada como **XY(Scatter)**. En el cuadro de diálogo introduzca el rango de entrada de los datos, como A1:B8. Haga clic en **Next** y proceda a utilizar los cuadros de diálogo para modificar la gráfica como lo deseé.

También puede emplearse el complemento Data Desk XL. Haga clic en **DDXL** y seleccione **Regression**, después haga clic en el cuadro del tipo de función y seleccione **Correlation**. En el cuadro de diálogo, haga clic en el ícono del lápiz para la Variable del eje X e introduzca el rango de valores para la variable x , como A1:A8. Haga clic en el ícono del lápiz para la Variable del eje Y e introduzca el rango de valores para y . Haga clic en **OK**. Aparecerán el diagrama de dispersión y el coeficiente de correlación. También hay una opción para realizar una prueba t , tal como se describió en el método 1 de esta sección.

TI-83/84 PLUS Introduzca los datos apareados en las listas L1 y L2, después presione **STAT** y seleccione **TESTS**. Si utiliza la opción de **LinRegTTest** resultarán diversos valores, incluyendo el valor del coeficiente de correlación lineal r .

Para obtener un diagrama dispersión, presione **2nd**, después **Y=** (para **STAT PLOT**). Presione **Enter** dos veces para activar Plot 1, después seleccione el primer tipo de gráfica, que representa un diagrama de dispersión. Coloque las etiquetas de la lista X y Y para L1 y L2 y presione la tecla **ZOOM**, finalmente elija **ZoomStat** y presione la tecla **Enter**.

lineal entre la duración de las erupciones y los intervalos de tiempo posteriores a las erupciones. Una vez que se concluyó que existe una relación, nos gustaría determinar de qué relación se trata, de manera que podamos predecir el intervalo de tiempo posterior a una erupción para una duración dada. Es decir, podemos usar el tiempo que dura una erupción para predecir el intervalo hasta la próxima erupción. En la siguiente sección se estudiará la etapa posterior del análisis.

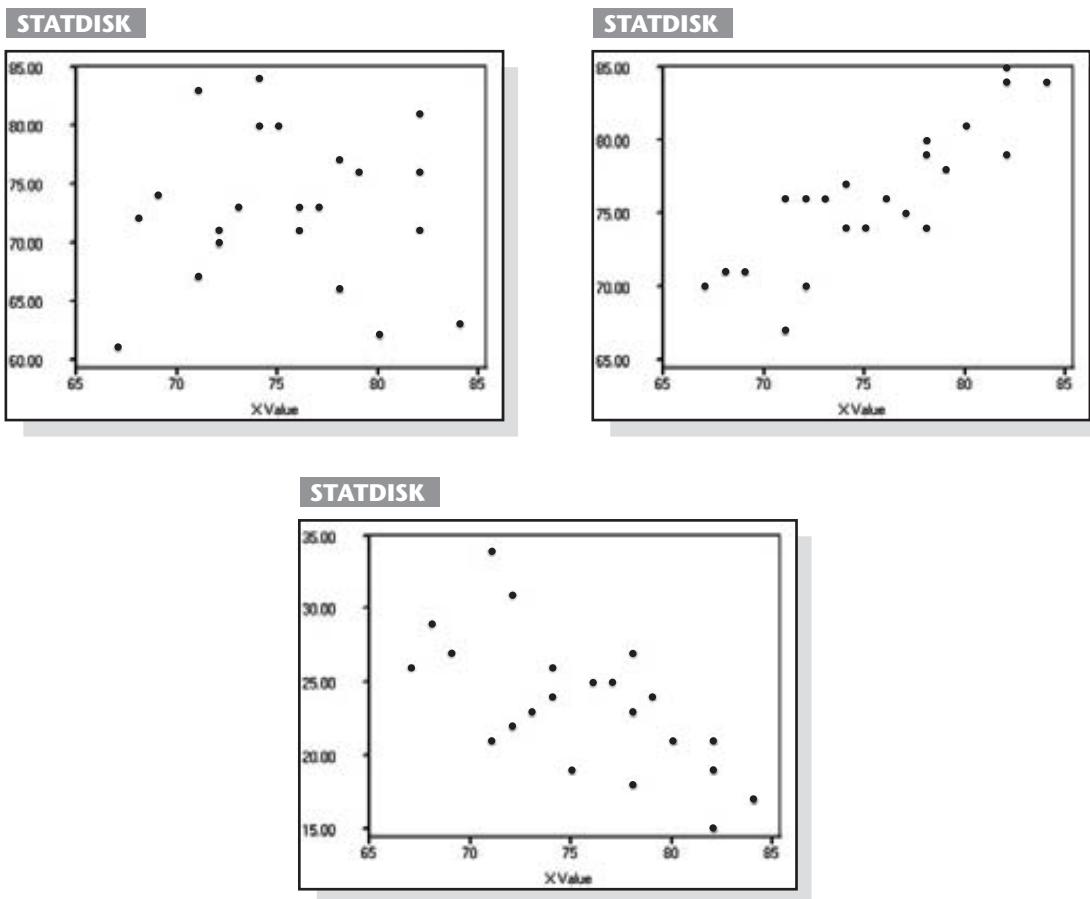
10-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- NSS e ingreso.** Un investigador del gobierno desea realizar un estudio para determinar si existe una correlación entre los números del seguro social y el ingreso. El investigador reúne los datos apareados de una muestra aleatoria de 100 personas. ¿Deberá emplear los métodos de esta sección con el coeficiente de correlación lineal? ¿Por qué?
- Reacción adversa de un fármaco.** Se realizó un estudio clínico para investigar la eficacia del fármaco Dozenol para tratar el insomnio. Se descubrió que existe una correlación entre la cantidad ingerida de Dozenol y la duración del sueño. Con base en este análisis estadístico, ¿podemos concluir que el Dozenol es la causa del sueño? ¿Por qué?

3. Correlación y variable intervadora. ¿Qué es una correlación? ¿Qué es una variable intervadora?

4. Identificación de diagramas de dispersión. A continuación se presentan tres diagramas de dispersión generados por STATDISK. Determine cuál de los diagramas de dispersión corresponde a los siguientes valores del coeficiente de correlación lineal: $r = 0.857$, $r = -0.658$, $r = 0.012$.



Interpretación de r. En los ejercicios 5 a 8, utilice un nivel de significancia de $\alpha = 0.05$.

5. Tamaño del pecho y peso de osos. Mientras ocho osos se encontraban anestesiados, unos investigadores midieron las distancias (en pulgadas) alrededor del pecho de los osos y los pesaron (en libras). Se utilizó Minitab para calcular el valor del coeficiente de correlación lineal, que resultó ser $r = 0.993$.

- ¿Existe una correlación lineal significativa entre tamaño del pecho y el peso? Explique.
- ¿Qué proporción de la variación del peso puede explicarse por la relación lineal entre el peso y el tamaño del pecho?

6. Armas de fuego y tasa de homicidios. Con datos reunidos del FBI y de la Oficina de alcohol, tabaco y armas de fuego, se obtuvo el número de armas automáticas registradas y la tasa de homicidios (en asesinatos por 10,000 individuos) de cada uno de ocho estados seleccionados al azar. Por medio de STATDISK se calculó el valor del coeficiente de correlación lineal $r = 0.885$.

continúa

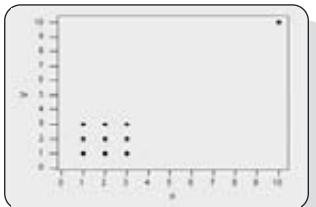
- a. ¿Existe una correlación lineal entre el número de armas automáticas registradas y la tasa de homicidios? Explique.
- b. ¿Qué proporción de la variación de la tasa de homicidios puede explicarse por la relación lineal entre la tasa de homicidios y el número de armas automáticas registradas?
- 7. Acciones y el Súper Bowl.** Se registraron los valores máximos del Promedio Industrial Dow Jones (*Dow Jones Industrial Average*, DJIA) y el número total de puntos anotados en el Súper Bowl en 21 años. Se utilizó Excel para calcular el valor del coeficiente de correlación lineal $r = -0.133$.
- a. ¿Existe una correlación lineal entre el valor máximo del DJIA y los puntos en el Súper Bowl? Explique.
- b. ¿Qué proporción de la variación de los puntos del Súper Bowl puede explicarse por la variación del valor máximo del DJIA?
- 8. Edades de corredores de maratón.** Se registraron las edades y los tiempos de 150 corredores elegidos al azar que completaron la maratón de la ciudad de Nueva York. Se utilizó una calculadora TI-83/84 Plus para calcular el valor del coeficiente de correlación lineal $r = 0.144$. (*Sugerencia:* Considere que para $n = 150$, los valores críticos son ± 0.160).
- a. ¿Existe una correlación lineal entre la edad y el tiempo? Explique.
- b. ¿Qué proporción de la variación del tiempo puede explicarse por la variación de la edad?

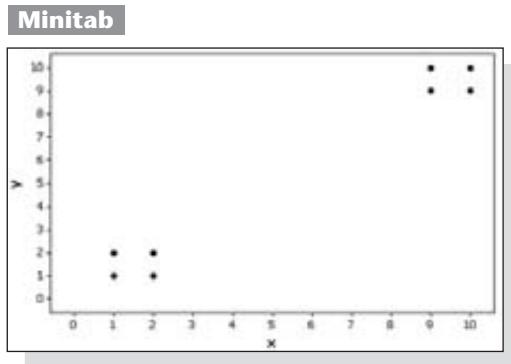
Prueba de una correlación lineal. En los ejercicios 9 y 10, utilice un diagrama de dispersión y el coeficiente de correlación lineal r para determinar si existe una correlación entre las dos variables.

9.	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 20px; text-align: center;">x</td><td style="width: 20px; text-align: center;">1</td><td style="width: 20px; text-align: center;">0</td><td style="width: 20px; text-align: center;">5</td><td style="width: 20px; text-align: center;">2</td><td style="width: 20px; text-align: center;">3</td></tr> <tr> <td style="text-align: center;">y</td><td style="text-align: center;">3</td><td style="text-align: center;">1</td><td style="text-align: center;">15</td><td style="text-align: center;">6</td><td style="text-align: center;">8</td></tr> </table>	x	1	0	5	2	3	y	3	1	15	6	8
x	1	0	5	2	3								
y	3	1	15	6	8								
10.	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 20px; text-align: center;">x</td><td style="width: 20px; text-align: center;">0</td><td style="width: 20px; text-align: center;">3</td><td style="width: 20px; text-align: center;">3</td><td style="width: 20px; text-align: center;">1</td><td style="width: 20px; text-align: center;">4</td></tr> <tr> <td style="text-align: center;">y</td><td style="text-align: center;">1</td><td style="text-align: center;">7</td><td style="text-align: center;">2</td><td style="text-align: center;">5</td><td style="text-align: center;">5</td></tr> </table>	x	0	3	3	1	4	y	1	7	2	5	5
x	0	3	3	1	4								
y	1	7	2	5	5								

- 11. Efectos de un valor extremo** Remítase al diagrama de dispersión generado por Minitab que aparece al margen.
- a. Examine el patrón de los 10 puntos y determine de forma subjetiva si parece existir una correlación entre x y y .
- b. Después de identificar los 10 pares de coordenadas correspondientes a los 10 puntos, calcule el valor del coeficiente de correlación r y determine si existe una correlación lineal.
- c. Ahora elimine el punto con las coordenadas (10, 10) y repita los incisos a) y b).
- d. ¿Qué concluye cerca del posible efecto de un solo par de valores?
- 12. Efectos de aglomerados.** Remítase al diagrama de dispersión generado por Minitab que aparece en la siguiente página. Los cuatro puntos de la parte inferior izquierda son mediciones de mujeres, y los cuatro puntos de la parte superior derecha son de hombres.
- a. Examine únicamente el patrón de los cuatro puntos en la parte inferior izquierda (de mujeres) y determine subjetivamente si parece haber una correlación entre x y y para las mujeres.
- b. Examine únicamente el patrón de los cuatro puntos en la parte superior derecha (de hombres) y determine subjetivamente si parece haber una correlación entre x y y para los hombres.
- c. Calcule el coeficiente de correlación lineal utilizando únicamente los cuatro puntos de la parte inferior izquierda (para mujeres). ¿Los cuatro puntos de la parte superior izquierda (para hombres) tendrán el mismo coeficiente de correlación lineal?

Minitab





- d. Calcule el valor del coeficiente de correlación lineal utilizando los ocho puntos. ¿Qué sugiere ese valor acerca de la relación entre x y y ?
e. Con base en los resultados anteriores, ¿qué concluye? ¿Se deben considerar en conjunto los datos de hombres y mujeres, o al parecer representan dos poblaciones diferentes que deben analizarse por separado?

Prueba de una correlación lineal. En los ejercicios 13 a 28, construya un diagrama de dispersión, calcule el valor del coeficiente de correlación lineal r , calcule el valor crítico de r a partir de la tabla A-6 utilizando una nivel de significancia de $\alpha = 0.05$ y determine si existe una correlación lineal entre las dos variables. Guarde su trabajo, ya que utilizaremos los mismos conjuntos de datos en los ejercicios de la sección 10-3.

- 13. Old Faithful.** Algunos ejemplos de esta sección analizaron la correlación entre la duración y los intervalos de tiempo posteriores a las erupciones del géiser Old Faithful. Utilice los datos que aparecen a continuación (de la tabla 10-1). ¿Existe una correlación lineal entre la altura de una erupción y el intervalo de tiempo posterior a la erupción?

Altura	140	110	125	120	140	120	125	150
Intervalo posterior	92	65	72	94	83	94	101	87

- 14. Oyentes y ventas de canciones.** En la siguiente tabla se listan los números de impresiones de oyentes (en cientos de millones) de canciones y los números correspondientes de álbumes vendidos (en cientos de miles). El número de impresiones de oyentes es un conteo del número de veces que la gente ha escuchado la canción. La tabla se basa en datos de *USA Today*. ¿Parece que las ventas del álbum están muy afectadas por el número de impresiones de oyentes?

Impresiones de oyentes	28	13	14	24	20	18	14	24	17
Álbumes vendidos	19	7	7	20	6	4	5	25	12

- 15. Presupuestos e ingresos brutos de películas.** En la siguiente tabla se muestran los presupuestos (en millones de dólares) y los ingresos brutos (en millones de dólares) de películas seleccionadas al azar (según datos de la Motion Picture Association of America). ¿Parece existir una correlación lineal entre el dinero gastado para filmar la película y la cantidad recuperada en las salas de cine? Además del monto del

presupuesto, identifique otro factor importante que puede afectar la cantidad de dinero que obtiene la película.

Presupuesto	62	90	50	35	200	100	90
Ingresos brutos	65	64	48	57	601	146	47

- 16. Pesos de automóvil y consumo de combustible.** A continuación se presentan los pesos (en libras) y las cantidades de combustible consumidas en carretera (en mi/gal) de automóviles elegidos al azar (Chrysler, Sebring, Ford Mustang, BMW Series 3, Ford Crown Victoria, Honda Civic, Mazda Protégé, Hyundai Accent). ¿Existe una correlación lineal entre el peso y el consumo de combustible en carretera? ¿Qué sugiere el resultado sobre un programa nacional para reducir el consumo de petróleo importado?

Peso	3175	3450	3225	3985	2440	2500	2290
Consumo de combustible	27	29	27	24	37	34	37

- 17. Tamaño del pecho y peso de osos.** A continuación se presenta el tamaño del pecho (en pulgadas) y el peso (en libras) de osos elegidos al azar, que fueron anestesiados y medidos (según datos de Gary Alt y Minitab, Inc.). Como es mucho más difícil pesar un oso que medir el tamaño de su pecho, la presencia de una correlación podría conducir a un método para estimar el peso a partir del tamaño del pecho. ¿Existe una correlación lineal entre el tamaño del pecho y el peso?

Tamaño del pecho	26	45	54	49	35	41	41	49	38	31
Peso	80	344	416	348	166	220	262	360	204	144

- 18. Estaturas y pesos de supermodelos.** Abajo se presentan las estaturas (en pulgadas) y los pesos (en libras) de la supermodelos Michelle Alves, Nadia Avermann, Paris Hilton, Kelly Dyer, Christy Turlington, Bridget Hall, Naomi Campbell, Valerie Mazza y Kristy Hume. ¿Existe una correlación entre la estatura y el peso? Si existe una correlación, ¿eso significa que hay una correlación entre la estatura y el peso de todas las mujeres adultas?

Estatura (pulgadas)	70	70.5	68	65	70	70	70	70	71
Peso (libras)	117	119	105	115	119	127	113	123	115

- 19. Mediciones de presión sanguínea.** Catorce estudiantes de medicina del segundo año midieron la presión sanguínea del mismo paciente; los resultados se presentan a continuación (según datos del doctor Marc Triola). ¿Existe una correlación entre los valores sistólicos y diastólicos? Además de la correlación, ¿hay algún otro método que se podría utilizar para estudiar un aspecto importante que sugieren los datos?

Sistólica	138	130	135	140	120	125	120	130	130	144	143	140	130	150
Diastólica	82	91	100	100	80	90	80	80	80	98	105	85	70	100

- 20. Homicidios y tamaño de población.** En la siguiente tabla se listan los números de homicidios y los tamaños poblacionales (en cientos de miles) de ciudades grandes de Estados Unidos durante un año reciente (según datos del *New York Times*). ¿Qué concluye usted?

Homicidios	258	264	402	253	111	648	288	654	256	60	590
Población	4	6	9	6	3	29	15	38	20	6	81

- 21. Compra de una audiencia televisiva.** El *New York Post* publicó los salarios anuales (en millones de dólares) y el número de televidentes (en millones) correspondientes a Oprah Winfrey, David Letterman, Jay Leno, Kelsey Grammer, Barbara Walters, Dan

Rather, James Gandolfini y Susan Lucci, respectivamente. Los datos se presentan a continuación. ¿Existe una correlación entre el salario y el número de televidentes? ¿Cuál de las estrellas listadas tiene el menor costo por televidente? ¿Y el mayor costo por televidente?

Salario	100	14	14	35.2	12	7	5	1
Televidentes	7	4.4	5.9	1.6	10.4	9.6	8.9	4.2

- 22. Tabaquismo y nicotina.** Cuando la nicotina es absorbida por el cuerpo, se produce cotinina. Por consiguiente, la medición de cotinina es un buen indicador de cuánto fuma una persona. A continuación se incluye el reporte del número de cigarrillos fumados por día y las cantidades medidas de nicotina (en ng/mL). (Los valores provienen de sujetos seleccionados al azar de la National Health Examination Survey). ¿Existe una correlación lineal? Explique el resultado.

x (cigarrillos por día)	60	10	4	15	10	1	20	8	7	10	10	20
y (cotinina)	179	283	75.6	174	209	9.51	350	1.85	43.4	25.1	408	344

- 23. Temperaturas y maratones.** En “The Effects of Temperature on Marathon Runner’s Performance”, de David Martin y John Buonchristiani (*Chance*, vol. 12, núm. 4), se incluyen las altas temperaturas registradas junto con los tiempos (en minutos) de mujeres que ganaron la maratón de la ciudad de Nueva York en años recientes. Los resultados se listan abajo. ¿Existe una correlación entre la temperatura y el tiempo de las triunfadoras? ¿Parece que los tiempos de las ganadoras se ven afectados por la temperatura?

x (temperatura)	55	61	49	62	70	73	51	57
y (tiempo)	145.283	148.717	148.300	148.100	147.617	146.400	144.667	147.533

- 24. Estatura de madre e hija.** A continuación se presentan las estaturas (en pulgadas) de madres e hijas (según datos de la National Health Examination Survey). ¿Al parecer existe una correlación lineal entre las estaturas de las madres y las estaturas de sus hijas?

Estatura de la madre	63	67	64	60	65	67	59	60
Estatura de la hija	58.6	64.7	65.3	61.0	65.4	67.4	60.9	63.1

- 25. Grillos y temperatura.** Una aplicación clásica de la correlación es la asociación entre la temperatura y el número de veces que un grillo chirría en un minuto. A continuación se indican los números de chirridos en un minuto y las temperaturas correspondientes en grados Fahrenheit (según datos de *The Song of Insects*, de George W. Pierce, Harvard University Press). ¿Existe evidencia suficiente para concluir que existe una relación entre el número de chirridos en un minuto y la temperatura?

Chirridos en un minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (°F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

- 26. Incendios y acres quemados.** Abajo se presenta la lista del número de incendios (en miles) y los acres que fueron quemados (en millones) en 11 estados occidentales de Estados Unidos, cada año de la década pasada (según datos de *USA Today*). ¿Existe una correlación? Los datos se listaron bajo el siguiente encabezado: “Leñadores se valen de los incendios para argumentar a favor de incrementar la tala”. ¿Los datos sustentan el argumento de que, a mayor número de árboles talados por los leñadores, menor es el riesgo de incendios porque los bosques son menos densos?

Incendios	73	69	58	48	84	62	57	45	70	63	48
Acres quemados	6.2	4.2	1.9	2.7	5.0	1.6	3.0	1.6	1.5	2.0	3.7

- 27. Estatura y pulso.** Un estudiante de medicina hipotetiza que las personas más altas tienen pulsos más rápidos porque la sangre tiene que viajar más lejos. A continuación se listan los pulsos (en latidos por minuto) y las estaturas (en pulgadas) de una muestra aleatoria de mujeres adultas (según datos de la National Health Examination Survey). Las estaturas y los pulsos aparecen en orden, de manera que corresponden a la misma persona. ¿Parecería que la hipótesis del estudiante de medicina es correcta?

Estatura						Pulso					
64.3	66.4	62.3	62.3	59.6	63.6	76	72	88	60	72	68
59.8	63.3	67.9	61.4	66.7	64.8	80	64	68	68	80	76

- 28. Presupuesto estatal y demora.** El estado de Nueva York ha ganado notoriedad por aprobar el presupuesto estatal después de la fecha límite anual del 1º de abril. A continuación se indican los montos del presupuesto (en miles de millones de dólares ajustados por la inflación) y el número de días de demora, ordenados de tal manera que corresponden al mismo año. (Los datos están en orden por renglón). Al parecer, ¿el monto del presupuesto afecta el número de días que se demora la aprobación? Utilice un nivel de significancia de 0.05.

Presupuesto									Número de días de demora					
101	96	91	85	80	73	71	66	63	63	133	44	45	124	34
62	58	55	52	49	46	44	40	37	35	125	13	125	103	67

Prueba de una correlación lineal. En los ejercicios 29 al 32, utilice los datos del apéndice B para construir un diagrama de dispersión; calcule el valor del coeficiente de correlación lineal r y utilice un nivel de significancia de $\alpha = 0.05$ para determinar si existe una correlación lineal entre las dos variables. Guarde su trabajo porque utilizaremos los mismos conjuntos de datos en la siguiente sección.

- 29. Conjunto de datos del apéndice B: Precios de lista y precios de venta.** Remítase al conjunto de datos 18 del apéndice B y utilice los precios de lista y los precios de venta de las casas vendidas.
- 30. Conjunto de datos del apéndice B: Plástico desechado y tamaño de casas.** Remítase al conjunto de datos 16 del apéndice B y utilice los pesos del plástico desechado y los tamaños correspondientes de las casas.
- 31. Conjunto de datos del apéndice B: Alquitrán y nicotina.** Remítase al conjunto de datos 3 del apéndice B.
- Utilice los datos apareados referentes a alquitrán y nicotina. Con base en el resultado, ¿parece existir una correlación lineal entre el alquitrán y la nicotina del cigarrillo? Si es así, ¿podrán los investigadores reducir sus gastos de laboratorio midiendo únicamente una de estas dos variables?
 - Utilice los datos apareados referentes a monóxido de carbono y nicotina. Con base en el resultado, ¿parece existir una correlación lineal entre el monóxido de carbono y la nicotina de los cigarrillos? Si es así, ¿podrán los investigadores reducir sus gastos de laboratorio midiendo únicamente una de estas dos variables?
 - Suponga que investigadores desean diseñar un método para predecir la cantidad de nicotina y sólo desean medir alguna otra variable. ¿Cuál será una mejor elección, el alquitrán o el monóxido de carbono? ¿Por qué?
- 32. Conjunto de datos del apéndice B: Temperaturas pronosticadas y reales.** Remítase al conjunto de datos 8 del apéndice B.
- Utilice el pronóstico de altas temperaturas para cinco días y las temperaturas máximas reales. ¿Existe correlación? ¿Una correlación lineal implica que las temperaturas pronosticadas a cinco días son exactas?

- b. Utilice el pronóstico de las temperaturas máximas para un día y las temperaturas máximas reales. ¿Existe correlación? ¿Una correlación lineal implica que las temperaturas pronosticadas para un día son exactas?
- c. ¿Cuál esperaría que tuviera una mayor correlación con las temperaturas máximas reales: el pronóstico de las temperaturas máximas para cinco días o el pronóstico las temperaturas máximas para un día? ¿Coincidieron los resultados de los incisos a) y b) con lo que usted esperaba? Si existe una correlación muy alta entre las temperaturas pronosticadas y las temperaturas reales, ¿se infiere que las temperaturas pronosticadas son precisas?

Identificación de errores de correlación. En los ejercicios 33 a 36, describa el error en la conclusión. (Vea la lista de errores comunes incluida en esta sección).

- 33. Se sabe que:** Los datos muestrales apareados de las edades de sujetos y sus puntuaciones en una prueba de razonamiento dan como resultado un coeficiente de correlación lineal muy cercano a 0.

Conclusión: Las personas más jóvenes tienden a obtener puntuaciones más altas.

- 34. Se sabe que:** Existe una correlación lineal significativa entre los ingresos personales y los años de escolaridad.

Conclusión: Una mayor escolaridad causa que se incrementen los ingresos de una persona.

- 35. Se sabe que:** Sujetos resuelven una prueba de habilidades verbales y una prueba de destreza manual, y esos pares de puntuaciones dan como resultado un coeficiente de correlación lineal muy cercano a 0.

Conclusión: Las puntuaciones en ambas pruebas no tienen ninguna relación.

- 36. Se sabe que:** Existe una correlación lineal entre las cargas del impuesto estatal promedio y los ingresos estatales promedio.

Conclusión: Existe una correlación lineal entre las cargas de impuestos individuales y los ingresos individuales.

10-2 MÁS ALLÁ DE LO BÁSICO

- 37. Correlaciones con datos transformados.** Además de probar una correlación lineal entre x y y , con frecuencia podemos utilizar *transformaciones* de datos para explorar otras relaciones. Por ejemplo, podríamos reemplazar cada valor de x por x^2 y emplear los métodos de esta sección para determinar si existe una correlación lineal entre y y x^2 . A partir de los datos apareados en la siguiente tabla, construya el diagrama de dispersión y luego realice una prueba de correlación lineal entre y y cada uno de los siguientes elementos. ¿Cuál de estos casos da por resultado el valor más grande de r ?

- a. x b. x^2 c. $\log x$ d. \sqrt{x} e. $1/x$

x	1.3	2.4	2.6	2.8	2.4	3.0	4.1
y	0.11	0.38	0.41	0.45	0.39	0.48	0.61

- 38. Cálculo de valores críticos de r .** Los valores críticos de r en la tabla A-6 se calculan despejando r a partir de la expresión

$$t = \frac{r}{\sqrt{\frac{1 - r^2}{n - 2}}}$$

continúa

para obtener

$$r = \frac{t}{\sqrt{t^2 + n - 2}}$$

donde el valor t se obtiene de la tabla A-3, suponiendo un caso de dos colas con $n - 2$ grados de libertad. La tabla A-6 lista los resultados para valores seleccionados de n y α . Aplique la fórmula para r dada aquí y consulte la tabla A-3 (con $n - 2$ grados de libertad) para calcular los valores críticos de r en los siguientes casos.

- a. $H_1: \rho \neq 0, n = 50, \alpha = 0.05$
- b. $H_1: \rho \neq 0, n = 75, \alpha = 0.10$
- c. $H_1: \rho < 0, n = 20, \alpha = 0.05$
- d. $H_1: \rho > 0, n = 10, \alpha = 0.05$
- e. $H_1: \rho > 0, n = 12, \alpha = 0.01$

- 39. Construcción de intervalos de confianza para ρ .** Cuando se obtienen muestras de n pares de datos de una población con un coeficiente de correlación lineal ρ , la distribución

de los coeficientes de correlación lineal r no es normal, sino que los valores de

$$z = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$$

tienen una distribución que es aproximadamente normal, con media $\frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right)$ y

desviación estándar $\sqrt{\frac{1}{n-3}}$. Esta conversión de valores r a valores z se conoce como

transformación de Fisher y sirve para construir un estimado de un intervalo de confianza del parámetro poblacional ρ . Utilice el siguiente procedimiento para construir un intervalo de confianza del 95% para ρ , dados 50 pares de datos para los que $r = 0.600$.

Paso a. Utilice la tabla A-2 para calcular $z_{\alpha/2}$ que corresponde al nivel de confianza deseado.

Paso b. Evalúe los límites w_L y w_D del intervalo:

$$w_L = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) - z_{\alpha/2} \cdot \frac{1}{\sqrt{n-3}}$$

$$w_D = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) + z_{\alpha/2} \cdot \frac{1}{\sqrt{n-3}}$$

Paso c. Ahora evalúe los límites del intervalo de confianza en la siguiente expresión.

$$\frac{e^{2w_L} - 1}{e^{2w_L} + 1} < \rho < \frac{e^{2w_D} - 1}{e^{2w_D} + 1}$$

- 40. Potencia de una prueba.** Se utiliza una muestra de tamaño $n = 36$, con un nivel de significancia de 0.05, para contrastar la hipótesis nula de $\rho = 0$ con la hipótesis alternativa de $\rho \neq 0$. Si la población realmente tiene un coeficiente de correlación $\rho = 0.55$, el valor de β es 0.06. ¿Qué indica el valor de β ? ¿Cuál es la potencia de la prueba? Redacte una afirmación que interprete el valor de la potencia.



10-3 Regresión

Concepto clave La sección 10-2 nos dio las herramientas para determinar si existe una correlación lineal entre dos variables. El concepto clave de esta sección es describir la relación entre dos variables por medio del cálculo de la gráfica y la ecuación de la recta que representa mejor la relación. Esta recta se conoce como *recta de regresión* y su ecuación como *ecuación de regresión*. A partir de datos muestrales apareados, calcularemos valores estimados de b_0 , que es la intersección en y , y la pendiente b_1 , de manera que podamos identificar una línea recta con la ecuación $\hat{y} = b_0 + b_1x$. En condiciones adecuadas, esa ecuación resulta útil para hacer predicciones. Existen programas de cómputo y calculadoras para realizar los cálculos aritméticos que son hasta cierto punto engorrosos, de manera que nos enfocaremos en entender los conceptos más que en procesar los datos numéricos. Al igual que en la sección 10-2, esta sección se divide en dos partes: **1.** Conceptos básicos de regresión; **2.** Más allá de los conceptos básicos de regresión. La primera parte incluye conceptos fundamentales que deben quedar muy claros antes de pasar a la segunda parte.

Parte 1: Conceptos básicos de regresión

En algunos casos, dos variables están relacionadas de una forma *determinista*, es decir, dado un valor de una variable, el valor de la otra variable se determina automáticamente sin error. Por ejemplo, el costo total y de un artículo con un precio de lista x y un impuesto de venta del 5% se calcula utilizando la ecuación determinista $y = 1.05x$. Si un artículo tiene un precio de \$50, su costo total será de \$52.50. Este tipo de funciones se estudian ampliamente en los cursos de álgebra. En este capítulo estamos más interesados en los modelos *probabilísticos*, en los que una variable no está determinada por completo por la otra variable. Por ejemplo, la estatura de un niño no está completamente determinada por la estatura del padre (o de la madre). Sir Francis Galton (1822-1911) estudió el fenómeno de la herencia y demostró que cuando parejas altas o bajas tienen hijos, las estaturas de éstos tienden a *regresar* o a revertirse a la estatura media más común de las personas del mismo género. Continuaremos utilizando la terminología de “regresión” de Galton, aun cuando nuestros datos no incluyen el mismo fenómeno de estatura estudiado por Galton.

El recuadro de la página 544 incluye la definición de la ecuación de regresión y de la recta de regresión, así como la notación y las fórmulas que estamos utilizando. La ecuación de regresión expresa una relación entre x (llamada **variable explicativa, variable de predicción o variable independiente**) y \hat{y} (llamada **variable de respuesta o variable dependiente**). La ecuación típica de una línea recta $y = mx + b$ está expresada en la forma $\hat{y} = b_0 + b_1x$, donde b_0 es el intercepto y (o la intersección en y) y b_1 es la pendiente. La notación dada muestra que b_0 y b_1 son estadísticos muestrales utilizados para estimar los parámetros poblacionales β_0 y β_1 . Emplearemos datos muestrales apareados para estimar la ecuación de regresión. Si utilizamos únicamente datos muestrales no podemos calcular los valores exactos de los parámetros poblacionales β_0 y β_1 , pero podemos emplear los datos muestrales para estimarlos con b_0 y b_1 , que se calculan con las fórmulas 10-2 y 10-3.

LA ESTADÍSTICA EN LAS NOTICIAS

Error de pronóstico de 1°F = mil millones de dólares

A pesar de que el pronóstico de las temperaturas puede parecer una ciencia inexacta, muchas compañías están trabajando con fervor para obtener estimaciones más exactas. El reportero de *USA Today*, Del Jones, escribió que “el costo anual de la electricidad podría disminuir por lo menos \$1000 millones si se mejorara la exactitud de las predicciones del tiempo en 1 grado Fahrenheit”. Al referirse a las autoridades de Tennessee Valley, afirma que “los pronósticos sobre sus 80,000 millas cuadradas han fallado un promedio de 2.35 grados durante los últimos 2 años, que es bastante representativo de los pronósticos que se hacen a nivel nacional. Si se mejorara en 1.35 grados, esto ahorraría al Tennessee Valley tanto como \$100,000 diarios y tal vez más”. El pronóstico de temperaturas se utiliza para determinar la distribución de la energía proveniente de generadores, plantas nucleares, plantas hidroeléctricas, de carbón, de gas natural y eólicas. Las técnicas de pronóstico estadístico se encuentran en proceso de refinamiento, de manera que se pueda ahorrar dinero y recursos naturales.

Requisitos

1. La muestra de datos apareados (x, y) es una muestra *aleatoria* de datos cuantitativos.
2. El examen visual del diagrama de dispersión indica que los puntos se aproximan al patrón de una línea recta.
3. Se debe eliminar cualquier valor extremo, si se sabe que es un error. Es importante tomar en cuenta los efectos de cualquier valor extremo que no sea un error conocido.

Nota: Los requisitos 2 y 3 representan una verificación simplificada de los siguientes requisitos formales del análisis de regresión:

- Para cada valor fijo de x , los valores correspondientes de y tienen una distribución en forma de campana.
- Para los distintos valores fijos de x , las distribuciones de los valores correspondientes de y tienen la misma varianza. (Esto se viola si parte del diagrama de dispersión presenta puntos muy cercanos a la línea de regresión, mientras otra porción del diagrama presenta puntos que se alejan mucho de la línea de regresión. Consulte la explicación de los puntos residuales casi al final de esta sección).
- Para los distintos valores fijos de x , las distribuciones de los valores correspondientes de y tienen medias que se ubican en la misma línea recta.
- Los valores de y son independientes.

Los resultados no se ven muy afectados si la distribución no se aleja demasiado de la normalidad y si las varianzas no son demasiado diferentes.

Definiciones

Dado un conjunto de datos muestrales apareados, la **ecuación de regresión**

$$\hat{y} = b_0 + b_1x$$

describe algebraicamente la relación entre las dos variables. La gráfica de la ecuación de regresión se denomina **recta de regresión** (o *recta del mejor ajuste* o *recta de mínimos cuadrados*).

Notación para la ecuación de regresión

Parámetro poblacional Estadístico muestral

Intercepto y de la ecuación de regresión	β_0	b_0
Pendiente de la ecuación de regresión	β_1	b_1
Ecuación de la recta de regresión	$y = \beta_0 + \beta_1x$	$\hat{y} = b_0 + b_1x$

Cálculo de la pendiente b_1 y de b_0 (el intercepto y) en la ecuación de regresión $\hat{y} = b_0 + b_1x$

Fórmula 10-2 Pendiente: $b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$

Fórmula 10-3 Intercepto y : $b_0 = \bar{y} - b_1\bar{x}$

El intercepto y , b_0 , también puede calcularse por medio de la siguiente fórmula, pero es mucho más fácil utilizar la fórmula 10-3.

$$b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Tal vez las fórmulas 10-2 y 10-3 resulten intimidantes, pero están programadas en muchas calculadoras y programas de cómputo, de manera que los valores de b_0 y b_1 se calculan con facilidad. (Véase el recuadro “Uso de la tecnología” al final de esta sección). Una vez que hemos evaluado b_1 y b_0 , podemos identificar la ecuación estimada de regresión, que tiene la siguiente propiedad especial: *La recta de regresión es la que se ajusta mejor a los puntos muestrales.* (El criterio específico utilizado para determinar cuál recta se ajusta “mejor” es la propiedad de los mínimos cuadrados, que se describirá más adelante). Ahora estudiaremos brevemente el redondeo y después ejemplificaremos el procedimiento del cálculo y la aplicación de la ecuación de regresión.

Redondeo de la pendiente b_1 y de b_0 (el intercepto y)

Redondeo de b_1 y b_0 a tres dígitos significativos. Es difícil dar una regla universal sencilla para redondear los valores de b_1 y b_0 , pero esta regla servirá en la mayor parte de las situaciones de este libro. Dependiendo de la forma de redondeo, las respuestas a los ejemplos y ejercicios de este libro pueden variar un poco de las respuestas de usted.

EJEMPLO Cálculo de la ecuación de regresión En la sección 10-2 empleamos la muestra aleatoria simple de los valores listados abajo para calcular el coeficiente de correlación lineal de $r = -0.956$. (Al usar los métodos de la sección 10-2, concluimos que existe una correlación lineal entre x y y). Use los datos muestrales indicados para calcular la ecuación de regresión.

x	3	1	3	5
y	5	8	6	4

SOLUCIÓN

REQUISITO ✓ Se trata de una muestra aleatoria simple. El diagrama de dispersión generado por SPSS, que se presenta en la sección 10-2, sugiere un patrón de puntos similar al patrón de una recta. No hay valores extremos. Procedemos a calcular la pendiente y el intercepto de la recta de regresión. ✓

Calculamos la ecuación de regresión por medio de las fórmulas 10-2 y 10-3, y los siguientes valores ya se encuentran en la tabla 10-2 de la sección 10-2:

$$\begin{array}{lll} n = 4 & \Sigma x = 12 & \Sigma y = 23 \\ \Sigma x^2 = 44 & \Sigma y^2 = 141 & \Sigma xy = 61 \end{array}$$

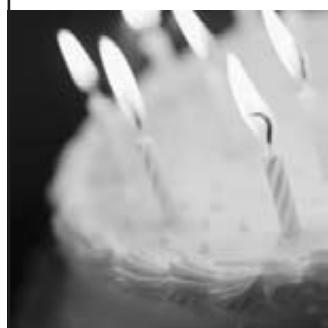
Primero calcule b_1 usando la fórmula 10-2:

$$\begin{aligned} b_1 &= \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2} \\ &= \frac{4(61) - (12)(23)}{4(44) - (12)^2} = \frac{-32}{32} = -1 \end{aligned}$$

Luego, encuentre b_0 , el intercepto en y , utilizando la fórmula 10-3 (con $\bar{y} = 23/4 = 5.75$ y $\bar{x} = 12/4 = 3$):

$$\begin{aligned} b_0 &= \bar{y} - b_1 \bar{x} \\ &= 5.75 - (-1)(3) = 8.75 \end{aligned}$$

continúa



Posposición de la muerte

Varios estudios tratan sobre la capacidad de las personas para retrasar su muerte hasta después de un suceso importante. Por ejemplo, el sociólogo David Phillips analizó las tasas de mortalidad de hombres judíos que murieron cerca de la Pascua judía, y descubrió que la tasa de mortalidad disminuía drásticamente una semana antes de esa festividad, pero que aumentaba la semana posterior a ella. Un estudio más reciente sugiere que las personas no tienen esta capacidad de posponer la muerte. Con base en los registros de 1.3 millones de muertes, este estudio no encontró una relación entre el momento de la muerte y la Navidad, el Día de Acción de Gracias o el cumpleaños del individuo. El doctor Donn Young, uno de los investigadores, dijo que “el hecho es que, la muerte no toma en cuenta el calendario. Usted no puede anotar en su Palm Pilot ‘Bien, cenemos el viernes y agendaré mi muerte para el domingo’”. David Phillips rebatió estos resultados y dijo que el estudio se enfocó en pacientes con cáncer, quienes tienen menos probabilidades de presentar efectos psicosomáticos.

Conociendo la pendiente b_1 y b_0 (el intercepto y), ahora podemos expresar la ecuación estimada de la recta de regresión como

$$\hat{y} = 8.75 - 1x$$

Debemos estar conscientes de que esta ecuación es un *estimado* de la verdadera ecuación de regresión $y = \beta_0 + \beta_1 x$. Este estimado se basa en un conjunto particular de datos muestrales, pero otra muestra obtenida de la misma población quizás generaría una ecuación ligeramente diferente.



EJEMPLO Old Faithful Con los datos de las duraciones y los intervalos posteriores a las erupciones de la tabla de 10-1, calculamos que el coeficiente de correlación lineal es $r = 0.926$. Utilice los mismos datos muestrales y calcule la recta de regresión.

SOLUCIÓN

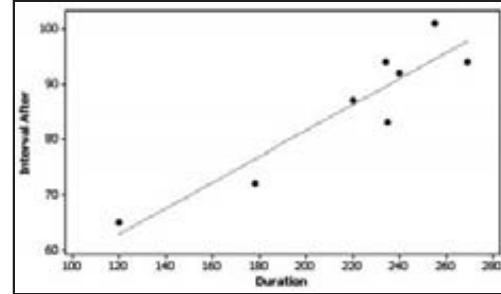
REQUISITO ✓ La figura 10-1a) es un diagrama de dispersión de esos puntos, el cual muestra un patrón que se aproxima a una línea recta. No hay valores extremos y los datos muestrales se obtuvieron al azar. Ahora podemos continuar con el cálculo de la ecuación de la recta de regresión. ✓

Si utilizamos el mismo procedimiento ilustrado en el ejemplo anterior o empleamos herramientas tecnológicas, podemos calcular que los 8 pares de datos duración/intervalos posteriores a la erupción de la tabla 10-1 dan como resultado $b_0 = 34.8$ y $b_1 = 0.234$. Abajo se presentan los resultados de Minitab. Sustituyendo los valores calculados para b_0 y b_1 , expresamos la ecuación de regresión como $\hat{y} = 34.8 + 0.234x$. A continuación se presenta también el diagrama de dispersión generado por Minitab, con la recta de regresión incluida. Podemos ver que la recta de regresión se ajusta bien a los datos.

Minitab

The regression equation is Interval After = 34.8 + 0.234 Duration					
Predictor	Coeff	SE Coef	T	P	
Constant	34.778	0.732	3.98	0.007	
Duration	0.23406	0.03908	5.99	0.001	
	$S = 4.97392$	R-Sq = 85.7%	R-Sq(adj) = 83.3%		

Minitab



Uso de la ecuación de regresión para hacer predicciones

Las ecuaciones de regresión a menudo se utilizan para *predecir* el valor de una variable, dado algún valor particular de la otra variable. Si la recta de regresión se ajusta bastante bien a los datos, entonces es sensato utilizar su ecuación para hacer predicciones, siempre y cuando no vayamos más allá del alcance de los valores disponibles. No haga predicciones con base en valores que rebasan las fronteras de los datos muestrales conocidos. Por ejemplo, si utilizamos los datos muestrales de la tabla 10-1, un diagrama de dispersión sugiere que las duraciones y los intervalos de tiempo posteriores a las erupciones se ajustan bastante bien al patrón de una línea recta, de manera que si observamos una erupción de 180 segundos, podemos prede-

cir el intervalo posterior a la erupción (hasta la siguiente erupción) al sustituir $x = 180$ en la ecuación de regresión $\hat{y} = 34.8 + 0.234x$. El siguiente resultado indica que si una erupción tiene una duración de 180 segundos, el mejor intervalo de tiempo predicho después de la erupción es de 76.9 minutos. (Recuerde, las duraciones están en segundos y los intervalos posteriores a las erupciones están en minutos).

$$\begin{aligned}\hat{y} &= 34.8 + 0.234x \\ &= 34.8 + 0.234(180) = 76.9 \text{ min (redondeado)}\end{aligned}$$

Sin embargo, *debemos utilizar la ecuación de la recta de regresión sólo si la ecuación de regresión es un buen modelo para los datos*. La adecuación de la ecuación de regresión se puede juzgar al probar la significancia del coeficiente de correlación lineal r . Observe el siguiente procedimiento en el que se utiliza la ecuación de regresión para hacer predicciones. Vea también el siguiente ejemplo, que formaliza el procedimiento para utilizar una duración de 180 segundos para predecir el intervalo posterior a la erupción.

Parte 2: Más allá de los conceptos básicos de regresión

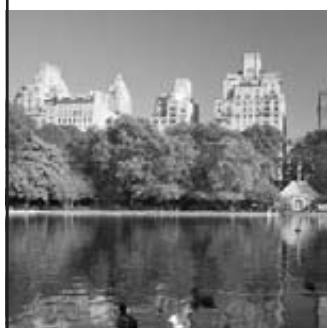
Uso de la ecuación de regresión para hacer predicciones (continuación)

Ya señalamos que debemos utilizar la ecuación de la recta de regresión para hacer predicciones sólo si la ecuación de regresión es un buen modelo para los datos. Para ser más precisos, *debemos utilizar la ecuación de regresión para hacer predicciones sólo si existe una correlación lineal. En ausencia de una correlación lineal, no debemos usar la ecuación de regresión para proyectar o predecir; en vez de ello, nuestro mejor estimado de la segunda variable es simplemente su media muestral*.

Al predecir un valor de y con base en algún valor dado de x . . .

- Si no existe una correlación lineal, el mejor valor predicho de y es \bar{y} .**
- Si existe una correlación lineal, el mejor valor predicho de y se calcula sustituyendo el valor de x en la ecuación de regresión.**

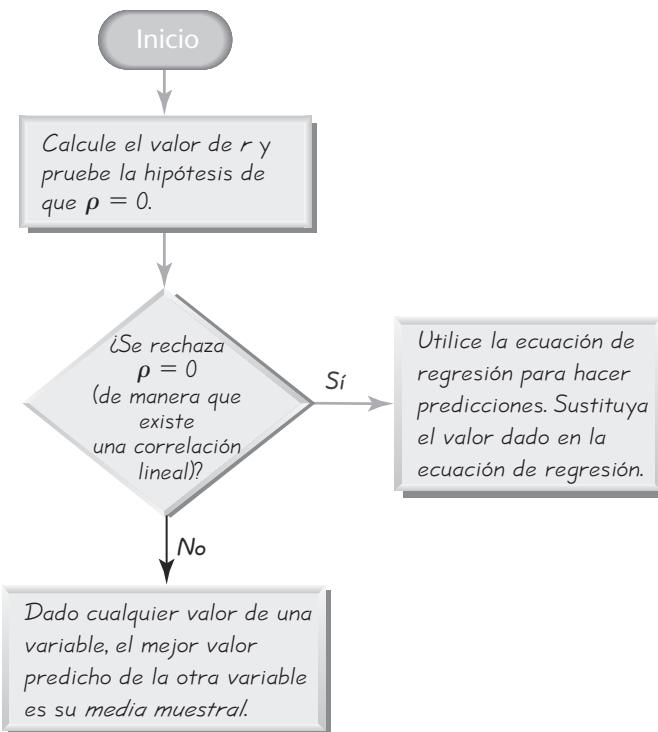
¿Qué tan buena es la ecuación de regresión como modelo para los datos poblacionales? En la figura 10-7 se resume el proceso para hacer predicciones, el cual se comprende con mayor facilidad si pensamos en r como una medida de lo bien que se ajusta la recta de regresión a los datos muestrales. Además de reflexionar sobre la presencia o ausencia de una correlación lineal, también debemos considerar lo siguiente: las ecuaciones de regresión obtenidas de datos muestrales apareados con r muy cercana a -1 o $+1$ (porque los puntos en el diagrama de dispersión se acercan mucho a la recta de regresión) tienen mayores probabilidades de ser mejores modelos para los datos poblacionales que las ecuaciones de regresión de conjuntos de datos con valores de r que no son tan cercanos a -1 o $+1$ (aun cuando r resulte significativa). Con algunas muestras muy grandes de datos apareados, podríamos encontrar que r es significativa, aun cuando tenga un valor relativamente pequeño como 0.200. En este caso, la ecuación de regresión podría ser un modelo aceptable, pero las predicciones no serían tan exactas porque r no se acerca tanto a -1 o $+1$. Si r no es significativa, entonces la recta de regresión no se ajusta bien a los datos, y la ecuación de regresión no debe utilizarse para hacer predicciones.



Predicción de precios de condominios

Un estudio masivo incluyó 99,491 operaciones de venta de condominios y cooperativas en Manhattan. El estudio utilizó 41 variables diferentes para predecir el valor del condominio o de la cooperativa. Las variables incluían la condición de la vivienda, el vecindario, la antigüedad, el tamaño y si contaba con portero. He aquí algunas conclusiones: si todos los factores permanecen igual, un condominio vale 15.5% más que una cooperativa; una chimenea incrementa el valor de un condominio en un 9.69% e incrementa el valor de una cooperativa en un 11.36%; una recámara adicional en un condominio aumenta su valor en un 7.11% y en una cooperativa en un 18.53%. Este uso de los métodos estadísticos permite a los compradores y a los vendedores estimar el valor con mucha mayor exactitud. Los métodos de regresión múltiple (sección 10-5) se utilizan cuando existe más de una variable de predicción, como en este estudio. (Según datos de "So How Much Is That . . . Worth", de Dennis Havesi, *New York Times*).

Figura 10-7
Procedimiento para hacer predicciones



PROBLEMA
10
DEL CAPÍTULO

EJEMPLO **Predicción de las erupciones del Old Faithful** Con los datos de las duraciones y los intervalos posteriores a las erupciones de la tabla de 10-1, encontramos que existe una correlación lineal entre las dos variables y también que la ecuación de regresión es $\hat{y} = 34.8 + 0.234x$. Suponiendo que la erupción actual tiene una duración de $x = 180$ segundos, calcule el mejor valor de predicción de y , el intervalo posterior a esta erupción (que es el tiempo predicho hasta la siguiente erupción).

SOLUCIÓN

REQUISITO ✓ Quizá nos sintamos tentados a apresurarnos y sustituir x por 180 en la ecuación de regresión, pero primero debemos considerar si existe una correlación lineal que justifique el uso de esa ecuación. En este ejemplo, el diagrama de dispersión indica que los puntos se aproximan al patrón de una línea recta, por lo que sí tenemos una correlación lineal (con $r = 0.926$), de manera que el valor predicho se calcula como se indica a continuación. ✓

Puesto que existe una correlación lineal, el valor predicho se calcula sustituyendo en la ecuación de regresión como sigue:

$$\begin{aligned}\hat{y} &= 34.8 + 0.234x \\ &= 34.8 + 0.234(180) = 76.9 \text{ min (redondeado)}\end{aligned}$$

El intervalo de tiempo predicho desde la erupción actual hasta la siguiente erupción es de 76.9 minutos. (Si no hubiera una correlación lineal, el mejor valor predicho sería $\bar{y} = 688/8 = 86.0$ min).

EJEMPLO Medida de sombrero y CI Obviamente no existe una correlación lineal entre la medida de sombrero de las personas y sus puntuaciones de CI. Dado que un individuo utiliza un sombrero de tamaño 7, calcule el mejor valor predicho de la puntuación de CI de esta persona.

SOLUCIÓN

REQUISITO ✓ Puesto que no existe una correlación lineal, no empleamos la ecuación de la recta de regresión para hacer predicciones. ✓

Puesto que no hay una correlación lineal, no usamos una ecuación de regresión. No hay necesidad de reunir datos muestrales apareados consistentes en las medidas de sombrero y las puntuaciones de CI de una muestra de adultos seleccionados al azar. En vez de ello, la mejor puntuación de CI predicha es sencillamente la media del CI de todos los adultos, que es de 100.

Compare con cuidado las soluciones de los dos ejemplos anteriores y note que utilizamos la ecuación de regresión cuando existía una correlación lineal; sin embargo, en ausencia de esa correlación, el mejor valor predicho de y es sencillamente el valor de la media muestral \bar{y} . Un error común es el uso de la ecuación de regresión para hacer una predicción cuando no existe una correlación lineal. Este error viola el primero de los siguientes lineamientos.

Lineamientos para el uso de la ecuación de regresión

1. *Si no existe una correlación lineal, no utilice la ecuación de regresión para hacer predicciones.*
2. *Cuando utilice la ecuación de regresión para hacer predicciones, permanezca en el ámbito de los datos muestrales disponibles.* Si usted calcula una ecuación de regresión que relaciona la estatura y el número de calzado de mujeres, es absurdo predecir el número de calzado de una mujer que mide 10 pies de estatura.
3. *Una ecuación de regresión que está basada en datos antiguos no necesariamente es válida ahora.* La ecuación de regresión que relaciona precios de automóviles usados con la antigüedad de los automóviles ya no es útil si está basada en datos de la década de 1990.
4. *No haga predicciones acerca de una población distinta de la población de donde se obtuvieron los datos muestrales.* Si reunimos datos muestrales de hombres y desarrollamos una ecuación de regresión que relaciona la edad con el uso del control remoto del televisor, los resultados no necesariamente se aplican a las mujeres. Si empleamos *promedios* estatales para desarrollar una ecuación de regresión que relaciona las calificaciones de matemáticas del SAT con las calificaciones verbales del SAT, los resultados no necesariamente se aplican a los *individuos*.

Interpretación de la ecuación de regresión: Cambio marginal

Podemos utilizar la ecuación de regresión para observar el efecto en una variable, cuando la otra variable cambia una cantidad específica.



Definición

Cuando se trabaja con dos variables relacionadas por una ecuación de regresión, el **cambio marginal** en una variable es la cantidad que cambia cuando la otra variable cambia exactamente una unidad. La pendiente b_1 en la ecuación de regresión representa el cambio marginal que ocurre en y cuando x cambia una unidad.

Para los datos de la tabla 10-1 sobre las duraciones y los intervalos de tiempo posteriores a las erupciones del géiser Old Faithful, la recta de regresión tiene una pendiente de 0.234, lo que demuestra que si incrementamos x (las duraciones) en un segundo, el intervalo posterior a la erupción que se predijo se incrementará 0.234 minutos. Es decir, por cada segundo adicional de duración, esperamos que el intervalo posterior a la erupción (hasta la siguiente erupción) aumente 0.234 minutos más.

Valores extremos y puntos de influencia

Un análisis de correlación/regresión de datos bivariados (apareados) debe incluir la investigación de *valores extremos* y *puntos de influencia*, que se definen a continuación.

Definiciones

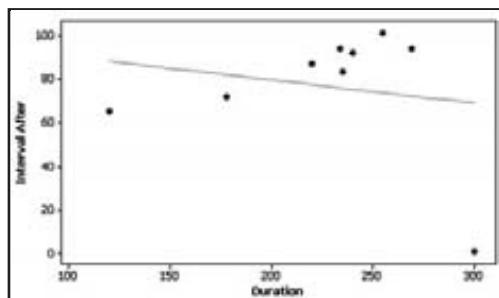
En un diagrama de dispersión, un **valor extremo** es un punto que aparece muy lejos de los otros puntos de datos.

Los datos muestrales apareados pueden incluir uno o más **puntos de influencia**, que son puntos que afectan fuertemente la gráfica de la recta de regresión.

Es fácil identificar un valor extremo: examine el diagrama de dispersión e identifique un punto que se aleja de los otros puntos. He aquí cómo determinamos un punto de influencia: grafique la recta de regresión que resulta de los datos con el punto incluido, después grafique la recta de regresión resultante de los datos sin incluir el punto. Si la gráfica cambia de forma considerable, se trata de un punto de influencia. Los puntos de influencia a menudo se encuentran al identificar los valores extremos que están alejados *horizontalmente* de los demás puntos.

Por ejemplo, observe el diagrama de dispersión de la figura 10-1a) en la página 516. Suponga que incluimos el siguiente par adicional de datos: $x = 300$, $y = 1$ (una erupción dura 300 segundos y el intervalo posterior a la erupción es de 1 minuto). Este punto adicional sería un punto de influencia porque la gráfica de la recta de regresión cambiaría considerablemente, tal como se observa en la siguiente pantalla de Minitab. Compare esta recta de regresión con la que se presentó en la imagen previa de Minitab, y observará con claridad que la incorporación de ese par de valores tiene un efecto drástico en la recta de regresión.

Minitab



Residuales y la propiedad de los mínimos cuadrados

Hemos establecido que la ecuación de regresión representa la recta que se ajusta “mejor” a los datos, y ahora describiremos el criterio utilizado para determinar

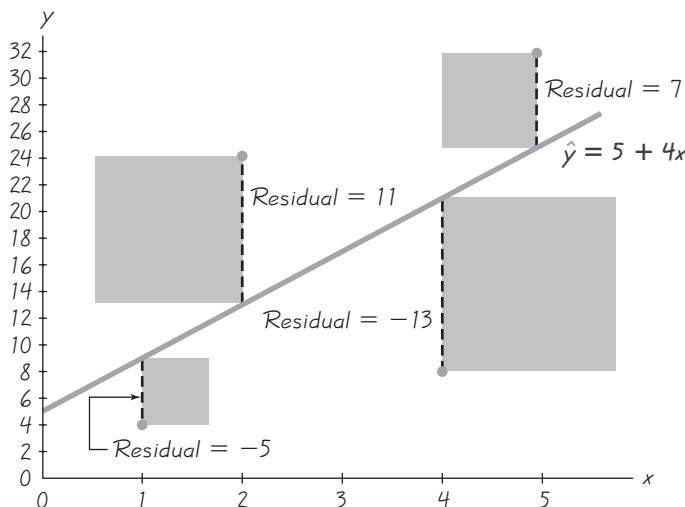


Figura 10-8
Residuales y cuadrados
de los residuales

cuál recta es mejor que todas las demás. Este criterio se basa en las distancias verticales entre los puntos de datos originales y la recta de regresión. Tales distancias se denominan *residuales*.

Definición

Para una muestra de datos apareados (x, y) , un **residual** es la diferencia $(y - \hat{y})$ entre un valor y muestral observado y el valor de \hat{y} , que es el valor de y predicho por medio de la ecuación de regresión. Es decir,

$$\text{residual} = \text{observada } y - \text{predicha } y = y - \hat{y}$$

Esta definición puede parecer tan clara como las instrucciones de un formulario fiscal, pero usted comprenderá fácilmente los residuales si se remite a la figura 10-8, que corresponde a los datos muestrales apareados que se presentan al margen. En la figura 10-8, los residuales están representados por las líneas punteadas. Para tener un ejemplo específico, observe el residual indicado como 7, que se encuentra directamente por arriba de $x = 5$. Si sustituimos $x = 5$ en la ecuación de regresión $\hat{y} = 5 + 4x$, obtenemos un valor predicho de $\hat{y} = 25$. Cuando $x = 5$, el valor *predicho* de y es $\hat{y} = 25$, pero el valor muestral real *observado* es $y = 32$. La diferencia $y - \hat{y} = 32 - 25 = 7$ es un residual.

La ecuación de regresión representa la recta que se ajusta “mejor” a los puntos, de acuerdo con la siguiente *propiedad de mínimos cuadrados*.

x	1	2	4	5
y	4	24	8	32

Definición

Una recta satisface la **propiedad de mínimos cuadrados** si la suma de los cuadrados de los residuales es la menor suma posible.

En la figura 10-8 podemos observar que los residuales son $-5, 11, -13$ y 7 , de manera que la suma de sus cuadrados es

$$(-5)^2 + 11^2 + (-13)^2 + 7^2 = 364$$

Podemos visualizar la propiedad de mínimos cuadrados si nos remitimos a la figura 10-8, donde los cuadrados de los residuales están representados por las áreas de los cuadrados sombreados. La suma de las áreas sombreadas cuadradas es 364, que es la menor suma posible. Utilice cualquier otra recta y los cuadrados sombreados se combinarán para producir una área mayor que el área sombreada combinada de 364.

Por fortuna, no necesitamos lidiar directamente con la propiedad de mínimos cuadrados cuando deseamos obtener la ecuación de la recta de regresión. Ya se realizaron los cálculos para satisfacer la propiedad de mínimos cuadrados en las fórmulas 10-2 y 10-3. Puesto que la obtención de estas fórmulas requiere del cálculo, no las incluimos en este libro.

Gráficas residuales

En esta sección, al igual que en la anterior, describimos requisitos simplificados para el análisis eficaz de los resultados de correlación y regresión. Indicamos que siempre debemos comenzar con un diagrama de dispersión y que debemos verificar que el patrón de puntos se aproxime a una línea recta. También es necesario tomar en cuenta los valores extremos. La *gráfica residual* puede ser otra herramienta útil para analizar resultados de correlación y regresión, así como para verificar los requisitos necesarios para hacer inferencias sobre una correlación y una regresión.



Definición

Una **gráfica residual** es un diagrama de dispersión de los valores (x, y) una vez que cada uno de los valores de la coordenada y han sido reemplazados por el valor residual $y - \hat{y}$ (donde \hat{y} denota el valor predicho de y). Es decir, una gráfica residual es una gráfica de los puntos $(x, y - \hat{y})$.

Para construir una gráfica residual, utilice el mismo eje x que en el diagrama de dispersión, pero use un eje vertical de valores residuales. Dibuje una línea horizontal de referencia a partir del valor residual de 0 y luego grafique los valores apareados de $(x, y - \hat{y})$. Puesto que la construcción manual de gráficas residuales podría convertirse en lo que los matemáticos denominan “tedioso”, se recomienda el uso de un programa de cómputo. Cuando analice una gráfica residual, busque un patrón en la configuración de los puntos y utilice los siguientes criterios:

Si una gráfica residual no revela ningún patrón, la ecuación de regresión es una buena representación de la asociación entre las dos variables.

Si una gráfica residual revela algún patrón sistemático, la ecuación de regresión no es una buena representación de la asociación entre las dos variables.

Considere los siguientes ejemplos de pantallas de Minitab. En el caso 1 todo está bien; los puntos se aproximan a la recta de regresión, de manera que la ecuación de regresión es un buen modelo para describir la asociación entre las dos variables. La gráfica residual correspondiente no revela un patrón diferente.

El caso 2 genera un diagrama de dispersión que señala una asociación entre las dos variables, pero la relación no es lineal. La gráfica residual correspondiente indica un patrón diferente, lo que confirma que el modelo lineal no es un buen modelo en este caso.

El caso 3 tiene un diagrama de dispersión en el que los puntos se van alejando de la recta de regresión, y la gráfica residual revela un patrón de variación creciente,

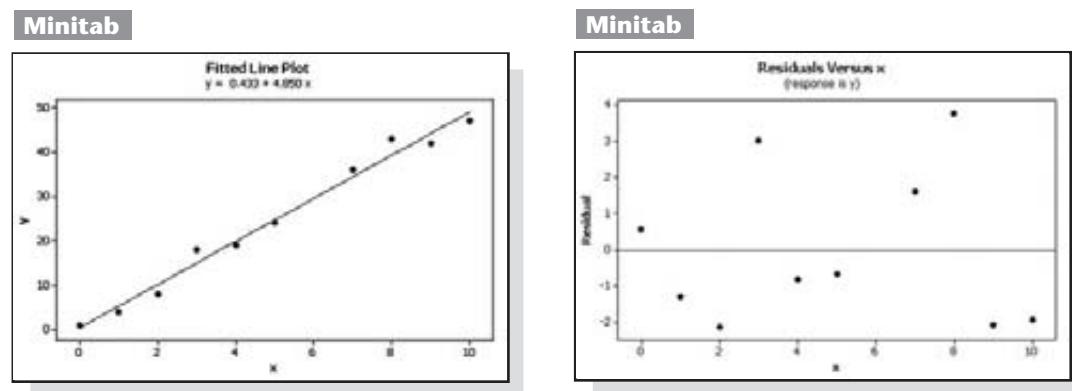
lo cual viola el requisito de que, para diferentes valores de x , las distribuciones de los valores de y deben tener la misma varianza. En este caso, probablemente la ecuación de regresión no sea un buen modelo.

Después de adquirir la habilidad de calcular y analizar gráficas residuales junto con diagramas de dispersión, estamos más preparados para verificar los requisitos necesarios para asegurarnos de la validez de las inferencias que hacemos a partir de los procedimientos de correlación y regresión.

Caso 1

x	0	1	2	3	4	5	7	8	9	10
y	1	4	8	18	19	24	36	43	42	47

La recta de regresión se ajusta bien a los datos. La gráfica residual no revela ningún patrón.

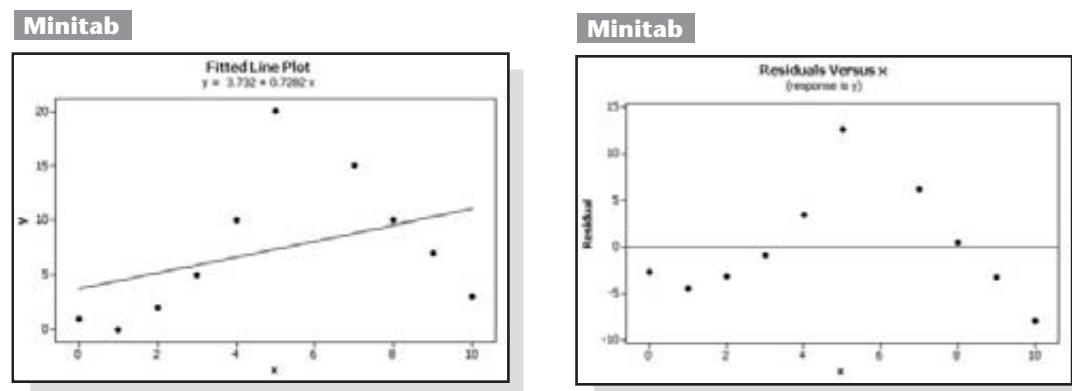


Caso 2

x	0	1	2	3	4	5	7	8	9	10
y	1	0	2	5	10	20	15	10	7	3

El diagrama de dispersión muestra que la asociación no es lineal.

La gráfica residual revela un patrón diferente.



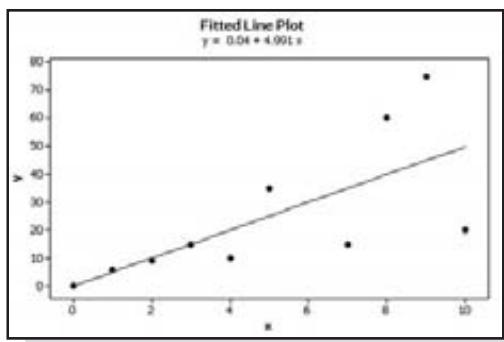
Caso 3

x	0	1	2	3	4	5	7	8	9	10
y	0	6	9	15	10	35	15	60	75	20

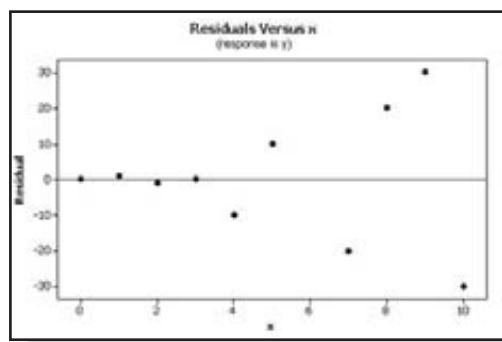
El diagrama de dispersión muestra la variación creciente de los puntos, alejándose de la recta de regresión.

La gráfica residual revela este patrón: de izquierda a derecha los puntos están más dispersos. (Esto va en contra del requisito de que, para diferentes valores de x , las distribuciones de los valores de y deben tener la misma varianza).

Minitab



Minitab



Uso de la tecnología

A causa de los complejos cálculos implicados, el coeficiente de correlación lineal r , la pendiente y el intercepto y de la recta de regresión suelen calcularse por medio de una calculadora o un programa de cómputo.

STATDISK Primero ingrese los datos apareados en columnas de la ventana de datos de Statdisk. Seleccione **Analysis** de la barra del menú principal, después utilice la opción **Correlation and Regression**. Introduzca el valor para el nivel de significancia y seleccione las columnas de datos. Haga clic en el botón de **Evaluate**. Los resultados incluyen el valor del coeficiente de correlación lineal, junto con el valor crítico de r , la conclusión de la correlación, el intercepto y la pendiente de la ecuación de regresión, así como otros resultados. Haga clic en **Scatterplot** para obtener una gráfica del diagrama de dispersión con la recta de regresión incluida.

MINITAB Primero introduzca los valores de x en la columna C1 y los valores de y en la columna C2. En la sección 10-2 vimos que

podemos obtener el valor del coeficiente de correlación lineal r seleccionando **Stat/Basic Statistics/Correlation**. Para obtener la ecuación de la recta de regresión, seleccione **Stat/Regression/Regresión**, introduzca C2 en “respuesta” y C1 en “predictor”. Para obtener la gráfica del diagrama de dispersión con la recta de regresión, seleccione **Stat/Regression/Fitted Line Plot**, después introduzca C2 en la variable de respuesta y C1 en la variable de predicción. Seleccione el modelo “lineal”.

EXCEL Introduzca los datos apareados en las columnas A y B. Utilice el complemento de análisis de datos de Excel seleccionando **Tools** del menú principal, después seleccione **Data Analysis y Regression**, luego haga clic en **OK**. Introduzca el rango para los valores de y , como B1:B10. Introduzca el rango para los valores de x , como A1:A10. Haga clic en el recuadro adyacente a Line Fit Plots, después haga clic en **OK**. De toda la información proporcionada por Excel, la pendiente y el intercepto de la ecuación de regresión aparecen en la tabla con el encabezado “Coefficient”. La gráfica presentada incluirá un diagrama de dispersión de los puntos muestrales origina-

les, junto con los puntos que serían predichos por la ecuación de regresión. La recta de regresión se obtiene fácilmente conectando los puntos “predichos de y ”.

Para emplear el complemento Data Desk XL, haga clic en **DDXL** y seleccione **Regression**, luego haga clic en el recuadro **Function Type** y seleccione **Simple Regression**. Haga clic en el ícono del lápiz para la variable de respuesta e introduzca el rango de valores para la variable y (o dependiente). Haga clic en el ícono del lápiz para la variable de respuesta e introduzca el rango de valores para la variable x (o independiente). Haga clic en **OK**. La pendiente y el intercepto de la ecuación de regresión se encuentran en la tabla con el encabezado “Coefficient”.

TI-83/84 PLUS Introduzca los datos apareados en las listas L1 y L2, luego presione **STAT** y seleccione **TESTS**; después elija la opción **LinRegTTest**. El despliegue de resultados incluirá el intercepto y la pendiente de la ecuación de regresión. La calculadora TI-83/84 Plus representa los valores b_0 y b_1 como a y b .

10-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Variables de predicción y de respuesta.** ¿Qué es una variable de predicción? ¿Qué es una variable de respuesta? ¿Por qué cree usted que se les nombró de esa forma específica?
- Recta con el mejor ajuste.** ¿En qué sentido la recta de regresión es la línea recta que ajusta “mejor” los puntos en un diagrama de dispersión?
- Predicciones.** ¿Por qué no es adecuado utilizar una ecuación de regresión para predecir un valor de y utilizando un valor de x que está más allá del ámbito de los datos muestrales disponibles?
- Requisitos.** ¿Los requisitos para el análisis de regresión son idénticos a los requisitos para el análisis de correlación? Si no es así, ¿en qué difieren?

Haciendo predicciones. En los ejercicios 5 a 8, utilice los datos indicados para calcular el mejor valor predicho de la variable dependiente. Asegúrese de seguir el procedimiento para predicciones descrito en esta sección.

- Puntuaciones de CI de gemelos separados al nacer.** Se obtuvieron las puntuaciones de CI de gemelos separados al nacer, elegidos al azar. Para 20 de estos gemelos, el coeficiente de correlación lineal es 0.870 y la ecuación de la recta de regresión es $\hat{y} = -3.22 + 1.02x$, donde x representa la puntuación de CI del gemelo que nació primero. Asimismo, los 20 valores de x tienen una media de 104.2 y los 20 valores de y tienen una media de 103.1. ¿Cuál es el mejor CI predicho de un gemelo que nació en segundo lugar, dado que el gemelo que nació primero tiene un CI de 110?
- Puntuaciones de CI de niños adoptados que viven juntos.** Se obtuvieron las puntuaciones de CI de niños adoptados que viven juntos, elegidos al azar. Para 20 de estos niños, el coeficiente de correlación lineal es 0.027, y la ecuación de la recta de regresión es $\hat{y} = -3.22 + 1.02x$, donde x representa la puntuación de CI del niño más grande. Asimismo, los 20 valores de x tienen una media de 104.2 y los 20 valores de y tienen una media de 103.1. ¿Cuál es el mejor CI predicho del niño adoptado más pequeño, dado que el niño adoptado más grande en la misma familia tiene un CI de 110?
- Azúcar y calorías en el cereal.** El autor reunió datos de 16 cereales, que consisten en los contenidos de azúcar (en gramos por gramo de cereal) y las calorías (por gramo de cereal). Se utilizó STATDISK para calcular el coeficiente de correlación lineal $r = 0.765$ y la ecuación de la recta de regresión $\hat{y} = 3.46 + 1.01x$, donde x representa el contenido de azúcar. Asimismo, la cantidad media de calorías es 3.76 calorías por gramo de cereal. ¿Cuál es el mejor contenido de calorías predicho de un cereal con 0.40 gramos de azúcar por gramo de cereal?
- Acciones y Súper Bowl.** Se registró el valor máximo del Promedio Industrial de Dow Jones (DJIA) y el número total de puntos anotados en el Súper Bowl durante un periodo que comprende 21 años. Se utilizó Excel para calcular que el valor del coeficiente de correlación lineal es $r = -0.133$ y que la ecuación de regresión es $\hat{y} = 53.3 - 0.000442x$, donde x es el valor máximo del DJIA. Además, la media de los puntos anotados en el Súper Bowl es de 51.4. ¿Cuál es el mejor valor predicho para el número total de puntos anotados en el Súper Bowl en un año, con un valor máximo del DJIA de 1200?

Cálculo de la ecuación de la recta de regresión. En los ejercicios 9 y 10, utilice los datos indicados para calcular la ecuación de la recta de regresión.

9.	x	1	1	2	5	3
	y	0	2	2	3	5

10.	x	0	2	1	4	8	
	y	6	5	4	0	6	

11. **Efectos de un valor extremo.** Remítase al diagrama de dispersión generado por Minitab, que se presenta en el ejercicio 11 de la sección 10-2.

- a. Utilice los pares de valores de los 10 puntos y calcule la ecuación de la recta de regresión.
- b. Despues de eliminar el punto con las coordenadas (10, 10), utilice los pares de valores de los nueve puntos restantes y calcule la ecuación de la recta de regresión.
- c. Compare los resultados de los incisos a) y b).

12. **Efectos de aglomerados.** Remítase al diagrama de dispersión generado por Minitab, que se presenta en el ejercicio 12 de la sección 10-2.

- a. Utilice los pares de valores de los 8 puntos y calcule la ecuación de la recta de regresión.
- b. Utilice únicamente los pares de valores de los cuatro puntos en la esquina inferior izquierda y calcule la ecuación de la recta de regresión.
- c. Utilice únicamente los pares de valores de los cuatro puntos en la esquina superior derecha y calcule la ecuación de la recta de regresión.
- d. Compare los resultados de los incisos a), b) y c).

Cálculo de la ecuación de la recta de regresión y predicciones. En los ejercicios 13 a 32, utilice los mismos conjuntos de datos que en los ejercicios de la sección 10-2. En cada caso, calcule la ecuación de regresión, permitiendo que la primera variable sea la variable independiente (x). Calcule los valores predichos indicados. Sugerencia: Al calcular los valores predichos, asegúrese de seguir los procedimientos para predicciones descritos en esta sección.

13. **Old Faithful.** Calcule el mejor tiempo predicho del intervalo posterior a una erupción (hasta la siguiente erupción) dado que la erupción actual tiene una altura de 100 pies.

Altura	140	110	125	120	140	120	125	150	
Intervalo posterior	92	65	72	94	83	94	101	87	

14. **Oyentes y ventas de canciones.** Calcule el mejor número predicho de álbumes vendidos por una canción con 20 (en cientos de millones) impresiones de oyentes. (En la siguiente tabla, las impresiones de oyentes están en cientos de millones y el número de álbumes vendidos está en cientos de miles).

Impresiones de oyentes	28	13	14	24	20	18	14	24	17	
Álbumes vendidos	19	7	7	20	6	4	5	25	12	

15. **Presupuestos e ingresos brutos de películas.** Calcule la mejor cantidad predicha de los ingresos brutos de una película con un presupuesto de \$40 millones. (En la siguiente tabla, todas las cantidades están en millones de dólares).

Presupuesto	62	90	50	35	200	100	90	
Ingresos brutos	65	64	48	57	601	146	47	

16. **Pesos de automóviles y consumo de combustible.** Calcule la mejor cantidad predicha de combustible consumido en carretera (en mi/gal), para un automóvil que pesa 3000 libras.

Peso	3175	3450	3225	3985	2440	2500	2290	
Consumo de combustible	27	29	27	24	37	34	37	

17. **Tamaño del pecho y peso de osos.** Calcule el mejor peso predicho (en libras) de un oso cuyo pecho mide 50 pulgadas.

Tamaño del pecho	26	45	54	49	35	41	41	49	38	31
Peso	80	344	416	348	166	220	262	360	204	144

- 18. Estaturas y pesos de supermodelos.** Calcule el mejor peso predicho de una supermodelo que mide 72 pulgadas.

Estatura (in)	70	70.5	68	65	70	70	70	70	71
Peso (lb)	117	119	105	115	119	127	113	123	115

- 19. Mediciones de presión sanguínea.** Calcule la mejor presión sanguínea diastólica predicha para una persona con una lectura sistólica de 140.

Sistólica	138	130	135	140	120	125	120	130	130	144	143	140	130	150
Diastólica	82	91	100	100	80	90	80	80	80	98	105	85	70	100

- 20. Homicidios y tamaño de la población.** Calcule el mejor tamaño poblacional predicho para una ciudad con 120 homicidios. (El tamaño de las poblaciones está en cientos de miles).

Homicidios	258	264	402	253	111	648	288	654	256	60	590
Población	4	6	9	6	3	29	15	38	20	6	81

- 21. Compra de una audiencia televisiva.** Calcule el mejor número predicho de televidentes para una estrella de televisión que tiene un salario de \$2 millones. (En la siguiente tabla, los salarios están en millones de dólares y los números de televidentes en millones).

Salario	100	14	14	35.2	12	7	5	1
Televidentes	7	4.4	5.9	1.6	10.4	9.6	8.9	4.2

- 22. Tabaquismo y nicotina.** Calcule el mejor nivel de cotinina predicho para una persona que fuma 40 cigarrillos al día.

x (cigarrillos por día)	60	10	4	15	10	1	20	8	7	10	10	20
y (cotinina)	179	283	75.6	174	209	9.51	350	1.85	43.4	25.1	408	344

- 23. Temperaturas y maratones.** Calcule el mejor tiempo ganador predicho para la maratón de 1990, cuando la temperatura era de 73°F. ¿Qué diferencia hay entre el tiempo ganador predicho y el tiempo ganador real de 150.750 minutos?

x (temperatura)	55	61	49	62	70	73	51	57
y (tiempo)	145.283	148.717	148.300	148.100	147.617	146.400	144.667	147.533

- 24. Estatura de madre e hija.** Calcule la mejor estatura predicha de la hija de una mujer que mide 66 pulgadas de estatura.

Estatura de la madre	63	67	64	60	65	67	59	60
Estatura de la hija	58.6	64.7	65.3	61.0	65.4	67.4	60.9	63.1

- 25. Grillos y temperatura.** Calcule la mejor temperatura predicha para un momento en el que un grillo está emitiendo 1000 chirridos por minuto.

Chirridos en un minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (en °F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

- 26. Incendios y acres quemados.** Calcule el mejor número predicho de acres quemados, dado que hubo 50,000 incendios. (En la siguiente tabla, los números de incendios están en miles y los acres en millones).

Incendios	73	69	58	48	84	62	57	45	70	63	48
Acres quemados	6.2	4.2	1.9	2.7	5.0	1.6	3.0	1.6	1.5	2.0	3.7

- 27. Estatura y pulso.** Calcule el mejor pulso predicho para una mujer que mide 66 pulgadas.

Estatura (in)	Pulso (latidos/min)
64.3 66.4 62.3 62.3 59.6 63.6	76 72 88 60 72 68
59.8 63.3 67.9 61.4 66.7 64.8	80 64 68 68 80 76

- 28. Presupuesto estatal y demora.** Calcule el mejor número predicho de días que se retrasa un presupuesto del estado de Nueva York, dado que el monto del presupuesto es de \$104 mil millones. (En la siguiente tabla, los montos del presupuesto están en miles de millones de dólares).

Presupuesto	Número de días de retraso
101 96 91 85 80 73 71 66 63 63	133 44 45 124 34 125 13 125 103 67
62 58 55 52 49 46 44 40 37 35	68 4 1 64 48 18 19 10 4 4

- 29. Conjunto de datos del apéndice B: Precios de lista y precios de venta.** Remítase al conjunto de datos 18 del apéndice B y utilice los precios de lista y los precios de venta de las casas. Calcule el mejor precio de venta predicho para una casa que tiene un precio de lista de \$400,000.

- 30. Conjunto de datos del apéndice B: Plástico desecharo y tamaño de casas.** Remítase al conjunto de datos 16 del apéndice B y utilice los pesos del plástico desecharo y los tamaños correspondientes de las casas. Calcule el mejor tamaño predicho de una casa, dado que la casa desecha 5.00 libras de plástico.

- 31. Conjunto de datos del apéndice B: Elementos dañinos en cigarrillos.** Remítase al conjunto de datos 3 del apéndice B.

- a. Utilice los datos apareados referentes a alquitrán (x) y nicotina (y). ¿Cuál es el mejor contenido predicho de nicotina de un cigarrillo con 15 mg de alquitrán?
- b. Utilice los datos apareados referentes a monóxido de carbono (x) y nicotina (y). ¿Cuál es el mejor nivel predicho de nicotina para un cigarrillo con 15 mg de monóxido de carbono?

- 32. Conjunto de datos del apéndice B: Temperaturas pronosticadas y reales.** Remítase al conjunto de datos 8 del apéndice B.

- a. Utilice el pronóstico de las temperaturas máximas para cinco días (x) y las temperaturas máximas reales (y). ¿Cuál es la mejor temperatura máxima real predicha si la temperatura máxima pronosticada para cinco días es de 70 grados?
- b. Utilice el pronóstico de las temperaturas máximas para un día (x) y las temperaturas máximas reales (y). ¿Cuál es la mejor temperatura máxima predicha si la temperatura máxima pronosticada para un día es de 70 grados?

10-3 MÁS ALLÁ DE LO BÁSICO

- 33. Prueba de hipótesis equivalentes.** Explique por qué la prueba de la hipótesis nula $H_0: \rho = 0$ es equivalente a la prueba de la hipótesis nula $H_0: \beta_1 = 0$ donde ρ es el coeficiente de correlación lineal para una población de datos apareados, y β_1 es la pendiente de la recta de regresión para la misma población.

- 34. Prueba de la propiedad de mínimos cuadrados.** Según la propiedad de mínimos cuadrados, la recta de regresión minimiza la suma de los cuadrados de los residuales. Señalamos que, con los datos apareados al margen, la ecuación de regresión es $\hat{y} = 5 + 4x$ y la suma de cuadrados de los residuales es 364. Demuestre que la ecuación $\hat{y} = 8 + 3x$ da por resultado una suma de cuadrados de residuales mayor que 364.

- 35. Uso de logaritmos para transformar datos.** Si un diagrama de dispersión revela un patrón no lineal (es decir, sin una recta), que usted reconoce como otro tipo de curva, podría aplicar los métodos de esta sección. Para los datos presentados al margen, calcule la ecuación lineal ($y = b_0 + b_1x$) que se ajuste mejor a los datos muestrales y calcule la

x	1	2	4	5
y	4	24	8	32

ecuación logarítmica ($y = a + b \ln x$) que se ajuste mejor a los datos muestrales. (Sugerencia: Inicie reemplazando cada valor de x por $\ln x$). ¿Cuál de estas dos ecuaciones se ajusta mejor a los datos? ¿Por qué?

x	2.0	2.5	4.2	10.0
y	12.0	18.7	53.0	225.0

36. Gráfica residual. Considere los datos de la siguiente tabla.

- Examine los datos e identifique la relación entre x y y .
- Calcule el coeficiente de correlación lineal y úselo para determinar si existe una correlación lineal significativa entre x y y .
- Construya un diagrama de dispersión. ¿Qué sugiere el diagrama acerca de la relación entre x y y ?
- Construya una gráfica residual. ¿Existe algún patrón evidente? ¿Qué sugiere la gráfica residual acerca de la relación entre x y y ?

x	0	1	2	3	4	5	6	7	8	9	10
y	0	1	4	9	16	25	36	49	64	81	100

10-4 Variación e intervalos de predicción

Concepto clave Utilizando datos apareados (x, y) describiremos la variación que puede explicarse por la correlación lineal entre x y y y la variación que no puede explicarse. Luego, procederemos a considerar un método para construir un *intervalo de predicción*, que es un estimado del intervalo de un valor predicho de y . (Los estimados de intervalos de parámetros se conocen como *intervalos de confianza*, en tanto que los estimados de intervalos de variables suelen denominarse *intervalos de predicción*).

Variación explicada y sin explicar

Ahora examinaremos medidas de *desviación* y *variación* para un par de valores (x, y). En vez de contemplar abstracciones, consideraremos el caso específico descrito en la figura 10-9. Imagine una muestra de datos apareados (x, y) que incluye (5, 19). Suponga que utilizamos esta muestra de datos apareados para calcular los siguientes resultados:

- Existe una correlación lineal (con r significativamente diferente de 0).
- La ecuación de la recta de regresión es $\hat{y} = 3 + 2x$.
- La media de los valores de y está dada por $\bar{y} = 9$.
- Uno de los pares de datos muestrales es $x = 5$ y $y = 19$.
- El punto (5, 13) es uno de los puntos sobre la recta de regresión, ya que la sustitución de $x = 5$ en la ecuación de regresión produce $\hat{y} = 13$.

La figura 10-9 indica que el punto (5, 13) está sobre la recta de regresión, pero el punto (5, 19) del conjunto de datos original no se ubica en la recta de regresión. Si ignoramos por completo los conceptos de correlación y regresión, y deseamos predecir un valor de y dado un valor de x y un conjunto de datos apareados (x, y), nuestra mejor conjectura sería la media \bar{y} . Pero en este caso, con una correlación lineal significativa, la forma de predecir el valor de y cuando $x = 5$ consiste en usar la ecuación de regresión para obtener $\hat{y} = 13$. Podemos explicar la discrepancia entre $\bar{y} = 9$ y $\hat{y} = 13$ al señalar que existe una relación lineal mejor descrita por medio de la recta de regresión. Como consecuencia, cuando $x = 5$, el valor predicho de y es 13 y no el valor medio de 9. Para $x = 5$, el valor predicho de y es 13, pero el valor muestral observado de y es en realidad 19. La discrepancia entre $\hat{y} = 13$ y $y = 19$ no puede explicarse por medio de la recta de regresión y se le denomina *desviación sin explicación* o *residual*. Esta desviación sin explicar se expresa en símbolos como $y - \hat{y}$.



El Súper Bowl como pronosticador del mercado de valores

La “superstición del Súper Bowl” afirma que una victoria por parte de un equipo de la NFL original es seguida por un año en el que aumenta el índice de cotizaciones del mercado de valores de Nueva York; si no resulta victorioso un equipo que originalmente formó parte de la NFL, el índice cae. (En 1970, la NFL y la AFL se unieron para formar la NFL actual). Después de los primeros 29 juegos del Súper Bowl, la predicción fue correcta el 90% de las veces, pero ha tenido mucho menos éxito en años recientes. Hasta el momento en que se escribe esto, ha sido correcta en 29 de 38 juegos del Súper Bowl, con una tasa de éxito del 76%. Los pronósticos y las predicciones son objetivos importantes de los especialistas en estadística y de los consejeros de inversiones, pero el sentido común sugiere que nadie debe basar sus inversiones en el resultado de un juego de fútbol. Otros indicadores que se utilizan para pronosticar el desempeño del mercado de valores incluyen el aumento del tamaño del dobladillo de las faldas, las ventas de aspirina, las limosinas en Wall Street, los pedidos de cajas de cartón, las ventas de cerveza contra las ventas de vino y el tránsito de los elevadores en la Bolsa de Valores de Nueva York.

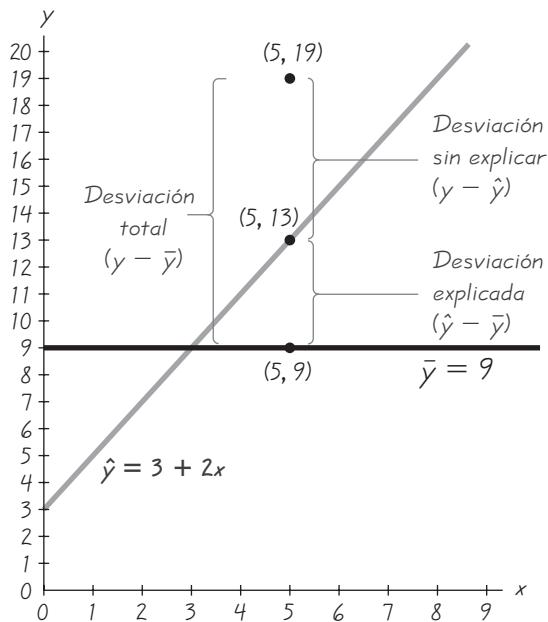


Figura 10-9 Desviación sin explicar, explicada y total

Igual que en la sección 3-3, donde definimos la desviación estándar, nuevamente considere que una *desviación* es la diferencia entre un valor y la media. (En este caso, la media es $\bar{y} = 9$.) Examine con atención la figura 10-9 y observe las siguientes desviaciones específicas a partir de $\bar{y} = 9$:

Desviación total (a partir de $\bar{y} = 9$) del punto $(5, 19) = y - \bar{y} = 19 - 9 = 10$

Desviación explicada (a partir de $\bar{y} = 9$) del punto $(5, 19) = \hat{y} - \bar{y} = 13 - 9 = 4$

Desviación sin explicar (a partir de $\bar{y} = 9$) del punto

$$(5, 19) = y - \hat{y} = 19 - 13 = 6$$

Estas desviaciones a partir de la media se generalizan y definen formalmente como sigue.

Definiciones

Suponga que tenemos un conjunto de datos apareados que contienen el punto muestral (x, y) , que \hat{y} es el valor predicho de y (obtenido por medio de la ecuación de regresión), y que la media de los valores y muestrales es \bar{y} .

La **desviación total** de (x, y) es la distancia vertical $y - \bar{y}$, que es la distancia entre el punto (x, y) y la recta horizontal que pasa por la media muestral \bar{y} .

La **desviación explicada** es la distancia vertical $\hat{y} - \bar{y}$, que es la distancia entre el valor predicho \hat{y} y la recta horizontal que pasa por la media muestral \bar{y} .

La **desviación sin explicar** es la distancia vertical $y - \hat{y}$, que es la distancia vertical entre el punto (x, y) y la recta de regresión. (La distancia $y - \hat{y}$ también se conoce como *residual*, tal como se definió en la sección 10-3).

En la figura 10-9 podemos apreciar la siguiente relación:

$$\begin{aligned} \text{(desviación total)} &= \text{(desviación explicada)} + \text{(desviación sin explicar)} \\ (y - \bar{y}) &= (\hat{y} - \bar{y}) + (y - \hat{y}) \end{aligned}$$

Esta última expresión implica desviaciones a partir de la media y se aplica a cualquier punto (x, y) particular. Si sumamos los cuadrados de las desviaciones utilizando todos los puntos (x, y) , obtenemos cantidades de **variación**, y la misma relación se aplica a las sumas de cuadrados que se muestran en la fórmula 10-4, aunque esta última expresión no es algebraicamente equivalente a la fórmula 10-4. En esa fórmula, la **variación total** se expresa como la suma de los cuadrados de los valores de desviación totales, la **variación explicada** es la suma de los cuadrados de los valores de desviación explicados, y la **variación sin explicar** es la suma de los cuadrados de los valores de desviación sin explicar.

Fórmula 10-4

$$\begin{aligned} \text{(variación total)} &= \text{(variación explicada)} + \text{(variación sin explicar)} \\ \text{o } \Sigma(y - \bar{y})^2 &= \Sigma(\hat{y} - \bar{y})^2 + \Sigma(y - \hat{y})^2 \end{aligned}$$

En la sección 10-2 vimos que el coeficiente de correlación lineal r se utiliza para calcular la proporción de la variación total en y que puede explicarse por medio de la correlación lineal. En la sección 10-2 hicimos la siguiente afirmación:

El valor de r^2 es la proporción de la variación en y que está explicada por la relación lineal entre x y y .

Esta aseveración sobre la variación explicada se formaliza en la siguiente definición.

Definición

El **coeficiente de determinación** es la cantidad de variación en y que está explicada por la recta de regresión. Se calcula como

$$r^2 = \frac{\text{variación explicada}}{\text{variación total}}$$

Podemos calcular r^2 por medio de la definición dada con la fórmula 10-14 o podemos simplemente elevar al cuadrado el coeficiente de correlación lineal r .



EJEMPLO Old Faithful En la sección 10-2 utilizamos las duraciones y los intervalos posteriores a las erupciones de la tabla 10-1 para calcular que $r = 0.926$. Calcule el coeficiente de determinación. También, obtenga el porcentaje de la variación total en y (intervalo posterior a la erupción) que puede explicarse por medio de la relación lineal entre la duración y el intervalo posterior a una erupción.

SOLUCIÓN El coeficiente de determinación es $r^2 = 0.926^2 = 0.857$. Como r^2 es la proporción de la variación total que está explicada, concluimos que aproximadamente el 86% de la variación total de los intervalos posteriores a las erupciones se pueden explicar por las duraciones. Esto significa que alrededor del 14% de la variación total de los intervalos posteriores a las erupciones está explicada por otros factores y no por las duraciones. Sin embargo, recuerde que esos resultados son estimaciones que se basan en los datos muestrales con que se cuenta. Es probable que otros datos muestrales produzcan estimaciones diferentes.

Intervalos de predicción

En la sección 10-3 empleamos los datos muestrales de la tabla 10-1 para calcular la ecuación de regresión $\hat{y} = 34.8 + 0.234x$, donde \hat{y} representa el intervalo de tiempo predicho (en minutos) después de una erupción y hasta la siguiente, en tanto que x representa la duración de una erupción (en segundos). Después utilizamos esa ecuación para predecir el valor de y , dado que una erupción tiene una duración de $x = 180$ segundos. Encontramos que el mejor intervalo de tiempo predicho después de una erupción es de 76.9 minutos. Puesto que 76.9 es un único valor, se le conoce como *estimado puntual*. En el capítulo 7 aprendimos que los estimados puntuales tienen la grave desventaja de no darnos ninguna información acerca de su exactitud. Aquí, sabemos que 76.9 es el mejor valor predicho, pero no sabemos qué tan exacto es este valor. En el capítulo 7 elaboramos estimados del intervalo de confianza para superar esa desventaja, y en esta sección seguiremos el mismo método; utilizaremos un **intervalo de predicción**, que es un estimado del intervalo de un valor predicho de y . [Un estimado del intervalo de un *parámetro* (como la media de todos los intervalos posteriores a las erupciones) suele denominarse *intervalo de confianza*, mientras que el estimado de un intervalo de una *variable* (como el intervalo estimado posterior a una erupción con una duración de 180 segundos) se conoce generalmente como *intervalo de predicción*].

La creación de un intervalo de predicción requiere una medida de la dispersión de los puntos muestrales alrededor de la recta de regresión. Recuerde que la desviación sin explicar (o residual) es la distancia vertical entre un punto muestral y la recta de regresión, tal como se ilustra en la figura 10-9. El *error estándar del estimado* es una medida colectiva de la dispersión de los puntos muestrales alrededor de la recta de regresión, y se define de manera formal como sigue.

Definición

El **error estándar del estimado**, denotado por s_e , es una medida de las diferencias (o distancias) entre los valores muestrales observados de y y los valores predichos \hat{y} que se obtienen por medio de la ecuación de regresión. Está dado por

$$s_e = \sqrt{\frac{\sum(y - \hat{y})^2}{n - 2}} \quad (\text{donde } \hat{y} \text{ es el valor predicho de } y)$$

o por medio de la siguiente fórmula equivalente:

$$\text{Fórmula 10-5} \quad s_e = \sqrt{\frac{\sum y^2 - b_0 \sum y - b_1 \sum xy}{n - 2}}$$

STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus están diseñados para calcular de manera automática el valor de s_e . Consulte el recuadro “Uso de la tecnología” al final de esta sección.

El cálculo del error estándar del estimado s_e se asemeja mucho al de la desviación estándar ordinaria que se explicó en la sección 3-3. Así como la desviación estándar es una medida de la desviación de los valores a partir de su media, el error estándar del estimado s_e es una medida de la desviación de los puntos de los datos muestrales a partir de su recta de regresión. La lógica que subyace en la división entre $n - 2$ es similar a la lógica que condujo a la división entre $n - 1$ para la desviación estándar ordinaria. Es importante señalar que valores relativamente pequeños de s_e reflejan puntos que están cercanos a la recta de regresión, y los valores relativamente grandes se presentan cuando hay puntos que se alejan de la recta de regresión.

La fórmula 10-5 es algebraicamente equivalente a la otra ecuación en la definición, pero la fórmula 10-5 suele ser más fácil ya que no requiere que calculemos cada uno de los valores predichos \hat{y} por medio de sustitución en la ecuación de regresión. Sin embargo, la fórmula 10-5 sí requiere que calculemos el intercepto y , b_0 , y la pendiente b_1 de la recta de regresión estimada.

EJEMPLO Cálculo de s_e Utilice la fórmula 10-5 para calcular el error estándar del estimado s_e para los datos de las duraciones y los intervalos posteriores a las erupciones del géiser Old Faithful, que aparecen en la tabla 10-1.

SOLUCIÓN Con los datos muestrales de la tabla 10-1, calculamos estos valores:

$$n = 8 \quad \Sigma y^2 = 60,204 \quad \Sigma y = 688 \quad \Sigma xy = 154,378$$

En la sección 10-3 empleamos los datos muestrales de la tabla 10-1 para obtener el intercepto y y la pendiente de la recta de regresión. Esos valores se presentan aquí con más decimales para una mayor precisión.

$$b_0 = 34.7698041 \quad b_1 = 0.2340614319$$

Ahora podemos usar estos valores en la fórmula 10-5 para calcular el error estándar del estimado s_e .

$$\begin{aligned} s_e &= \sqrt{\frac{\Sigma y^2 - b_0 \Sigma y - b_1 \Sigma xy}{n - 2}} \\ &= \sqrt{\frac{60,204 - (34.7698041)(688) - (0.2340614319)(154,378)}{8 - 2}} \\ &= 4.973916052 = 4.97 \quad (\text{redondeado}) \end{aligned}$$

Es posible medir la dispersión de los puntos muestrales alrededor de la recta de regresión con el error estándar del estimado $s_e = 4.97$. Podemos emplear el error estándar del estimado s_e para construir estimados de intervalo que nos ayuden a ver qué tan confiables son realmente nuestros estimados puntuales de y . Suponga que para cada valor fijo de x , los valores muestrales correspondientes de y se distribuyen normalmente alrededor de la recta de regresión, y que estas distribuciones normales tienen la misma varianza. El siguiente estimado del intervalo se aplica a un valor y *individual*. (Consulte el ejercicio 26 para ver un intervalo de confianza utilizado para predecir la *media* de todos los valores de y , para algún valor dado de x).

Intervalo de predicción para una y individual

Dado el valor fijo x_0 , el intervalo de predicción para una y individual es

$$\hat{y} - E < y < \hat{y} + E$$

donde el margen de error E es

$$E = t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

y x_0 representa el valor dado de x , $t_{\alpha/2}$ tiene $n - 2$ grados de libertad, y s_e se calcula a partir de la fórmula 10-5.



EJEMPLO Old Faithful Para los pares de datos de la tabla 10-1 sobre las duraciones y los intervalos posteriores a las erupciones, encontramos que para una duración de 180 segundos el mejor intervalo posterior a la erupción predicho es 76.9 min. Construya un intervalo de predicción del 95% para el intervalo posterior a una erupción, dado que la duración de la erupción es de 180 segundos (de manera que $x = 180$). Esto nos dará una idea de qué tan exacto es el valor predicho de 76.9 min.

SOLUCIÓN En secciones anteriores hemos demostrado que $r = 0.926$, de manera que existe una correlación lineal (a un nivel de significancia de 0.05), y la ecuación de regresión es $\hat{y} = 34.8 + 0.234x$. En el ejemplo anterior encontramos que $s_e = 4.973916052$, y los siguientes estadísticos se obtienen de los datos de la tabla 10-1 sobre las duraciones y los intervalos posteriores a las erupciones:

$$n = 8 \quad \bar{x} = 218.875 \quad \Sigma x = 1751 \quad \Sigma x^2 = 399,451$$

En la tabla A-3 encontramos que $t_{\alpha/2} = 2.447$. (Utilizamos $8 - 2 = 6$ grados de libertad con $\alpha = 0.05$ en dos colas). Primero calculamos el margen de error E permitiendo que $x_0 = 180$, ya que buscamos el intervalo de predicción del intervalo posterior a la erupción, dado que la duración es $x = 180$.

$$\begin{aligned} E &= t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}} \\ &= (2.447)(4.973916052) \sqrt{1 + \frac{1}{8} + \frac{8(180 - 218.875)^2}{8(399,451) - (1751)^2}} \\ &= (2.447)(4.973916052)(1.103758562) = 13.4 \quad (\text{redondeado}) \end{aligned}$$

Con $\hat{y} = 76.9$ y $E = 13.4$, obtenemos el intervalo de predicción de la siguiente forma:

$$\hat{y} - E < y < \hat{y} + E$$

$$76.9 - 13.4 < y < 76.9 + 13.4$$

$$63.5 < y < 90.3$$

Es decir, para una erupción con una duración de 180 segundos, tenemos una certeza del 95% de que el intervalo posterior a la erupción está entre 63.5 y 90.3 minutos. Se trata de un rango relativamente grande. (Un factor que contribuye a la gran dimensión del rango es que el tamaño muestral es muy pequeño, pues estamos utilizando únicamente 8 pares de datos muestrales).

Se puede emplear Minitab para calcular los límites del intervalo de predicción. Si utilizamos Minitab, obtendremos el resultado de (63.47, 90.33), bajo el encabezado “95.0% P.I.”. Éste corresponde al mismo intervalo de predicción calculado antes.

Además de saber que para una erupción que dura $x = 180$ segundos, el intervalo predicho desde la última erupción hasta la siguiente es de 76.9 minutos, ahora tenemos una idea de qué tan confiable es realmente el estimado. El intervalo de predicción del 95% calculado en este ejemplo indica que el valor real de y puede variar sustancialmente del valor predicho de 76.9 minutos.

Uso de la tecnología

STATDISK STATDISK puede utilizarse para calcular el coeficiente de correlación lineal r , la ecuación de la recta de regresión, el error estándar del estimado s_e , la variación total, la variación explicada, la variación sin explicar y el coeficiente de determinación. Introduzca los datos apareados en las columnas de la ventana de datos de STATDISK, seleccione **Analysis** de la barra del menú principal, después utilice la opción **Correlation and Regression**. Introduzca un valor para el nivel de significancia y seleccione las columnas de datos. Haga clic en el botón **Evaluate**. Los resultados de STATDISK incluirán el coeficiente

de correlación lineal, el coeficiente de determinación, la ecuación de regresión y el valor del error estándar del estimado s_e .

MINITAB Minitab puede utilizarse para calcular la ecuación de regresión, el error estándar del estimado s_e (denotado por S), el valor del coeficiente de determinación (denotado por R-sq) y los límites del intervalo de predicción. Ingrese los datos de x en la columna C1 y los datos de y en la columna C2, luego seleccione las opciones **Stat**, **Regression** y **Regression**. Introduzca C2 en el recuadro denominado "Response" e introduzca C1 en el recuadro denominado "Predictors". Si busca un intervalo de predicción para algún valor dado de x , haga clic en **Options** e introduzca el valor deseado de x_0 en el recuadro denominado "Prediction intervals for new observations".

EXCEL Excel sirve para calcular la ecuación de regresión, el error estándar del estimado s_e y el coeficiente de determinación (denotado por R square). Introduzca los datos apareados en las columnas A y B.

Para emplear el complemento Data Analysis, seleccione **Tools** del menú principal, después

elija **Data Analysis**, seguido por **Regression** y después haga clic en **OK**. Ingrese el rango para los valores de y , como B1:B8. Ingrese el rango para los valores de x , como A1:A8. Haga clic en **OK**.

Para emplear el complemento Data Desk XL haga clic en **DDXL** y seleccione **Regression**, luego haga clic en el recuadro **Function Type** y seleccione **Sample Regression**. Haga clic en el ícono del lápiz para la variable de respuesta e introduzca el rango de valores para la variable y . Haga clic en el ícono del lápiz para la variable explicativa e introduzca el rango de valores para la variable x . Haga clic en **OK**.

TI-83/84 PLUS La calculadora TI-83/84 Plus se puede utilizar para calcular el coeficiente de correlación lineal r , la ecuación de la recta de regresión, el error estándar del estimado s_e y el coeficiente de determinación (denominado por r^2). Ingrese los datos apareados en las listas L1 y L2, después presione **STAT** y seleccione **TESTS**, luego elija la opción **LinRegTTest**. Para Xlist ingrese L1, para Ylist ingrese L2, utilice un valor Freq (frecuencia) de 1 y seleccione $\neq 0$. Desplace la pantalla hacia abajo hasta Calculate y luego presione la tecla **Enter**.



10-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Intervalo de predicción.** Con sus propias palabras, describa un intervalo de predicción.
- Coeficiente de determinación.** En general, ¿que es un coeficiente de determinación? ¿Cuál es el coeficiente de determinación para $r = 0.400$? ¿Qué información práctica nos brinda?
- Desviación y variación explicadas.** ¿Cuál es la diferencia entre la *desviación explicada* y la *variación explicada*?
- Variación explicada y sin explicar.** ¿Cuál es la diferencia entre la variación explicada y la variación sin explicar?

Interpretación del coeficiente de determinación. En los ejercicios 5 a 8, utilice el valor del coeficiente de correlación lineal r para calcular el coeficiente de determinación y el porcentaje de la variación total que se explica por medio de la relación lineal entre las dos variables.

5. $r = 0.3$

6. $r = -0.2$

7. $r = -0.901$

8. $r = 0.747$

Interpretación de resultados de un programa de cómputo. En los ejercicios 9 a 12, remítase a los resultados de Minitab que se obtuvieron utilizando datos apareados de alquitrán y nicotina de una muestra de 29 cigarrillos, como se listan en el conjunto de datos 3 del apéndice B. Además de los datos muestrales apareados, se le indicó a Minitab que usara una cantidad de alquitrán de 17 mg para predecir la cantidad de nicotina.

Minitab						
The regression equation is NICOTINE = 0.154 + 0.0651 TAR						
Predictor Coeff SE Coef T P						
Constant	0.15403	0.04635	3.32	0.003		
TAR	0.065052	0.003583	18.15	0.000		
S = 0.0078548 R-Sq = 92.4% R-Sq(adj) = 92.1%						
Predicted Values for New Observations						
New						
Obs Fit SE Fit 95% CI 95% PI						
1	1.2599	0.0249	(1.2197, 1.3091)	(1.0731, 1.4468)		

9. **Prueba de correlación.** Utilice la información de la pantalla para determinar el valor del coeficiente de correlación lineal. Dado que hay 29 pares de datos, ¿existe una correlación lineal significativa entre la cantidad de alquitrán y la cantidad de nicotina en un cigarrillo?
10. **Identificación de la variación total.** ¿Qué porcentaje de la variación total de nicotina puede explicarse por la relación lineal entre alquitrán y nicotina?
11. **Predicción de la cantidad de nicotina.** Si un cigarrillo contiene 17 mg de alquitrán, ¿cuál es el valor que predice mejor la cantidad de nicotina? (Suponga que existe una correlación lineal significativa entre alquitrán y nicotina).
12. **Cálculo del intervalo de predicción.** Para una cantidad de alquitrán dada de 17 mg, identifique el estimado del intervalo de predicción del 95% para la cantidad de nicotina y redacte una afirmación que interprete ese intervalo.

Cálculo de medidas de variación. En los ejercicios 13 a 16, calcule a) la variación explicada, b) la variación no explicada, c) la variación total, d) el coeficiente de determinación y e) el error estándar del estimado s_e . En cada caso existe una correlación lineal significativa, de manera que es razonable utilizar la ecuación de regresión para hacer predicciones. (Los resultados se utilizan en los ejercicios 17 a 20).

13. **Pesos de automóvil y consumo de combustible.** A continuación se presentan los pesos (en libras) y las cantidades de combustible consumidas en carretera (en mi/gal) de automóviles elegidos al azar (Chrysler, Sebring, Ford Mustang, BMW Serie 3, Ford Crown Victoria, Honda Civic, Mazda Protégé, Hyundai Accent).

Peso	3175	3450	3225	3985	2440	2500	2290
Consumo de combustible	27	29	27	24	37	34	37

14. **Presupuestos e ingresos brutos de películas.** En la siguiente tabla se muestran los presupuestos (en millones de dólares) y los ingresos brutos (en millones de dólares) de películas seleccionadas al azar (según datos de la Motion Picture Association of America).

Presupuesto	62	90	50	35	200	100	90
Ingresos brutos	65	64	48	57	601	146	47

15. **Mediciones de presión sanguínea.** Catorce estudiantes de medicina del segundo año midieron la presión sanguínea del mismo paciente, y los resultados se presentan a continuación (según datos del doctor Marc Triola).

Sistólica	138	130	135	140	120	125	120	130	130	144	143	140	130	150
Diastólica	82	91	100	100	80	90	80	80	80	98	105	85	70	100

- 16. Grillos y temperatura.** Una aplicación clásica de la correlación es la asociación entre la temperatura y el número de veces que un grillo chirría en un minuto. A continuación se listan los números de chirridos en un minuto y las temperaturas correspondientes en grados Fahrenheit (según datos de *The Song of Insects*, de George w. Pierce, Harvard University Press).

Chirridos en un minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (en °F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

- 17. Efecto de la variación en un intervalo de predicción.** Remítase a los datos del ejercicio 13 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
- Calcule la tasa de consumo de combustible predicha para un automóvil que pesa 3700 lb.
 - Calcule un estimado de un intervalo de predicción del 95% de la proporción del consumo de combustible para un automóvil que pesa 3700 lb.
- 18. Cálculo del valor predicho y del intervalo de predicción.** Remítase al ejercicio 14 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
- Calcule el ingreso bruto predicho para una película que tiene un presupuesto de \$100 millones.
 - Calcule un estimado de un intervalo de predicción del 95% del ingreso bruto para una película con un presupuesto de \$100 millones.
- 19. Cálculo del valor predicho y del intervalo de predicción.** Remítase a los datos del ejercicio 15 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
- Calcule la lectura diastólica predicha, dado que la lectura sistólica es de 120.
 - Calcule un estimado de un intervalo de predicción del 95% de la lectura diastólica, dado que la lectura sistólica es de 120.
- 20. Cálculo del valor predicho y del intervalo de predicción.** Remítase a los datos del ejercicio 16 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
- Calcule la temperatura predicha cuando un grillo chirría 1000 veces en un minuto.
 - Calcule un estimado de un intervalo de predicción del 99% de la temperatura, cuando un grillo chirría 1000 veces en un minuto.

Cálculo de un intervalo de predicción. En los ejercicios 21 a 24, remítase a los datos muestrales de la tabla 10-1. Permita que x represente la duración (en segundos) y permita que y represente el intervalo (en minutos) entre la erupción actual y la siguiente erupción. Utilice la duración y el nivel de confianza dados para construir un estimado del intervalo de predicción del intervalo de tiempo entre la erupción actual y la siguiente erupción. (Véase el ejemplo en esta sección).

21. $x = 180$ segundos; 99% de confianza 22. $x = 180$ segundos; 90% de confianza
 23. $x = 200$ segundos; 95% de confianza 24. $x = 120$ segundos; 99% de confianza

10-4 MÁS ALLÁ DE LO BÁSICO

- 25. Intervalos de confianza para β_0 y β_1 .** Los intervalos de confianza para β_0 , el intercepto y, y la pendiente β_1 de una recta de regresión ($y = \beta_0 + \beta_1x$) se pueden obtener evaluando los límites en los intervalos que siguen.

$$\beta_0 - E < \beta_0 < \beta_0 + E$$

donde

$$E = t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum x^2 - \frac{(\sum x)^2}{n}}}$$

continúa

$$b_1 - E < \beta_1 < b_1 + E$$

donde

$$E = t_{\alpha/2} \cdot \frac{s_e}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}}}$$

En estas expresiones b_0 (el intercepto y) y la pendiente b_1 se calculan a partir de los datos muestrales, y $t_{\alpha/2}$ se obtiene de la tabla A-3 utilizando $n - 2$ grados de libertad. Con los datos de la tabla 10-1 sobre las duraciones y los intervalos posteriores a las erupciones, calcule los estimados del intervalo de confianza del 95% de β_0 y β_1 .

- 26. Intervalo de confianza para un valor predicho de la media.** A partir de la expresión dada en esta sección para el margen de error correspondiente a un intervalo de predicción para y , obtenemos

$$s_{\hat{y}} = s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\sum x^2) - (\sum x)^2}}$$

que es el *error estándar de la predicción* cuando se predice para una sola y , dado que $x = x_0$. Cuando se predice la *media* de todos los valores de y para los que $x = x_0$, el estimado puntual \hat{y} es el mismo, pero $s_{\hat{y}}$ se obtiene de la siguiente manera:

$$s_{\hat{y}} = s_e \sqrt{\frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\sum x^2) - (\sum x)^2}}$$

Utilice los datos de la tabla 10-1 para calcular un estimado puntual y un estimado del intervalo de confianza del 95% de la media de los intervalos de tiempo posteriores a las erupciones que tienen una duración de 180 segundos.

10-5 Regresión múltiple

Concepto clave Aunque las secciones previas de este capítulo se aplican a una relación entre *dos* variables, en esta sección presentamos un método para analizar una relación lineal que incluye *más de dos* variables. Nos enfocamos en tres elementos fundamentales: **1.** la ecuación de regresión múltiple, **2.** el valor de R^2 ajustada y **3.** el valor P . Debido a la naturaleza tan compleja de las operaciones requeridas, los cálculos manuales son poco prácticos y constituyen una amenaza para la salud mental; así que en esta sección se destaca el uso y la interpretación de los resultados obtenidos con un programa estadístico de cómputo o una calculadora TI-83/84 Plus.

Ecuación de regresión múltiple

Al igual que en las secciones anteriores de este capítulo, sólo estudiaremos relaciones *lineales*. Utilizamos la siguiente *ecuación de regresión múltiple* para describir relaciones lineales que incluyen más de dos variables.



Definición

Una **ecuación de regresión múltiple** expresa una relación lineal entre una variable de respuesta y y dos o más variables de predicción (x_1, x_2, \dots, x_k). La forma general de una ecuación de regresión múltiple es

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k.$$

Emplearemos la siguiente notación, que surge de manera natural de la notación utilizada en la sección 9-3.

Notación

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k \quad (\text{forma general de la ecuación de regresión múltiple estimada})$$

n = tamaño de la muestra

k = número de variables de predicción. (Las variables de predicción también se conocen como *variables independientes* o variables x).

\hat{y} = valor predicho de y (se calcula por medio de la ecuación de regresión múltiple)

x_1, x_2, \dots, x_k son las variables de predicción

β_0 = intercepto y , o el valor de y cuando todas las variables de predicción son 0. (Este valor es un parámetro poblacional).

b_0 = estimado de β_0 basado en los datos muestrales (b_0 es un estadístico muestral).

$\beta_1, \beta_2, \dots, \beta_k$ son los coeficientes de las variables de predicción x_1, x_2, \dots, x_k

b_1, b_2, \dots, b_k son estimados muestrales de los coeficientes $\beta_1, \beta_2, \dots, \beta_k$



Factores de predicción de éxito

Cuando una universidad acepta a un nuevo estudiante, desearía tener algunos indicadores positivos de que éste será exitoso en sus estudios. Los decanos universitarios de admisiones toman en cuenta las calificaciones de la prueba SAT, las pruebas estándar de aprovechamiento, el lugar que ocupa el alumno en la clase, la dificultad de los cursos de preparatoria, las calificaciones de preparatoria y las actividades extra-curriculares. En un estudio de las características que suelen ser buenos factores de predicción de éxito en la universidad, se encontró que el lugar que se tiene en la clase y las puntuaciones en pruebas estándar de aprovechamiento son mejores factores de predicción que las calificaciones del SAT. Una ecuación de regresión múltiple con el promedio general en la universidad, que se predijo por el lugar que ocupa el alumno en la clase y su puntuación en pruebas de aprovechamiento, no mejoró al incluir la calificación de la prueba SAT como otra variable. Este estudio en particular sugiere que las calificaciones de la prueba SAT no deben incluirse entre los criterios de admisión, aunque otros argumentan que las calificaciones de esta prueba son útiles para comparar estudiantes de diferentes lugares y de distintas preparatorias de procedencia.

Para cualquier conjunto específico de valores de x , la ecuación de regresión está asociada con un error aleatorio que suele denominarse por ε , y suponemos que estos errores se distribuyen normalmente, con una media de 0 y una desviación estándar de σ , y que los errores aleatorios son independientes. Es difícil verificar estos supuestos. A lo largo de esta sección suponemos que los requisitos necesarios se satisfacen.

Los cálculos que se requieren para la regresión múltiple son tan complicados que *debe* utilizarse un programa de cómputo de estadística, así que nos concentraremos en *interpretar* las pantallas de resultados de los programas de cómputo. Al final de esta sección se incluyen instrucciones para el uso de STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus.

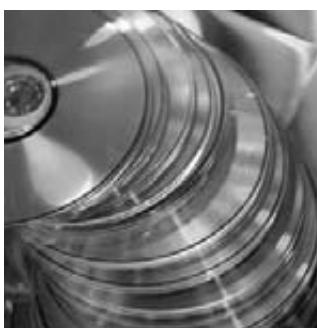


EJEMPLO Old Faithful En la sección 10-3 analizamos métodos para predecir el intervalo posterior a una erupción del géiser Old Faithful, pero en esa sección sólo se incluyó una variable de predicción. Utilice los datos muestrales de la tabla 10-1, incluida en el problema del capítulo, y calcule la ecuación de regresión múltiple en la que la variable de respuesta (y) es el intervalo posterior a una erupción y las variables de predicción (x) son la duración y la altura de la erupción. A continuación se presentan los resultados de Minitab.

continúa

Minitab

The regression equation is Interval After = 45.1 + 0.245 Duration - 0.098 Height						
Predictor	Coeff	SE Coef	T	P		
Constant	45.10	19.41	2.32	0.068		
Duration	0.24464	0.04486	5.45	0.003		
Height	-0.0983	0.1623	-0.61	0.571		
$S = 5.25937 \quad R-Sq = 86.7\% \quad R-Sq(\text{adj}) = 81.3\%$						
Analysis of Variance						
Source	DF	SS	MS	F	P	
Regression	2	897.69	448.85	16.23	0.007	
Residual Error	5	138.31	27.66			
Total	7	1036.00				



Fabricación de CD con regresión múltiple

Sony fabrica millones de discos compactos en Terre Haute, Indiana. En un punto del proceso de fabricación, se expone una placa fotográfica a un láser, de manera que una señal musical es transferida a una señal digital codificada con ceros y unos. Este proceso se analizó estadísticamente para identificar los efectos de diferentes variables, tales como el tiempo de exposición y el grosor de la emulsión fotográfica. Métodos de regresión múltiple demostraron que, entre todas las variables consideradas, cuatro eran las más significativas. El proceso fotográfico se ajustó, con base en estas cuatro variables, para obtener resultados óptimos. Esto permitió disminuir el número de discos defectuosos y mantener la calidad. El uso de métodos de regresión múltiple condujo a costos más bajos de producción y a un mejor control del proceso de fabricación.

SOLUCIÓN Con Minitab obtenemos los resultados que se presentan a continuación. La ecuación de regresión múltiple que aparece en los resultados anteriores de Minitab es

$$\text{Interval After} = 45.1 + 0.245 \text{ Duration} - 0.098 \text{ Height}$$

Si utilizamos la notación presentada anteriormente en esta sección, podemos escribir esta ecuación de la siguiente forma

$$\hat{y} = 45.1 + 0.245x_1 - 0.098x_2$$

Si una ecuación de regresión múltiple se ajusta bien a los datos muestrales, se puede emplear para hacer predicciones. Por ejemplo, si determinamos que la ecuación es adecuada para hacer predicciones, y tenemos una erupción con una duración de 180 segundos y una altura de 130 pies, podemos predecir el intervalo posterior a la erupción sustituyendo esos valores en la ecuación de regresión, para obtener una duración predicha de 76.5 minutos. (Recuerde, las duraciones están en segundos, las alturas en pies y los intervalos posteriores a las erupciones en minutos). Además, los coeficientes $b_1 = 0.245$ y $b_2 = -0.098$ pueden emplearse para determinar el cambio marginal, como se describió en la sección 10-3. Por ejemplo, el coeficiente $b_1 = 0.245$ indica que, cuando la altura de una erupción permanece constante, el intervalo posterior a la erupción predicho aumenta 0.245 minutos por cada incremento de un segundo en la duración de la erupción.

R² ajustada

R^2 denota el **coeficiente múltiple de determinación**, que es una medida de lo bien que se ajusta la ecuación de regresión múltiple a los datos muestrales. Un ajuste perfecto daría como resultado $R^2 = 1$, y un ajuste muy bueno daría por resultado un valor cercano a 1. Un ajuste muy deficiente se relaciona con un valor de R^2 cercano a 0. El valor de $R^2 = 86.7\%$ en los resultados de Minitab, indica que el 86.7% de la variación de los intervalos posteriores a las erupciones puede explicarse por la duración x_1 y la altura x_2 . Sin embargo, el coeficiente múltiple de determinación R^2 tiene una grave desventaja: a mayor número de variables incluidas, R^2 se incrementa. (R^2 podría permanecer igual, pero suele incrementarse). La R^2 más grande se obtiene por el simple hecho de incluir todas las variables disponibles, pero la mejor ecuación de regresión múltiple no necesariamente utiliza todas las variables disponibles. A causa de esta desventaja, la comparación de diferentes ecuaciones de regresión múltiple se logra mejor con el coeficiente ajustado de determinación, que es R^2 ajustada para el número de variables y el tamaño de la muestra.

Definición

El **coeficiente ajustado de determinación** es el coeficiente múltiple de determinación R^2 modificado para justificar el número de variables y el tamaño de la muestra. Se calcula por medio de la fórmula 10-6.

$$\text{Fórmula 10-6} \quad R^2 \text{ ajustada} = 1 - \frac{(n - 1)}{[n - (k + 1)]} (1 - R^2)$$

donde

n = tamaño muestral

k = número de variables de predicción (x)

Los resultados anteriores de Minitab para los datos indican que el coeficiente ajustado de determinación es $R\text{-sq(adj)} = 81.3\%$. Si utilizamos la fórmula 10-6 con el valor de $R^2 = 0.867$, $n = 8$ y $k = 2$, encontramos que el valor ajustado de R^2 es 0.813, lo que confirma el valor de 81.3% de los resultados de Minitab. (En realidad obtuvimos 0.814, pero obtenemos 0.813 si utilizamos más dígitos para minimizar el error de redondeo). El valor de R^2 de 86.7% indica que el 86.7% de la variación de los intervalos posteriores a las erupciones puede explicarse por la duración x_1 y la altura x_2 , pero cuando comparamos esta ecuación de regresión múltiple con otras, es mejor utilizar la R^2 ajustada de 81.3% (o 0.813).

Valor P

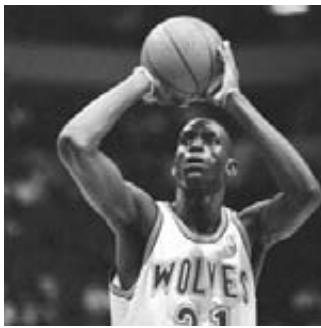
El valor P es una medida de la significancia general de la ecuación de regresión múltiple. El valor P de 0.007 de los resultados de Minitab es pequeño, lo que indica que la ecuación de regresión múltiple tiene una buena significancia general y es útil para hacer predicciones. Es decir, tiene sentido predecir intervalos posteriores a las erupciones con base en las duraciones y las alturas de las erupciones. Al igual que la R^2 ajustada, este valor P es una buena medida de qué tan bien se ajusta la ecuación a los datos muestrales. El valor de 0.007 resulta de una prueba de la hipótesis nula de que $\beta_1 = \beta_2 = 0$. El rechazo de $\beta_1 = \beta_2 = 0$ implica que al menos uno de β_1 y β_2 no es 0, lo que indica que esta ecuación de regresión es eficaz para determinar los intervalos posteriores a las erupciones. Un análisis completo de los resultados de Minitab podría llevarnos a incluir otros elementos importantes, como la significancia de los coeficientes individuales, pero limitaremos nuestra explicación a los tres componentes principales: la ecuación de regresión múltiple, la R^2 ajustada y el valor P .

Cálculo de la mejor ecuación de regresión múltiple

El resultado anterior de Minitab se basa en el uso de las variables de predicción de duración y altura con los datos muestrales de la tabla 10-1. Pero si deseamos predecir el intervalo posterior a una erupción, ¿existe alguna otra combinación de variables que podría ser mejor que la duración y la altura? La tabla 10-3 lista

Tabla 10-3 Búsqueda de la mejor ecuación de regresión múltiple			
	R^2	R^2 ajustada	Significancia general
DURACIÓN	0.857	0.833	0.001
INTERVALO PREVIO	0.011	0.000	0.802
ALTURA	0.073	0.000	0.519
DURACIÓN e INTERVALO PREVIO	0.872	0.820	0.006
DURACIÓN y ALTURA	0.867	0.813	0.007
INTERVALO PREVIO y ALTURA	0.073	0.000	0.828
DURACIÓN e INTERVALO PREVIO y ALTURA	0.875	0.781	0.028

← R^2 ajustada más alta y valor P más bajo



Salarios de la NBA y desempeño

El investigador Matthew Weeks estudió la correlación entre los salarios de la NBA y las estadísticas del juego de básquetbol. Además del salario (S), consideró los minutos jugados (M), las intervenciones (I), los rebotes (R) y los puntos anotados (P); utilizó datos de 30 jugadores. La ecuación de regresión múltiple es $S = -0.716 - 0.0756M - 0.425I + 0.0536R + 0.742P$ con $R^2 = 0.458$. Debido a una alta correlación entre los minutos jugados (M) y los puntos anotados (P), y puesto que estos últimos tuvieron una alta correlación con el salario, la variable de minutos jugados se eliminó de la ecuación de regresión múltiple. Además, no se encontró que las variables de intervenciones (I) y rebotes (R) fueran significativas, por lo que también se eliminaron. La variable de los puntos anotados pareció ser la mejor elección para predecir los salarios de la NBA, pero se encontró que las predicciones no eran muy exactas ya que no se consideraron otras variables, tales como la popularidad del jugador.

distintas combinaciones de variables. Ahora nos enfrentamos al importante objetivo de calcular la *mejor* ecuación de regresión múltiple. Puesto que la determinación de la mejor ecuación de regresión múltiple requiere de una buena dosis de juicio, no existe un procedimiento exacto y automático para esto. *La determinación de la mejor ecuación de regresión múltiple suele ser bastante difícil y rebasa los alcances de este libro*, pero los siguientes lineamientos servirán de guía.

Lineamientos para el cálculo de la mejor ecuación de regresión múltiple

- Utilice el sentido común y consideraciones prácticas para incluir o excluir variables.* Por ejemplo, podríamos excluir la variable de la estatura después de saber que esta variable es un estimado visual y no una medida exacta.
- Considere el valor P .* Seleccione una ecuación que tenga significancia general, tal como lo determina el valor P indicado en los resultados del programa de cómputo. Por ejemplo, observe los valores de significancia general en la tabla 10-3. Los valores P de 0.802, 0.519 y 0.828 corresponden a las combinaciones de variables que no dan como resultado una significancia general, de manera que esas combinaciones deben excluirse.
- Considere ecuaciones con valores altos de R^2 ajustada y trate de incluir sólo unas cuantas variables.* En vez de incluir casi todas las variables disponibles, trate de incluir relativamente pocas variables de predicción (x). Utilice los siguientes lineamientos:
 - Seleccione una ecuación que tenga un valor de R^2 ajustada con esta propiedad: si se incluye una variable independiente adicional, el valor de R^2 ajustada no se incrementa de manera sustancial.
 - Para un número dado de variables de predicción (x), seleccione la ecuación con el valor más grande de la R^2 ajustada.
 - Para eliminar las variables de predicción (x) que no tienen mucho efecto sobre la variable de respuesta (y), sería útil calcular el coeficiente de correlación lineal r para cada par de variables en consideración. Si dos valores de predicción tienen un coeficiente de correlación lineal muy alto, no es necesario incluir a ambos, y debemos excluir la variable con el valor de r más bajo.

Si seguimos estos lineamientos al intentar calcular la mejor ecuación para predecir los intervalos posteriores a las erupciones del géiser Old Faithful, encontramos que, para los datos de la tabla 10-3, la mejor ecuación de regresión utiliza únicamente la variable de predicción (x) de duración. Parece que la mejor ecuación de regresión es

$$\text{INTERVALO POSTERIOR} = 34.8 + 0.234 \text{ DURACIÓN}$$

o

$$\hat{y} = 34.8 + 0.234x_1$$

Los lineamientos anteriores se basan en la R^2 ajustada y en el valor P , pero también podemos hacer pruebas de hipótesis individuales, basadas en los valores

de los coeficientes de regresión. Considere el coeficiente de regresión de β_1 . Una prueba de la hipótesis nula $\beta_1 = 0$ nos puede indicar si la variable de predicción correspondiente debe incluirse en la ecuación de regresión. El rechazo de $\beta_1 = 0$ sugiere que β_1 tiene un valor diferente de cero y, por lo tanto, sirve para predecir el valor de la variable de respuesta. En el ejercicio 17 se describen procedimientos para este tipo de pruebas.

Algunos programas estadísticos de cómputo incluyen un programa para realizar la **regresión por pasos**, en la cual los cálculos se realizan con distintas combinaciones de variables de predicción (x), aunque existen graves problemas asociados con ella, incluyendo los siguientes: la regresión por pasos no necesariamente produce el mejor modelo si algunas variables de predicción tienen una alta correlación; produce valores inflados de R^2 ; utiliza demasiado papel y no nos permite *pensar* acerca del problema. Como siempre, debemos ser cuidadosos de emplear los resultados de las computadoras como una herramienta que nos ayude a tomar decisiones inteligentes; no debemos permitir que la computadora tome las decisiones. En vez de confiar únicamente en los resultados de una regresión por pasos realizada por un programa de cómputo, considere los factores anteriores cuando trate de identificar la mejor ecuación de regresión múltiple.

Si corremos el programa de regresión por pasos de Minitab utilizando los datos de la tabla 10-1, obtendremos una pantalla de resultados que sugiere que la mejor ecuación de regresión es aquella en la que la duración es la única variable de predicción. Parece que podemos estimar el intervalo posterior a una erupción, y la ecuación de regresión nos conduce a esta regla: se estima que el intervalo (en minutos) posterior a una erupción es 34.8 más 0.234 veces la duración (en segundos).

Variables ficticias o indicadoras y regresión logística

En esta sección todas las variables han sido de naturaleza continua. El intervalo posterior a una erupción puede ser cualquier valor en un rango continuo de minutos, por lo que es un buen ejemplo de una variable continua. Sin embargo, muchas aplicaciones incluyen una **variable dicotómica**, que sólo tiene *dos* valores discretos posibles (como hombre/mujer, vivo/muerto o sano/no sano). Un procedimiento común consiste en representar las dos variables discretas posibles con 0 y 1, donde 0 representa un “fracaso” (como muerte) y 1 representa un éxito. A una variable dicotómica con los dos valores posibles de 0 y 1 se le llama **variable ficticia o indicadora**.

Los procedimientos de análisis difieren de manera drástica, dependiendo de si la variable ficticia es una variable de predicción (x) o la variable de respuesta (y). Si incluimos una variable ficticia como otra variable de predicción (x), podemos utilizar los métodos de esta sección, como se ilustra en el siguiente ejemplo.

EJEMPLO Uso de una variable ficticia Utilice la estatura, el peso, la medida de la cintura y el pulso del conjunto de datos combinados de 80 mujeres y hombres, tal como aparecen en el conjunto de datos 1 del apéndice B. Permita que la variable de respuesta y represente la estatura y , para la primera variable de predicción, utilice la variable ficticia de *género* (codificada como 0 = mujer, 1 = hombre). Dado un peso de 150 lb, una cintura de 80 cm y un pulso de 75 latidos por minuto, calcule la ecuación de regresión múltiple y úsela para predecir la estatura de *a*) una mujer y *b*) un hombre.

continúa



Congelando al pateador

Una estrategia común en el fútbol americano consiste en que, justo en el momento en que un pateador está a punto de tratar de anotar un gol de campo, el entrenador del equipo opuesto pide un tiempo fuera para “congelar” al pateador. La teoría sostiene que el pateador tiene tiempo para pensar, sentirse nervioso y perder la confianza. Sin embargo, ¿realmente funciona esta práctica? En el artículo “The Cold-Foot Effect” publicado por la revista *Chance*, Scott M. Berry reportó su análisis estadístico de los resultados de dos temporadas de la NFL. Utilizó un modelo de regresión logística con variables tales como el viento, las nubes, la precipitación pluvial, la temperatura, la presión de realizar la patada y si hubo o no una petición de tiempo fuera antes de la patada. El autor escribió que “la conclusión a partir del modelo es que congelar al pateador sí funciona: parece que congelar al pateador reduce la probabilidad de una patada exitosa”.

SOLUCIÓN Si utilizamos los métodos de esta sección con un programa de cómputo, obtenemos la siguiente ecuación de regresión:

$$\begin{aligned} \text{ESTATURA} = & 64.4 + 3.47(\text{GÉNERO}) + 0.118(\text{PESO}) - 0.222(\text{CINTURA}) \\ & + 0.00602(\text{PULSO}) \end{aligned}$$

Para calcular la estatura predicha de una mujer, sustituimos el género de la variable por 0. Si también sustituimos el peso por 150, la cintura por 80 y el pulso por 75, resulta una estatura predicha de 64.8 pulgadas (o 5 ft 5 in) para una mujer.

Para calcular la estatura predicha de un hombre, sustituimos el género de la variable por 1. Si también sustituimos los otros valores, obtenemos una estatura predicha de 68.3 pulgadas (o 5 ft 8 in) para un hombre. Observe que cuando todas las demás variables permanecen igual, un hombre tendrá una estatura predicha de 3.47 pulgadas más que la estatura de una mujer.

En el ejemplo anterior pudimos utilizar los métodos de esta sección porque la variable ficticia de género es una variable de predicción. Si la variable ficticia es la variable de respuesta (y), no podemos emplear los métodos de esta sección, sino que debemos utilizar un método conocido como **regresión logística**. Suponga, por ejemplo, que utilizamos la estatura, el peso, la medida de cintura y el pulso de los hombres y las mujeres que aparecen en el conjunto de datos 1 del apéndice B. Permita que la variable de respuesta y represente el género (0 = mujer, 1 = hombre). Si usamos los 80 valores de y (con los valores de las mujeres codificados con 0 y los de los hombres codificados con 1) y la lista combinada de estaturas, pesos, tamaños de cintura y pulsos correspondientes, podemos aplicar la regresión logística para obtener el siguiente modelo:

$$\ln \frac{p}{1-p} = -41.8193 + 0.679195(\text{ESTATURA}) - 0.0106791(\text{PESO}) + 0.0375373(\text{CINTURA}) - 0.0606805(\text{PULSO})$$

En esta expresión, p representa una probabilidad. Un valor de $p = 0$ indica que la persona es mujer y $p = 1$ que es hombre. Un valor de $p = 0.2$ indica que existe una probabilidad de 0.2 de que la persona sea hombre, y se infiere que existe la probabilidad de 0.8 de que la persona sea mujer. Si utilizamos el modelo anterior y sustituimos una estatura de 72 in, un peso de 200 lb, una circunferencia de cintura de 90 cm y un pulso de 85 latidos por minuto, podemos despejar p y así obtener $p = 0.960$, lo que indica que lo más probable es que una persona tan grande sea hombre. En contraste, una persona pequeña, con una estatura de 60 in, un peso de 90 lb, una circunferencia de cintura de 68 cm y un pulso de 85 latidos por minuto da por resultado un valor de $p = 0.00962$, y esto indica que lo más probable es que se trate de una mujer. En este libro no se incluyen procedimientos detallados para el uso de la regresión logística, pero existen muchos libros dedicados al tema, y muchos otros libros de texto incluyen información detallada sobre este modelo.

Cuando estudiamos la regresión en la sección 10-3, indicamos cuatro errores comunes que deben evitarse al utilizar ecuaciones de regresión para hacer predicciones. Estos mismos errores deben evitarse cuando se emplean ecuaciones de regresión múltiple. Sea especialmente cuidadoso al concluir que existe una relación causa-efecto.

Uso de la tecnología

STATDISK Primero ingrese los datos muestrales en las columnas de la ventana de datos de STATDISK, luego seleccione **Analysis** y después **Multiple Regression**. Seleccione las columnas que se incluirán y también identifique la columna que corresponde a la variable dependiente y (de predicción). Haga clic en **Evaluate** y obtendrá la ecuación de regresión múltiple y otros elementos, incluyendo el coeficiente múltiple de determinación R^2 , la R^2 ajustada y el valor P .

MINITAB Primero introduzca los valores en las distintas columnas. Para evitar confusiones entre las diferentes variables, escriba un nombre para cada variable en el cuadro que se encuentra en la parte superior de la columna de datos. Seleccione **Stat** del menú principal, después **Regression** y luego **Regression** una vez más. En el cuadro de diálogo, ingrese la variable que se empleará como variable de respuesta (y) y las variables que desea incluir como variables x . Haga clic en **OK**. Los resultados incluirán la ecuación de regresión

múltiple y otros elementos, incluyendo el coeficiente de determinación múltiple R^2 y la R^2 ajustada.

EXCEL Primero ingrese los datos muestrales en las columnas. Seleccione **Tools** del menú principal, después **Data Analysis** y **Regression**. En el cuadro de diálogo introduzca el rango de valores para la variable dependiente Y , después el rango de valores para las variables independientes X , que deben estar en columnas adyacentes. (Utilice copiar/pegar para mover las columnas como deseé). Los resultados incluirán el coeficiente múltiple de determinación R^2 , la R^2 ajustada y una lista de los valores del intercepto y el coeficiente utilizados para la ecuación de regresión múltiple.

TI-83/84 PLUS El programa A2MULREG de la calculadora TI-83/84 Plus puede descargarse del CD-ROM incluido con este libro. Seleccione el archivo de software, luego elija el archivo con los programas de TI; el programa se cargará en su calculadora.

Los datos muestrales se deben introducir primero como columnas de una matriz D, en donde la primera columna contenga los valores de la variable de respuesta (y). Para ingresar manualmente los datos en la matriz D, presione **2nd** y la tecla x^{-1} , gire hacia la derecha hasta **EDIT**, gire hacia abajo hasta **[D]**, después presione **ENTER** y proceda a introducir las dimensiones de la matriz en el formato de renglones por columnas. Para el número de renglones introduzca el número de valores muestrales listados para cada variable. Para el número de columnas ingrese el número total de variables

x y y . Proceda a ingresar los valores muestrales. Si los datos ya están almacenados en forma de listas, éstas se pueden combinar y almacenar en la matriz D. Presione **2nd** y la tecla x^{-1} , seleccione **MATH**, el elemento que aparece en la parte superior del menú, y elija **List→matr**; luego, introduzca los nombres de la lista, de manera que el primero corresponda a la variable y , y también introduzca el nombre de la matriz de **[D]**, todos separados con comas. Por ejemplo, **List→matr(NICOT,TAR,CO, [D])** crea una matriz D con los valores de NICOT en la primera columna, los valores de TAR en la segunda columna y los valores de CO en la tercera columna.

Ahora presione **PRGM**, seleccione **A2MULREG** y luego **ENTER** tres veces, luego seleccione **MULT REGRESSION** y presione **ENTER**. Cuando se le solicite, ingrese el número de variables independientes (x), después introduzca los números de las columnas de las variables independientes que desea incluir. La pantalla dará un resultado que incluye al valor P y el valor de R^2 ajustada. Presione **ENTER** para ver los valores que se utilizarán en la ecuación de regresión múltiple. Presione **ENTER** de nuevo para obtener el menú que incluye opciones para generar intervalos de confianza, intervalos de predicción, residuales o para salir. Si usted desea generar intervalos de confianza y de predicción, utilice el número de grados de libertad que aparecen, consulte la tabla A-3 y busque el valor t crítico correspondiente, intodúzcalo y proceda a ingresar los valores que se emplearán para las variables de predicción (x). Presione **ENTER** para seleccionar la opción **QUIT**.



10-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Regresión múltiple.** ¿Qué es la *regresión múltiple* y en qué difiere de la regresión analizada en la sección 10-3?
- 2. Coeficiente ajustado de discriminación.** Al comparar diferentes ecuaciones de regresión múltiple, ¿por qué la R^2 ajustada es una mejor medida que R^2 ?
- 3. Predicción del color de ojos.** Un genetista desea desarrollar un método para predecir el color de ojos de un bebé, dado el color de ojos de cada uno de los padres. ¿Se pueden emplear los métodos de esta sección? ¿Por qué?
- 4. Variables.** ¿Cuál es la diferencia entre una *variable de respuesta* y una *variable de predicción*?

Interpretación de resultados de programas de cómputo. En los ejercicios 5 a 8, remítase a los resultados de Minitab que se presentan aquí y responda las preguntas formuladas o identifique los elementos indicados. Los resultados de Minitab están basados en la muestra de 54 osos incluida en el conjunto de datos 6 del apéndice B.

Minitab

```
The regression equation is
WEIGHT = -372 - 0.87 HEADLEN + 0.55 LENGTH + 12.2 CHEST

Predictor      Coef    SE Coef      T      P
Constant     -371.71   31.62  -11.59  0.000
HEADLEN       -0.870   5.676  -0.15  0.879
LENGTH        0.554   1.259   0.44  0.662
CHEST         12.153   1.116  10.89  0.000

S = 33.6565  R-Sq = 92.8%  R-Sq(adj) = 92.4%
Analysis of Variance
Source          DF      SS      MS      F      P
Regression       3  729445  243148  214.71  0.000
Residual Error  50  56638   1133
Total            53  786183
```

5. **Mediciones de osos.** Identifique la ecuación de regresión múltiple que expresa el peso en términos de la longitud de la cabeza, la estatura y el tamaño del pecho.
6. **Mediciones de osos.** Identifique lo siguiente:
 - a. El valor P correspondiente a la significancia general de la ecuación de regresión múltiple
 - b. El valor del coeficiente múltiple de determinación R^2
 - c. El valor ajustado de R^2
7. **Mediciones de osos.** ¿Es útil la ecuación de regresión múltiple para predecir el peso de un oso con base en la longitud de su cabeza, la estatura y el tamaño del pecho? ¿Por qué?
8. **Mediciones de osos.** Se encuentra que un oso tiene una longitud de cabeza de 14.0 in, una estatura de 70.0 in y un tamaño del pecho de 50.0 in.
 - a. Calcule el peso predicho del oso.
 - b. El oso en cuestión en realidad pesaba 320 lb. ¿Qué tan exacto es el peso predicho en el inciso a)?

Datos de salud: cálculo de la mejor ecuación de regresión múltiple. En los ejercicios 9 a 12, remítase a la siguiente tabla, que se obtuvo utilizando los datos de hombres del conjunto de datos 1 del apéndice B. La variable de respuesta (y) es el peso (en libras) y las variables de predicción (x) son EST (estatura en pulgadas), CINT (circunferencia de la cintura en cm) y COL (colesterol en mg).

Variables de predicción (x)	Valor P	R^2	R^2 ajustada	Ecuación de regresión
EST, CINT, COL	0.000	0.880	0.870	$\hat{y} = -199 + 2.55 \text{EST} + 2.18 \text{CINT} - 0.00534 \text{COL}$
EST, CINT	0.000	0.877	0.870	$\hat{y} = -206 + 2.66 \text{EST} + 2.15 \text{CINT}$
EST, COL	0.002	0.277	0.238	$\hat{y} = -148 + 4.65 \text{EST} + 0.00589 \text{COL}$
CINT, COL	0.000	0.804	0.793	$\hat{y} = -42.8 + 2.41 \text{CINT} - 0.0106 \text{COL}$
EST	0.001	0.273	0.254	$\hat{y} = -139 + 4.55 \text{EST}$
CINT	0.000	0.790	0.785	$\hat{y} = -44.1 + 2.37 \text{CINT}$
COL	0.874	0.001	0.000	$\hat{y} = 173 - 0.00233 \text{COL}$

9. Si sólo se utiliza una variable de predicción (x) para predecir el peso, ¿cuál variable es mejor? ¿Por qué?
10. Si se utilizaran exactamente dos variables de predicción (x) para predecir el peso, ¿cuáles dos variables deberían elegirse? ¿Por qué?

11. ¿Cuál ecuación de regresión es mejor para predecir el peso? ¿Por qué?
12. Si un hombre tiene una estatura de 72 in, una circunferencia de cintura de 105 cm y un nivel de colesterol de 250 mg, ¿cuál es el mejor valor predicho de su peso? ¿Es posible que ese valor predicho constituya un buen estimado? ¿Es posible que el valor predicho sea muy exacto?
13. **Conjunto de datos del apéndice B: Predicción de nicotina en cigarrillos.** Remítase al conjunto de datos 3 del apéndice B.
- Calcule la ecuación de regresión que exprese la variable de respuesta (y) de la cantidad de nicotina en términos de la variable de predicción (x) de la cantidad de alquitrán.
 - Calcule la ecuación de regresión que exprese la variable de respuesta (y) de la cantidad de nicotina en términos de la variable de predicción (x) de la cantidad de monóxido de carbono.
 - Calcule la ecuación de regresión que exprese la variable de respuesta (y) de la cantidad de nicotina en términos de las variables de predicción (x) de la cantidad de alquitrán y la cantidad de monóxido de carbono.
 - Para las ecuaciones de regresión calculadas en los incisos *a*, *b* y *c*, ¿cuál es la mejor ecuación para predecir la cantidad de nicotina?
 - ¿La mejor ecuación de regresión identificada en el inciso *d*) es una *buenas* ecuación para predecir la cantidad de nicotina? ¿Por qué?
14. **Conjunto de datos del apéndice B: Uso de la basura para predecir el tamaño de la población.** Remítase al conjunto de datos 16 del apéndice B.
- Calcule la ecuación de regresión que exprese la variable de respuesta (y) del tamaño de la familia en términos de la variable de predicción del peso de la comida desechara.
 - Calcule la ecuación de regresión que exprese la variable de respuesta (y) del tamaño de la familia en términos de la variable de predicción (x) del peso del plástico desecharo.
 - Calcule la ecuación de regresión que exprese la variable de respuesta (y) del tamaño de la familia en términos de las variables de predicción (x) del peso de la comida desechara y del peso del plástico desecharo.
 - Para las ecuaciones de regresión calculadas en los incisos *a*, *b* y *c*, ¿cuál es la mejor ecuación para predecir el tamaño de la familia? ¿Por qué?
 - ¿La mejor ecuación de regresión identificada en el inciso *d*) es una *buenas* ecuación para predecir el tamaño de la familia? ¿Por qué?
15. **Conjunto de datos del apéndice B: Precio de venta de casas.** Remítase al conjunto de datos 18 del apéndice B y calcule la mejor ecuación de regresión múltiple con el precio de venta como variable de respuesta (y). ¿Es esta “mejor” ecuación buena para predecir el precio de venta de una casa?
16. **Conjunto de datos del apéndice B: Old Faithful.** En esta sección se utilizaron los datos de 8 erupciones del géiser Old Faithful, tal como se listan en la tabla 10-1. Remítase al conjunto de datos 11 del apéndice B y utilice el conjunto completo de datos de las 40 erupciones. Determine la mejor ecuación de regresión múltiple que exprese la variable de respuesta (y) del intervalo posterior a una erupción, en términos de una o más de las otras variables. Explique su elección.

10-5 MÁS ALLÁ DE LO BÁSICO

17. **Prueba de hipótesis sobre coeficientes de regresión.** Si el coeficiente β_1 tiene un valor que no es cero, entonces sirve para predecir el valor de la variable de respuesta. Si $\beta_1 = 0$, no es útil para predecir el valor de la variable de respuesta y puede eliminarse de la ecuación de regresión. Para poner a prueba la aseveración de que $\beta_1 = 0$, utilice el estadístico de prueba $t = (b_1 - 0)/s_{b_1}$. Los valores críticos o los valores P se pueden obtener utilizando la distribución t con $n - (k + 1)$ grados de libertad, donde k es el número de variables de predicción (x) y n es el número de observaciones en

la muestra. (Por ejemplo, $n = 8$ para la tabla 10-1). Con frecuencia los programas de cómputo proporcionan la desviación estándar s_{b_1} . Por ejemplo, los resultados de Minitab incluidos en esta sección (véase la página 567) indican que $s_{b_1} = 0.04486$ se localiza en la columna con el encabezado de SE Coeff y en el renglón correspondiente a la primera variable de predicción de la duración. Utilice los datos muestrales de la tabla 10-1 y los resultados de Minitab incluidos en esta sección para poner a prueba la aseveración de que $\beta_1 = 0$. También pruebe la aseveración de que $\beta_2 = 0$. ¿Qué implican los resultados sobre la ecuación de regresión?

- 18. Intervalo de confianza para un coeficiente de regresión.** Un intervalo de confianza para el coeficiente de regresión β_1 se expresa como

$$b_1 - E < \beta_1 < b_1 + E$$

donde

$$E = t_{\alpha/2} s_{b_1}$$

La puntuación crítica t se calcula utilizando $n - (k + 1)$ grados de libertad, donde k, n y s_{b_1} se describen en el ejercicio 17. Utilice los datos muestrales de la tabla 10-1 y los resultados de Minitab incluidos en esta sección (véase la página 567) para construir estimados de un intervalo de confianza del 95% de β_1 (el coeficiente de la variable que representa la duración) y β_2 (el coeficiente de la variable que representa la altura). ¿Alguno de los intervalos de confianza incluye a 0, lo que sugeriría que se elimine la variable de la ecuación de regresión?

- 19. Variable ficticia.** Remítase al conjunto de datos 6 del apéndice B y considere el sexo, la edad y el peso de los osos. Para el sexo, permita que 0 represente una hembra y que 1 represente un macho. (En el conjunto de datos 6 los machos ya están representados por 1, pero para las hembras cambie los valores de sexo de 2 a 0). Permita que la variable de respuesta (y) sea el peso y utilice la variable de edad y la variable ficticia de sexo para calcular la ecuación de regresión múltiple; luego utilice la ecuación para calcular el peso predicho de un oso con las siguientes características. ¿Parecería que el sexo tiene un gran efecto sobre el peso de un oso?

- a. Oso hembra de 20 años de edad
- b. Oso macho de 20 años de edad

- 20. Uso de la regresión múltiple para la ecuación de la parábola.** En algunos casos, la ecuación de regresión múltiple que se ajusta mejor tiene la forma $\hat{y} = b_0 + b_1x + b_2x^2$. La gráfica de una ecuación como ésta es una parábola. Utilice el conjunto de datos listado al margen, permita que $x_1 = x$, y que $x_2 = x^2$, y calcule la ecuación de regresión múltiple para la parábola que se ajuste mejor a los datos. Con base en el valor del coeficiente múltiple de determinación, ¿qué tan bien se ajusta esta ecuación a los datos?

x	1	3	4	7	5
y	5	14	19	42	26

10-6 Elaboración de modelos

Concepto clave En esta sección se introducen algunos conceptos básicos para el desarrollo de un **modelo matemático**, el cual es una función matemática que se “ajusta” o describe datos del mundo real. Por ejemplo, podríamos buscar un modelo matemático consistente en una ecuación que relaciona una variable del tamaño poblacional con otra variable que representa el tiempo. A diferencia de la sección 10-3, no estamos restringidos a un modelo que deba ser lineal. Además, en vez de utilizar datos muestrales seleccionados al azar, consideraremos datos reunidos periódicamente a través del tiempo o alguna otra unidad básica de medición. Existen algunos métodos estadísticos poderosos que podemos estudiar (tales como las *series de tiempo*), pero el principal objetivo de esta sección es describir brevemente la manera en que puede utilizarse la tecnología para obtener un buen modelo matemático.

A continuación se presentan algunos modelos genéricos como aparecen en un menú de la calculadora TI-83/84 Plus (presione **STAT** y luego seleccione **CALC**):

Lineal: $y = a + bx$

Cuadrático: $y = ax^2 + bx + c$

Logarítmico: $y = a + b \ln x$

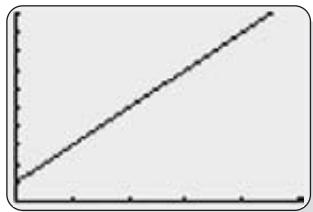
Exponencial: $y = ab^x$

Potencia: $y = ax^b$

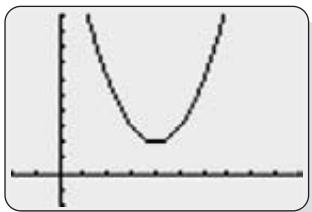
El modelo particular que usted seleccione depende de la naturaleza de los datos muestrales, y un diagrama de dispersión resultará muy útil para tomar esta determinación. Las ilustraciones que aparecen a continuación son gráficas de algunos modelos comunes elaborados en una calculadora TI-83/84 Plus.

TI-83/84 Plus

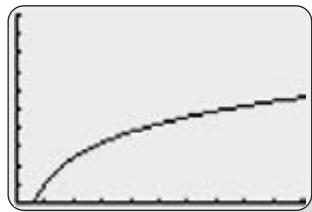
Lineal: $y = 1 + 2x$



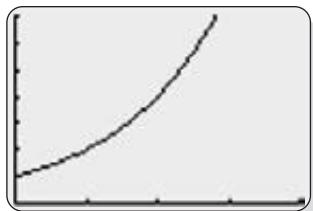
Cuadrático: $y = x^2 - 8x + 18$



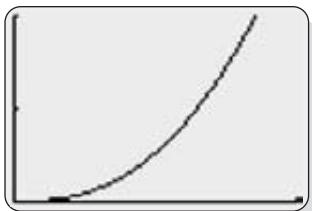
Logarítmico: $y = 1 + 2 \ln x$



Exponencial: $y = 2^x$



Potencia: $y = 3x^{2.5}$



Éstas son las reglas básicas para la creación de un buen modelo matemático:

- 1. Busque un patrón en la gráfica.** Examine la gráfica con los puntos y compare el patrón básico de las gráficas genéricas conocidas de una función lineal, una función cuadrática, una función exponencial, una función potencial, etcétera. (Remítase a las gráficas de los resultados de la calculadora TI-83/84 Plus que se presentan en los ejemplos). Cuando trate de seleccionar un modelo, considere sólo aquellas funciones que parecen ajustarse visualmente a los puntos observados, de una forma razonablemente adecuada.
- 2. Calcule y compare valores de R^2 .** Para cada modelo que considere, utilice programas de cómputo o una calculadora TI-83/84 Plus para obtener el valor del coeficiente de determinación R^2 . Los valores de R^2 se pueden interpretar aquí de la misma forma que se interpretaron en la sección 10-5. Al delimitar sus posibles modelos, seleccione funciones que den como resultado valores más grandes de R^2 , ya que corresponden a funciones que se ajustan mejor a los puntos observados. Sin embargo, no dé demasiada importancia a las diferencias pequeñas, como la diferencia entre $R^2 = 0.984$ y $R^2 = 0.989$. (Otra medición utilizada para evaluar la calidad de un modelo es la suma de cuadrados de los residuales. Véase el ejercicio 15).



Estadística: Empleos y empleadores

A continuación se describe una muestra pequeña de anuncios de empleos en el campo de la estadística: pronosticador del tiempo, analista de bases de datos, científico de marketing, gerente de riesgos de crédito, investigador y evaluador del cáncer, analista de riesgos de seguros, investigador de pruebas educativas, bioestadístico, estadístico para productos farmacéuticos, criptólogo, programador estadístico.

La siguiente es una muestra pequeña de empresas que ofrecen empleos en el campo de la estadística: Centers for Disease Control and Prevention, Cardiac Pacemakers, Inc., National Institutes of Health, National Cancer Institute, CNA Insurance Company, Educational Testing Service, Roswell Park Cancer Institute, Cleveland Clinic Foundation, National Security Agency, Quantiles, 3M, IBM, Nielsen Media Research, AT&T Labs, Bell Labs, Hewlett Packard, Johnson & Johnson, Smith Hanley.

3. Reflexione. Aplique el sentido común. No utilice un modelo que conduzca a valores predichos que se sabe son poco realistas. Utilice el modelo para calcular valores futuros, valores pasados y valores de años faltantes; luego determine si los resultados son realistas.

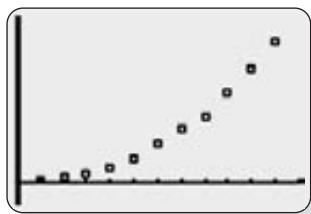
EJEMPLO La tabla 10-4 lista la población de Estados Unidos en diferentes años. Encuentre un buen modelo matemático para el tamaño poblacional, después haga una predicción del tamaño de la población de Estados Unidos para el año 2020.

SOLUCIÓN Primero “codificamos” los valores del año utilizando 1, 2, 3 . . . , en vez de 1800, 1820, 1840. . . . La razón de esta codificación es que, de esta forma, los valores de x son mucho más pequeños y tienen menos posibilidades de causar problemas de cálculo como los que ocurrirían con valores realmente grandes de x .

Busque un patrón en la gráfica. Examine el patrón de los valores de los datos en los resultados de la calculadora TI-83/84 Plus (mostrados al margen) y compare el patrón con los modelos genéricos presentados antes en esta sección. El patrón de estos puntos no es una recta, por lo que descartamos un modelo lineal. Parece que los buenos candidatos para el modelo son las funciones cuadrática, exponencial y potencial.

Calcule y compare valores de R^2 . Las siguientes pantallas muestran resultados de la calculadora TI-83/84 Plus basados en los modelos cuadrático, exponencial y potencial. Al comparar los valores del coeficiente R^2 , parece que el modelo cuadrático es el mejor, ya que tiene el valor más alto de 0.9992, aunque los otros valores mostrados también son bastante altos. Si seleccionamos la función cuadrática como el mejor modelo, concluimos que la ecuación $y = 2.77x^2 - 6.00x + 10.01$ describe mejor la relación entre el año x (codificado de manera que $x = 1$ representa 1800, $x = 2$ representa 1820, y así sucesivamente) y la población y (en millones).

TI-83/84 Plus



TI-83/84 Plus

```
QuadReg
y=ax^2+bx+c
a=2.766899767
b=-6.002797203
c=10.01212121
R^2=.9991688446
```

TI-83/84 Plus

```
ExpReg
y=a*b^x
a=5.236195756
b=1.48297613
r^2=.9631105179
r=.9813819429
```

TI-83/84 Plus

```
PwrReg
y=a*x^b
a=3.353115397
b=1.766059823
r^2=.976406226
r=.9881326966
```

Para predecir la población de Estados Unidos para el año 2020, primero observe que el año 2020 está codificado como $x = 12$ (véase la tabla 10-4). Sustituyendo $x = 12$ en el modelo cuadrático de $y = 2.77x^2 - 6.00x + 10.01$, obtenemos el resultado $y = 337$, lo cual indica que se estima que la población de Estados Unidos será de 337 millones en el año 2020.

Tabla 10-4 Población (en millones) de Estados Unidos

Año	1800	1820	1840	1860	1880	1900	1920	1940	1960	1980	2000
Año codificado	1	2	3	4	5	6	7	8	9	10	11
Población	5	10	17	31	50	76	106	132	179	227	281

Reflexione. El resultado predicho de 337 millones en 2020 parece razonable (una proyección del U.S. Bureau of the Census sugiere que la población en 2020 será de alrededor de 325 millones). Sin embargo, existe un gran riesgo al hacer estimados de tiempos que están más allá del alcance de los datos disponibles. Por ejemplo, el modelo cuadrático sugiere que en 1492 la población de Estados Unidos era de 671 millones, un resultado absurdo. El modelo cuadrático parece ser bueno para los datos disponibles (1800-2000), pero otros modelos podrían ser mejores si es absolutamente necesario hacer estimados poblacionales más allá de este periodo.

En su artículo “Modeling the U.S. Population” (*AMATYC Review*, vol. 20, núm. 2), Sheldon Gordon emplea más datos que los de la tabla 10-4 y utiliza técnicas mucho más avanzadas para obtener mejores modelos poblacionales. En ese artículo, comenta algo importante:

“La mejor opción (de un modelo) depende del conjunto de datos que se analiza y requiere no sólo de cálculos, sino también de ejercitarse el juicio”.

Uso de la tecnología

Cualquier sistema capaz de realizar regresión múltiple puede emplearse para generar algunos de los modelos descritos en esta sección. Por ejemplo, STATDISK no está diseñado para trabajar directamente con el modelo

cuadrático, pero su función de regresión múltiple puede emplearse con los datos de la tabla 10-4 para generar el modelo cuadrático de la siguiente manera: primero introduzca los valores poblacionales en la columna 1 de la ventana de datos de STATDISK. Introduzca 1, 2, 3, . . . , 11 en la columna 2 e introduzca 4, 9, . . . , 121 en la columna 3. Haga clic en **Analysis** y luego seleccione **Multiple Regression**. Utilice las columnas 1, 2, 3, con la columna 1 como variable dependiente. Después de hacer clic en **Evaluate**, STATDISK genera la ecuación $y = 10.012 - 6.0028x + 2.7669x^2$, junto con $R^2 = 0.99917$, que son los mismos resultados obtenidos con la calculadora TI-83/84 Plus.

MINITAB Primero ingrese los datos apareados en las columnas C1 y C2, después seleccione **Stat**, **Regression** y **Fitted**

Line Plot. Usted puede elegir un modelo lineal, un modelo cuadrático o un modelo cúbico. Los resultados incluyen la ecuación, el valor de R^2 y la suma de cuadrados de los residuales.

TI-83/84 PLUS Primero inicie la función diagnóstica de la siguiente manera: presione **2nd CATALOG**, después baje hasta **DiagnosticON** y presione la tecla **ENTER** dos veces. Introduzca los datos apareados en las listas L1 y L2. Presione **STAT**, seleccione **CALC** y luego elija el modelo deseado de las opciones disponibles. Presione **ENTER**, luego ingrese L1, L2 (con la coma) y presione **ENTER** nuevamente. Los resultados incluyen el formato de la ecuación junto con los coeficientes utilizados en la ecuación; también se incluye el valor de R^2 en muchos de los modelos.



10-6 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Modelo.** ¿Qué es un modelo matemático?
- 2. R^2 .** ¿De qué manera se emplean los valores de R^2 para comparar los diferentes modelos que se están considerando?
- 3. Proyecciones.** En esta sección utilizamos los valores poblacionales del año 1800 al año 2000, y encontramos que el mejor modelo es el que está descrito por $y = 2.77x^2 - 6.00x + 10.01$, donde el valor poblacional de y está dado en millones. ¿Por qué sería erróneo utilizar este modelo para proyectar el tamaño de la población para el año 3000?

- 4. El mejor modelo.** Suponga que utilizamos una muestra con los métodos de esta sección para calcular que, de los cinco modelos posibles, el mejor modelo es $y = 4x^{1.2}$ con $R^2 = 0.200$. ¿Este mejor modelo parece ser un buen modelo? ¿Por qué?

Obtención del mejor modelo. En los ejercicios 5 a 12, construya un diagrama de dispersión e identifique el modelo matemático que se ajusta mejor a los datos indicados. Suponga que el modelo se va a emplear únicamente para el alcance que tienen los datos y considere sólo los modelos lineal, cuadrático, logarítmico, exponencial y potencial.

5.
$$\begin{array}{c|cccccc} x & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline y & 5 & 7 & 9 & 11 & 13 & 15 \end{array}$$

6.
$$\begin{array}{c|cccccc} x & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline y & 2 & 4 & 8 & 16 & 32 & 64 \end{array}$$

7.
$$\begin{array}{c|cccccc} x & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline y & 1 & 7 & 17 & 31 & 49 & 71 \end{array}$$

8.
$$\begin{array}{c|cccccc} x & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline y & 3 & 8.485 & 15.588 & 24 & 33.541 & 44.091 \end{array}$$

- 9. Muertes de manatíes por barcos.** La siguiente tabla lista el número de muertes de manatíes en Florida, relacionadas con encuentros con embarcaciones (según datos de Florida Fish and Wildlife Conservation).

Año	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Muertes	15	34	33	33	39	43	50	47	53	38	35
Año	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Muertes	49	42	60	54	67	82	78	81	95	73	69

- 10. Muertes de manatíes por causas naturales.** La siguiente tabla lista el número de muertes de manatíes en Florida por causas naturales (según datos de Florida Fish and Wildlife Conservation). ¿El mejor modelo parece ser un modelo razonablemente bueno?

Año	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Muertes	6	24	19	1	10	15	18	21	13	20	22
Año	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Muertes	33	35	101	42	12	37	37	34	59	102	25

- 11. Experimento de física.** Un experimento para una clase de física implica dejar caer una pelota de golf y registrar la distancia (en metros) que cae en diferentes tiempos (en segundos) después de ser soltada. Los datos se incluyen en la siguiente tabla. Proyecte la distancia para un tiempo de 12 segundos, dado que la pelota de golf se dejó caer de un edificio con una altura de 50 m.

Tiempo	0	0.5	1	1.5	2	2.5	3
Distancia	0	1.2	4.9	11.0	19.5	30.5	44.0

- 12. Mercado bursátil.** A continuación se listan, por renglón, los valores máximos anuales del Promedio Industrial Dow Jones para cada año a partir de 1980. ¿Cuál es el mejor valor predicho para el año 2004? Dado que el valor real máximo en 2004 fue 10,855, ¿qué tan bueno fue el valor predicho? ¿Qué sugiere el patrón acerca del mercado

bursátil para fines de inversión? (Actos de terrorismo y malas condiciones económicas causaron grandes pérdidas en el mercado bursátil en 2002).

1000	1024	1071	1287	1287	1553	1956	2722	2184	2791	3000	3169	3413
3794	3978	5216	6561	8259	9374	11,568	11,401	11,350	10,635	10,454		

10-6 MÁS ALLÁ DE LO BÁSICO

- 13. Ley de Moore.** En 1965 el cofundador de Intel, Gordon Moore, creó lo que ahora se conoce como *ley de Moore*: el número de transistores por pulgada cuadrada, en circuitos integrados, se duplica aproximadamente cada 18 meses. A continuación se incluyen datos que describen el número de transistores (en miles) para distintos años.

Año	1971	1974	1978	1982	1985	1989	1993	1997	1999	2000	2002	2003
Transistores	2.3	5	29	120	275	1180	3100	7500	24,000	42,000	220,000	410,000

- a. Suponiendo que la ley de Moore es correcta y que los transistores se duplican cada 18 meses, ¿cuál modelo matemático describe mejor esta ley: lineal, cuadrático, logarítmico, exponencial, potencial? ¿Qué función específica describe la ley de Moore?
 - b. ¿Cuál modelo matemático se ajusta mejor a los datos muestrales listados?
 - c. Compare los resultados de los incisos a) y b). ¿Parece que la ley de Moore funciona razonablemente bien?
- 14. Población en 2050.** Cuando se redactaron los ejercicios de esta sección, las Naciones Unidas utilizaron su propio modelo para predecir una población de 394 millones de habitantes en Estados Unidos para el año 2050. Con base en los datos de la tabla 10-4, ¿cuál de los modelos estudiados en la sección 10-6 da por resultado una población proyectada cercana a los 394 millones en 2050?

- 15. Uso del criterio de suma de cuadrados.** Además del valor de R^2 , otra medición utilizada para evaluar la calidad de un modelo es la *suma de cuadrados de los residuales*. Un residual es la diferencia entre un valor observado y y el valor \hat{y} predicho a partir del modelo, y se denota por \hat{y} . Los mejores modelos poseen las sumas de cuadrados más pequeñas. Remítase al ejemplo de esta sección.
- a. Calcule $\sum(y - \hat{y})^2$, la suma de cuadrados de los residuales que resultan del modelo lineal.
 - b. Calcule la suma de cuadrados de los residuales que resultan del modelo cuadrático.
 - c. Compruebe que, de acuerdo con el criterio de la suma de cuadrados, el modelo cuadrático es mejor que el modelo lineal.

Repaso

Este capítulo presentó métodos básicos para investigar relaciones o correlaciones entre dos o más variables.

- La sección 10-2 empleó diagramas de dispersión y el coeficiente de correlación lineal para decidir si existe una correlación lineal entre dos variables.
- La sección 10-3 presentó métodos para el cálculo de la ecuación de la recta de regresión que (por medio del criterio de los mínimos cuadrados) se ajusta mejor a los datos apareados. Cuando existe una correlación lineal significativa, la ecuación de regresión puede utilizarse para predecir el valor de una variable, dado algún valor de la otra variable.

- En la sección 10-4 se estudió el concepto de variación total, con componentes de variación explicada y sin explicar. El coeficiente de determinación r^2 nos brinda la proporción de la variación en la variable de respuesta (y) que puede explicarse por medio de la correlación lineal entre x y y . Desarrollamos métodos para construir intervalos de predicción, los cuales sirven para juzgar la exactitud de valores predichos.
- En la sección 10-5 consideramos la regresión múltiple, la cual nos permite investigar relaciones que implican más de una variable de predicción (x). Estudiamos procedimientos para obtener una ecuación de regresión múltiple, así como el valor del coeficiente múltiple de determinación R^2 , la R^2 ajustada y el valor P para la significancia general de la ecuación.
- En la sección 10-6 exploramos conceptos básicos para el desarrollo de un modelo matemático, consistente en una función que puede emplearse para describir una relación entre dos variables. A diferencia de las secciones anteriores de este capítulo, la sección 10-6 incluyó varias funciones no lineales.

Conocimientos estadísticos y pensamiento crítico

- 1. Correlación y regresión.** Con sus propias palabras, describa la correlación, la regresión y la diferencia entre ellas.
- 2. Correlación.** Dado un conjunto de datos apareados, se obtiene un coeficiente de correlación lineal de $r = 0$. ¿Significa esto que no existe una relación entre las dos variables?
- 3. Causalidad.** Un investigador médico descubre que existe una correlación lineal significativa entre la cantidad consumida de un fármaco y el nivel de colesterol del sujeto. ¿Se justifica el hecho de afirmar en un artículo científico que el fármaco reduce los niveles de colesterol? ¿Por qué?
- 4. Predicciones.** Después de descubrir que existe una correlación lineal significativa entre dos variables, se obtiene un valor predicho de y por medio de la ecuación de regresión. Dado que existe una correlación lineal significativa, ¿el valor proyectado será muy exacto?

Ejercicios de repaso

- 1. Muerte de manatíes.** La tabla que se incluye a continuación lista el número de muertes de manatíes en Florida, relacionadas con encuentros con embarcaciones y provocadas por causas naturales durante varios años (según datos de Florida Fish and Wildlife Conservation).
 - Calcule el valor del coeficiente de correlación lineal y determine si existe una correlación lineal significativa entre las dos variables.
 - Calcule la ecuación de la recta de regresión. Permita que el número de muertes naturales represente la variable de respuesta (y). ¿Cuál es el mejor número predicho de muertes naturales en un año con 50 muertes por encuentros con embarcaciones?

Embarcación	49	42	60	54	67	82	78	81	95	73	69
Natural	33	35	101	42	12	37	37	34	59	102	25

- 2. Old Faithful.** Utilice los datos que se presentan a continuación (de la tabla 10-1). Las duraciones están en segundos y las alturas en pies.
 - ¿Existe una correlación lineal significativa entre la duración de una erupción del géiser Old Faithful y la altura de la erupción?

- b.** Calcule la ecuación de la recta de regresión, donde la altura representa la variable de respuesta (y).
c. ¿Cuál es la mejor altura predicha de una erupción que tiene una duración de 180 segundos?

Duración	240	120	178	234	235	269	255	220
Altura	140	110	125	120	140	120	125	150

Predicción del costo de la electricidad. A continuación se muestran algunas mediciones de la casa del autor, tomadas del conjunto de datos 9 en el apéndice B. Utilice esos datos para los ejercicios 3 a 5.

kWh	3375	2661	2073	2579	2858	2296	2812	2433	2266	3128
Grado día para calentar	2421	1841	438	15	152	1028	1967	1627	537	26
Tempera- tura diaria promedio	26	34	58	72	67	48	33	39	66	71
Costo (en dólares)	321.94	221.11	205.16	251.07	279.8	183.84	244.93	218.59	213.09	333.49

- 3. a.** Utilice un nivel de significancia de 0.05 para probar una correlación lineal entre el costo de la electricidad y los kWh de electricidad consumidos.
b. ¿Qué porcentaje de la variación en el costo se puede explicar por medio de la relación lineal entre el consumo de electricidad (en kWh) y el costo?
c. Calcule la ecuación de la recta de regresión que exprese el costo (y) en términos de la cantidad de electricidad consumida (en kWh).
d. ¿Cuál es el mejor costo predicho para un tiempo en el que se utilizan 3000 kWh de electricidad?
- 4. a.** Utilice un nivel de significancia de 0.05 para probar una correlación lineal entre la temperatura diaria promedio y el costo.
b. ¿Qué porcentaje de la variación en el costo se puede explicar por medio de la relación lineal entre el costo y la temperatura diaria promedio?
c. Calcule la ecuación de la recta de regresión que exprese el costo (y) en términos de la temperatura diaria promedio.
d. ¿Cuál es el mejor costo predicho para un tiempo en el que la temperatura diaria promedio es de 40?
- 5.** Utilice un programa de cómputo como STATDISK, Minitab o Excel para calcular la ecuación de regresión múltiple con la forma $\hat{y} = b_0 + b_1x_1 + b_2x_2$, donde la variable de respuesta y representa el costo, x_1 representa el consumo de electricidad en kWh y x_2 representa la temperatura diaria promedio. Además, identifique el valor del coeficiente múltiple de determinación R^2 , la R^2 ajustada y el valor P que representa la significancia general de la ecuación de regresión múltiple. ¿Se puede usar la ecuación de regresión para predecir el costo? ¿Son mejores las ecuaciones de regresión del ejercicio 3 y del ejercicio 4?

Ejercicios de repaso acumulativo

Súper Bowl y DJIA. A continuación se listan los números totales de puntos anotados en juegos del Súper Bowl y el valor máximo del Promedio Industrial Dow Jones (DJIA). Los datos están apareados de acuerdo al año y representan años recientes y consecutivos. Utilice estos datos muestrales para los ejercicios 1 a 8.

Puntos en el Súper Bowl	56	55	53	39	41	37	69	61
DJIA	6561	8259	9374	11,568	11,401	11,350	10,635	10,454

1. Pruebe si existe una correlación entre los puntos del Súper Bowl y el DJIA. ¿El resultado es el que usted esperaba?
2. Calcule la ecuación de regresión en la que el valor máximo del DJIA sea la variable de respuesta (y). ¿Cuál es el mejor valor del DJIA predicho para un año en el que se anotaron 50 puntos en el Súper Bowl?
3. ¿Es posible probar la aseveración de que el número medio de puntos anotados en el Súper Bowl es igual al valor medio del DJIA? ¿Tendría sentido una prueba como ésta?
4. Construya un estimado de un intervalo de confianza del 95% para el número medio de puntos anotados en juegos del Súper Bowl.
5. ¿Por qué sería una mala idea tratar de estimar el siguiente valor máximo consecutivo del DJIA construyendo un estimado de un intervalo de confianza para los valores del DJIA?
6. Al parecer, ¿los puntos del Súper Bowl provienen de una población con una distribución normal? ¿Por qué?
7. Calcule la media y la desviación estándar de la muestra de puntos del Súper Bowl.
8. La media y la desviación estándar del ejercicio 7 son estadísticos muestrales, pero trátelos como parámetros poblacionales para una población distribuida normalmente, y calcule la probabilidad de que en un juego del Súper Bowl seleccionado al azar se anoten menos de 40 puntos totales.

Actividades de cooperación en equipo

1. **Actividad en clase** Organicen grupos de 8 a 12 personas. Para cada miembro de cada grupo, midan su estatura y también midan su estatura umbilical, que es la altura desde el piso hasta el ombligo. ¿Existe una correlación entre la estatura y la estatura umbilical? Si es así, calcule la ecuación de regresión con la estatura expresada en términos de la estatura umbilical. Según una antigua teoría, la proporción de la estatura respecto a la estatura umbilical de la persona promedio es la proporción áurea: $(1 + \sqrt{5})/2 \approx 1.6$. ¿Parece ser razonablemente exacta esta teoría?
2. **Actividad en clase** Formen grupos de 8 a 12 personas. Para cada miembro de cada grupo, *midan* la estatura y la envergadura de los brazos. Para la envergadura de los brazos, el sujeto debe estar de pie con los brazos extendidos, como las alas de un avión. Es fácil marcar la estatura y la envergadura de los brazos en el pizarrón y después medir las distancias desde ahí. Con los datos muestrales apareados, ¿existe la correlación entre estatura y la envergadura de los brazos? Si es así, calcule la ecuación de regresión con la estatura expresada en términos de la envergadura de los brazos. ¿Puede emplearse la envergadura de los brazos como un factor de predicción razonablemente bueno de la estatura?
3. **Actividad en clase** Formen grupos de 8 a 12 personas. Para cada sujeto, utilicen un hilo y una regla para medir la circunferencia de la cabeza y la longitud del antebrazo.

¿Existe relación entre estas dos variables? Si es así, ¿cuál es?

4. **Actividad en clase** Use una regla para medir el tiempo de reacción. Una persona debe suspender la regla sosteniéndola de un extremo, mientras el sujeto coloca sus dedos pulgar e índice en el extremo inferior, preparado para atrapar la regla cuando sea soltada. Registre la distancia que cae la regla antes de ser atrapada. Convierta esa distancia en el tiempo (en segundos) que tardó el sujeto en reaccionar y atrapar la regla. (Si la distancia se mide en pulgadas, utilice $t = \sqrt{d}/192$. Si la distancia se mide en centímetros, utilice $t = \sqrt{d}/487.68$.) Pruebe a cada sujeto una vez con la mano derecha y una vez con la mano izquierda, y luego registre los datos apareados. Haga una prueba de correlación. Calcule la ecuación de la recta de regresión. ¿La ecuación de la recta de regresión sugiere que la mano dominante tiene un tiempo de reacción más veloz?
5. **Actividad en clase** Formen grupos de 8 a 12 personas. Para cada miembro del grupo, registren el pulso contando el número de latidos por minuto. También registren la estatura. ¿Existe una relación entre el pulso y la estatura? Si es así, ¿cuál es?
6. **Actividad en clase** Reúna datos de cada estudiante referentes al número de tarjetas de crédito y el número de llaves que posee cada uno. ¿Existe una correlación? Si es así, ¿cuál es? Traten de identificar al menos una

explicación razonable para la presencia o ausencia de una correlación.

7. **Actividad en clase** Dividan la clase en grupos de tres o cuatro personas. El apéndice B incluye muchos conjuntos de datos que aún no se han utilizado en los ejemplos o ejercicios de este capítulo. Busque en el apéndice B un par de variables de interés, y luego investigue la correlación y la regresión. Establezca sus conclusiones y trate de identificar aplicaciones prácticas.
8. **Actividad fuera de clase** Dividan la clase en grupos de tres o cuatro personas. Investiguen la relación entre dos variables reuniendo sus propios datos muestrales apareados y utilizando los métodos de este capítulo para determinar si existe una correlación lineal significativa. También identifiquen la ecuación de regresión y describan

un procedimiento para predecir valores de una de las variables, cuando se tienen valores de la otra variable.
Temas sugeridos:

- ¿Existe una relación entre el sabor y el costo de distintas marcas de galletas de chocolate (o bebidas de cola)? El sabor puede medirse con base en una escala numérica, como del 1 al 10.
- ¿Existe una relación entre los salarios de los jugadores profesionales de béisbol (básquetbol o fútbol) y sus logros por temporada?
- ¿Existe una relación entre el largo de los pies de hombres (o mujeres) y su estatura?
- ¿Existe una relación entre el promedio de calificaciones de los estudiantes y la cantidad de tiempo que ven televisión? Si es así, ¿cuál es?

Proyecto tecnológico

Se dedica un gran esfuerzo a estudiar gemelos idénticos que fueron separados al nacer y que se criaron aparte uno de otro. Los gemelos idénticos se producen cuando un solo óvulo fertilizado se separa en dos, de manera que ambos comparten la misma configuración genética. Al obtener las puntuaciones de CI de gemelos idénticos separados al nacer, los investigadores esperan identificar los efectos de la herencia y del ambiente sobre la inteligencia. En este proyecto simularemos 100 conjuntos de nacimientos de gemelos, pero generaremos sus puntuaciones de CI de manera que no existan influencias genéticas o ambientales comunes. Con un generador de números aleatorios de un paquete de cómputo o de una calculadora, genere una lista de 100 puntuaciones de CI simuladas, seleccionadas al azar de una población distribuida normalmente, con una media de 100 y una desviación estándar de 15. Ahora use el mismo procedimiento para generar una segunda lista de 100 puntuaciones de CI simuladas, que también se seleccionen al azar de una población distribuida normalmente, con una media de 100 y una desviación estándar de 15. Aun cuando las dos listas se generan de manera independiente, considérelas como datos apareados,

de manera que la primera puntuación de cada lista represente el primer conjunto de gemelos, la segunda puntuación de cada lista al segundo conjunto de gemelos y así sucesivamente. Antes de realizar cualquier cálculo, primero estime un valor del coeficiente de correlación lineal que usted esperaría. Ahora aplique los métodos de la sección 10-2 con un nivel de significancia de 0.05 para probar una correlación lineal significativa y establezca sus resultados.

Consideré que el procedimiento anterior es un ensayo. Dada la forma en que los datos muestrales fueron generados, ¿qué proporción de esos ensayos debe conducir a la conclusión incorrecta de que existe una correlación lineal significativa? Si repetimos los ensayos, podemos verificar que la proporción es aproximadamente correcta. Repita el ensayo o combine sus resultados con otros para verificar que la proporción es aproximadamente correcta. Recuerde que el error tipo I es aquel que se comete al rechazar una hipótesis nula verdadera, lo que, en este caso, implicaría concluir que existe una correlación lineal significativa, cuando en realidad no existe tal correlación.

De los datos a la decisión

Pensamiento crítico: ¿Duragesic es eficaz para reducir el dolor?

A continuación se listan las medidas de la intensidad del dolor antes y después de utilizar el medicamento patentado Duragesic (según datos de Janssen Pharmaceutical Products, L.P.). Los datos aparecen en orden por renglón, y las medidas correspondientes son del mismo sujeto antes y después del tratamiento. Por ejemplo, el primer sujeto tuvo una medida de 1.2 antes del tratamiento, y una medida de 0.4 después del tratamiento. Cada par de mediciones corresponde a un sujeto, y la intensidad del dolor se midió utilizando la puntuación análoga visual estándar.

Intensidad del dolor antes del tratamiento con Duragesic

1.2 1.3 1.5 1.6 8.0 3.4 3.5 2.8 2.6 2.2
3.0 7.1 2.3 2.1 3.4 6.4 5.0 4.2 2.8 3.9
5.2 6.9 6.9 5.0 5.5 6.0 5.5 8.6 9.4 10.0
7.6

Intensidad del dolor después del tratamiento con Duragesic

0.4 1.4 1.8 2.9 6.0 1.4 0.7 3.9 0.9 1.8
0.9 9.3 8.0 6.8 2.3 0.4 0.7 1.2 4.5 2.0
1.6 2.0 2.0 6.8 6.6 4.1 4.6 2.9 5.4 4.8
4.1

Análisis de los resultados

- Utilice los datos indicados para construir un diagrama de dispersión y luego aplique los métodos de la sección 10-2 para hacer una prueba de correlación lineal entre la intensidad del dolor antes del tratamiento y después del tratamiento. Si existe una correlación lineal significativa, ¿se infiere que el tratamiento con el fármaco es eficaz?
- Utilice los datos para calcular la ecuación de la recta de regresión. Permita que la variable de respuesta (y) sea la intensidad del dolor después del tratamiento. ¿Cuál sería la ecuación de la recta de regresión para un tratamiento que no tiene ningún efecto?
- Los métodos de la sección 9-3 se pueden emplear para probar la aseveración de que dos poblaciones tienen la misma media. Identifique la aseveración específica de que el tratamiento es eficaz y luego utilice los métodos de la sección 9-3 para probar esa aseveración. Los métodos de la sección 9-3 se basan en el requisito de que las muestras son independientes. ¿Son independientes en este caso?
- Los métodos de la sección 9-4 permiten poner a prueba una aseveración acerca de datos apareados. Identifique la aseveración específica de que el tratamiento es eficaz y luego aplique los métodos de la sección 9-4 para probar esa aseveración.

ración específica de que el tratamiento es eficaz y luego aplique los métodos de la sección 9-4 para probar esa aseveración.

- ¿Cuál de los resultados anteriores es mejor para determinar si el tratamiento con el fármaco es eficaz para reducir el dolor? ¿Cuál de los resultados anteriores es el menos efectivo para determinar si el tratamiento con el fármaco es eficaz para reducir el dolor? Con base en los resultados anteriores, ¿parece que el fármaco es eficaz?



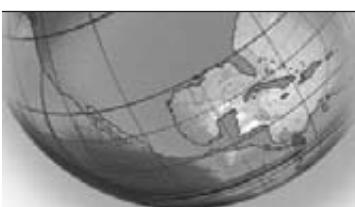
Regresión lineal

El coeficiente de correlación lineal es una herramienta que se utiliza para medir la fuerza de una relación lineal entre dos conjuntos de mediciones. Desde el punto de vista de los cálculos, el coeficiente de correlación puede obtenerse para cualesquier dos conjuntos de datos de valores apareados, sin importar lo que éstos representen. Por esta razón se deben plantear ciertas preguntas cuando se investiga una correlación. ¿Es razonable esperar una correlación lineal? ¿Podría una correlación obtenida ser causada por una tercera cantidad rela-

cionada con cada una de las variables estudiadas? Visite la página de Internet de este libro de texto:

<http://www.pearsoneducacion.net/triola>

El proyecto de Internet para este capítulo lo guiará hasta varios conjuntos de datos apareados en las áreas de deportes, medicina y economía. Entonces usted aplicará los métodos de este capítulo, calculará coeficientes de correlación y determinará rectas de regresión mientras considera las verdaderas relaciones que existen entre las variables implicadas.



Proyecto de Internet

La estadística en el trabajo

"En un mundo de negocios que se muestra fascinado con los números y con los datos, la estadística es fundamental para poder analizar y resumir apropiadamente grandes cantidades de datos".



Mark D. Haskell

*Director de pronósticos y análisis
Walt Disney World Resort*

Como director de pronósticos y análisis de Walt Disney World Resort, Mark dirige un equipo de personas responsables de planear y pronosticar valores como la asistencia, la ocupación de hoteles y las ganancias proyectadas. Al analizar diversos factores, Mark y su equipo ayudan a que Disney continúe trabajando para asegurarse de que cada huésped tenga una experiencia divertida e inolvidable en Walt Disney World Resort.

¿En qué consiste su trabajo?

Dirijo un equipo de personas responsables de planear y pronosticar medidas tales como la asistencia al parque, la ocupación en cada uno de nuestros hoteles y las utilidades que Walt Disney World obtendrá por estos negocios básicos.

¿Cómo usa usted la estadística y qué conceptos específicos de esta materia emplea?

La estadística es fundamental para el proceso de pronóstico. Muchas de nuestras herramientas de pronóstico se basan en técnicas de regresión múltiple, y algunos de esos modelos son más complejos que otros. También empleamos cotidianamente muchos conceptos estadísticos básicos, como el reporte del "error medio del porcentaje absoluto" de nuestros pronósticos, la comprensión de las medidas de tendencia central, las distribuciones y las técnicas de muestreo cuando realizamos la investigación de mercado; o la aplicación de correlaciones para entender la manera en que diferentes variables se asocian con nuestros principales negocios. Se dispone de muchos enfoques para crear pronósticos de alta calidad, pero la estadística es un bloque de construcción básico para casi cualquiera de esos enfoques.

Describa un ejemplo específico de cómo el uso de la estadística sirvió para mejorar un producto o servicio.

Recientemente, mi equipo utilizó el análisis de correlación para entender qué fuentes de datos serían más útiles para predecir la asistencia y los gastos en uno de nuestros centros de venta al detalle. Con base en ese trabajo, desarrollamos un modelo de regresión que sirve para que los líderes de la empresa conozcan las ganancias potenciales, determinen las necesidades de personal, establezcan las horas de operación, identifiquen nuevas oportunidades de productos y nuevas necesidades de inversión de capital, sólo por nombrar algunas aplicaciones.

¿Qué conocimientos de estadística se requieren para obtener un empleo como el suyo?

Yo tengo una maestría en economía, con especialidad en métodos de análisis cuantitativo. Por lo general se requiere de algún título de posgrado con énfasis en análisis estadístico para tener éxito en un puesto como el mío.

¿Considera que las personas que solicitan empleo en su compañía son vistas de forma más favorable si tienen algunos estudios de estadística?

Se requiere de cierto nivel de experiencia con la estadística para tener un puesto en el equipo de pronóstico y análisis. Hay muchos otros puestos en Walt Disney World que verían de manera más favorable a los solicitantes que tienen estudios de estadística.

¿Recomendaría a los estudiantes universitarios de hoy que estudien estadística? ¿Por qué?

Definitivamente sí. En un mundo de negocios que se muestra fascinado con los números y con los datos, la estadística es fundamental para poder analizar y resumir apropiadamente grandes cantidades de datos. Incluso si uno no es el responsable de realizar el análisis, necesita una comprensión básica para utilizar adecuadamente la información en la toma de decisiones. Es necesario aprender a utilizar la estadística de manera apropiada, o se corre el riesgo de que los individuos que saben más de estadística la utilicen en contra de uno.

¿Qué otras habilidades son importantes para los estudiantes universitarios de hoy?

Las habilidades de comunicación, tanto verbales como escritas. Se considera muy valiosas a las personas que saben analizar información compleja, que luego la simplifican y la comunican con claridad para su uso sencillo.



Experimentos multinomiales y tablas de contingencia

11



- 11-1 Panorama general**
- 11-2 Experimentos multinomiales: Bondad de ajuste**
- 11-3 Tablas de contingencia: Independencia y homogeneidad**
- 11-4 Prueba de McNemar para datos apareados**

Uso de la estadística para detectar fraudes

Malcolm Browne, en un artículo publicado en el *New York Times* (“Following Benford’s Law, or Looking Out for No. 1”), afirma que “las oficinas de recaudación de impuestos de varias naciones y de varios estados, entre ellos California, al igual que diversas compañías grandes y negocios contables, están utilizando programas de cómputo de detección basados en la ley de Benford”. De acuerdo con la ley de Benford, una variedad de conjuntos diferentes de datos incluyen números que tienen dígitos líderes (iniciales) que siguen la distribución que aparece en los primeros dos renglones de la tabla 11-1. Los conjuntos de datos con valores que tienen dígitos líderes que se ajustan a la ley de Benford incluyen valores de acciones bursátiles, tamaños poblacionales, números que aparecen en la primera página de un periódico, montos de las declaraciones de impuestos, longitudes de ríos y montos de cheques.

Cuando trabajaba para la fiscalía del distrito de Brooklyn, el investigador Robert Burton utilizó la ley

de Benford para identificar fraudes analizando los dígitos líderes en 784 cheques. Si los 784 cheques siguen la ley de Benford perfectamente, el 30.1% de ellos deberían tener montos con un dígito líder de 1. El número esperado de cheques con montos que tienen un dígito líder de 1 es 235.984 (puesto que el 30.1% de 784 es 235.984). Las otras frecuencias esperadas se listan en el tercer renglón de la tabla 11-1. El último renglón de la tabla 11-1 lista las frecuencias de los dígitos líderes de los montos de 784 cheques expedidos por siete compañías diferentes. Una rápida comparación visual indica que ahí parecen estar las principales discrepancias entre las frecuencias esperadas por la ley de Benford y las frecuencias observadas en los montos de los cheques, pero, ¿cómo medimos esta discordancia? ¿Son significativas estas discrepancias? ¿Existe evidencia suficiente para justificar la conclusión de que se cometió un fraude? ¿La evidencia está más allá de una “duda razonable”? En este capítulo examinaremos estas preguntas.

Tabla 11-1 Ley de Benford: Distribución de dígitos líderes

Dígito líder	1	2	3	4	5	6	7	8	9
Ley de Benford: Distribución de frecuencias de dígitos líderes	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%
Frecuencias esperadas de dígitos líderes de 784 cheques según la ley de Benford	235.984	137.984	98.000	76.048	61.936	52.528	45.472	39.984	36.064
Dígitos líderes observados de 784 cheques que se analizaron por fraude	0	15	0	76	479	183	8	23	0

11-1 Panorama general

Este capítulo incluye datos categóricos (cualitativos o de atributo) que pueden separarse en diferentes celdas. Por ejemplo, podemos separar una muestra de dulces M&M en las categorías de color rojo, anaranjado, amarillo, café, azul y verde. Después de calcular el conteo de frecuencia para cada categoría, probamos la aseveración de que las frecuencias se ajustan a (o concuerdan con) la distribución de color aseverada por el fabricante (Mars Inc.). El objetivo principal de este capítulo es probar aseveraciones acerca de datos categóricos que consisten en conteos de frecuencias para diferentes categorías. En la sección 11-2 consideraremos experimentos multinomiales, que consisten en conteos de frecuencias observadas en un solo renglón o columna (llamado tabla de frecuencias de un factor), y probaremos la aseveración de que los conteos de frecuencia observados concuerdan con alguna distribución aseverada. En la sección 11-3 consideraremos tablas de contingencia (o tablas de frecuencias de dos factores), que consisten en conteos de frecuencias ordenados en una tabla con al menos dos renglones y dos columnas. En la sección 11-4 estudiaremos tablas de dos factores con datos apareados.

Los métodos de este capítulo utilizan la misma distribución χ^2 (chi cuadrada) que se estudió en la sección 7-5. A manera de repaso, las propiedades importantes de la distribución chi cuadrada se listan a continuación:

1. La distribución chi cuadrada no es simétrica. (Véase la figura 11-1).
2. Los valores de la distribución chi cuadrada pueden ser 0 o positivos, pero no pueden ser negativos. (Véase la figura 11-1).
3. La distribución chi cuadrada es diferente para cada número de grados de libertad. (Véase la figura 11-2).

Los valores críticos de la distribución chi cuadrada se encuentran en la tabla A-4.

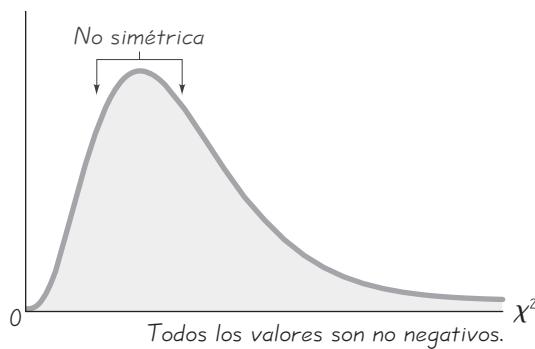


Figura 11-1 La distribución chi cuadrada

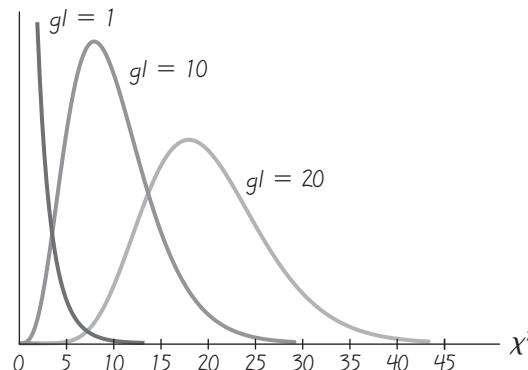


Figura 11-2 Distribución chi cuadrada para 1, 10 y 20 grados de libertad

Experimentos multinomiales:

11-2 Bondad de ajuste

Concepto clave Con datos separados en diferentes categorías, someteremos a prueba la hipótesis de que la distribución de los datos coincide con alguna distribución aseverada (es decir, “se ajusta” a ella). La prueba de hipótesis empleará la distribución chi cuadrada con los conteos de frecuencias observados y los conteos de frecuencias que se esperarían con la distribución aseverada. El estadístico de prueba chi cuadrada es una medida de la discrepancia entre las frecuencias observadas y las esperadas.

Comenzamos con la definición de un *experimento multinomial*, que es muy similar a la definición de un experimento binomial que se estudió en la sección 5-3, excepto que un experimento multinomial tiene más de dos categorías (a diferencia de un experimento binomial, que tiene exactamente dos categorías).

Definición

Un **experimento multinomial** es un experimento que satisface las siguientes condiciones:

1. El número de ensayos es fijo.
2. Los ensayos son independientes.
3. Todos los resultados de cada ensayo deben clasificarse exactamente en una de varias categorías diferentes.
4. Las probabilidades para las diferentes categorías permanecen constantes en cada ensayo.

EJEMPLO Últimos dígitos de pesos Miles de sujetos se estudian de manera rutinaria como parte del National Health Examination Survey. Los procedimientos de examen son muy exactos. Por ejemplo, cuando se obtienen pesos de sujetos, es sumamente importante pesarlos de verdad en vez de pedirles que reporten su peso. Se sabe que cuando la gente reporta su peso, generalmente da un peso más bajo que el real. Entonces, ¿cómo pueden verificar los investigadores que los pesos se obtuvieron por medio de mediciones reales y no por el reporte de los sujetos? Un método consiste en analizar los *últimos dígitos* de los pesos. Cuando la gente reporta su peso, tiende a redondear la cifra, a menudo *hacia* el entero inferior. Los últimos dígitos de los pesos reportados suelen tener un número desproporcionado de ceros y cinco, en comparación con los últimos dígitos de los pesos obtenidos a través de un proceso de medición. En contraste, cuando realmente se pesa a las personas, los últimos dígitos tienden a distribuirse de manera uniforme, de modo que 0, 1, 2,..., 9 se presentan aproximadamente con la misma frecuencia. El autor obtuvo los pesos de 80 estudiantes elegidos al azar, cuyos últimos dígitos se resumen en la tabla 11-2. Más adelante analizaremos los datos; por ahora, simplemente verificaremos que se satisfagan las cuatro condiciones de un experimento multinomial.

continúa

Tabla 11-2
Últimos dígitos de pesos

Último dígito	Frecuencia
0	35
1	0
2	2
3	1
4	4
5	24
6	1
7	4
8	7
9	2

SOLUCIÓN He aquí la verificación de que se satisfacen las cuatro condiciones del experimento multinomial:

1. El número de ensayos (últimos dígitos) es el número fijo 80.
2. Los ensayos son independientes, puesto que el último dígito de cualquier peso individual no afecta al último dígito de cualquier otro peso.
3. Cada resultado (último dígito) se clasifica exactamente en una de 10 categorías diferentes. Las categorías se identifican como 0, 1, 2,..., 9.
4. Al poner a prueba la aseveración de que los 10 dígitos son igualmente probables, cada dígito posible tiene la probabilidad de 1/10, y se supone que esa probabilidad permanece constante para cada sujeto.

En esta sección presentamos un método para probar la aseveración de que en un experimento multinomial, las frecuencias observadas en las diferentes categorías se ajustan a una distribución en particular. Puesto que hacemos una prueba de qué tan bien se ajusta una frecuencia de distribución observada a alguna distribución teórica especificada, este método suele llamarse *prueba de bondad de ajuste*.

Definición

 La **prueba de bondad de ajuste** se utiliza para probar la hipótesis de que una distribución de frecuencias se ajusta a (o coincide con) alguna distribución aseverada.

Por ejemplo, utilizando los datos de la tabla 11-2, podemos probar la hipótesis de que los datos se ajustan a una distribución uniforme, con igual probabilidad para todos los dígitos. Nuestras pruebas de bondad de ajuste incorporarán la siguiente notación.

Notación

- O representa la *frecuencia observada* de un resultado.
- E representa la *frecuencia esperada* de un resultado.
- k representa el número de *categorías diferentes* o resultados.
- n representa el *número total de ensayos*.

Cálculo de frecuencias esperadas

En la tabla 11-2 las frecuencias observadas O se denotan por 35, 0, 2, 1, 4, 24, 1, 4, 7 y 2. La suma de las frecuencias observadas es 80, de manera que $n = 80$. Si suponemos que los 80 dígitos se obtuvieron de una población en la que todos los dígitos son igualmente probables, entonces *esperamos* que cada dígito se presente en 1/10 de los 80 ensayos, de manera que cada una de las 10 frecuencias esperadas está dada por $E = 8$. Si generalizamos este resultado, obtenemos un procedimiento sencillo para calcular las frecuencias esperadas, siempre y cuando supongamos que todas las frecuencias esperadas son iguales: simplemente divide el número total de observaciones entre el número de categorías diferentes ($E = n/k$). En otros casos en los que no todas las frecuencias esperadas son iguales, a menudo podemos calcular las frecuencias esperadas para cada categoría multiplicando la suma de todas las frecuencias observadas por la probabilidad p de la categoría, de manera que $E = np$. Aquí resumimos estos dos procedimientos.

- Si todas las frecuencias esperadas son iguales, entonces cada frecuencia esperada es la suma de todas las frecuencias observadas dividida entre el número de categorías, de manera que $E = n/k$.
- Si las frecuencias esperadas no son todas iguales, entonces cada frecuencia esperada se calcula multiplicando la suma de todas las frecuencias observadas por la probabilidad para la categoría, de manera que $E = np$ para cada categoría.

Aun cuando estas dos fórmulas para E pueden ser muy buenas, sería mejor utilizar un método informal basado en la comprensión de las circunstancias. Sólo pregúntese: “¿Cómo se pueden repartir las frecuencias observadas entre las diferentes categorías, de manera que exista una coincidencia perfecta con la distribución aseverada?” Además, reconozca que todas las frecuencias *observadas* deben ser números enteros, puesto que representan conteos reales, en tanto que las frecuencias *esperadas* no requieren ser números enteros. Por ejemplo, cuando se tira un dado 33 veces, la frecuencia esperada para cada posible resultado es $33/6 = 5.5$. Se espera que el número 3 se presente con una frecuencia de 5.5, aunque es imposible obtener el resultado de que el 3 se presente exactamente 5.5 veces.

Sabemos que las frecuencias muestrales por lo regular se desvían un poco de los valores que esperamos teóricamente, y ahora planteamos la pregunta clave: ¿Son estadísticamente significativas las diferencias entre los valores *observados O* reales y los valores teóricos *esperados E*? Necesitamos una medida de la discrepancia entre los valores O y E , así que utilizamos el estadístico de prueba dado con los supuestos y los valores críticos. (Más adelante explicaremos cómo se desarrolló este estadístico de prueba, pero usted puede ver que incluye diferencias de $O - E$ como componente clave).

Requisitos

1. Los datos se seleccionaron al azar.
2. Los datos muestrales consisten en conteos de frecuencias para cada una de las diferentes categorías.
3. Para cada categoría, la frecuencia *esperada* es al menos de 5. (La frecuencia esperada para una categoría es la frecuencia que ocurriría si los datos realmente tuvieran la distribución que se asevera. No existe ningún requisito de que la frecuencia *observada* para cada categoría deba ser al menos de 5).

Estadístico de prueba para pruebas de bondad de ajuste en experimentos multinomiales

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Valores críticos

1. Los valores críticos se encuentran en la tabla A-4, utilizando $k - 1$ grados de libertad, donde k = número de categorías.
2. Las pruebas de hipótesis por bondad de ajuste siempre son de *cola derecha*.

LA ESTADÍSTICA EN LAS NOTICIAS

Los asientos de avión más seguros

Muchos de nosotros creemos que, en un accidente aéreo, los asientos de la parte de atrás son los más seguros. Los expertos en seguridad no están de acuerdo en que alguna parte específica de un avión sea más segura que las otras. Algunos aviones chocan primero con la nariz cuando caen, pero otros chocan con la cola al despegar. Matt McCormick, un experto en supervivencia del National Transportation Safety Board, dijo a la revista *Travel* que “no existe ningún lugar seguro para sentarse”. Se pueden utilizar pruebas de bondad de ajuste con la hipótesis nula de que todas las secciones de un avión son igualmente seguras. Los aviones accidentados podrían dividirse en las secciones frontal, media y posterior. Entonces, las frecuencias observadas de decesos podrían compararse con las frecuencias que se esperarían con una distribución de decesos uniforme. El estadístico de prueba χ^2 refleja el tamaño de las discrepancias entre las frecuencias observadas y las esperadas, y revelaría si algunas secciones son más seguras que las otras.

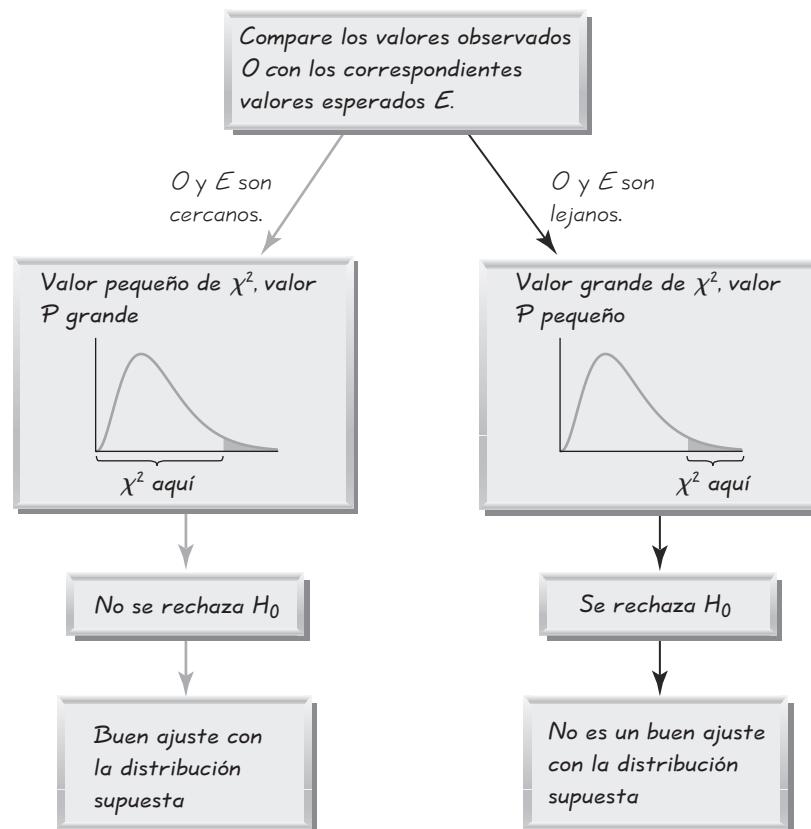


Figura 11-3 Relaciones entre el estadístico de prueba χ^2 , el valor P y la bondad de ajuste

El estadístico de prueba χ^2 se basa en las diferencias entre valores observados y esperados, de manera que una *concordancia cercana* entre los valores observados y esperados conducirá a un valor de χ^2 *pequeño* y un valor P *grande*. Una discrepancia grande entre los valores observados y esperados conducirá a un valor de χ^2 *grande* y un valor P *pequeño*. De esta forma, las pruebas de hipótesis de esta sección siempre son de cola derecha, puesto que el valor crítico y la región crítica se localizan en el extremo derecho de la distribución. Estas relaciones se resumen e ilustran en la figura 11-3.

Una vez que sabemos calcular el valor del estadístico de prueba y el valor crítico, podemos probar hipótesis utilizando el mismo procedimiento general que se describió en el capítulo 8.

EJEMPLO Análisis de los últimos dígitos de pesos: frecuencias iguales esperadas Consulte los últimos dígitos de los 80 pesos en la tabla 11-2. Ponga a prueba la aseveración de que los dígitos *no* se presentan con la misma frecuencia. Con base en los resultados, ¿qué concluye acerca del procedimiento usado para obtener los pesos?

SOLUCIÓN

REQUISITO Es necesario que los datos muestrales se seleccionen al azar, que consistan en conteos de frecuencias, que provengan de un experimento multinomial y que cada frecuencia esperada sea al menos de 5. Antes señalamos

que los datos provienen de estudiantes elegidos al azar. Los datos consisten en conteos de frecuencias. En el ejemplo anterior se estableció que las condiciones para un experimento multinomial se satisfacen. El análisis previo de los valores esperados incluyó el resultado de que cada frecuencia esperada es 8, de manera que cada frecuencia esperada sí satisface el requisito de ser al menos de 5. Todos los requisitos se satisfacen y podemos proceder con la prueba de hipótesis.

La aseveración de que los dígitos no ocurren con la misma frecuencia es equivalente a la aseveración de que las frecuencias relativas o probabilidades de las 10 celdas (p_0, p_1, \dots, p_9) no son todas iguales. Aplicaremos el método tradicional para prueba de hipótesis (véase la figura 8-9).

- Paso 1: La aseveración original es que los dígitos no se presentan con la misma frecuencia. Es decir, al menos una de las probabilidades p_0, p_1, \dots, p_9 es diferente de las otras.
- Paso 2: Si la aseveración original es falsa, entonces todas las probabilidades son las mismas. Es decir, $p_0 = p_1 = \dots = p_9$.
- Paso 3: La hipótesis nula debe contener la condición de igualdad, así que tenemos

$$H_0: p_0 = p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9$$

$$H_1: \text{Al menos una de las probabilidades es diferente de las otras.}$$

- Paso 4: No se especificó un nivel de significancia, así que seleccionamos $\alpha = 0.05$, una elección muy común.
- Paso 5: Puesto que probamos la aseveración de que la distribución de los últimos dígitos es una distribución uniforme, utilizamos la prueba de bondad de ajuste descrita en esta sección. Se utiliza la distribución χ^2 con el estadístico de prueba que se dio al principio.
- Paso 6: Las frecuencias observadas O se listan en la tabla 11-2. Cada frecuencia esperada E correspondiente es igual a 8 (porque los 80 dígitos estarían distribuidos de manera uniforme a lo largo de las 10 categorías). La tabla 11-3 muestra el cálculo del estadístico de prueba χ^2 . El estadístico de prueba es $\chi^2 = 156.500$. El valor crítico es $\chi^2 = 16.919$ (obtenido de la tabla A-4 con $\alpha = 0.05$ en la cola derecha y con grados de libertad iguales a $k - 1 = 9$). El estadístico de prueba y el valor crítico se muestran en la figura 11-4.
- Paso 7: Puesto que el estadístico de prueba cae dentro de la región crítica, existe evidencia suficiente para rechazar la hipótesis nula.
- Paso 8: Existe suficiente evidencia para sustentar la aseveración de que los últimos dígitos no se presentan con la misma frecuencia relativa. Ahora tenemos evidencia muy fuerte que sugiere que los pesos realmente no se midieron. Es razonable especular que se trata de valores reportados y no de mediciones reales.

El ejemplo anterior incluyó la hipótesis nula de que las probabilidades para las diferentes categorías son todas iguales. Los métodos de esta sección también pueden utilizarse cuando las probabilidades (o frecuencias) hipotéticas son diferentes, como en el ejemplo siguiente.

Tabla 11-3 Cálculo del estadístico de prueba χ^2 para los últimos dígitos de pesos					
Último dígito	Frecuencia observada O	Frecuencia esperada E	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
0	35	8	27	729	91.1250
1	0	8	-8	64	8.0000
2	2	8	-6	36	4.500
3	1	8	-7	49	6.125
4	4	8	-4	16	2.000
5	24	8	16	256	32.000
6	1	8	-7	49	6.125
7	4	8	-4	16	2.000
8	7	8	-1	1	0.125
9	2	8	-6	36	4.500
		80	80	$\chi^2 = \sum \frac{(O - E)^2}{E} = 156.500$	
(Con excepción de los errores por redondeo, estos dos totales deben coincidir)					

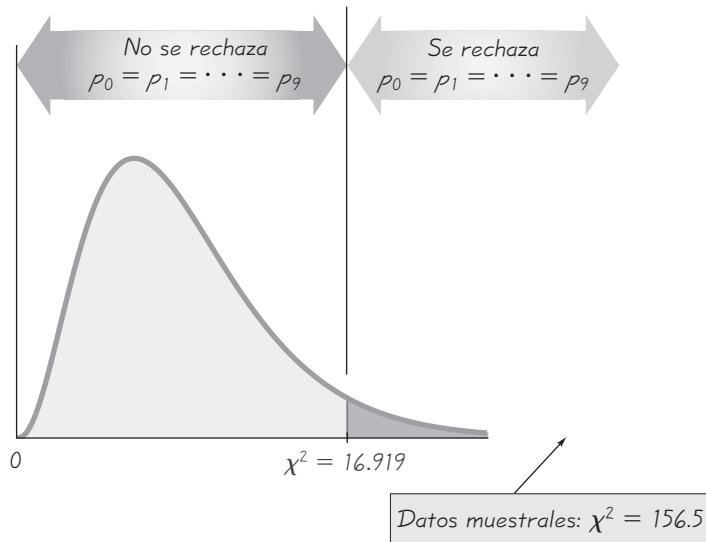


Figura 11-4 Prueba de $p_0 = p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9$



EJEMPLO Detección de fraude: frecuencias desiguales esperadas

En el problema del capítulo se señaló que algunas veces se utiliza la estadística para detectar fraudes. El segundo renglón de la tabla 11-1 lista porcentajes para dígitos líderes, tal como se esperarían según la ley de Benford, y el tercer renglón lista los conteos de frecuencias esperadas cuando los porcentajes de la ley de Benford se aplican a 784 dígitos líderes. El último renglón de la tabla 11-1 lista las frecuencias observadas de los dígitos líderes esperados de los montos de 784 cheques expedidos por siete compañías diferentes. Pruebe la aseveración de que existe una discrepancia significativa entre los dígitos líderes esperados según la ley de Benford y los dígitos líderes observados en los 784 cheques. Utilice un nivel de significancia de 0.01.

SOLUCIÓN

REQUISITOS ✓ Para verificar los requisitos indicados anteriormente, comenzaremos señalando que los dígitos líderes de los cheques en realidad no son aleatorios. Sin embargo, los manejamos como si fueran aleatorios para determinar si son los resultados típicos que se obtendrían de una muestra aleatoria, según la ley de Benford. Los datos aparecen como conteos de frecuencias y satisfacen los requisitos de un experimento multinomial. Cada frecuencia esperada (como se observa en la tabla 11-1) es al menos de 5. Todos los requisitos se satisfacen y podemos proceder con la prueba de hipótesis. ✓

Paso 1: La aseveración original dice que los dígitos líderes no tienen la misma distribución que plantea la ley de Benford. Es decir, al menos una de las siguientes ecuaciones es incorrecta: $p_1 = 0.301$ y $p_2 = 0.176$ y $p_3 = 0.125$ y $p_4 = 0.097$ y $p_5 = 0.079$ y $p_6 = 0.067$ y $p_7 = 0.058$ y $p_8 = 0.051$ y $p_9 = 0.046$. (Las proporciones son los valores decimales equivalentes de los porcentajes listados para la ley de Benford en la tabla 11-1).

Paso 2: Si la aseveración original es falsa, entonces lo siguiente es verdadero: $p_1 = 0.301$ y $p_2 = 0.176$ y $p_3 = 0.125$ y $p_4 = 0.097$ y $p_5 = 0.079$ y $p_6 = 0.067$ y $p_7 = 0.058$ y $p_8 = 0.051$ y $p_9 = 0.046$.

Paso 3: La hipótesis nula debe contener la condición de igualdad, entonces tenemos

$$H_0: p_1 = 0.301 \text{ y } p_2 = 0.176 \text{ y } p_3 = 0.125 \text{ y } p_4 = 0.097 \text{ y } p_5 = 0.079 \text{ y } p_6 = 0.067 \text{ y } p_7 = 0.058 \text{ y } p_8 = 0.051 \text{ y } p_9 = 0.046.$$

$H_1:$ Al menos una de las proporciones es diferente del valor aseverado.

Paso 4: Se especificó un nivel de significancia de $\alpha = 0.01$.

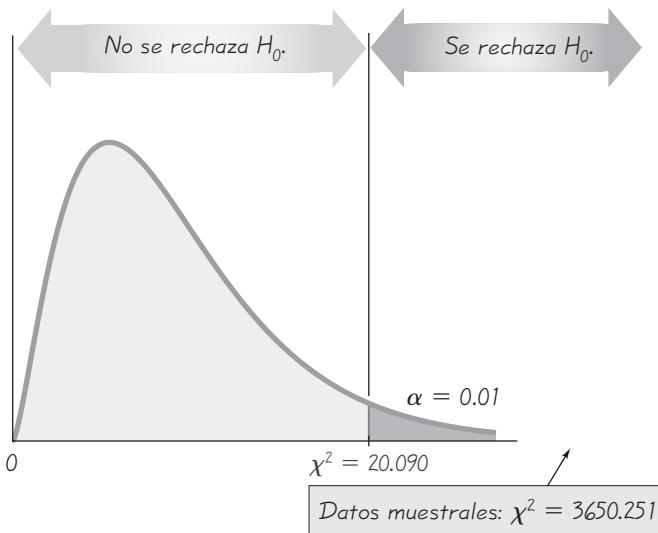
Paso 5: Puesto que estamos probando la aseveración de que la distribución de dígitos coincide con la distribución planteada por la ley de Benford, utilizamos la prueba de bondad de ajuste descrita en esta sección. Se utiliza la distribución χ^2 con el estadístico de prueba dado anteriormente.

Paso 6: Las frecuencias observadas O y las frecuencias esperadas E se listan en la tabla 11-1. La suma de los nueve valores $(O - E)^2/E$ da como resultado el estadístico de prueba $\chi^2 = 3650.251$. El valor crítico es $\chi^2 = 20.090$ (obtenido de la tabla A-4 con $\alpha = 0.01$ en la cola derecha y con grados de libertad iguales a $k - 1 = 8$). El estadístico de prueba y el valor crítico se muestran en la figura 11-5.

continúa

Figura 11-5

Prueba de concordancia entre frecuencias observadas y frecuencias esperadas con la ley de Benford



- Paso 7: Puesto que el estadístico de prueba cae dentro de la región crítica, existe suficiente evidencia para rechazar la hipótesis nula.
- Paso 8: Hay evidencia suficiente para sustentar la aseveración de que existen discrepancias entre la distribución esperada según la ley de Benford y la distribución observada de los dígitos líderes de los cheques.

En la figura 11-6a), graficamos las proporciones aseveradas de 0.301, 0.176, 0.125, 0.097, 0.079, 0.067, 0.058, 0.051 y 0.046 junto con las proporciones observadas de 0.000, 0.019, 0.000, 0.097, 0.611, 0.233, 0.010, 0.029 y 0.000, para poder visualizar la discrepancia entre la distribución aseverada de la ley de Benford y las frecuencias que se observaron. Los puntos a lo largo de la línea gris oscuro representan las proporciones aseveradas, y los puntos a lo largo de la línea gris claro representan las proporciones observadas. Los pares de puntos correspondientes están muy alejados, lo cual indica que las frecuencias esperadas son muy diferentes de las frecuencias observadas correspondientes. La gran disparidad entre la línea gris claro para las frecuencias observadas y la línea gris oscuro para las frecuencias esperadas sugiere que los montos de los cheques no son el resultado de transacciones típicas; al parecer, podría haber un fraude. De hecho, la fiscalía del distrito de Brooklyn levantó cargos por fraude utilizando esta línea de razonamiento. Para hacer una comparación, observe la figura 11-6b), que se basa en los dígitos líderes de las cantidades de los últimos 200 cheques expedidos por el autor. Note cómo las proporciones observadas de los cheques del autor concuerdan bastante bien con las proporciones esperadas con la ley de Benford. Los cheques del autor parecen ser típicos y no muestran un patrón que podría sugerir un fraude. En general, gráficas como la de la figura 11-6 son muy útiles para comparar visualmente las frecuencias esperadas y las frecuencias observadas, así como para sugerir cuáles categorías dan por resultado las principales discrepancias.

Valores ***P***

Los ejemplos de esta sección emplearon el método tradicional de prueba de hipótesis, pero también se puede utilizar el método del valor *P*. Los valores *P* se obtienen automáticamente con STATDISK o la calculadora TI-83/84 Plus, o bien, mediante los métodos descritos en el capítulo 8. Observe cómo, el ejemplo anterior

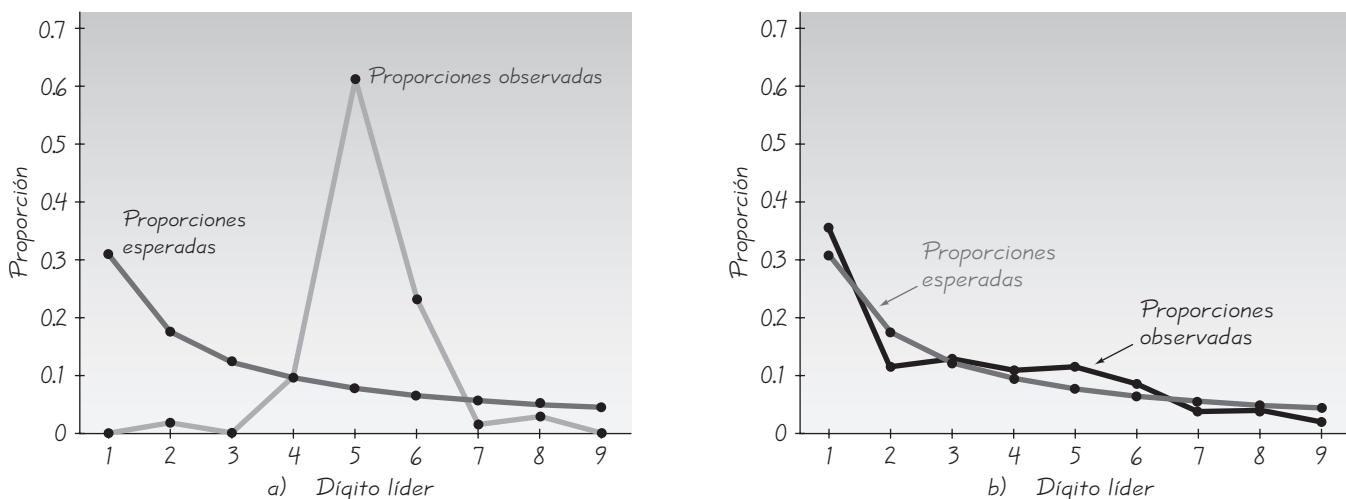


Figura 11-6 Comparación de las frecuencias observadas y las frecuencias esperadas con la ley de Benford

dio como resultado un estadístico de prueba de $\chi^2 = 3650.251$. Ese ejemplo tenía $k = 9$ categorías, por lo tanto, había $k - 1 = 8$ grados de libertad. Remitiéndonos a la tabla A-4, vemos que para el renglón con 8 grados de libertad, el estadístico de prueba de 3650.251 es mayor que el valor más alto del renglón (21.955). Puesto que el estadístico de prueba de $\chi^2 = 3650.251$ está más a la derecha que 21.955, el valor P es menor que 0.005. Si se ejecutan los cálculos para el ejemplo anterior en STATDISK, la pantalla de resultados incluirá un valor P de 0.0000. El valor P pequeño sugiere que la hipótesis nula debería rechazarse. (Recuerde, rechazamos la hipótesis nula cuando el valor P es igual o menor que el nivel de significancia). Mientras que el método tradicional de prueba de hipótesis nos llevó a rechazar la aseveración de que las 784 cantidades de los cheques tienen dígitos líderes que cumplen con la ley de Benford, el valor P de 0.0000 indica que la probabilidad de obtener dígitos líderes como los que se obtuvieron es extremadamente pequeña. Esto parece ser evidencia, “más allá de una duda razonable”, de que los montos de los cheques no son el resultado de transacciones típicas honestas.

Fundamentos del estadístico de prueba: Los ejemplos anteriores deben ser útiles para tener una idea de la función del estadístico de prueba χ^2 . Debe quedar claro que queremos medir la cantidad de discordancia entre las frecuencias observadas y las esperadas. Sumar simplemente las diferencias entre los valores observados y esperados no da como resultado una medida eficaz, puesto que esa suma siempre es 0. Se obtiene un mejor estadístico al elevar al cuadrado los valores $O - E$, lo cual refleja las diferencias entre las frecuencias observadas y esperadas. (Las razones para elevar al cuadrado los valores $O - E$ son esencialmente las mismas que las razones para elevar al cuadrado los valores $x - \bar{x}$ en la fórmula de la desviación estándar). El valor de $\sum(O - E)^2$ sólo mide la magnitud de las diferencias, pero necesitamos calcular la magnitud relativa de las diferencias en relación con lo que se esperaba. Esta magnitud relativa se calcula a través de la división entre las frecuencias esperadas, como en el estadístico de prueba.

La distribución teórica de $\sum(O - E)^2/E$ es una distribución discreta, puesto que el número de valores posibles está limitado a un número finito. La distribu-

Uso de la tecnología

STATDISK Primero registre las frecuencias observadas en la primera columna de la ventana de datos. Si las frecuencias esperadas no son todas iguales, también incluya una segunda columna que indique las proporciones esperadas o las frecuencias esperadas reales. Seleccione **Analysis** de la barra del menú principal, luego seleccione la opción **Multinomial Experiments**. Elija entre “equal expected frequencies” y “unequal

expected frequencies” e ingrese los datos en el cuadro de diálogo; luego, haga clic en **Evaluate**.

EXCEL Para usar DDXL, ingrese los nombres de las categorías en una columna, registre las frecuencias observadas en una segunda columna y utilice una tercera columna para ingresar las *proporciones* esperadas en forma decimal (como 0.20, 0.25 y 0.30). Haga clic en **DDXL** y seleccione **Tables** del menú. En el menú denominado **Function Type**, seleccione **Goodness-of-Fit**. Haga clic en el ícono del lápiz para **Category Names** e indique el rango de celdas que contenga los nombres de la categoría, como A1:A5. Haga clic en el ícono del lápiz para **Observed Counts** e indique el rango de celdas que contengan las frecuencias observadas, como B1:B5. Haga clic en el ícono del lápiz para **Test Distribution** e indique el rango de celdas que contenga las proporciones esperadas en forma decimal, como C1:C5. Haga clic en **OK** para obtener el estadístico de prueba chi cuadrada y el valor *P*.

TI-83/84 PLUS Los métodos de esta sección no están disponibles como procedimiento directo en la calculadora TI-83/84 Plus, pero se puede utilizar el programa X2GOF de Michael Lloyd. (El programa está incluido en el CD-ROM que viene con este libro; también se puede descargar del sitio del texto en Internet en www.pearsoneducacion.net/Triola). Primero anote las frecuencias observadas en la lista L1. Después, calcule las frecuencias esperadas y regístrelas en la lista L2. Presione la tecla **PRGM**, luego corra el programa X2GOF y haga lo que se le pide. Los resultados incluyen el estadístico de prueba y el valor *P*.

MINITAB Anote las frecuencias observadas en la columna C1. Si las frecuencias esperadas no son todas iguales, ingréselas como proporciones en la columna C2. Seleccione **STAT**, **TABLES** y **Chi-Square Goodness-of-Fit**. Complete la información que se le pide en la ventana y haga clic en **OK**.

ción puede aproximarse por medio de una distribución chi cuadrada, que es continua. Esta aproximación por lo regular se considera aceptable, siempre y cuando todos los valores esperados *E* sean al menos de 5. (Existen maneras de superar el problema de una frecuencia esperada menor que 5, como combinar las categorías para que todas las frecuencias esperadas sean al menos de 5).

El número de grados de libertad refleja el hecho de que podemos asignar libremente frecuencias a $k - 1$ categorías, antes de que se determine la frecuencia para cada categoría. (Aunque decimos que podemos asignar “con libertad” frecuencias a $k - 1$ categorías, no podemos tener frecuencias negativas ni podemos tener frecuencias tan grandes que su suma exceda el total de las frecuencias observadas de todas las categorías combinadas).

11-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Bondad de ajuste.** ¿A qué nos referimos cuando decimos que hacemos una prueba de la “bondad de ajuste”?
- 2. Prueba de cola derecha.** ¿Por qué la prueba de hipótesis para la bondad de ajuste siempre es una prueba de cola derecha?
- 3. Frecuencias observadas y esperadas.** ¿Qué es una frecuencia observada? ¿Qué es una frecuencia esperada?
- 4. Pesos de estudiantes.** Un investigador reúne los pesos de 20 estudiantes varones elegidos al azar de cuatro clases diferentes, luego calcula el total de los pesos y los resume en la siguiente tabla (según datos del National Health Examination Survey). ¿Se



pueden utilizar los métodos de esta sección para probar la aseveración de que los pesos provienen de poblaciones con la misma media? ¿Por qué?

	Primer grado	Segundo grado	Tercer grado	Cuarto grado
Peso total (en lb)	1034	1196	1440	1584

En los ejercicios 5 y 6, identifique los componentes de la prueba de hipótesis.

- 5. Prueba para categorías igualmente probables.** Las siguientes son las frecuencias observadas de tres categorías: 5, 5, 20. Suponga que queremos utilizar un nivel de significancia de 0.05 para probar la aseveración de que las tres categorías son igualmente probables.
- ¿Cuál es la hipótesis nula?
 - ¿Cuál es la frecuencia esperada para cada una de las tres categorías?
 - ¿Cuál es el valor del estadístico de prueba?
 - ¿Cuál es el valor crítico?
 - ¿Qué concluye usted acerca de la aseveración dada?
- 6. Prueba para categorías con proporciones diferentes.** Las siguientes son las frecuencias observadas para cuatro categorías: 5, 10, 10, 20. Suponga que queremos utilizar un nivel de significancia de 0.05 para probar la aseveración de que las cuatro categorías tienen proporciones de 0.20, 0.25, 0.25 y 0.30, respectivamente.
- ¿Cuál es la hipótesis nula?
 - ¿Cuáles son las frecuencias esperadas para las cuatro categorías?
 - ¿Cuál es el valor del estadístico de prueba?
 - ¿Cuál es el valor crítico?
 - ¿Qué concluye usted acerca de la aseveración dada?
- 7. Prueba de balance de rueda de ruleta.** El autor observó 500 giros de una rueda de ruleta en el Mirage Resort and Casino. (Para el sistema fiscal IRS: ¿No es cierto que ahora un viaje a Las Vegas es deducible de impuestos?). Para cada giro, la bola puede detenerse en cualquiera de 38 ranuras diferentes que se supone son igualmente probables. Cuando se utilizó STATDISK para probar la aseveración de que las ranuras son de hecho igualmente probables, se obtuvo el estadístico de prueba $\chi^2 = 38.232$.
- Calcule el valor crítico suponiendo que el nivel de significancia es 0.10.
 - STATDISK produjo un valor P de 0.41331, pero ¿qué sabe usted acerca del valor P si sólo debe utilizar la tabla A-4 junto con el estadístico de prueba dado de 38.232, que resulta de los 500 giros?
 - Escriba una conclusión acerca de la aseveración de que los 38 resultados son igualmente probables.
- 8. Prueba de una máquina tragamonedas.** El autor compró una máquina tragamonedas (Bally modelo 809) y la probó jugando 1197 veces. Al probar la aseveración de que los resultados observados coinciden con las frecuencias esperadas, se obtuvo el estadístico de prueba $\chi^2 = 8.185$. Existen 10 categorías de resultados diferentes, incluyendo no ganar, ganar el premio mayor, ganar con tres campanas, etcétera.
- Calcule el valor crítico suponiendo que el nivel de significancia es de 0.05.
 - ¿Qué concluye usted acerca del valor P de la tabla A-4, si sabe que el estadístico de prueba es $\chi^2 = 8.185$ y que existen 10 categorías?
 - Plantee una conclusión acerca de la aseveración de que los resultados observados coinciden con las frecuencias esperadas. ¿Parece que la máquina tragamonedas del autor funciona correctamente?
- 9. Dado cargado.** El autor taladró un hoyo en un dado, lo rellenó con plomo y procedió a lanzarlo 200 veces. Las siguientes son las frecuencias observadas para los resultados de 1, 2, 3, 4, 5 y 6, respectivamente: 27, 31, 42, 40, 28 y 32. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los resultados no son igualmente probables. ¿Parece que el dado cargado se comporta de forma diferente que un dado legal?

- 10. El neumático desinflado y la clase perdida.** Un cuento clásico se refiere a cuatro estudiantes que van juntos en un automóvil y no llegan a un examen; como excusa, dijeron al profesor que un neumático se desinfló en el camino. En el examen de recuperación, el profesor pidió a los estudiantes que identificaran el neumático en particular que se desinfló. Si en realidad no tuvieron un neumático desinflado, ¿serían capaces de identificar el mismo neumático? El autor pidió a otros 41 estudiantes que identificaran el neumático que ellos seleccionarían. Los resultados están listados en la siguiente tabla (excepto el de un estudiante que seleccionó el neumático de refacción). Utilice un nivel de significancia de 0.05 para probar la aseveración del autor de que los resultados se ajustan a una distribución uniforme. ¿Qué sugiere el resultado acerca de la capacidad de los cuatro estudiantes de seleccionar el mismo neumático cuando en realidad su excusa fue una mentira?

Neumático	Frontal izquierdo	Frontal derecho	Trasero izquierdo	Trasero derecho
Número seleccionado	11	15	8	6

- 11. Muertes por choques de automóviles.** Se seleccionaron al azar muertes por choques de automóviles y los resultados se incluyen en la siguiente tabla (según datos del Insurance Institute for Highway Safety). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las muertes por choques de automóviles ocurren con la misma frecuencia en los diferentes días de la semana. ¿Cómo se explicarían los resultados? ¿Por qué parece haber un número excepcionalmente grande de muertes por choques de automóviles los sábados?

Día	Dom	Lun	Mar	Mié	Jue	Vie	Sáb
Número de muertes	132	98	95	98	105	133	158

Según datos del Insurance Institute for Highway Safety.

- 12. Nacimientos.** Se obtuvieron registros de nacimientos elegidos al azar; los resultados se presentan en la siguiente tabla (según datos del *National Vital Statistics Report*, vol 49, núm. 1). Utilice un nivel de significancia de 0.05 para probar la razonable aseveración de que los nacimientos ocurren con la misma frecuencia los diferentes días de la semana. ¿Cómo se podrían explicar las aparentes bajas frecuencias del sábado y del domingo?

Día	Dom	Lun	Mar	Mié	Jue	Vie	Sáb
Nacimientos	36	55	62	60	60	58	48

- 13. Muertes en motocicleta.** Los datos de las muertes de conductores de motocicleta seleccionadas al azar se resumen en la siguiente tabla (según datos del Insurance Institute of Highway Safety). Utilice un nivel de significancia de 0.05 para probar la aseveración de que este tipo de decesos ocurren con igual frecuencia durante los diferentes meses. ¿Cómo se podrían explicar los resultados?

Mes	Ene	Feb	Mar	Abr	Mayo	Jun	Jul	Ago	Sept	Oct	Nov	Dic
Número	6	8	10	16	22	28	24	28	26	14	10	8

- 14. Calificaciones y lugar para sentarse.** ¿Los estudiantes con calificación “A” (o 10) tienden a sentarse en una zona particular del salón de clases? El autor registró los lugares de los estudiantes que recibieron calificaciones de “A”, con estos resultados: 17 se sentaron al frente, 9 se sentaron en medio y 5 se sentaron en la parte de atrás del salón. ¿Existe suficiente evidencia para sustentar la aseveración de que los estudiantes de calificación “A” no están distribuidos de manera uniforme en la totalidad del salón? Si esto fuera así, ¿significa que usted puede aumentar su probabilidad de obtener una A si se sienta al frente?

- 15. Actrices ganadoras del Óscar.** El autor reunió datos del mes de nacimiento de actrices ganadoras del Óscar. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las actrices ganadoras del Óscar nacen en los distintos meses con la misma frecuencia. ¿Existe alguna razón por la que las actrices ganadoras del Óscar podrían nacer con mayor frecuencia en ciertos meses que en otros?

Mes	Ene	Feb	Mar	Abr	Mayo	Jun	Jul	Ago	Sept	Oct	Nov	Dic
Número	7	3	7	7	8	7	6	6	5	6	9	5

- 16. Actores ganadores del Óscar.** El autor reunió datos del mes de nacimiento de actores ganadores del Óscar. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los actores ganadores del Óscar nacen en los distintos meses con la misma frecuencia. Compare los resultados con los del ejercicio 15.

Mes	Ene	Feb	Mar	Abr	Mayo	Jun	Jul	Ago	Sept	Oct	Nov	Dic
Número	9	5	7	14	8	1	7	6	4	5	1	9

- 17. Novia de junio.** Un organizador de banquetes para bodas selecciona al azar clientes de los últimos años y registra los meses en que se celebraron las recepciones. Los resultados se presentan abajo (según datos de *The Amazing Almanac*). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las bodas se realizan en los diferentes meses con la misma frecuencia. ¿Los resultados sustentan o desmienten la creencia de que la mayoría de las bodas se realizan en junio?

Mes	Ene	Feb	Mar	Abr	Mayo	Jun	Jul	Ago	Sept	Oct	Nov	Dic
Número	5	8	6	8	11	14	10	9	10	12	8	9

- 18. Experimento de color de ojos.** Un investigador desarrolló un modelo teórico para predecir el color de los ojos. Después de examinar una muestra aleatoria de padres, predice el color de ojos de su primer hijo. La siguiente tabla lista el color de ojos de descendientes. Con base en su teoría, el investigador predijo que el 87% de los descendientes tendrían ojos cafés, que el 8% tendría ojos azules y que el 5% tendría ojos verdes. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las frecuencias reales corresponden a la distribución que predijo.

	Ojos cafés	Ojos azules	Ojos verdes
Frecuencia	132	17	0

- 19. Juegos de la Serie Mundial.** El encabezado de *USA Today* “La serie de siete juegos desafía las probabilidades” se refería a la aseveración de que una Serie Mundial de siete juegos ocurre con mayor frecuencia de lo esperado por el azar. A continuación se indican los números de juegos de series mundiales (se omiten dos que duraron ocho juegos) junto con la proporción que se esperaría si los equipos tuvieran la misma capacidad. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las frecuencias observadas coinciden con las proporciones teóricas. Con base en los resultados, ¿al parecer existe evidencia que sustente la aseveración de que una Serie Mundial de siete juegos ocurre con mayor frecuencia de lo esperado?

Juegos	4	5	6	7
Serie Mundial real	18	20	22	37
Proporción esperada	2/16	4/16	5/16	5/16

- 20. Experimento de genética.** Con base en los genotipos de los padres, se espera que sus descendientes tengan genotipos distribuidos de tal forma que el 25% tenga genotipos denotados por AA, que el 50% tenga genotipos denotados por Aa, y que el 25% tenga genotipos denotados por aa. Cuando se seleccionan 145 descendientes, se descubre que 20 de ellos tienen genotipos AA, 90 tienen genotipos Aa y 35 tienen genotipos aa. Ponga a prueba la aseveración de que las frecuencias observadas de los genotipos de los descendientes coinciden con la distribución esperada del 25% para AA, 50% para Aa y 25% para aa. Utilice un nivel de significancia de 0.05.
- 21. Dulces M&M.** Mars Inc. asevera que sus dulces M&M clásicos se distribuyen con los siguientes porcentajes de color: 16% verdes, 20% anaranjados, 14% amarillos, 24% azules, 13% rojos y 13% cafés. Remítase al conjunto de datos 13 del apéndice B y utilice los datos muestrales para probar la aseveración de que la distribución de color es como lo afirma Mars Inc. Utilice un nivel de significancia de 0.05.
- 22. Medición del pulso.** Un ejemplo de esta sección se basó en el principio de que cuando se miden ciertas cantidades, los últimos dígitos tienden a estar distribuidos de manera uniforme, pero que si son estimados o reportados, los últimos dígitos tienden a tener desproporcionadamente más ceros o cincos. Remítase al conjunto de datos 1 del apéndice B y utilice los últimos dígitos de los pulsos de los 80 hombres y mujeres. Estos pulsos se obtuvieron como parte de la National Health Examination Survey. Pruebe la aseveración de que los últimos dígitos de 0, 1, 2,..., 9 se presentan con la misma frecuencia. Con base en los dígitos observados, ¿qué se infiere acerca del procedimiento utilizado para obtener los pulsos?
- 23. Participación en pruebas clínicas según la raza.** Se realizó un estudio para investigar la disparidad racial en pruebas clínicas de cáncer. De los participantes seleccionados al azar, 644 eran caucásicos, 23 eran hispanos, 69 eran afroestadounidenses, 14 eran asiáticos o de las islas del Pacífico, y 2 eran indígenas norteamericanos o nativos de Alaska. Las proporciones de estos grupos en la población estadounidense son 0.757, 0.091, 0.108, 0.038 y 0.007, respectivamente. (Según datos de “Participation in Clinical Trials”, de Murthy, Krumholz y Gross, *Journal of the American Medical Association*, vol. 291, núm. 22). Utilice un nivel de significancia de 0.05 para probar la aseveración de que los participantes se ajustan a la distribución de la población estadounidense. ¿Por qué es importante tener una representación proporcional en este tipo de pruebas clínicas?
- 24. ¿Se ajustan los impactos de las bombas de la Segunda Guerra Mundial a una distribución de Poisson?** En el análisis de los impactos por bombas V-1 en la Segunda Guerra Mundial, el sur de Londres se subdividió en regiones, cada una con una área de 0.25 km^2 . En la sección 5-5 presentamos un ejemplo e incluimos una tabla de frecuencias reales y las frecuencias esperadas de impactos con la distribución de Poisson. Utilice los valores que se listan aquí y pruebe la aseveración de que las frecuencias reales se ajustan a una distribución de Poisson. Utilice un nivel de significancia de 0.05.

Número de impactos de bomba	0	1	2	3	4 o más
Número real de regiones	229	211	93	35	8
Número esperado de regiones (de la distribución de Poisson).	227.5	211.4	97.9	30.5	8.7

- 25. Montos de cheques del autor y la ley de Benford.** La figura 11-6b) ilustra las frecuencias observadas de los dígitos líderes de las cantidades de los últimos 200 cheques que expidió el autor. Las frecuencias observadas de estos dígitos líderes se listan abajo. Con un nivel de significancia de 0.05, pruebe la aseveración de que provienen de una población de dígitos líderes que cumple con la ley de Benford. (Consulte los primeros dos renglones de la tabla 11-1, incluidos en el problema del capítulo).

Dígito líder	1	2	3	4	5	6	7	8	9
Frecuencia	72	23	26	20	21	18	8	8	4

11-2 MÁS ALLÁ DE LO BÁSICO

- 26. Prueba de efectos de valores extremos.** Al realizar una prueba para la bondad de ajuste, como se describe en esta sección, ¿un valor extremo tendrá un gran efecto sobre el valor del estadístico de prueba χ^2 ? Pruebe el efecto de un valor extremo repitiendo el ejercicio 10, después de cambiar las frecuencias para el neumático trasero derecho de 6 a 60. Describa el efecto general de un valor extremo.
- 27. Detección de datos experimentales alterados.** Cuando Gregor Mendel realizó sus famosos experimentos de hibridación con guisantes, aparentemente su asistente de jardinería conocía los resultados que Mendel esperaba y alteró los resultados para ajustarlos a las expectativas de Mendel. Un análisis posterior de los resultados condujo a la conclusión de que existe una probabilidad de sólo 0.00004 de que los resultados esperados y los resultados reportados coincidieran tanto. ¿Cómo pueden utilizarse los métodos de esta sección para detectar resultados como éste, que son demasiado perfectos para ser realistas?
- 28. Prueba equivalente.** En este ejercicio mostraremos que una prueba de hipótesis que implica un experimento binomial con sólo dos categorías es equivalente a una prueba de hipótesis para una proporción (sección 8-3). Suponga que un experimento multinomial en particular tiene sólo dos resultados posibles, A y B, con frecuencias observadas de f_1 y f_2 , respectivamente.
- Calcule una expresión para el estadístico de prueba χ^2 y calcule el valor crítico para un nivel de significancia de 0.05. Suponga que estamos probando la aseveración de que ambas categorías tienen la misma frecuencia, $(f_1 + f_2)/2$.
 - El estadístico de prueba $z = (\hat{p} - p)/\sqrt{pq/n}$ se utiliza para probar la aseveración de que una proporción poblacional es igual a algún valor p . Con la aseveración de que $p = 0.5$, $\alpha = 0.05$ y $\hat{p} = f_1/(f_1 + f_2)$, demuestre que z^2 es equivalente al valor crítico χ^2 [del inciso a)]. Además, demuestre que el cuadrado de la puntuación crítica z es igual al valor crítico χ^2 del inciso a).
- 29. Prueba de bondad de ajuste con una distribución binomial.** La distribución de una frecuencia observada es la siguiente:
- | Número de éxitos | 0 | 1 | 2 | 3 |
|------------------|----|-----|----|----|
| Frecuencia | 89 | 133 | 52 | 26 |
- Suponiendo que una distribución binomial tiene $n = 3$ y $p = 1/3$, utilice la fórmula de la probabilidad binomial para calcular la probabilidad que corresponde a cada categoría de la tabla.
 - Utilice las probabilidades calculadas en el inciso a) para calcular la frecuencia esperada para cada categoría.
 - Utilice un nivel de significancia de 0.05 para probar la aseveración de que las frecuencias observadas se ajustan a una distribución binomial para la que $n = 3$ y $p = 1/3$.
- 30. Prueba de bondad de ajuste con una distribución normal.** La distribución de frecuencias observada de una muestra de puntuaciones de CI es la siguiente:

Puntuación de CI	Menor que				Mayor que
	80	80–95	96–110	111–120	
Frecuencia	20	20	80	40	40

continúa

- Suponga una distribución normal con $\mu = 100$ y $\sigma = 15$ y utilice los métodos del capítulo 6 para calcular la probabilidad de que un sujeto seleccionado al azar pertenezca a cada clase. (Utilice fronteras de clase de 79.5, 95.5, 110.5 y 120.5).
- Utilice las probabilidades calculadas del inciso a) y calcule la frecuencia esperada para cada categoría.
- Utilice un nivel de significancia de 0.01 para probar la aseveración de que las puntuaciones de CI fueron seleccionadas al azar de una población distribuida normalmente con $\mu = 100$ y $\sigma = 15$.

Tablas de contingencia: 11-3 Independencia y homogeneidad

Concepto clave En esta sección estudiaremos *tablas de contingencia* (o *tablas de frecuencias de dos factores*), que incluyen conteos de frecuencia para datos categóricos ordenados en una tabla, con al menos dos renglones y al menos dos columnas. Presentamos un método para poner a prueba la aseveración de que las variables de renglón y de columnas son independientes unas de otras. Utilizaremos el mismo método para una prueba de homogeneidad, en la que ponemos a prueba la aseveración de que distintas poblaciones tienen la misma proporción de algunas características.

Comenzamos con la definición de una tabla de contingencia.



Definición

Una **tabla de contingencia** (o **tabla de frecuencias de dos factores**) es una tabla en la que las frecuencias corresponden a dos variables. (Una variable se utiliza para categorizar renglones, y una segunda variable se utiliza para categorizar columnas).

La tabla 11-4 es un ejemplo de una tabla de contingencia con dos renglones y tres columnas, y los datos en las celdas son conteos de frecuencias. Los datos de la tabla 11-4 corresponden a un estudio retrospectivo (o de casos y controles). La variable de renglón tiene dos categorías: controles y casos. Los sujetos del grupo de control eran motociclistas elegidos al azar en ciertos lugares de la carretera. Los sujetos del grupo de casos eran motociclistas gravemente heridos o fallecidos. La variable de columna se utiliza para el color del casco que usaban. La pregunta importante es la siguiente: ¿El color del casco del motociclista se relaciona de alguna forma con el riesgo de lesiones relacionadas con choques? (Los datos se basan en “Motorcycle Rider Conspicuity and Crash Related Injury: Case-Control Study”, de Wells *et al.*, *BMJ USA*, vol. 4).

Esta sección presenta dos tipos de prueba de hipótesis basadas en tablas de contingencia. Primero consideraremos las pruebas de independencia, las cuales se usan

Tabla 11-4 Estudio de casos y controles de conductores de motocicleta

	Color del caso		
	Negro	Blanco	Amarillo/anaranjado
Controles (sin lesiones)	491	377	31
Casos (lesionados o fallecidos)	213	112	8

para determinar si una variable de renglón de una tabla de contingencia es independiente de su variable de columna. Luego, consideramos las pruebas de homogeneidad, que se utilizan para determinar si situaciones diferentes tienen las mismas proporciones de alguna característica. Ambos tipos de pruebas de hipótesis utilizan los *mismos* métodos básicos. Comenzamos con las pruebas de independencia.

Prueba de independencia

Una de las dos pruebas incluidas en esta sección es la *prueba de independencia* entre la variable de renglón y la variable de columna.

Definición

Una **prueba de independencia** pone a prueba la hipótesis nula de que no existe asociación entre la variable de renglón y la variable de columna en una tabla de contingencia. (Para la hipótesis nula, utilizaremos la afirmación de que “las variables de renglón y de columna son independientes”).

Es muy importante reconocer que, en este contexto, la palabra *contingencia* se refiere a dependencia, pero sólo se trata de una dependencia estadística, y no puede utilizarse para establecer un vínculo directo de causa-efecto entre las dos variables en cuestión. Cuando se prueba la hipótesis nula de independencia entre las variables de renglón y de columna en una tabla de contingencia, los requisitos, el estadístico de prueba y los valores críticos son como se describe en el siguiente cuadro.

Requisitos

1. Los datos muestrales son seleccionados al azar y se representan como conteos de frecuencias en una tabla de dos factores.
2. La hipótesis nula H_0 es la afirmación de que las variables de renglón y de columna son *independientes*; la hipótesis alternativa H_1 es la afirmación de que las variables de renglón y de columna son dependientes.
3. Para cada celda de la tabla de contingencia, la frecuencia *esperada* E es al menos de 5. (No existe el requisito de que cada frecuencia *observada* deba ser al menos de 5. Además, no existe el requisito de que la población deba tener una distribución normal o cualquier otra distribución específica).

Estadístico de prueba para una prueba de independencia

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Valores críticos

1. Los valores críticos se encuentran en la tabla A-4, utilizando

$$\text{grados de libertad} = (r - 1)(c - 1)$$

donde r es el número de renglones y c es el número de columnas.

2. En una prueba de independencia de una tabla de contingencia, la región crítica se localiza *sólo en la cola derecha*.

El estadístico de prueba nos permite medir el grado de discordancia entre las frecuencias observadas en la realidad y aquellas que se esperarían teóricamente cuando las dos variables son independientes. Los valores grandes del estadístico de prueba χ^2 están en la región de la extrema derecha de la distribución chi cuadrada y reflejan diferencias significativas entre las frecuencias observadas y esperadas. En muestreos grandes repetidos, la distribución del estadístico de prueba χ^2 puede aproximarse por medio de la distribución chi cuadrada, siempre y cuando todas las frecuencias esperadas sean al menos de 5. El número de grados de libertad $(r - 1)(c - 1)$ refleja el hecho de que, puesto que conocemos el total de las frecuencias en una tabla de contingencia, podemos asignar con libertad frecuencias a sólo $r - 1$ renglones y $c - 1$ columnas antes de que se determine la frecuencia para cada celda. [Sin embargo, no podemos tener frecuencias negativas o frecuencias tan grandes que la suma de cualquier renglón (o columna) exceda al total de las frecuencias observadas para ese renglón (o columna)].

La frecuencia esperada E para cada celda puede calcularse multiplicando el total de las frecuencias de renglón por el total de las frecuencias de columna, y dividiendo el producto entre el gran total de todas las frecuencias, como se ilustra a continuación.

Frecuencia esperada para una tabla de contingencia

$$\text{frecuencia esperada} = \frac{(\text{total de renglón})(\text{total de columna})}{(\text{gran total})}$$

EJEMPLO Cálculo de la frecuencia esperada Remítase a la tabla 11-4 y calcule la frecuencia esperada para la primera celda, cuando la frecuencia es 491.

SOLUCIÓN La primera celda se ubica en el primer renglón (con un total de 899) y la primera columna (con un total de 704), y la suma de todas las frecuencias en la tabla es 1232.

$$E = \frac{(\text{total por renglón})(\text{total por columna})}{(\text{gran total})} = \frac{(899)(704)}{1232} = 513.714$$

INTERPRETACIÓN Al interpretar este resultado para la primera celda, podemos decir que, aunque 491 motociclistas del grupo de control usaron cascos negros, habríamos esperado que 513.714 de ellos usaran cascos negros si el grupo (de controles o de casos) es independiente del color del casco usado. Existe una discrepancia entre $O = 491$ y $E = 513.714$, y este tipo de discrepancias son componentes fundamentales del estadístico de prueba.

Para comprender mejor las frecuencias esperadas, suponga que conocemos sólo los totales del renglón y de la columna, como en la tabla 11-5, y que debemos llenar la celda de las frecuencias esperadas suponiendo independencia (o ausencia de relación) entre las variables de renglón y de columna. En el primer renglón, 899 de los 1232 sujetos están en el grupo de control, de manera que $P(\text{grupo de control}) = 899/1232$. En la primera columna, 704 de los 1232 conductores usaron cascos negros, de manera que $P(\text{casco negro}) = 704/1232$. Puesto que estamos suponiendo independencia entre el grupo y el color del casco, la regla de la multiplicación para sucesos independientes [$P(A \text{ y } B) = P(A) \cdot P(B)$] se expresa como

$$\begin{aligned} P(\text{grupo de control y casco negro}) &= P(\text{grupo de control}) \cdot P(\text{casco negro}) \\ &= \frac{899}{1232} \cdot \frac{704}{1232} \end{aligned}$$

		Estudio de casos y controles de conductores de motocicleta			Totales de renglón
		Color del caso			
		Negro	Blanco	Amarillo/anaranjado	
Controles					899
Casos					333
Totales de columna	704	489	39		Gran total: 1232



Conociendo la probabilidad de estar en la celda superior izquierda, ahora podemos calcular el *valor esperado* para esa celda, el cual obtenemos multiplicando la probabilidad para esa celda por el número total de sujetos, como en la siguiente ecuación:

$$E = n \cdot p = 1232 \left[\frac{899}{1232} \cdot \frac{704}{1232} \right] = 513.714$$

La forma de este producto sugiere una manera general para obtener la frecuencia esperada de una celda:

$$\text{Frecuencia esperada } E = (\text{gran total}) \cdot \frac{(\text{total de renglón})}{(\text{gran total})} \cdot \frac{(\text{total de columna})}{(\text{gran total})}$$

Esta expresión se simplifica para obtener

$$E = \frac{(\text{total de renglón}) \cdot (\text{total de columna})}{(\text{gran total})}$$

Como ya sabemos calcular los valores esperados, ahora procedemos a utilizar los datos de la tabla de contingencia para probar las hipótesis, como en el siguiente ejemplo.

EJEMPLO Lesiones y color del casco para motocicleta Remítase a los datos de la tabla 11-4. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el grupo (de control o de casos) es independiente del color del casco.

SOLUCIÓN

REQUISITO Como se requiere: los datos se seleccionaron al azar; los datos consisten en conteos de frecuencias en una tabla de dos factores; se está poniendo a prueba la hipótesis nula de que las variables son independientes y las frecuencias esperadas son al menos de 5. (Las frecuencias esperadas son 513.714, 356.827, 28.459, 190.286, 132.173, 10.541). Puesto que todos los requisitos se satisfacen, procedemos a la prueba de hipótesis.

La hipótesis nula y la hipótesis alternativa son las siguientes:

H_0 : El hecho de que el sujeto pertenezca al grupo de control o al grupo de casos es independiente del color del casco. (Esto equivale a decir que las lesiones son independientes del color del casco).

H_1 : El grupo y el color del casco son dependientes.

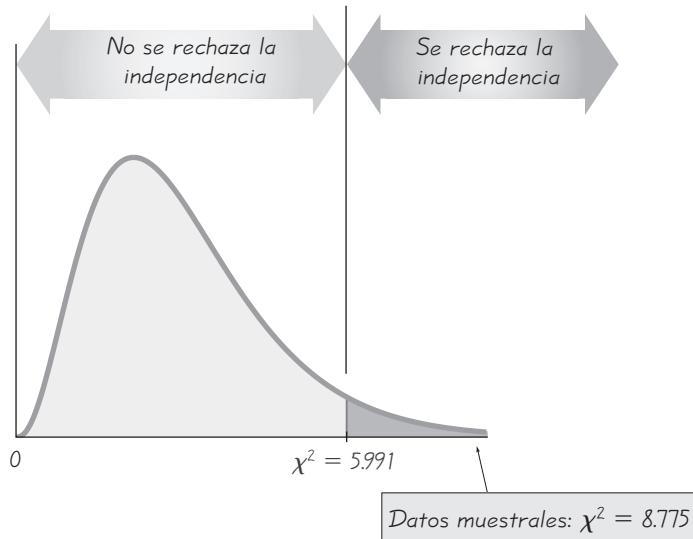
El nivel de significancia es $\alpha = 0.05$.

Un falso positivo durante ocho años

La Associated Press publicó recientemente una nota informativa acerca de Jim Malone, quien había recibido un resultado de prueba positivo para una infección de VIH. Durante ocho años Jim acudió a reuniones de grupos de apoyo, luchó contra la depresión y perdió peso mientras temía morir de SIDA. Finalmente, se le informó que la prueba original era errónea: él no estaba infectado con VIH. Se realizó una prueba de seguimiento después del primer resultado positivo, y la prueba de confirmación demostró que no tenía una infección de VIH, pero nadie le informó al señor Malone el nuevo resultado. Jim Malone sufrió durante ocho años por el resultado de una prueba que en realidad era un falso positivo.

continúa

Figura 11-7
Prueba de independencia para los datos de motociclistas



Puesto que los datos se presentan en una tabla de contingencia, utilizamos la distribución χ^2 con este estadístico de prueba:

$$\begin{aligned}\chi^2 &= \sum \frac{(O - E)^2}{E} = \frac{(491 - 513.714)^2}{513.714} + \dots + \frac{(8 - 10.541)^2}{10.541} \\ &= 8.775\end{aligned}$$

El valor crítico es $\chi^2 = 5.991$ y se encuentra en la tabla A-4, observando que $\alpha = 0.05$ en la cola derecha y que el número de grados de libertad está dado por $(r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$. El estadístico de prueba y el valor crítico se muestran en la figura 11-7. Puesto que el estadístico de prueba está dentro de la región crítica, rechazamos la hipótesis nula de independencia entre el grupo y el color del casco. Al parecer, el color del casco y el grupo (de control o de casos) son dependientes. Puesto que los controles no estaban lesionados y los casos tenían lesiones o habían fallecido, parece que existe una relación entre el color del casco y la seguridad en la motocicleta. Los autores del artículo científico afirmaron que el estudio sustenta la promulgación de leyes que exijan que los motociclistas sean más visibles.

Valores P

En el ejemplo anterior se utilizó el método tradicional de prueba de hipótesis, pero podemos utilizar con facilidad el método del valor P . STATDISK, Minitab, Excel y la calculadora TI-83/84 Plus proporcionan valores P para pruebas de independencia de tablas de contingencia. Si usted no tiene una calculadora o un programa de cómputo adecuados, puede estimar los valores P con la tabla A-4, localizando el estadístico de prueba en el renglón correspondiente al número apropiado de grados de libertad. En el ejemplo anterior, consulte el renglón para 2 grados de libertad y observe que el estadístico de prueba de 8.775 se ubica entre 7.378 y 9.210. Por lo tanto, el valor P debe estar entre 0.025 y 0.01, de manera que concluimos que $0.01 < P < 0.025$. (El valor P real es 0.0124). Sabiendo que el valor P es menor que el nivel de significancia de 0.05, rechazamos la hipótesis nula como lo hicimos en el ejemplo anterior.

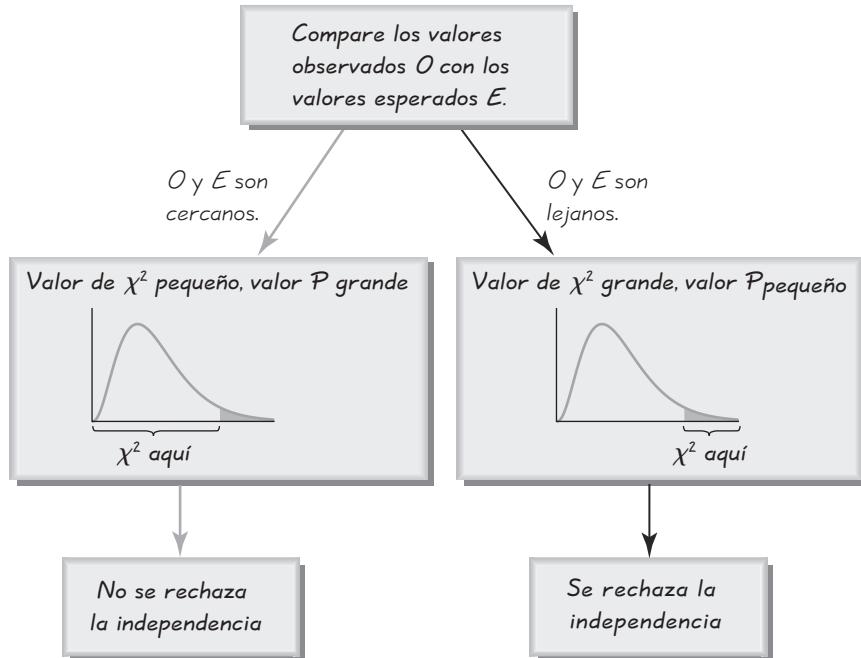
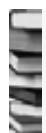


Figura 11-8 Relaciones entre componentes clave en la prueba de independencia

A igual que en la sección 11-2, si las frecuencias observada y esperada son cercanas, el estadístico de prueba χ^2 será pequeño y el valor P será grande. Si las frecuencias observada y esperada se alejan mucho, el estadístico de prueba χ^2 será grande y el valor P será pequeño. Estas relaciones se resumen e ilustran en la figura 11-8.

Prueba de homogeneidad

En el ejemplo anterior, ilustramos una prueba de independencia entre dos variables y utilizamos una población de motociclistas. Sin embargo, algunas otras muestras se obtienen de poblaciones *diferentes*, y tal vez deseemos determinar si esas poblaciones tienen las mismas proporciones de las características en consideración. En estos casos se utiliza la *prueba de homogeneidad*. (La palabra *homogéneo* significa “que tiene la misma calidad” y, en este contexto, estamos haciendo una prueba para determinar si las proporciones son las mismas).



Definición

En una **prueba de homogeneidad** probamos la aseveración de que *poblaciones diferentes* tienen las mismas proporciones de algunas características.

Al hacer una prueba de homogeneidad, podemos utilizar los mismos requisitos, estadístico de prueba, valor crítico y procedimientos descritos en esta sección, con una excepción: en vez de probar la hipótesis nula de independencia entre las variables de renglón y de columna, probamos la hipótesis nula de que las poblaciones diferentes tienen las mismas proporciones de algunas características.



Ventaja del equipo local

En un artículo de la revista *Chance* titulado “Predicting Professional Sports Game Outcomes from Intermediate Game Scores”, los autores Harris Cooper, Kristina DeNeve y Frederick Mosteller utilizaron la estadística para analizar dos creencias comunes: **1.** Los equipos tienen una ventaja cuando juegan en casa y **2.** En realidad sólo cuenta el último cuarto de los partidos profesionales de básquetbol. Utilizando una muestra aleatoria de cientos de partidos, encontraron que, en los cuatro deportes más populares, el equipo local gana aproximadamente el 58.6% de los partidos. Además, los equipos de básquetbol que van ganando después de tres cuartos del juego, ganan aproximadamente cuatro de cada cinco ocasiones, aunque los equipos de béisbol que van ganando después de 7 entradas ganan alrededor de 19 de cada 20 ocasiones. Los métodos de análisis estadístico incluyeron la distribución chi cuadrada aplicada a una tabla de contingencia.

Tabla 11-6 Género y respuestas de encuesta

		Género del entrevistador	
		Hombre	Mujer
Hombres que están de acuerdo		560	308
Hombres que están en desacuerdo		240	92

EJEMPLO Influencia del género ¿Tiene efecto el género del encuestador en las respuestas de encuesta de varones? En un artículo del *U.S. News & World Report* acerca de encuestas se afirmó que “en temas sensibles, las personas tienden a dar respuestas ‘aceptables’ en vez de respuestas honestas; sus respuestas podrían depender del género o el origen étnico del entrevistador”. Para sustentar esta aseveración, el Eagleton Institute proporcionó los datos de una encuesta en la cual se preguntó a hombres si estaban de acuerdo con esta afirmación: “El aborto es un asunto privado que la mujer debe decidir sin intervención gubernamental”. Analizaremos el efecto del género sólo en los hombres encuestados. La tabla 11-6 está basada en estas respuestas de hombres encuestados. Suponga que la encuesta se diseñó de manera que los entrevistadores varones recibieron instrucciones para obtener 800 respuestas de sujetos varones, y las entrevistadoras mujeres recibieron instrucciones para obtener 400 respuestas de sujetos varones. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las proporciones de las respuestas de acuerdo/en desacuerdo son las mismas para los sujetos entrevistados por hombres y los sujetos entrevistados por mujeres.

SOLUCIÓN

REQUISITO ✓ Los datos consisten en conteos de frecuencias independientes; cada observación se puede categorizar de acuerdo con dos variables; y las frecuencias esperadas (que en la tabla de resultados de Minitab aparecen como 578.67, 289.33, 221.33 y 110.67) son al menos de 5. [Las dos variables son: **1.** género del entrevistador, y **2.** si el sujeto estuvo de acuerdo o en desacuerdo]. Puesto que se trata de una prueba de homogeneidad, ponemos a prueba la aseveración de que las proporciones de respuestas de acuerdo/en desacuerdo son iguales para los sujetos entrevistados por hombres y para los sujetos entrevistados por mujeres. Todos los requisitos se satisfacen, así que procedemos con la prueba de hipótesis. ✓

Puesto que tenemos dos poblaciones separadas (sujetos entrevistados por hombres y sujetos entrevistados por mujeres), probamos la homogeneidad con estas hipótesis:

H_0 : Las proporciones de las respuestas acuerdo/en desacuerdo son iguales para los sujetos entrevistados por hombres y los sujetos entrevistados por mujeres.

H_1 : Las proporciones son diferentes.

El nivel de significancia es $\alpha = 0.05$. Utilizamos el mismo estadístico de prueba χ^2 descrito antes, el cual calculamos por medio del mismo procedimiento. En vez de hacer una lista de los detalles de este cálculo, presentamos la pantalla de Minitab que resulta de los datos de la tabla 11-6.

Minitab

Expected counts are printed below observed counts
 Chi-Square contributions are printed below expected counts

	C1	C2	Total
1	560	300	860
	578.67	289.33	
	0.602	1.204	
2	240	92	332
	221.33	110.67	
	1.574	3.149	
Total	800	400	1200

Chi-Sq = 6.529, DF = 1, P-Value = 0.011

La pantalla de Minitab indica las frecuencias esperadas de 578.67, 289.33, 221.33 y 110.67. Los resultados también incluyen el estadístico de prueba $\chi^2 = 6.529$ y el valor P de 0.011. Utilizando el método del valor P para probar la hipótesis, rechazamos la hipótesis nula de proporciones iguales (homogéneas) (puesto que el valor P de 0.011 es menor que 0.05). Existe suficiente evidencia para sustentar el rechazo de la aseveración de que las proporciones son iguales. Parece que la respuesta y el género del entrevistador son dependientes. Aunque este análisis estadístico no puede utilizarse para justificar ninguna afirmación acerca de la causalidad, parece que los hombres se ven influidos por el género del entrevistador.

EJEMPLO Lanzamiento y giro de monedas de un centavo Cuando se lanza o se hace girar una moneda de un centavo de dólar, ¿existe la misma probabilidad de obtener caras? Utilice los datos de la tabla 11-7, con un nivel de significancia de 0.05, para probar la aseveración de que la proporción de caras es la misma al lanzar y al girar la moneda. (Los datos provienen de resultados experimentales presentados en *Chance News*).

SOLUCIÓN

REQUISITOS ✓ Como se requiere, los datos son aleatorios y consisten en conteos de frecuencia en una tabla de dos factores. Aquí estamos probando la hipótesis nula de que la proporción de caras con lanzamientos es igual a la proporción de caras con giros. Las frecuencias esperadas son al menos de 5. (Las frecuencias esperadas son 2007.291, 2032.709, 993.709 y 1006.291). Puesto que todos los requisitos se satisfacen, procedemos con la prueba de hipótesis. ✓

Puesto que tenemos dos poblaciones separadas (en un experimento se lanzaron monedas y en un experimento diferente se hicieron girar monedas), deseamos hacer una prueba de homogeneidad con las siguientes hipótesis:

H_0 : La proporción de caras es igual con lanzamientos que con giros.

H_1 : Las proporciones son diferentes.

El nivel de significancia es $\alpha = 0.05$. Utilizamos el mismo estadístico de prueba χ^2 descrito antes, el cual calculamos por medio del mismo procedimiento.



El medio de encuesta puede afectar los resultados

En una encuesta de católicos en Boston, se preguntó a los sujetos si los anticonceptivos debían ponerse a la disposición de mujeres solteras. En entrevistas personales, el 44% de los encuestados dijo que sí. Sin embargo, en un grupo similar encuestado por correo o por teléfono, el 75% de los individuos respondió afirmativamente a la misma pregunta.

TABLA 11-7
Experimento con monedas

	Caras	Cruces
Lanzamiento	2048	1992
Rotación	953	1047

ccontinúa

En vez de hacer una lista de los detalles de este cálculo, presentamos la pantalla de Minitab que resulta de los datos de la tabla 11-7.

Minitab

Chi-Square contributions are printed below expected counts

	Heads	Tails	Total
1	2048	1992	4040
	2007.29	2032.71	
	0.826	0.815	
2	950	1047	2000
	993.71	1006.29	
	1.668	1.647	
Total	3001	3039	6040

Chi-Sq = 4.955, DF = 1, P-Value = 0.026

La pantalla de Minitab indica las frecuencias esperadas de 2007.29, 2032.71, 993.71 y 1006.29. Los resultados incluyen también el estadístico de prueba $\chi^2 = 4.955$ y el valor P de 0.026. Utilizando el método del valor P para prueba de hipótesis, rechazamos la hipótesis nula de proporciones iguales (homogéneas) (puesto que el valor P de 0.026 es menor que 0.05). Existe suficiente evidencia para sustentar el rechazo de la aseveración de que las proporciones son iguales. Parece que el lanzamiento de una moneda de un centavo y el giro de una moneda de un centavo produce proporciones diferentes de caras.

Prueba exacta de Fisher

Para el análisis de tablas 2×2 , incluimos el requisito de que cada celda debe tener una frecuencia esperada de 5 o más. Este requisito es necesario para que la distribución χ^2 sea una aproximación adecuada para la distribución exacta del estadístico de prueba $\sum \frac{(O - E)^2}{E}$. En consecuencia, si una tabla 2×2 tiene una celda con una frecuencia esperada menor que 5, los procedimientos anteriores no deben utilizarse, porque la distribución no es una aproximación adecuada. A menudo se utiliza la *prueba exacta de Fisher* para una tabla 2×2 de este tipo, ya que proporciona un valor P exacto y no requiere de una técnica de aproximación.

Considere los datos de la tabla 11-8, en la cual las frecuencias esperadas aparecen en paréntesis, debajo de las frecuencias observadas. La primera celda tiene una frecuencia esperada menor que 5, por lo que no debemos utilizar los métodos

Tabla 11-8 Cascos y lesiones faciales en accidentes de bicicleta (las frecuencias esperadas están en paréntesis)

	Con casco	Sin casco
Lesiones faciales recibidas	2 (3)	13 (12)
Todas las lesiones no faciales	6 (5)	19 (20)

anteriores. Con la prueba exacta de Fisher, calculamos la probabilidad de obtener por azar los resultados observados (suponiendo que el uso de un casco y ser víctima de lesiones faciales son sucesos independientes), y también calculamos la probabilidad de cualquier resultado que sea *más extremo*. (El hecho de referirnos a resultados “más extremos” podría generar confusión, por lo que sería útil repasar el apartado de la sección 5-2 “Uso de las probabilidades para determinar resultados infrecuentes”). Cuando se pone a prueba la hipótesis nula de independencia entre el uso de casco y ser víctima de lesiones faciales, las frecuencias de 2, 13, 6 y 19 pueden reemplazarse por 1, 14, 7 y 18, respectivamente, para obtener resultados *más extremos* con los mismos totales de renglón y de columna. (En ocasiones se critica la prueba exacta de Fisher porque el uso de totales fijos de renglón y de columna a menudo es poco realista). La prueba exacta de Fisher requiere que calculemos las probabilidades de las frecuencias observadas y cada conjunto de frecuencias más extremas. Luego, esas probabilidades se suman para obtener un valor *P* exacto.

Como por lo general los cálculos son muy complejos, es mejor utilizar un programa de cómputo. Para los datos de la tabla 11-8, la prueba exacta de Fisher por medio de STATDISK, SPSS, SAS y Minitab produjo un valor *P* exacto de 0.686. Puesto que este valor *P* exacto no es pequeño (como menor que 0.05), no rechazamos la hipótesis nula de que el uso del casco y ser víctima de lesiones faciales son sucesos independientes.

Datos apareados Además del requisito de que cada celda debe tener una frecuencia esperada al menos de 5, los métodos de esta sección también requieren que las observaciones individuales sean *independientes*. Si una tabla 2×2 consiste en conteos de frecuencias que provienen de *datos apareados*, no contamos con la independencia requerida. En estos casos, podemos utilizar la prueba de McNemar, que se explica en la siguiente sección.

Uso de la tecnología

STATDISK Primero registre las frecuencias observadas en columnas de la ventana de datos. Seleccione **Analysis** de la barra del menú principal, luego seleccione **Contingency Tables** y proceda a identificar las columnas que contienen las frecuencias. Haga clic en **Evaluate**. El resultado de STATDISK incluye el estadístico de prueba, el valor crítico, el valor *P* y la conclusión, tal como se muestra en los resultados de la tabla 11-4.

STATDISK

```
Degrees of freedom: 2
Test Statistic,  $\chi^2$ : 8.7747
Critical  $\chi^2$ : 5.991471
P-Value: 0.0124

Reject the Null Hypothesis
Data provides evidence that the rows and
columns are related
```

MINITAB Primero ingrese las frecuencias observadas en columnas, luego seleccione **Stat** de la barra del menú principal. Después seleccione la opción **Tables**, luego seleccione **Chi Square Test** y proceda a ingresar los nombres de las columnas que contienen las frecuencias observadas, como son C1 C2 C3 C4. Minitab proporciona el estadístico de prueba y el valor *P*.

TI-83/84 PLUS Primero ingrese la tabla de contingencia como una *matriz* presionando **2nd x⁻¹** para obtener el menú **MATRIX** (o el botón **MATRIX** del teclado de la TI-83/84). Seleccione **EDIT** y presione **ENTER**. Ingrese las dimensiones de la matriz (renglones por columnas) y proceda a ingresar las frecuencias individuales. Cuando termine, oprima **STAT**, seleccione **TESTS**, luego seleccione la opción χ^2 -Test. Asegúrese de haber ingresado la matriz observada, como la matriz A. Las frecuen-

cias esperadas se calcularán automáticamente y se guardarán en la matriz separada identificada como esperada (“expected”). Descienda con el cursor hasta **Calculate** y oprima **ENTER** para obtener el estadístico de prueba, el valor *P* y el número de grados de libertad.

EXCEL

Usted debe ingresar las frecuencias observadas y también debe determinar e ingresar las frecuencias esperadas. Cuando termine, haga clic en el ícono **fx** en la barra del menú, seleccione la categoría de función **Statistical** y luego seleccione el nombre de la función **CHITEST**. Usted debe ingresar el rango de valores para las frecuencias observadas y el rango de valores para las frecuencias esperadas. Sólo aparecerá el valor *P*. (También se puede utilizar DDXL seleccionando **Tables**, luego **Indep. Test for Summ Data**).

11-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Estadístico de prueba chi cuadrada.** Con sus propias palabras, describa lo que mide el estadístico de prueba chi cuadrada, tal como se utilizó en esta sección.
- Prueba de cola derecha.** ¿Por qué las pruebas de hipótesis descritas en esta sección siempre son de cola derecha?
- Contingencia.** ¿Qué significa la palabra “contingencia” en el contexto de esta sección?
- Causalidad.** Suponga que rechazamos la hipótesis nula de independencia entre la variable de renglón de si un sujeto fuma y la variable de columna de si el sujeto puede pasar una prueba estándar de resistencia física. ¿Podemos concluir que el tabaquismo causa que la gente no pase esa prueba? ¿Por qué?

En los ejercicios 5 y 6, pruebe la aseveración dada utilizando los resultados del programa de cómputo.

- ¿Existe discriminación racial?** La *discriminación racial* es la práctica polémica de señalar que alguien tiene una conducta criminal con base en su raza, país de origen u origen étnico. La tabla adjunta resume resultados de conductores seleccionados al azar, detenidos por la policía en un año reciente (según datos del Departamento de Justicia de Estados Unidos, *Bureau of Justice Statistics*). El uso de los datos de esta tabla dio como resultado una pantalla de Minitab. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el hecho de ser detenido es independiente de la raza y del origen étnico. Con base en la evidencia disponible, ¿podemos concluir que hay discriminación racial?

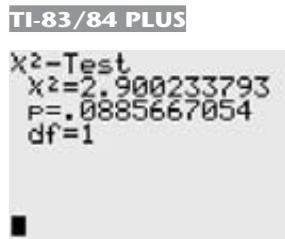
		Raza y grupo étnico	
		Afroestadounidenses y no hispanos	Caucásicos y no hispanos
Detenidos por la policía		24	147
No detenidos por la policía		176	1253

Minitab

Chi-Sq = 0.413, DF = 1, P-Value = 0.521

- No fumar.** La tabla adjunta resume éxitos y fracasos de sujetos que utilizaron diferentes métodos para tratar de dejar de fumar. Cinco meses después de comenzar el tratamiento, se determinó si los sujetos fumaban o no fumaban; los datos están basados en resultados de los Centros para el Control y Prevención de Enfermedades. Utilice los resultados de la calculadora TI-83/84 Plus (de la siguiente página) con un nivel de significancia de 0.05 para probar la aseveración de que el éxito es independiente del método utilizado. Si alguien quiere dejar de fumar, ¿la elección del método provoca una diferencia?

		Goma de mascar de nicotina	Parche de nicotina
Fuman		191	263
No fuman		59	57



- 7. ¿La vacuna es efectiva?** En un artículo de *USA Today*, sobre una vacuna experimental para niños, se publicó la siguiente aseveración: “En una prueba con 1602 niños, sólo 14 (el 1%) de los 1070 que recibieron la vacuna desarrollaron gripe, comparados con 95 (el 18%) de los 532 que recibieron placebo”. Los datos se incluyen en la siguiente tabla. Utilice un nivel de significancia de 0.05 para hacer una prueba de independencia entre la variable de tratamiento (vacuna o placebo) y la variable que representa la gripe (desarrolló gripe, no desarrolló gripe). ¿Parece que la vacuna es efectiva?

		¿Desarrolló gripe?	
		Sí	No
Tratamiento con la vacuna	Sí	14	1056
	No	95	437

- 8. Muertes de peatones.** Se realizó un estudio de la relación entre la intoxicación y la muerte de peatones; los resultados se presentan en la siguiente tabla (según datos de la National Highway Traffic Safety Administration). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las muertes de peatones son independientes de la intoxicación de los conductores y de la intoxicación de los peatones.

	Peatón intoxicado	Peatón no intoxicado
Conductor intoxicado	59	79
Conductor no intoxicado	266	581

- 9. Zurdos y género.** La siguiente tabla se basa en datos de una encuesta de Scripps Survey Research Center. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el género y la dominancia de la mano izquierda son independientes.

	Zurdos	No zurdos
Hombre	83	17
Mujer	184	16

- 10. Peso al nacer y graduación.** Los datos de la siguiente tabla se basan en datos de un artículo de la revista *Time*. Utilice en un nivel de significancia de 0.05 para probar la aseveración de que el hecho de que un sujeto haya nacido con bajo peso o con peso normal es independiente de que el sujeto se gradúe de la preparatoria a los 19 años. ¿Los resultados indican que nacer con bajo peso causa que la gente no se haya graduado de preparatoria a los 19 años?

	Graduado de preparatoria a los 19 años	No graduado de preparatoria a los 19 años
Bajo peso al nacer	8	42
Peso normal al nacer	86	64

- 11. Exactitud de pruebas de polígrafo.** Los datos en la tabla adjunta resumen resultados de pruebas de exactitud de polígrafos (según datos de la Oficina de Evaluación Tecnológica, *Office of Technology Assessment*). Utilice un nivel de significancia de 0.05 para probar la aseveración de que el hecho de que el sujeto minta es independiente de la indicación del polígrafo. ¿Qué sugieren los resultados acerca de la eficacia de los polígrafos?

	El polígrafo indicó verdad	El polígrafo indicó mentira
El sujeto realmente dijo la verdad	65	15
El sujeto realmente dijo una mentira	3	17

- 12. ¿Los perros pueden detectar el cáncer?** Se llevó a cabo un experimento para poner a prueba la habilidad de los perros para detectar cáncer de vejiga. Los perros fueron entrenados con muestras de orina de pacientes con cáncer de vejiga y de personas en un grupo de control que no tenían cáncer de vejiga. Los resultados se presentan en la siguiente tabla (según datos del *New York Times*). Utilice un nivel de significancia de 0.01 para probar la aseveración de que la fuente de la muestra (sujeto saludable o sujeto con cáncer de vejiga) es independiente de las selecciones del perro. ¿Qué sugieren los resultados acerca de la habilidad de los perros para detectar el cáncer de vejiga? Si los perros tuvieron más aciertos que lo esperado por el azar, ¿su desempeño fue lo suficiente bueno para justificar su uso en diagnósticos exactos?

	Muestra de sujetos con cáncer de vejiga	Muestra de sujetos sin cáncer de vejiga
El perro identificó al sujeto como canceroso	22	32
El perro no identificó al sujeto como canceroso	32	282

- 13. ¿La sentencia de un acusado depende de su declaración?** Muchas personas creen que los criminales que se declaran culpables tienden a obtener sentencias más cortas que aquellos que son sentenciados en un juicio. La tabla adjunta resume datos muestrales, seleccionados al azar, de casos de acusados de robo en San Francisco. Todos los sujetos tenían sentencias previas. Con un nivel de significancia de 0.05, pruebe la aseveración de que la sentencia (enviar a prisión o no enviar a prisión) es independiente de la declaración de culpabilidad. Si usted fuera el abogado defensor de un acusado culpable, ¿sugieren estos resultados que usted debe fomentar una declaración de culpabilidad?

	Declaración de culpabilidad	Declaración de inocencia
Enviados a prisión	392	58
No enviados a prisión	564	14

Según datos de “Does It Play to Plead Guilty? Differential Sentencing and the Functioning on the Criminal Courts”, de Brereton y Casper, *Law and Society Review*, vol. 16, núm. 1.

- 14. ¿Cuál tratamiento es mejor?** Se diseñó una prueba controlada y aleatorizada para comparar la eficacia del entablillado y de la cirugía en el tratamiento del síndrome del túnel carpiano. Los resultados se presentan en la siguiente tabla (según datos de “Splinting vs. Surgery in the Treatment of Carpal Tunnel Syndrome”, de Gerristen *et al.*, *Journal of the American Medical Association*, vol. 288, núm. 10). Los resultados están basados en evaluaciones efectuadas un año después del tratamiento. Con un nivel de significancia de 0.01, pruebe la aseveración de que el éxito es independiente del tipo de tratamiento. ¿Qué sugieren los resultados acerca del tratamiento del síndrome del túnel carpiano?

	Tratamiento exitoso	Tratamiento sin éxito
Tratamiento con entablillado	60	23
Tratamiento con cirugía	67	6

- 15. Lanzamiento y giro de monedas de un centavo.** Cuando se lanza o se hace girar una moneda de un centavo de dólar, ¿existe la misma probabilidad de obtener caras? Utilice los datos de la siguiente tabla, con un nivel de significancia de 0.05, para probar la aseveración de que la proporción de caras es igual en un lanzamiento que en un giro. (Los datos provienen de resultados experimentales del profesor Robin Lock, tal como aparecieron en *Chance News*).

	Caras	Cruces
Lanzamiento	14,709	14,306
Giro	9197	11,225

- 16. Prueba de la influencia del género.** En la tabla 11-6 se resumen los datos de hombres encuestados, pero la siguiente tabla resume los datos de una muestra de mujeres (según datos de una encuesta del Eagleton Institute). Utilice un nivel de significancia de 0.01 y, suponiendo que los tamaños muestrales de 800 hombres y 400 mujeres están predeterminados, pruebe la aseveración de que la proporción de respuestas de acuerdo/en desacuerdo es igual para los sujetos entrevistados por hombres y para los sujetos entrevistados por mujeres. Al parecer, ¿el género del entrevistador afectó las respuestas de las mujeres?

	Género del entrevistador	
	Hombre	Mujer
Mujeres que están de acuerdo	512	336
Mujeres que están en desacuerdo	288	64

- 17. Riesgos de trabajo.** Utilice los datos en la tabla para probar la aseveración de que la ocupación es independiente de que la causa de muerte sea un homicidio. La tabla está basada en datos del Departamento del Trabajo de Estados Unidos, *Bureau of Labor Statistics*. Al parecer, ¿una ocupación en particular es más proclive a los homicidios? De ser así, ¿cuál es?

	Policías	Cajeros	Taxistas	Guardias
Homicidio	82	107	70	59
Otra causa de muerte que no es homicidio	92	9	29	42

- 18. ¿La exactitud del escáner es la misma para las ofertas?** En un estudio de sistemas de cobro por escáner en almacenes, se utilizaron muestras de compras para comparar las lecturas por escáner de los precios con los precios etiquetados. La tabla adjunta resume resultados de una muestra de 819 artículos. Cuando los almacenes utilizan escáner para cobrar los artículos, ¿las tasas de error son las mismas para los artículos con precio normal que para los artículos en oferta? ¿Cómo podría cambiar la conducta de los consumidores si creen que ocurren desproporcionadamente más cobros excesivos en los artículos en oferta?

	Artículos con precio normal	Artículos en oferta
Cobros de menos	20	7
Cobros de más	15	29
Precio correcto	384	364

Según datos de “UPC Scanner Pricing Systems: Are They Accurate?”, de Ronald Goodstein, *Journal of Marketing*, vol. 58.

- 19. ¿Es independiente el uso de cinturón de seguridad del tabaquismo?** Un estudio de usuarios y no usuarios del cinturón de seguridad produjo los datos seleccionados al azar que se resumen en la siguiente tabla. Pruebe la aseveración de que la cantidad de

cigarrillos fumados es independiente del uso del cinturón de seguridad. Una hipótesis plausible es que la gente que fuma mucho está menos preocupada por su salud y su seguridad y, por lo tanto, es menos propensa a utilizar el cinturón de seguridad. ¿Sustentan los datos muestrales esta hipótesis?

		Número de cigarrillos fumados por día			
		0	1–14	15–34	35 o más
Utilizan cinturones de seguridad	175	20	42	6	
	149	17	41	9	

Según datos de “What Kinds of People Do Not Use Seat Belts?”, de Helsing y Comstock, *American Journal of Public Health*, vol. 67, núm. 11.

- 20. ¿La ventaja de ser un equipo local es independiente del deporte?** Se reunieron datos del equipo ganador en diferentes deportes, con los resultados que se presentan en la tabla adjunta. Utilice un nivel de significancia de 0.10 para probar la aseveración de que los triunfos de equipos locales y de visitantes son independientes del tipo de deporte. De los cuatro deportes incluidos aquí, el béisbol es el único en el cual el equipo local puede modificar las dimensiones del campo a favor de sus propios jugadores. ¿Parece que los equipos de béisbol son eficientes al utilizar esta ventaja?

		Básquetbol	Béisbol	Hockey	Fútbol
Triunfos del equipo local	127	53	50	57	
	71	47	43	42	

Según datos de “Predicting Professional Sports Game Outcomes from Intermediate Game Scores”, de Cooper, DeNeve y Mosteller, *Chance*, vol. 5, núm. 3-4.

- 21. Lesiones y color del casco de motociclistas.** Un ejemplo de esta sección incluyó datos de un estudio de casos y controles acerca de las lesiones y el color del casco de los motociclistas. Utilice los datos adicionales que se incluyen en la siguiente tabla y pruebe la aseveración de que las lesiones son independientes del color del casco. ¿Estos datos conducen a la misma conclusión que los datos del ejemplo de esta sección?

		Color del casco				
		Negro	Blanco	Amarillo/anaranjado	Rojo	Azul
Controles (sin lesiones)	491	377	31	170	55	
	213	112	8	70	26	

- 22. Rechazo de encuestas y grupo de edad.** Un estudio de personas que se rehusaron a responder preguntas de encuestas arrojó los datos muestrales seleccionados al azar que aparecen en la siguiente tabla. Con un nivel de significancia de 0.01, pruebe la aseveración de que la cooperación del sujeto (respuesta o negativa) es independiente del grupo de edad. ¿Parece que un grupo de edad se niega especialmente a cooperar?

		Edad					
		18–21	22–29	30–39	40–49	50–59	60 o mayores
Respondieron	73	255	245	136	138	202	
	11	20	33	16	27	49	

Según datos de “I Hear You Knocking but You Can’t Come In”, de Fitzgerald y Fuller, *Sociological Methods and Research*, vol. 11, núm. 1.

11-3 MÁS ALLÁ DE LO BÁSICO

- 23. Uso de la corrección de Yates por continuidad.** La distribución chi cuadrada es continua, mientras que el estadístico que se utilizó en esta sección es discreto. Algunos estadísticos utilizan la *corrección por continuidad de Yates* en celdas con una frecuencia esperada menor de 10 o en todas las celdas de una tabla de contingencia con dos renglones y dos columnas. Con la corrección de Yates, reemplazamos

$$\sum \frac{(O - E)^2}{E} \text{ con } \sum \frac{(|O - E| - 0.5)^2}{E}$$

Dada la tabla de contingencia del ejercicio 5, calcule el valor del estadístico de prueba χ^2 con y sin la corrección de Yates. ¿Qué efecto tiene la corrección de Yates?

- 24. Pruebas equivalentes.** Suponga que una tabla de contingencia tiene dos renglones y dos columnas con las frecuencias de a y b en el primer renglón y las frecuencias de c y d en el segundo renglón.

a) Verifique que el estadístico de prueba pueda expresarse como

$$\chi^2 = \frac{(a + b + c + d)(ad - bc)^2}{(a + b)(c + d)(b + d)(a + c)}$$

b) Permita que $\hat{p}_1 = a/(a + c)$ y permita que $\hat{p}_2 = b/(b + d)$. Demuestre que el estadístico de prueba

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

donde

$$\bar{p} = \frac{a + b}{a + b + c + d}$$

y

$$\bar{q} = 1 - \bar{p}$$

es tal que $z^2 = \chi^2$ [el mismo resultado del inciso a)]. Este resultado indica que la prueba chi cuadrada que implica una tabla de 2×2 es equivalente a la prueba para la diferencia entre dos proporciones, como se describe en la sección 9-2.

11-4 Prueba de McNemar para datos apareados

Concepto clave Los procedimientos de la sección 11-3 para la tabla de contingencia están basados en datos *independientes*. Las tablas de contingencia 2×2 que consisten en conteos de frecuencias que resultan de *datos apareados* no tienen independencia y, en tales casos, podemos utilizar la prueba de McNemar para datos apareados. Probaremos la hipótesis nula de que las frecuencias de las categorías discordantes (diferentes) ocurren en la misma proporción.

Suponga que varios sujetos de prueba padecen pie de atleta en los dos pies, y que cada sujeto recibe un tratamiento X en un pie y un tratamiento Y en el otro pie. La tabla 11-9 resume de manera general los conteos de frecuencias que resultan de los datos apareados de los pies con los dos tratamientos diferentes. Si $\alpha = 12$ en la tabla 11-9, entonces 12 sujetos fueron curados en los dos pies. Si $b = 8$ en la tabla 11-9, entonces uno de los pies de 8 sujetos no se curó con el tratamiento X, mientras que el otro pie se curó con el tratamiento Y. *Importante:* Observe que los datos de la tabla 11-9 son conteos de frecuencias de *personas*, no de pies.

		Tratamiento X	
		Curados	No curados
Tratamiento Y	Curados	<i>a</i>	<i>b</i>
	No curados	<i>c</i>	<i>d</i>

Puesto de los conteos de frecuencias de la tabla 11-9 resultan de los datos *apareados* de pies, los datos no son independientes y no podemos utilizar los procedimientos de la sección 11-3 para tablas de contingencia. En vez de ello, usamos la prueba de McNemar.

Definición

La **prueba de McNemar** utiliza conteos de frecuencias de *datos apareados* nominales de dos categorías para probar la hipótesis nula de que, para una tabla como la 11-9, las frecuencias *b* y *c* ocurren en la misma proporción.

Requisitos

1. Los datos muestrales se eligieron al azar.
2. Los datos muestrales consisten en *datos apareados* o conteos de frecuencias.
3. Los datos tienen un nivel de medición nominal, y cada observación se puede clasificar de dos maneras: 1) según la categoría que distingue valores con cada dato apareado (como pie izquierdo y pie derecho) y 2) según otra categoría con dos valores posibles (como curado y no curado).
4. Para tablas como la 11-9, las frecuencias son tales que $b + c \geq 10$.

Estadístico de prueba (para poner a prueba la hipótesis nula de que, para tablas como la 11-9, las frecuencias *b* y *c* ocurren en la misma proporción):

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

donde las frecuencias de *b* y *c* se obtienen de la tabla 2×2 , con un formato similar al de la tabla 11-9. (Las frecuencias *b* y *c* deben provenir de pares “discordantes”, como se describe más adelante en esta sección).

Valores críticos

1. La región crítica se localiza *únicamente en la cola derecha*.
2. Los valores críticos se encuentran en la tabla A-4, utilizando **grados de libertad = 1**.

EJEMPLO Comparación de tratamientos Se utilizan dos cremas diferentes para tratar el pie de atleta. A cada sujeto con esta infección micótica en ambos pies se le trata un pie con Pedacream y el otro con Fungacream. Los resultados muestrales se presentan en la tabla 11-10. Utilice un nivel de significancia de 0.05 y aplique la prueba de McNemar para poner a prueba la hipótesis nula de que las siguientes proporciones son iguales:

- La proporción de sujetos cuyo pie tratado con Pedacream no se cura y cuyo pie tratado con Fungacream sí se cura.
- La proporción de sujetos cuyo pie tratado con Pedacream sí se cura y cuyo pie tratado con Fungacream no se cura.

Con base en los resultados, ¿parece que hay una diferencia entre los dos tratamientos? ¿Un tratamiento es mejor que otro?

SOLUCIÓN

REQUISITO ✓ Los datos son pares de conteos de frecuencias, a partir de sujetos elegidos al azar, y cada observación se puede categorizar de acuerdo con dos variables. (Una variable tiene los valores de “Pedacream” y “Fungacream”, y la otra variable tiene los valores de “curado” y “no curado”). Además, para tablas como la 11-9, las frecuencias deben ser tales que $b + c \geq 10$. Para la tabla 11-10, $b = 8$ y $c = 40$, de manera que $b + c = 48$, que es al menos 10. Por lo tanto, todos los requisitos se satisfacen. Aunque la tabla 11-10 podría parecer una tabla de contingencia 2×2 , no podemos usar los procedimientos de la sección 11-3 porque se trata de *datos apareados* (en vez de ser independientes). En su lugar, empleamos la prueba de McNemar. ✓

Después de comparar los conteos de frecuencia de la tabla 11-9 con los de la tabla 11-10, observamos que $b = 8$ y $c = 40$, y el estadístico de prueba se calcula de la siguiente manera:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c} = \frac{(|8 - 40| - 1)^2}{8 + 40} = 20.021$$

Con un nivel de significancia de 0.05 y grados de libertad dados por $gl = 1$, nos remitimos a la tabla A-4 para encontrar el valor crítico de $\chi^2 = 3.841$ para esta prueba de cola derecha. El estadístico de prueba $\chi^2 = 20.021$ excede al va-

continued

Tabla 11-10 Pruebas clínicas de tratamientos para el pie de atleta

		Tratamiento con Pedacream	
		Curados	No curados
Tratamiento con Fungacream	Curados	12	8
	No curados	40	20

80 sujetos tratados en 160 pies:
 12 se les curaron ambos pies.
 20 no se les curó ningún pie.
 8 se curaron con Fungacream, pero no con Pedacream.
 40 se curaron con Pedacream, pero no con Fungacream.

lor crítico $\chi^2 = 3.841$, de manera que rechazamos la hipótesis nula. Parece que las dos cremas producen resultados diferentes. Al analizar las frecuencias de 8 y 40, vemos que Pedacream cura muchos más pies que Fungacream, de manera que el tratamiento con Pedacream parece ser más efectivo.

Observe que en el cálculo del estadístico de prueba del ejemplo anterior, no incluimos a los 12 sujetos que se curaron de ambos pies (cada pie tratado con una de las cremas) y tampoco incluimos a los 20 sujetos que no se curaron de ningún pie. En vez de incluir los resultados de curado/curado y los resultados de no curado/no curado, sólo usamos los resultados de curado/no curado y no curado/curado. Es decir, sólo incluimos los resultados de las categorías que son *diferentes*. A este tipo de categorías diferentes se les denomina *pares discordantes*.

Definición

Los **pares discordantes** de resultados provienen de pares de categorías en las que ambas categorías son diferentes (como en curado/no curado o no curado/curado).

Al tratar de determinar si existe una diferencia significativa entre los dos tratamientos con las cremas en la tabla 11-10, no nos resultan útiles los sujetos que se curaron de ambos pies, ni tampoco los sujetos que no se curaron de ningún pie. Las diferencias se reflejan en los resultados discordantes de los sujetos que se curaron de un pie y no del otro. Como consecuencia, el estadístico de prueba sólo incluye las dos frecuencias que resultan de los dos pares discordantes (o diferentes) de categorías.

Advertencia: Cuando aplique la prueba de McNemar, tenga cuidado de usar únicamente las frecuencias de los pares de categorías que son diferentes. No utilice a ciegas las frecuencias en las esquinas superior derecha e inferior izquierda, porque no necesariamente representan los pares discordantes. Si reordenara la tabla 11-10 como se indica abajo, tendría un formato inconsistente, pero técnicamente la tabla sería correcta al resumir los mismos resultados que la tabla anterior; sin embargo, el uso irreflexivo de las frecuencias de 20 y 12 produciría el estadístico de prueba *incorrecto*.

		Tratamiento con Pedacream	
		Curados	No curados
Tratamiento con Fungacream	No curados	40	20
	Curados	12	8

En esta tabla reconfigurada, los pares discordantes de frecuencias son:

curado/no curado: 40

no curado/curado: 8

Con esta tabla reconfigurada, nuevamente debemos usar las frecuencias de 40 y 8, y no las de 20 y 12. En un mundo más perfecto, todas las tablas 2×2 estarían configuradas con un formato consistente y seríamos mucho menos proclives a utilizar las frecuencias incorrectas.

Uso de la tecnología

STATDISK Seleccione **Analysis** y luego **McNemar's Test**. Proceda a registrar las frecuencias en la tabla, indique el nivel de significancia y luego haga clic en **Evaluate**. Los resultados de STATDISK incluyen el estadístico de prueba, el valor crítico, el valor

P y la conclusión. MINITAB, EXCEL y calculadora TI-83/84 Plus: La prueba de McNemar no está disponible.



Además de comparar tratamientos aplicados a datos apareados (como en el ejemplo anterior), la prueba de McNemar a menudo se utiliza para probar la hipótesis nula de que no hay cambios en experimentos de antes y después. (Vea los ejercicios 5 a 12).

11-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Prueba de McNemar.** Cuando se realizan pruebas de hipótesis con tablas 2×2 , ¿qué circunstancias indican que la prueba de McNemar es la adecuada, en tanto que los métodos de la sección 11-3 no lo son?
- Prueba de McNemar.** ¿Se puede usar la prueba de McNemar en tablas de dos factores con más de dos renglones o más de dos columnas? ¿Por qué?
- Pares discordantes.** ¿Qué son los resultados de pares discordantes?
- Pares discordantes.** ¿Por qué la prueba de McNemar sólo incluye datos discordantes e ignora el resto de los datos?

En los ejercicios 5 a 12, remítase a la siguiente tabla, que resume los resultados de un experimento en el que primero se clasificó a los sujetos como fumadores o no fumadores, luego se les dio un tratamiento y después se les volvió a clasificar como fumadores o no fumadores.

		Antes del tratamiento	
		Fuman	No fuman
Después del tratamiento	Fuman	50	6
	No fuman	8	80

- Tamaño muestral.** ¿Cuántos sujetos se incluyeron en el experimento?
- Eficacia del tratamiento.** ¿Cuántos sujetos cambiaron su estatus de fumadores después del tratamiento?
- Ineficacia del tratamiento.** ¿A cuántos sujetos parece no haberles afectado el tratamiento de una u otra forma?
- ¿Por qué no usar una prueba *t*?** En la sección 9-4 se presentaron procedimientos para manejar datos apareados. ¿Por qué no podemos usar los procedimientos de la sección 9-4 para el análisis de los resultados que se resumen en la tabla?

9. Pares discordantes. ¿Cuál de los siguientes pares de resultados antes/después son discordantes?

- a) fuma/fuma
- b) fuma/no fuma
- c) no fuma/fuma
- d) no fuma/no fuma

10. Estadístico de prueba. Utilice las frecuencias apropiadas para calcular el valor del estadístico de prueba.

11. Valor crítico. Utilice un nivel de significancia de 0.01 para calcular el valor crítico.

12. Conclusión. Con base en los resultados anteriores, ¿qué concluye? ¿Qué sentido tiene la conclusión en términos de los resultados muestrales originales?

13. Tratamiento para el pie de atleta. Como en el ejemplo de esta sección, suponga que los sujetos tienen pie de atleta en ambos pies. También suponga que para cada sujeto, un pie es tratado con una solución fungicida mientras que el otro recibe un placebo. Los resultados se presentan en la siguiente tabla. Utilice un nivel de significancia de 0.05 para probar la eficacia del tratamiento.

		Tratamiento fungicida	
		Curados	No curados
Placebo	Curados	5	12
	No curados	22	55

14. Tratamiento para el pie de atleta. Repita el ejercicio 13 después de cambiar la frecuencia de 22 por 66.

15. Comparación de TEP/TC con IRM. En el artículo “Whole-Body Dual-Modality PET/TC and Whole Body MRI for Tumor Staging in Oncology” (Antoch *et al.*, *Journal of the American Medical Association*, vol. 290, núm. 24), los autores citan la importancia de identificar con exactitud la etapa de un tumor, ya que esto es crucial para determinar la terapia adecuada. El artículo analiza un estudio que compara la exactitud de la tomografía por emisión de positrones (TEP) y la tomografía computarizada (TC) con la imagen por resonancia magnética (IRM). Utilice los datos de la tabla sobre 50 tumores analizados con ambas tecnologías. Al parecer, ¿hay una diferencia en la exactitud? ¿Alguna de las tecnologías parece ser mejor?

		TEP/TC	
		Correcto	Incorrecto
IRM	Correcto	36	1
	Incorrecto	11	2

16. Prueba de un tratamiento. En el artículo “Eradication of Small Intestinal Bacterial Overgrowth Reduces Symptoms of Irritable Bowel Syndrome” (Pimentel, Chow, Lin, *American Journal of Gastroenterology*, vol. 95, núm. 12), los autores analizan si el tratamiento antibiótico del crecimiento excesivo de bacterias reduce las molestias intestinales. Se utilizó una prueba de McNemar para analizar los resultados de los sujetos con erradicación del crecimiento bacterial excesivo. Utilice los datos de la siguiente tabla. Al parecer, ¿el tratamiento es efectivo para el dolor abdominal?

		¿Dolor abdominal antes del tratamiento?	
		Sí	No
¿Dolor abdominal después del tratamiento?	Sí	11	1
	No	14	3

11-4 MÁS ALLÁ DE LO BÁSICO

- 17. Corrección por continuidad.** El estadístico de prueba que estudiamos en esta sección incluye una corrección por continuidad. El estadístico de prueba que se muestra a continuación no incluye la corrección por continuidad, y en ocasiones se utiliza como estadístico de prueba para la prueba de McNemar. Consulte el ejemplo de esta sección, calcule el valor del estadístico de prueba utilizando la siguiente expresión y compare el resultado con el que se obtuvo en el ejemplo.

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

- 18. Uso del sentido común.** Considere la siguiente tabla y utilice un nivel de significancia de 0.05.
- ¿Qué sugiere la prueba de McNemar sobre la eficacia del tratamiento?
 - Los valores de a y d no se utilizan en los cálculos, pero, ¿qué sugiere el sentido común si $a = 5000$ y $d = 4000$?

		Antes del tratamiento	
		Fuman	No fuman
Después del tratamiento	Fuman	a	5
	No fuman	20	d

- 19. Caso de muestra pequeña.** Los requisitos de la prueba de McNemar incluyen la condición de que $b + c \geq 10$, de manera que la distribución del estadístico de prueba se puede aproximar por medio de la distribución chi cuadrada. Remítase al ejemplo de esta sección y reemplace los datos de la tabla con los valores que se dan abajo. No se debe utilizar la prueba de McNemar, ya que la condición de $b + c \geq 10$ no se satisface con $b = 2$ y $c = 6$. En su lugar, utilice la distribución binomial para calcular la probabilidad de que, de 8 resultados igualmente probables, éstos consistan en 6 elementos en una categoría y 2 en la otra categoría, o de que los resultados sean más extremos. Esto es, utilice una probabilidad de 0.5 para calcular la probabilidad de que, en $n = 8$ ensayos, el número de éxitos x sea 6, 7 u 8. Duplique esa probabilidad para calcular el valor P para esta prueba. Compare el resultado con el valor P de 0.289, que se obtiene al utilizar la aproximación chi cuadrada, aun cuando se viole el requisito de $b + c \geq 10$. ¿Qué concluye acerca de los dos tratamientos?

		Tratamiento con Pedacream	
		Curados	No curados
Tratamiento con Fungacream	Curados	12	2
	No curados	6	20

Repaso

En este capítulo trabajamos con datos resumidos como conteos de frecuencias para diferentes categorías. En la sección 11-2 describimos métodos para probar la bondad de ajuste en un experimento multinomial, que es similar a un experimento binomial, excepto que existen más de dos categorías de resultados. Los experimentos multinomiales dan por resultado conteos de frecuencias acomodados en un solo renglón o columna, y realizamos pruebas para determinar si las frecuencias muestrales observadas concuerdan (o “se ajustan”) con alguna distribución aseverada.

En la sección 11-3 describimos métodos para probar aseveraciones que implican tablas de contingencia (o tablas de frecuencias de dos factores), que tienen al menos dos renglones y dos columnas. Las tablas de contingencia incorporan dos variables: una variable se utiliza para determinar el renglón que describe un valor muestral y la segunda variable se utiliza para determinar la columna que describe un valor muestral. La sección 11-3 incluyó dos tipos de prueba de hipótesis: **1.** una prueba de independencia entre las variables de renglón y de columna; **2.** una prueba de homogeneidad para decidir si diferentes poblaciones tienen las mismas proporciones de algunas características.

En la sección 11-4 se estudió la prueba de McNemar para probar la hipótesis nula de que una muestra de datos apareados proviene de una población en la que los pares discordantes (diferentes) ocurren en la misma proporción.

A continuación se mencionan algunos componentes clave de los métodos analizados en este capítulo:

- *Sección 11-2 (prueba de bondad de ajuste):*

$$\text{El estadístico de prueba es } \chi^2 = \sum \frac{(O - E)^2}{E}$$

La prueba es de cola derecha con $k - 1$ grados de libertad. Todas las frecuencias esperadas deben ser al menos de 5.

- *Sección 11-3 (prueba de independencia u homogeneidad de la tabla de contingencia):*

$$\text{El estadístico de prueba es } \chi^2 = \sum \frac{(O - E)^2}{E}$$

La prueba es de cola derecha con $(r - 1)(c - 1)$ grados de libertad. Todas las frecuencias esperadas deben ser al menos de 5.

- *Sección 11-4 (tabla 2 × 2 con frecuencias de datos apareados):*

$$\text{El estadístico de prueba es } \chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

donde las frecuencias de b y c deben provenir de pares “discordantes”. La prueba es de cola derecha con 1 grado de libertad.

Las frecuencias b y c deben ser de tal manera que $b + c \geq 10$.

Conocimientos estadísticos y pensamiento crítico

- 1. Datos categóricos.** En este capítulo se estudiaron algunos métodos diferentes para el análisis de datos *categóricos*. ¿Qué son los datos categóricos?
- 2. Realización de encuesta.** Un estudiante realiza un proyecto de investigación al preguntarle a 200 compañeros de clase si alguna vez les ha sido robada una tarjeta de crédito. El estudiante construye una tabla de contingencia con las categorías de renglón de género (hombre/mujer) y las categorías de columna de respuesta (sí, no, se rehusa a responder). Utiliza los métodos de la sección 11-3 para concluir que el género es independiente de la respuesta. ¿Qué es incorrecto en su proyecto?
- 3. Distribución chi cuadrada.** En este capítulo se presentaron diferentes métodos que incluyen la aplicación de la distribución chi cuadrada. ¿Cuáles de las siguientes propiedades de una distribución chi cuadrada son verdaderas?
 - a. Los valores de un estadístico de prueba chi cuadrada siempre son positivos o iguales a cero, pero nunca negativos.
 - b. Una distribución chi cuadrada es simétrica.
 - c. Existe una distribución chi cuadrada diferente para cada número de grados de libertad.

- d) Cuando se usa una distribución chi cuadrada, el número de grados de libertad siempre es igual al tamaño muestral menos 1.
- e) Cuando se usa una distribución chi cuadrada, los datos muestrales no necesitan ser aleatorios si el tamaño de la muestra es muy grande.
- 4. Verificación de requisitos.** Los métodos de prueba de la bondad de ajuste y los métodos de prueba de independencia entre dos variables para una tabla de contingencia requieren que todas las frecuencias esperadas sean al menos de 5. ¿Se pueden usar esos métodos si hay una celda con un conteo de frecuencia observado menor que 5? ¿Por qué?

Ejercicios de repaso

- 1. ¿Las muertes por conducción en estado de ebriedad son el resultado de beber el fin de semana?** Mucha gente cree que los accidentes fatales por conducción en estado de ebriedad se deben a bebedores casuales que se embriagan los viernes y los sábados por la noche, mientras que otros creen que los accidentes fatales por conducción en estado de ebriedad son causados por individuos que beben todos los días de la semana. En un estudio sobre accidentes automovilísticos fatales, se eligieron al azar 216 casos de un grupo en el que el conductor tenía un contenido de alcohol en la sangre mayor que 0.10. Esos casos se separaron de acuerdo con el día de la semana, y los resultados se presentan en la siguiente tabla (según datos del Programa STOP-DWI del condado Dutchess). Utilice un nivel de significancia de 0.05 para probar la aseveración de que este tipo de choques fatales ocurren los diferentes días de la semana con la misma frecuencia. ¿La evidencia sustenta la teoría de que los choques fatales por conducción en estado de ebriedad se deben a bebedores casuales o son causados por personas que beben todos los días?

Día	Dom	Lun	Mar	Mié	Jue	Vie	Sáb
Número	40	24	25	28	29	32	38

- 2. Correo electrónico y privacidad.** Se preguntó a empleados y a jefes de alto nivel si era muy poco ético vigilar el correo electrónico de los empleados, y los resultados se resumen en la tabla (según datos de una encuesta de Gallup). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la respuesta es independiente del hecho de que el sujeto sea un empleado o un jefe de alto nivel. ¿Cambia la conclusión si se utiliza un nivel de significancia de 0.01 en vez de 0.05? ¿Parecen estar de acuerdo los empleados y los jefes en este tema?

	Sí	No
Trabajadores	192	244
Jefes	40	81

- 3. El crimen y los extraños.** La tabla adjunta lista resultados de encuesta obtenidos de una muestra aleatoria de víctimas de diferentes crímenes (según datos del Departamento de Justicia de Estados Unidos). Con un nivel de significancia de 0.05, pruebe la aseveración de que el tipo de crimen es independiente del hecho de que el criminal sea un extraño. ¿Cómo podrían afectar los resultados la estrategia que los oficiales de policía utilizan cuando investigan crímenes?

	Homicidio	Robo	Asalto
El criminal era un extraño	12	379	727
El criminal era un conocido o un familiar	39	106	642

- 4. Comparación de tratamientos.** Se usan dos cremas diferentes para tratar sujetos con irritación por hiedra venenosa en ambas manos. Cada sujeto recibe tratamiento con Ivy Ease en una mano, mientras que su otra mano se trata con un placebo. Los resultados muestrales se resumen en la siguiente tabla. Utilice un nivel de significancia de 0.05 para probar la hipótesis nula de que las siguientes dos proporciones son iguales: 1) la

proporción de sujetos con mejoría en la mano tratada con Ivy Ease y sin mejoría en la mano tratada con un placebo; 2) la proporción de sujetos sin mejoría en la mano tratada con Ivy Ease y con mejoría en la mano tratada con un placebo. ¿Parece que el tratamiento con Ivy Ease es efectivo?

		Tratamiento con Ivy Ease	
		Mejoría	Sin mejoría
Placebo	Mejoría	12	8
	Sin mejoría	32	19

Tabla 11-11

	A	B	C	D
x	85	90	80	75
y	80	84	73	70

Ejercicios de repaso acumulativo

- Cálculo de estadísticos.** Suponga que en la tabla 11-11, los títulos del renglón y la columna no tienen significado, de manera que la tabla contiene calificaciones de pruebas de 8 prisioneros seleccionados al azar, convictos por quitar etiquetas a las almohadas. Calcule la media, la mediana, el rango, la varianza, la desviación estándar y el resumen de los 5 números.
- Cálculo de probabilidad.** Suponga que en la tabla 11-11 las letras A, B, C y D representan las opciones de la primera pregunta de un examen de opción múltiple. Suponga también que x representa hombres y y representa mujeres, y que los números de la tabla son conteos de frecuencias, de forma tal que 85 hombres eligen la respuesta A, 80 mujeres eligen la respuesta A, 90 hombres escogen la respuesta B, etcétera.
 - Si se selecciona al azar una respuesta, calcule la probabilidad de que sea la respuesta C.
 - Si se selecciona al azar una respuesta, calcule la probabilidad de que sea elegida por un hombre.
 - Si se selecciona al azar una respuesta, calcule la probabilidad de que sea C o que sea elegida por un hombre.
 - Si se seleccionan al azar dos respuestas diferentes, calcule la probabilidad de que ambas sean elegidas por una mujer.
 - Si se selecciona al azar una respuesta, calcule la probabilidad de que sea la respuesta B, dado que la respuesta fue elegida por una mujer.
- Prueba para proporciones iguales.** Utilice los mismos supuestos que en el ejercicio 2 y pruebe la aseveración de que los hombres y las mujeres eligen las diferentes respuestas en las mismas proporciones.
- Prueba para una relación.** Suponga que la tabla 11-11 lista puntuaciones de prueba de cuatro personas, donde la puntuación x corresponde a una prueba de memoria y la puntuación y a una prueba de razonamiento. Pruebe la aseveración de que existe una correlación lineal entre las puntuaciones x y y .
- Prueba de eficacia de entrenamiento.** Suponga que la tabla 11-11 lista puntuaciones de prueba para cuatro personas, donde la puntuación x corresponde a una prueba anterior aplicada antes de una sesión de entrenamiento para mejorar la memoria y la puntuación y corresponde a una prueba aplicada después del entrenamiento. Pruebe la aseveración de que la sesión de entrenamiento no tiene ningún efecto.
- Prueba para igualdad de medias.** Suponga que en la tabla 11-11 las letras A, B, C y D representan diferentes versiones de la misma prueba de razonamiento. Las puntuaciones x se obtuvieron de cuatro hombres seleccionados al azar y las puntuaciones y se obtuvieron de cuatro mujeres seleccionadas al azar. Pruebe la aseveración de que los hombres y las mujeres tienen la misma puntuación media.

Actividades de cooperación en grupos

- 1. Actividad fuera de clase** Formen grupos de cuatro o cinco estudiantes. Vean los primeros dos renglones de la tabla 11-1 en el problema del capítulo, sobre la distribución de dígitos líderes esperados de acuerdo con la ley de Benford. Reúnan datos y utilicen los métodos de la sección 11-2 para verificar que los datos coincidan razonablemente bien con la ley de Benford. Las siguientes son algunas posibilidades que podrían tomarse en cuenta:

 - La cantidad de cheques que usted ha expedido
 - Los precios de acciones
 - Las poblaciones de los condados en Estados Unidos
 - Los números de direcciones

2. Actividad fuera de clase Dividan la clase en grupos de cuatro o cinco estudiantes y reúnan resultados pasados de una lotería estatal. Este tipo de resultados suelen estar disponibles en sitios de Internet de las loterías estatales específicas. Utilice los métodos de la sección 11-2 para probar que los números se seleccionan de manera que todos los posibles resultados son igualmente probables.

3. Actividad fuera de clase Formen grupos de cuatro o cinco estudiantes. Cada miembro del grupo debe encuestar al menos a 15 estudiantes varones y a 15 estudiantes mujeres en la misma universidad, haciéndoles dos preguntas: 1) ¿Cuál partido político favorece más el sujeto? 2) Si el sujeto inventara la excusa de un neumático desinflado por no asistir a un examen, ¿cuál neumático diría que se desinfló ante la pregunta del profesor? (Vea el ejercicio 10 en la sección 11-2). Pida al sujeto que escriba las dos respuestas en una tarjeta; también registre el género del sujeto y si el sujeto escribió con la mano derecha o con la izquierda. Utilice los métodos de este capítulo para analizar los datos reunidos. Incluya estas pruebas:

 - Las cuatro respuestas posibles para un neumático desinflado se seleccionan en la misma proporción.
 - El neumático identificado como desinflado es independiente del género del sujeto.
 - La elección de un partido político es independiente del género del sujeto.
 - La elección de un partido político es independiente de si el sujeto es zurdo o diestro.

4. Actividad fuera de clase Formen grupos de cuatro o cinco estudiantes. Cada miembro del grupo debe seleccionar a 15 estudiantes y pedir primero a cada uno que elija “al azar” cuatro dígitos. Después de que los cuatro dígitos se hayan registrado, pida a cada sujeto que escriba los últimos cuatro dígitos de su número de seguro social. Tome los resultados muestrales “aleatorios” y mézclelos para formar una gran muestra, luego mezcle los dígitos del número del seguro social en una segunda gran muestra. Utilice el conjunto de datos “aleatorios” para probar la aseveración de que los estudiantes seleccionan dígitos al azar. Luego utilice los dígitos del seguro social para probar la aseveración de que provienen de una población de dígitos aleatorios. Compare los resultados. ¿Parece que los estudiantes pueden seleccionar dígitos al azar? ¿Son proclives a seleccionar algunos dígitos con más frecuencia que otros? ¿Son proclives a seleccionar algunos dígitos con menor frecuencia que otros? ¿Parece que los últimos dígitos de los números del seguro social fueron seleccionados al azar?

5. Actividad en clase Formen grupos de tres o cuatro estudiantes. A cada grupo se le debe entregar un dado junto con la instrucción de que debe probar su “balance”. ¿El dado es legal o está cargado? Describa el análisis y los resultados.

6. Actividad fuera de clase Formen grupos de dos o tres estudiantes. Algunos ejemplos y ejercicios de este capítulo se basaron en el análisis de los últimos dígitos de los valores. Se señaló que el análisis de los últimos dígitos en ocasiones revela si los valores son resultado de mediciones reales o si son estimados reportados. Remítase a un almanaque y encuentre las longitudes de los ríos en el mundo, luego analice los últimos dígitos para determinar si estas longitudes parecen ser mediciones reales o el reporte de estimados. (En vez de longitudes de ríos, se podrían utilizar alturas de montañas, alturas de los edificios más altos, longitudes de puentes, etcétera).

Proyecto tecnológico

Utilice STATDISK, Minitab, Excel, la calculadora TI-83/84 Plus o cualquier otro programa de cómputo o calculadora capaz de generar dígitos aleatorios igualmente probables entre 0 y 9, inclusive. Genere 500 dígitos y registre los resultados en la tabla adjunta. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los dígitos muestrales provienen

de una población con una distribución uniforme (de manera que todos los dígitos son igualmente probables). ¿El generador de números aleatorios parece funcionar como debería?

De los datos a la decisión

Pensamiento crítico: ¿El acusado es culpable de fraude?

En el juicio del *estado de Arizona vs. Wayne James Nelson*, el sujeto fue acusado de expedir cheques a un vendedor que no existía realmente. Las cantidades de los cheques se listan abajo, ordenadas por renglón.

\$1,927.48	\$27,902.31	\$86,241.90	\$72,117.46	\$81,321.75	\$97,473.96
\$93,249.11	\$89,658.16	\$87,776.89	\$92,105.83	\$79,949.16	\$87,602.93
\$96,879.27	\$91,806.47	\$84,991.67	\$90,831.83	\$93,766.67	\$88,336.72
\$94,639.49	\$83,709.26	\$96,412.21	\$88,432.86	\$71,552.16	

Análisis de los resultados

¿Los dígitos líderes cumplen con la ley de Benford descrita en el problema del capítulo? Cuando se prueba la bondad de ajuste con las proporciones esperadas por la ley de Benford, es necesario combinar categorías puesto que no todos los valores esperados son al menos de 5. Utilice una categoría con dígitos líderes de 1, una segunda categoría con dígitos líderes de 2, 3, 4 y 5, y una tercera categoría con dígitos líderes de 6, 7, 8 y 9. ¿Son todos los valores esperados para estas tres categorías al menos de 5? ¿Existe evidencia suficiente para concluir que los dígitos líderes en los cheques no cumplen la

ley de Benford? Además de los dígitos líderes, ¿existen otros patrones cualesquiera que sugieran que los montos de los cheques fueron creados por el acusado en vez de ser el resultado de transacciones típicas y reales? Con base en la evidencia, si usted fuera un miembro del jurado, ¿concluiría que los montos de los cheques son el resultado de un fraude? ¿Cuál sería un argumento que podría presentar si usted fuera el abogado defensor?



Tablas de contingencia

Una característica importante de las pruebas de independencia con tablas de contingencia es que los datos reunidos no requieren ser de naturaleza cuantitativa. Una tabla de contingencia resume observaciones por medio de las categorías o rótulos de los renglones y las columnas. Como resultado, características como el género, la raza y la afiliación política se convierten en información que puede someterse a los procedimientos formales de prueba de hipótesis. El

Proyecto de Internet

proyecto de Internet para este capítulo se encuentra en el sitio de Internet de *Estadística*:

<http://www.pearsoneducacion.net/triola>

Usted encontrará vínculos con una variedad de datos demográficos. Con estos conjuntos de datos, realizará pruebas en áreas tan diversas como la académica, la política y la industria del entretenimiento. En cada prueba usted obtendrá conclusiones relacionadas con la independencia de características interesantes.

La estadística en el trabajo

"Aun si usted no es un hábil operador de números, el conocimiento [estadístico] es útil en cualquier situación que requiera predicción, toma de decisiones o evaluación".



Nabil Lebbos

Ilustrador gráfico, *Published Image*

Como analista para *Published Image* de Standard & Poor, los estudios de Nabil en rendimiento de inversiones se publican en periódicos que son leídos por más de un millón de inversionistas.

Por favor, describa su ocupación.

Trabajo para Published Image donde utilicé la estadística para generar los gráficos y datos que utilizamos en nuestras publicaciones financieras; utilicé muchos estadísticos y sus aplicaciones. Escribímos notas informativas para bancos y sociedades de inversión.

¿Qué conceptos de estadística utiliza usted?

La desviación estándar para medir el riesgo, la regresión para medir la relación de la inversión con su punto de referencia, y la correlación para determinar el movimiento de una inversión en relación con otras inversiones.

¿Cómo utiliza la estadística en el trabajo?

Comienzo con un conjunto dado de datos en bruto. Por lo regular, se trata de rendimientos mensuales, diarios o anuales de las inversiones. Luego uso Excel para graficar los datos y así puedo obtener una imagen de lo que estoy examinando. A partir de esto, procedo a realizar un análisis. Algunas veces los resultados no respaldan un aspecto que el artículo adjunto quiere fortalecer. En situaciones como ésta, analizo otras posibilidades.

Por favor, describa un ejemplo específico que ilustre cómo el uso de la estadística tuvo éxito en mejorar un producto o servicio.

Uno de nuestros clientes quería señalar que aunque su sociedad de inversión no superaba a las otras, tenía éxito en evitar

consistentemente rendimientos negativos grandes. Ejecuté algunas pruebas de sesgo y riesgo, las cuales mostraron que, de hecho, los rendimientos de la inversión estaban sesgados positivamente. Creamos histogramas comparando este fondo de inversión con un promedio de todos los fondos de inversión, lo cual aclaró la cuestión.

En términos de estadística, ¿que le recomendaría a quienes buscan un empleo?

La estadística es una herramienta lógica que, cuando se utiliza con fines informativos, puede convencer a uno mismo y a su audiencia —con mucha más eficacia que las palabras— de aquello que se intenta comunicar. Aun si usted no es un hábil operador de números, el conocimiento [estadístico] es útil en cualquier situación que requiera predicción, toma de decisiones o evaluación.

¿Cree que las personas que solicitan un empleo reciben una evaluación más favorable si estudiaron algo de estadística?

Sí.

Cuando estudiaba en la universidad, ¿esperaba utilizar la estadística en el trabajo?

No. Estudié la carrera de arquitectura y un posgrado en negocios.



Análisis de varianza

12



12-1 Panorama general

12-2 ANOVA de un factor

12-3 ANOVA de dos factores

¿Tratamientos diferentes afectan los pesos de álamos?

El conjunto de datos 7 del apéndice B incluye los pesos (en kilogramos) de álamos que recibieron distintos tratamientos en terrenos diferentes. Sólo consideraremos los pesos del año 1 en el terreno 1, el cual tiene un suelo fértil y húmedo, y se localiza cerca de un arroyo. Los pesos que consideraremos se resumen en la tabla 12-1.

Con la intención de explorar los datos para investigar el centro, la variación, la distribución, los valores extremos y los patrones de cambio a través del tiempo (CVDVT), comenzamos calculando los estadísticos muestrales que aparecen en la parte inferior de la tabla 12-1. Al examinar las medias muestrales, vemos que parecen variar mucho, desde 0.164 kg hasta 1.334 kg. Además, las desviaciones estándar de las muestras varían considerablemente, desde 0.126 kg hasta 0.859 kg. Es difícil analizar las distribuciones porque cada muestra consiste únicamente en 5 valores, pero las gráficas cuantilares normales sugieren que tres de las muestras provienen de poblaciones con distribuciones aproximadamente normales. Sin embargo, el análisis de los pesos de los álamos que recibieron tratamiento con fertilizantes sugiere que el peso de 1.34 kg es un valor

extremo cuando se compara con los otros pesos de los árboles fertilizados. Con la presencia de un solo valor extremo, procederemos bajo el supuesto de que las muestras provienen de poblaciones con distribuciones aproximadamente normales. Podríamos realizar análisis adicionales posteriormente para determinar si el peso de 1.34 kg tiene un fuerte efecto en los resultados. (Véase el ejercicio 5 en la sección 12-2).

Parece que las diferencias entre las medias muestrales indican que las muestras provienen de poblaciones con medias diferentes, pero en vez de considerar únicamente las medias muestrales, también debemos considerar las cantidades de variación, los tamaños muestrales y la naturaleza de la distribución de las medias muestrales. Una forma de tomar en cuenta todos estos factores importantes consiste en realizar una prueba formal de hipótesis que los incluya de manera automática. En el presente capítulo se estudiará una prueba de este tipo, y la usaremos para determinar si existe evidencia suficiente para concluir que las medias no son iguales. Entonces sabremos si los distintos tratamientos tienen algún efecto.

Tabla 12-1 Pesos (en kg) de álamos

Tratamiento				
Ninguno	Fertilizante	Riego	Fertilizante y riego	
0.15	1.34	0.23	2.03	
0.02	0.14	0.04	0.27	
0.16	0.02	0.34	0.92	
0.37	0.08	0.16	1.07	
0.22	0.08	0.05	2.38	
n	5	5	5	5
\bar{x}	0.184	0.332	0.164	1.334
s	0.127	0.565	0.126	0.859

12-1 Panorama general

En la sección 12-2 explicamos un método importante para probar la igualdad de tres o más medias poblacionales. En la sección 9-3 estudiamos procedimientos para probar la hipótesis de que *dos* medias poblacionales son iguales, pero los métodos de esa sección no pueden aplicarse cuando se incluyen tres o más medias. En vez de referirnos al objetivo principal de probar medias iguales, el término *análisis de varianza* se refiere al *método* que empleamos, el cual está basado en un análisis de varianzas muestrales.

Definición

El **análisis de varianza (ANOVA)** es un método de prueba de igualdad de tres o más medias poblacionales, por medio del análisis de las varianzas muestrales.

¿Por qué no probar sencillamente dos muestras al mismo tiempo? ¿Por qué necesitamos un nuevo procedimiento, cuando podemos probar la igualdad de dos medias utilizando los métodos presentados en el capítulo 9? Por ejemplo, si deseamos utilizar los datos muestrales de la tabla 12-1 para probar la aseveración de que las tres poblaciones tienen la misma media, ¿por qué no simplemente tomamos dos a la vez y probamos $H_0: \mu_1 = \mu_2$, luego $H_0: \mu_2 = \mu_3$, y así sucesivamente? Para los datos de la tabla 12-1, el método de probar la igualdad de dos medias a la vez requiere de seis pruebas diferentes de hipótesis, de manera que el grado de confianza podría ser tan bajo como 0.95⁶ (o 0.735). En general, conforme incrementamos el número de pruebas de significancia individuales, aumentamos el riesgo de obtener una diferencia únicamente por el azar (en vez de una diferencia real en las medias). El riesgo de un error tipo I (encontrar una diferencia en uno de los pares cuando en realidad no existe tal diferencia) es demasiado alto. El método del análisis de varianza nos ayuda a evitar este problema en particular (rechazar una hipótesis nula verdadera) utilizando una prueba de igualdad de varias medias.

Distribución *F*

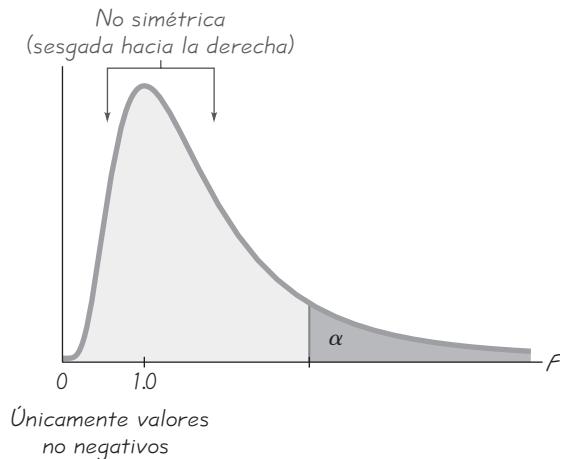
Los métodos del ANOVA de este capítulo requieren de la distribución *F* que se presentó por primera vez en la sección 9-5, en la cual señalamos que la distribución *F* tiene las siguientes propiedades importantes (véase la figura 12-1):

1. La distribución *F* no es simétrica; está sesgada hacia la derecha.
2. Los valores de *F* pueden ser 0 o positivos, pero no pueden ser negativos.
3. Existe una distribución *F* diferente para cada par de grados de libertad para el numerador y el denominador.

Los valores críticos de *F* se localizan en la tabla A-5.

El análisis de varianza (ANOVA) está basado en una comparación de dos estimados diferentes de la varianza común de las distintas poblaciones. Estos estimados (la *varianza entre muestras* y la *varianza dentro de las muestras*) se describirán en la sección 12-2. El término *un factor* se utiliza porque los datos muestrales están separados en grupos según una característica o factor. En la sección 12-3 estudiaremos el análisis de varianza de dos factores, el cual nos permite comparar poblaciones separadas en categorías por medio de dos características (o factores). Por ejemplo, podríamos separar la estatura de las personas utilizando los siguientes dos factores:

1. género (hombre o mujer) y 2. mano dominante derecha o izquierda.

**Figura 12-1****Distribución *F***

Existe una distribución *F* distinta para cada par de grados de libertad diferente para el numerador y el denominador.

Estrategia de estudio sugerida: Puesto que los procedimientos empleados en este capítulo requieren de cálculos complicados, enfatizaremos el uso y la interpretación de programas de cómputo, tales como STATDISK, Minitab y Excel, o de una calculadora TI-83/84 Plus. Sugerimos que inicie la sección 12-2 enfocándose en el siguiente concepto clave: estamos utilizando un procedimiento para probar la aseveración de que tres o más medias son iguales. A pesar de que los detalles de los cálculos son complicados, nuestro procedimiento será fácil debido a que está basado en un valor *P*. Si el valor *P* es pequeño, como 0.05 o menor, se rechaza la igualdad de las medias. De otra manera no se rechaza la igualdad de las medias. Después de comprender este procedimiento básico y sencillo, proceda a la comprensión de los fundamentos subyacentes.

12-2 ANOVA de un factor

Concepto clave En esta sección se presenta el método del *análisis de varianza de un factor*, que se utiliza para probar las hipótesis de que tres o más medias poblacionales son iguales, como en $H_0: \mu_1 = \mu_2 = \mu_3$. Como los cálculos son muy complicados, recomendamos interpretar los resultados obtenidos por medio de un programa de cómputo o de una calculadora TI-83/84 Plus. Sugerimos la siguiente estrategia de estudio:

1. Comprenda que un valor *P* pequeño (como 0.05 o menos) conduce al rechazo de la hipótesis nula de igualdad de medias. Con un valor *P* grande (como uno mayor que 0.05), no rechace la hipótesis nula de igualdad de medias.
2. Trate de comprender el fundamento subyacente al estudiar los ejemplos de esta sección.
3. Familiarícese con la naturaleza de los valores de la SC (suma de cuadrados) y de los CM (cuadrados medios), así como con el papel que desempeñan en la determinación del estadístico de prueba *F*, pero utilice programas estadísticos de cómputo o una calculadora para obtener esos valores.

El método que empleamos se denomina **análisis de varianza de un factor** (o **análisis de varianza de una entrada**) porque empleamos una sola propiedad o característica para categorizar las poblaciones. En ocasiones a esta característica se le llama *tratamiento o factor*.

Definición

Un **tratamiento** (o **factor**) es una propiedad o característica que nos permite distinguir entre sí a las distintas poblaciones.

Por ejemplo, los pesos de los álamos en la tabla 12-1 se distinguen de acuerdo con el tratamiento (ninguno, fertilizante, riego, fertilizante y riego). Se utiliza el término *tratamiento* porque las primeras aplicaciones del análisis de varianza implicaron experimentos de agricultura en los que distintas porciones de tierra se trataban con diferentes fertilizantes, tipos de semillas, insecticidas, etcétera. El siguiente recuadro incluye los requisitos y procedimientos que utilizaremos.

Requisitos

1. Las poblaciones tienen distribuciones que son aproximadamente normales. (Este requisito no es demasiado estricto, ya que el método funciona bien, a menos que la población tenga una distribución muy diferente de la normal. Si una población tiene una distribución muy diferente a la normal, utilice la prueba de Kruskal-Wallis, descrita en la sección 13-5).
2. Las poblaciones tienen la misma varianza σ^2 (o desviación estándar σ). (Este requisito no es demasiado estricto, ya que el método funciona bien a menos que las varianzas poblacionales difieran en grandes cantidades. El especialista en estadística de la Universidad de Wisconsin, George E. P. Box demostró que, siempre y cuando los tamaños muestrales sean iguales (o casi iguales), las varianzas pueden diferir de tal forma que la más grande sea hasta nueve veces el tamaño de la más pequeña, y los resultados del ANOVA continúan siendo esencialmente confiables).
3. Las muestras son aleatorias simples (es decir, muestras del mismo tamaño que tienen la misma probabilidad de ser elegidas).
4. Las muestras son independientes entre sí (es decir, no están aparejadas o asociadas de ninguna forma).
5. Las diferentes muestras provienen de poblaciones que están categorizadas de una sola forma. (Ésta es la base del nombre del método: análisis de varianza de *un factor*).

Procedimiento de prueba de $H_0: \mu_1 = \mu_2 = \mu_3 = \dots$

1. Utilice STATDISK, Minitab, Excel o una calculadora TI-83/84 Plus para obtener los resultados.
2. Identifique el valor P en los resultados.
3. Plantee una conclusión con base en estos criterios:
 - Si el valor $P \leq \alpha$, rechace la hipótesis nula de medias iguales y concluya que al menos una de las medias poblacionales es diferente de las otras.
 - Si el valor $P > \alpha$, no rechace la hipótesis nula de medias iguales.

Tenga cuidado al interpretar los resultados: Cuando concluimos que existe suficiente evidencia para rechazar la aseveración de medias poblacionales iguales, no podemos concluir a partir del ANOVA que cualquier media en particular es distinta de las demás. (Existen otras pruebas que pueden utilizarse para identificar las medias específicas que son diferentes, las cuales se conocen como *procedimientos de comparación múltiple* y que estudiaremos más adelante en esta sección).



EJEMPLO Pesos de álamos Dados los pesos de álamos listados en la tabla 12-1, y un nivel de significancia de $\alpha = 0.05$, utilice STAT-DISK, Minitab, Excel o una calculadora TI-83/84 Plus para probar la aseveración de que las cuatro muestras provienen de poblaciones con medias diferentes. Vea a las siguientes pantallas de resultados.

STATDISK

Source:	DF:	SS:	MS:	Test Stat. F:	Critical F:	P-Value:
Treatment:	3	4.682415	1.560805	5.731353	3.238871522	0.007348294
Error:	16	4.35724	0.2723275			
Total:	19	9.039655	0.4797713			

Reject the Null Hypothesis
Reject equality of means

Minitab

One-way ANOVA: None, Fert, Irrig, Fert&Irrig

Source	DF	SS	MS	F	P
Factor	3	4.682	1.561	5.73	0.007
Error	16	4.357	0.272		
Total	19	9.040			

S = 0.5219 R-Sq = 51.80% R-Sq(adj) = 42.76%

Excel

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	4.682415	3	1.560805	5.731352875	0.007348294	3.238871522
Within Groups	4.35724	16	0.2723275			
Total	9.039655	19				

TI-83/84 Plus

One-way ANOVA
 $F=5.731352875$
 $p=.0073482944$
 Factor
 $df=3$
 $SS=4.682415$
 $\downarrow MS=1.560805$

One-way ANOVA
 $\uparrow MS=1.560805$
 Error
 $df=16$
 $SS=4.35724$
 $MS=.2723275$
 $SxP=.521850074$

SOLUCIÓN

REQUISITO ✓ Cuando se investiga el requisito de normalidad de las cuatro poblaciones diferentes, la única muestra cuestionable es la segunda del grupo de tratamiento con fertilizante. Sólo para esa muestra, la gráfica cuantil normal y el histograma sugieren que tal vez no se trate de una distribución normal, y el problema es el valor de 1.34 kg, que al parecer es un valor extremo; por lo tanto, debemos ser cuidadosos. Lo más adecuado sería realizar el análisis con y sin la inclusión de ese valor. Consulte el ejercicio 5, en donde vemos que en este caso el valor extremo de 1.34 kg no tiene un efecto drástico en los resultados.

continúa

Las varianzas son muy diferentes, pero la más grande no es más de nueve veces el tamaño de la más pequeña, de manera que se satisface el requisito de varianzas iguales. Suponemos que tenemos muestras aleatorias simples. Sabemos que las muestras son independientes, y cada valor pertenece exactamente a un grupo. Por lo tanto, los requisitos se satisfacen y podemos proceder con la prueba de hipótesis. ✓

La hipótesis nula es $H_0: \mu_1 = \mu_2 = \mu_3$ y la hipótesis alternativa es la aseveración de que al menos una de las medias es diferente de las otras.

- Paso 1: Utilice algún recurso tecnológico para obtener los resultados del ANOVA, como alguno de los que se muestran en este ejemplo.
- Paso 2: Todas las pantallas de resultados indican que el valor P es 0.007, redondeado.
- Paso 3: Puesto que el valor P es menor que el nivel de significancia de $\alpha = 0.05$, rechazamos la hipótesis nula de igualdad de medias.

INTERPRETACIÓN Existe evidencia suficiente para sustentar la aseveración de que las cuatro medias poblacionales no son iguales. Con base en la muestra de pesos de la tabla 12-1. Con base en la muestra de pesos de la tabla 12-1, concluimos que esos pesos provienen de poblaciones con medias diferentes. Con base en esta prueba de ANOVA, no podemos concluir que cualquier media en particular sea diferente de las demás.

Fundamentos

Suponiendo que las poblaciones tienen la misma varianza σ^2 , el estadístico de prueba F es la razón o el cociente de los siguientes dos estimados de σ^2 : **1.** la variación *entre* muestras (con base en la variación entre medias muestrales), y **2.** variación *dentro* de muestras (con base en las varianzas muestrales). Un estadístico de prueba F significativamente *grande* (ubicado a la extrema derecha de la gráfica de distribución F) constituye evidencia en contra de medias poblacionales iguales, de manera que la prueba es de cola derecha. La figura 12-2 indica la relación entre el estadístico de prueba F y el valor P .

Estadístico de prueba del ANOVA de un factor

$$F = \frac{\text{varianza entre las muestras}}{\text{varianza dentro de las muestras}}$$

El numerador del estadístico de prueba F mide la variación entre medias muestrales. El estimado de la varianza en el denominador depende únicamente de las varianzas muestrales y no se ve afectado por las diferencias entre las medias muestrales. Como consecuencia, las medias muestrales que tienen valores cercanos dan como resultado un estadístico de prueba F pequeño y concluimos que no existen diferencias significativas entre las medias muestrales. Pero si el valor de F es excesivamente *grande*, entonces rechazamos la aseveración de igualdad de medias. (Los ambiguos términos “pequeño” y “excesivamente grande” se vuelven objetivos por medio del valor P correspondiente, el cual nos indica si el estadístico de prueba F está o no en la región crítica). Puesto que valores excesivamente grandes de F reflejan medias desiguales, la prueba es de cola derecha.

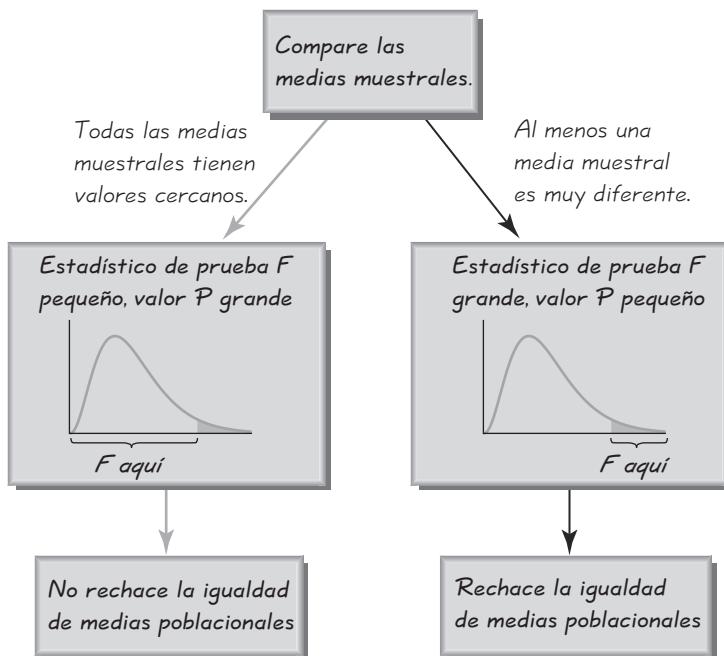


Figura 12-2

Relación entre el estadístico de prueba F y el valor P

Cálculos con tamaños muestrales n iguales

Para comprender realmente los efectos del análisis de varianza de un factor, vea la tabla 12-2. Compare el conjunto de datos A con el conjunto de datos B y observe que la única diferencia es que añadimos 10 a cada valor de la muestra 1 en el conjunto de datos A para obtener los valores de la muestra 1 en el conjunto de datos B. Si todos los conjuntos de datos tienen el mismo tamaño muestral (como en $n = 4$ para la tabla 12-2), los cálculos requeridos no son demasiado difíciles. Primero, calcule la varianza *entre* muestras al evaluar $ns_{\bar{x}}^2$, donde $s_{\bar{x}}$ es la varianza de las medias muestrales y n es el tamaño de cada una de las muestras. Es decir, considere las medias muestrales como un conjunto ordinario de valores y calcule la varianza. (Del teorema del límite central, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ se puede despejar σ para obtener $\sigma = \sqrt{n} \cdot \sigma_{\bar{x}}$, de manera que podemos estimar σ^2 con $ns_{\bar{x}}^2$). Por ejemplo, las medias muestrales del conjunto de datos A en la tabla 12-2 son 5.5, 6.0 y 6.0. Estos tres valores tienen una varianza de $s_{\bar{x}}^2 = 0.0833$, de manera que la

$$\text{varianza entre las muestras} = ns_{\bar{x}}^2 = 4(0.0833) = 0.3332$$

A continuación, estime la varianza *dentro* de las muestras, calculando s_p^2 , que es la varianza agrupada que se obtiene al calcular la media de las varianzas muestrales. Las varianzas muestrales en la tabla 12-2 son 3.0, 2.0 y 2.0, de forma que la

$$\begin{aligned}\text{varianza dentro de muestras} &= s_p^2 \\ &= \frac{3.0 + 2.0 + 2.0}{3} = 2.3333\end{aligned}$$

Finalmente, evalúe el estadístico de prueba F de la siguiente manera:

$$F = \frac{\text{varianza entre muestras}}{\text{varianza dentro de muestras}} = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{0.3332}{2.3333} = 0.1428$$



Resistencia a las encuestas

Las encuestas basadas en muestras relativamente pequeñas pueden ser bastante precisas, siempre y cuando la muestra sea aleatoria o representativa de la población. Sin embargo, el incremento en las tasas de rechazo a las encuestas está haciendo que sea más difícil obtener muestras aleatorias. El Council of American Survey Research Organizations reportó que, en un año reciente, el 38% de los consumidores se rehusaron a responder encuestas. El director de una compañía de investigación de mercados dijo que “las personas tienen temor de ser seleccionadas y les preocupa que las generalizaciones se realicen con base únicamente en aquellos que cooperan”. Los resultados de la industria de investigación de mercados, multimillonaria en dólares, afectan los productos que compramos, los programas de televisión que vemos y muchas otras facetas de nuestras vidas.

Tabla 12-2 Efecto de una media sobre el estadístico de prueba F

A añadir 10			B		
Muestra 1	Muestra 2	Muestra 3	Muestra 1	Muestra 2	Muestra 3
7	6	4	17	6	4
3	5	7	13	5	7
6	5	6	16	5	6
6	8	7	16	8	7
$n_1 = 4$	$n_2 = 4$	$n_3 = 4$	$n_1 = 4$	$n_2 = 4$	$n_3 = 4$
$\bar{x}_1 = 5.5$	$\bar{x}_2 = 6.0$	$\bar{x}_3 = 6.0$	$\bar{x}_1 = 15.5$	$\bar{x}_2 = 6.0$	$\bar{x}_3 = 6.0$
$s_p^2 = 3.0$	$s_p^2 = 2.0$	$s_p^2 = 2.0$	$s_p^2 = 3.0$	$s_p^2 = 2.0$	$s_p^2 = 2.0$
Varianza entre muestras	$ns_{\bar{x}}^2 = 4 (0.0833) = 0.3332$		$ns_{\bar{x}}^2 = 4 (30.0833) = 120.3332$		
Varianza dentro de muestras	$s_p^2 = \frac{3.0 + 2.0 + 2.0}{3} = 2.3333$		$s_p^2 = \frac{3.0 + 2.0 + 2.0}{3} = 2.3333$		
Estadístico de prueba F	$F = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{0.3332}{2.3333} = 0.1428$		$F = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{120.3332}{2.3333} = 51.5721$		
Valor P (obtenido con Excel)	$Valor P = 0.8688$		$Valor P = 0.0000118$		

El valor crítico de F se calcula suponiendo una prueba de cola derecha, ya que los valores grandes de F corresponden a diferencias significativas entre medias. Con k muestras, cada una con n valores, el número de grados de libertad se obtiene como sigue.

Grados de libertad:
(k = número de muestras y n = tamaño muestral)

grados de libertad del numerador = $k - 1$

grados de libertad del denominador = $k(n - 1)$

Para el conjunto de datos A de la tabla 12-2, $k = 3$ y $n = 4$, de manera que los grados de libertad son 2 para el numerador y $3(4 - 1) = 9$ para el denominador. Con $\alpha = 0.05$, 2 grados de libertad para el numerador y 9 grados de libertad para el denominador, el valor crítico F de la tabla A-5 es 4.2565. Si utilizáramos el método tradicional de prueba de hipótesis con el conjunto de datos A de la tabla 12-2, veríamos que esta prueba de cola derecha tiene un estadístico de prueba $F = 0.1428$ y un valor crítico de $F = 4.2565$, de manera que el estadístico de prueba no se encuentra en la región crítica y, por lo tanto, no rechazamos la hipótesis nula de igualdad de medias.

Para ver realmente cómo funciona el estadístico de prueba F , considere ambos conjuntos de datos muestrales en la tabla 12-2. Observe que las tres muestras de la parte A son idénticas a las tres muestras de la parte B, excepto que en la parte B añadimos 10 a cada valor de la muestra 1 de la parte A. Las tres medias muestrales de la parte A son muy cercanas, pero existen diferencias sustanciales en la parte B. Las tres varianzas muestrales de la parte A son idénticas a las de la parte B.

La suma de 10 a cada dato de la primera muestra de la tabla 12-2 tiene un efecto drástico en el estadístico de prueba, ya que F cambia de 0.1428 a 51.5721. La suma de 10 a cada dato de la primera muestra también tiene un efecto drástico en el valor P , el cual cambia de 0.8688 (no significativo) a 0.0000118 (significativo). Observe que la varianza entre muestras en la parte A es 0.3332, pero en la parte B es 120.3332 (lo que indica que las medias muestrales en la parte B están más separadas). Observe también que las varianzas dentro de las muestras son de 2.3333 en ambas partes, ya que la varianza dentro de una muestra no se ve afectada cuando sumamos una constante a cada valor muestral. *El cambio en el estadístico F y el valor P es atribuible únicamente a los cambios en \bar{x}_1 .* Esto ilustra que el estadístico de prueba F es muy sensible a las *medias* muestrales, aun cuando se obtiene a través de dos estimados distintos de la *varianza* poblacional común.

He aquí el punto clave de la tabla 12-2: los conjuntos de datos A y B son idénticos, excepto que en el conjunto de datos B se añadió 10 a cada valor de la primera muestra. La suma de 10 a cada valor de la primera muestra causa que las tres medias muestrales se aparten más, con el resultado de que el estadístico de prueba F se incrementa y el valor P disminuye.

Cálculos con tamaños muestrales desiguales

Mientras que los cálculos requeridos para los casos con tamaños muestrales iguales son razonables, las cosas se complican bastante cuando los tamaños muestrales no son iguales. Se aplica el mismo razonamiento básico, porque calculamos un estadístico de prueba F que es el cociente de dos estimados diferentes de la varianza poblacional común σ^2 , pero esos estimados implican medidas *ponderadas* que toman en cuenta los tamaños muestrales, tal como se indica a continuación.

$$F = \frac{\text{varianza entre muestras}}{\text{varianza dentro de muestras}} = \frac{\left[\frac{\sum n_i(\bar{x}_i - \bar{\bar{x}})^2}{k - 1} \right]}{\left[\frac{\sum (n_i - 1)s_i^2}{\sum (n_i - 1)} \right]}$$

donde

$\bar{\bar{x}}$ = media de todos los valores muestrales combinados

k = número de medias poblacionales que se están comparando

n_i = número de valores en la i -ésima muestra

\bar{x}_i = media de los valores en la i -ésima muestra

s_i^2 = varianza de los valores en la i -ésima muestra

El factor de n_i está incluido de manera que las muestras más grandes llevan más peso. El denominador del estadístico de prueba es sencillamente la media de las varianzas muestrales, pero se trata de una media ponderada cuyos pesos se basan en los tamaños muestrales.

Como el cálculo de este estadístico de prueba puede conducir a grandes errores de redondeo, los diferentes programas estadísticos de cómputo suelen emplear una expresión distinta (pero equivalente) que implica la notación de la SC (suma de



La ética en los reportes de encuestas

La American Association for Public Opinion Research creó un código de ética para aplicarse en los reportes de noticias de resultados de encuesta. Este código requiere que se incluya lo siguiente: 1. identificación del patrocinador, 2. fecha de la realización de la encuesta, 3. tamaño de la muestra, 4. naturaleza de la población muestreada, 5. tipo de encuesta utilizada y 6. redacción exacta de las preguntas de la encuesta. Las encuestas financiadas por el gobierno de Estados Unidos se someten a una verificación que evalúa el riesgo para los sujetos encuestados, el mérito científico de la encuesta y la garantía del consentimiento de los sujetos para participar.

cuadrados) y los CM (cuadrados medios). A pesar de que la siguiente notación y sus componentes son complicados y tediosos, la idea básica es la misma: el estadístico de prueba F es una razón con un numerador que refleja la variación *entre* las medias de las muestras y un denominador que refleja la variación *dentro* de las muestras. Si las poblaciones tienen medias iguales, el cociente F tiende a ser pequeño, pero si las medias poblacionales no son iguales, el cociente F tiende a ser significativamente grande. A continuación se describen los componentes más importantes del método ANOVA.

La SC (total) o suma total de cuadrados es una medida de la variación total \bar{x} en todos los datos muestrales combinados.

Fórmula 12-1

$$SC(\text{total}) = \sum(x - \bar{x})^2$$

La SC(total) se puede separar en los componentes de la SC(del tratamiento) y la SC(del error), descritas como sigue.

La SC(del tratamiento), también llamada SC(del factor), SC(entre grupos) o SC(entre muestras), es una medida de la variación entre las medias muestrales.

Fórmula 12-2

$$\begin{aligned} SC(\text{del tratamiento}) &= n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \cdots + n_k(\bar{x}_k - \bar{\bar{x}})^2 \\ &= \sum n_i(\bar{x}_i - \bar{\bar{x}})^2 \end{aligned}$$

Si las medias poblacionales ($\mu_1, \mu_2, \dots, \mu_k$) son iguales, entonces las medias muestrales $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ tenderán a acercarse entre sí y también a acercarse a $\bar{\bar{x}}$. El resultado será un valor de SC(del tratamiento) relativamente pequeño. Sin embargo, si las medias poblacionales no son todas iguales, entonces al menos una de $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ tenderá a estar lejos de las demás y también de $\bar{\bar{x}}$. El resultado será un valor relativamente grande de SC(del tratamiento).

La SC(del error), también conocida como SC(dentro de grupos) o SC(dentro de muestras), es una suma de cuadrados que representa la variación que se supone común a todas las poblaciones consideradas.

Fórmula 12-3

$$\begin{aligned} SC(\text{del error}) &= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2 \\ &= \sum(n_i - 1)s_i^2 \end{aligned}$$

Dadas las expresiones anteriores para SC(total), SC(del tratamiento) y SC(del error), siempre deben mantenerse las siguientes relaciones.

Fórmula 12-4

$$SC(\text{total}) = SC(\text{del tratamiento}) + SC(\text{del error})$$

SC(del tratamiento) y SC(del error) son ambas sumas de cuadrados, y si dividimos cada una de ellas entre su número correspondiente de grados de libertad, obtenemos los cuadrados *medios*. Algunas de las siguientes expresiones para los cuadrados medios incluyen la notación N :

$$N = \text{número total de valores en todas las muestras combinadas}$$

CM(del tratamiento) es un cuadrado medio de tratamiento, que se obtiene como sigue:

$$\text{Fórmula 12-5} \quad \text{CM}(\text{del tratamiento}) = \frac{\text{SC}(\text{del tratamiento})}{k - 1}$$

CM(del error) es un cuadrado medio del error, que se obtiene como sigue:

$$\text{Fórmula 12-6} \quad \text{CM}(\text{del error}) = \frac{\text{SC}(\text{del error})}{N - k}$$

CM(total) es un cuadrado medio de la variación total, que se obtiene como sigue:

$$\text{Fórmula 12-7} \quad \text{CM}(\text{total}) = \frac{\text{SC}(\text{total})}{N - 1}$$

Estadístico de prueba para ANOVA con tamaños muestrales desiguales

Al contrastar la hipótesis nula $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ con la hipótesis alternativa de que no todas estas medias son iguales, el estadístico de prueba

$$\text{Fórmula 12-8} \quad F = \frac{\text{CM}(\text{del tratamiento})}{\text{CM}(\text{del error})}$$

tiene una distribución F (cuando la hipótesis nula H_0 es verdadera) con grados de libertad dados por

$$\text{grados de libertad del numerador} = k - 1$$

$$\text{grados de libertad del denominador} = N - k$$

Este estadístico de prueba es esencialmente el mismo que se utilizó antes, y su interpretación también es igual a la descrita con anterioridad. El denominador depende únicamente de las varianzas muestrales que miden la variación dentro de los tratamientos y no se ve afectado por las diferencias entre las medias muestrales. En contraste, el numerador se ve afectado por las diferencias entre las medias muestrales. Si las diferencias entre las medias muestrales son extremas, causarán que el numerador sea excesivamente grande, por lo que F también será excesivamente grande. Como consecuencia, los valores muy grandes de F sugieren medias desiguales y, por lo tanto, la prueba ANOVA es de cola derecha.

Las tablas constituyen un formato adecuado para resumir los resultados más importantes en los cálculos del ANOVA, y la tabla 12-3 tiene un formato que suele utilizarse en el despliegue de resultados de las computadoras. (Véase las pantallas anteriores de STATDISK, Minitab y Excel). Las cifras de la tabla 12-3 resultan de los datos de la tabla 12-1.

Diseño del experimento En el análisis de varianza de un factor (o de una entraña) utilizamos un factor como base para separar los datos en diferentes categorías.

Tabla 12-3 Tabla de ANOVA para los datos de la tabla 12-1					
Fuente de variación	Suma de cuadrados (SC)	Grados de libertad	Cuadrado medio (CM)	Estadístico de prueba <i>F</i>	Valor <i>P</i>
Tratamientos	4.6824	3	1.5608	5.7314	0.0073
Error	4.3572	16	0.2723		
Total	9.0397	19			

Si concluimos que las diferencias entre las medias son significativas, no podemos estar completamente seguros de que las diferencias se puedan explicar por el factor utilizado. Es posible que la variación de algún otro factor desconocido sea el responsable. Una forma de reducir el efecto de factores extraños es planear el experimento de manera que tenga un **diseño completamente aleatorizado**, en el que se da a cada elemento la misma posibilidad de pertenecer a las diferentes categorías o tratamientos. Por ejemplo, usted podría asignar sujetos a dos diferentes grupos de tratamiento y a un grupo placebo por medio de un proceso de selección aleatoria equivalente a sacar papeles de un tazón. Otra forma de reducir el efecto de factores extraños es el uso de un **diseño rigurosamente controlado**, en el cual los elementos se eligen cuidadosamente de manera que el resto de los factores no tengan variabilidad. En general, los buenos resultados requieren que el experimento se diseñe y ejecute de forma cuidadosa.

Identificación de medias diferentes

Después de hacer una prueba con el análisis de varianza, podemos concluir que existe evidencia suficiente para rechazar una aseveración de igualdad de medias poblacionales, pero no podemos concluir a partir de un ANOVA que alguna media *en particular* sea diferente de las demás. Existen varios procedimientos formales e informales para identificar las medias específicas que son diferentes. Un método informal consiste en usar la misma escala para construir gráficas de cuadro con los conjuntos de datos, para ver si uno o más de ellos es muy diferente de los otros. Otro método consiste en construir estimados de intervalos de confianza de las medias a partir de los conjuntos de datos, y luego comparar esos intervalos de confianza para ver si uno o más de ellos no se traslanan con los demás.

Anteriormente señalamos que no es adecuado aparear las muestras y realizar pruebas de hipótesis individuales con los procedimientos descritos en la sección 9-3. Con cuatro poblaciones, este método (hacer dos a la vez) requiere de seis pruebas de hipótesis diferentes, de manera que si cada prueba se realiza con un nivel de significancia de 0.05, el nivel general de confianza para las seis pruebas sería demasiado bajo, como 0.95⁶ (o 0.735), de manera que el nivel de significancia sería tan alto como $1 - 0.735 = 0.265$. Este nivel de significancia indica que el riesgo de cometer un error tipo I (encontrar una diferencia en uno de los pares cuando en realidad no existe tal diferencia), es demasiado alto.

Existen varios procedimientos para identificar cuáles medias difieren de las demás. Algunas de las pruebas, llamadas **pruebas de rango**, nos permiten localizar subconjuntos de medias que no son significativamente diferentes de las demás.

Otras pruebas, denominadas **pruebas de comparación múltiple**, utilizan pares de medias, pero hacen ajustes para superar el problema de tener un nivel de significancia que aumenta conforme se incrementa el número de pruebas individuales. No existe consenso sobre cuál es la mejor prueba, pero algunas de las más comunes son la prueba de Duncan, la prueba de Student-Newman-Keuls (o prueba SNK), la prueba de Tukey (o prueba de diferencia significativa honesta de Tukey), la prueba de Scheffé, la prueba de Dunnett, la prueba de la diferencia mínima significativa y la prueba de Bonferroni. Ahora aplicaremos esta última con el fin de ver un ejemplo de uno de los métodos para identificar cuáles medias difieren de las demás. El procedimiento es el siguiente:

Prueba de comparación múltiple de Bonferroni

- Paso 1. Realice una prueba t separada para cada par de muestras, pero haga los ajustes que se describen en los siguientes pasos.
- Paso 2. Para obtener un estimado de la varianza σ^2 que es común a todas las poblaciones implicadas, utilice el valor del CM(del error), que utiliza todos los datos muestrales disponibles. El valor del CM(del error) generalmente se obtiene al realizar la prueba del análisis de varianza. Utilice el valor del CM(del error) para calcular el valor del estadístico de prueba t , como se indica abajo. Este estadístico de prueba en particular se basa en la opción de la muestra 1 y la muestra 4; cambie los subíndices y utilice otro par de muestras hasta haber probado todos los pares de muestras posibles.

$$t = \frac{\bar{x}_1 - \bar{x}_4}{\sqrt{\text{CM(error)} \cdot \left(\frac{1}{n_1} + \frac{1}{n_4} \right)}}$$

- Paso 3. Después de calcular el valor del estadístico de prueba t para un par específico de muestras, calcule el valor t crítico o el valor P , pero haga el siguiente ajuste para que el nivel de significancia general no se salga de control.

Valor P : Utilice el estadístico de prueba t con $gl = N - k$, donde N es el número total de valores muestrales y k es el número de muestras, y calcule el valor P de la manera acostumbrada, pero ajústelo multiplicándolo por el número de pares de muestras diferentes posibles. (Por ejemplo, con cuatro muestras, existen seis pares posibles diferentes, de manera que el valor P se ajusta multiplicándolo por 6).

Valor crítico: Cuando calcule el valor crítico, ajuste el nivel de significancia α dividiéndolo entre el número de pares de muestras posibles. (Por ejemplo, con cuatro muestras, existen seis pares posibles diferentes, de manera que el nivel de significancia se ajusta dividiéndolo entre 6).

Observe que en el paso 3 del procedimiento de Bonferroni se realiza una prueba individual con un nivel de significancia mucho más bajo, o bien, el valor P aumenta de manera importante. Por ejemplo, si tenemos cuatro muestras, existen seis pares diferentes de muestras posibles. Si utilizamos un nivel de significancia de 0.05, cada una de las seis comparaciones apareadas utiliza un nivel de significancia de $0.05/6 = 0.00833$. Por lo tanto, el rechazo de la igualdad de medias requiere de diferencias que estén muy separadas. El grado de confianza para las

seis pruebas puede ser tan bajo como $(1 - 0.00833)^6 = 0.95$, de manera que la posibilidad de cometer un error tipo I en una de las seis pruebas es de alrededor de 0.05. Este ajuste en el paso 3 compensa el hecho de que estamos realizando varias pruebas en vez de una sola.

EJEMPLO Uso de la prueba de Bonferroni En un ejemplo previo de esta sección se utilizó el análisis de varianza con los datos muestrales de la tabla 12-1. Concluimos que existe evidencia suficiente para justificar el rechazo de la aseveración de igualdad de medias. Utilice la prueba de Bonferroni, con un nivel de significancia de 0.05, para identificar cuál de las medias difiere de las demás.

SOLUCIÓN La prueba de Bonferroni requiere de una prueba t separada para cada diferente par de muestras posible. Las hipótesis nulas que deben probarse son las siguientes:

$$\begin{array}{lll} H_0: \mu_1 = \mu_2 & H_0: \mu_1 = \mu_3 & H_0: \mu_1 = \mu_4 \\ H_0: \mu_2 = \mu_3 & H_0: \mu_2 = \mu_4 & H_0: \mu_3 = \mu_4 \end{array}$$

Comenzamos con $H_0: \mu_1 = \mu_4$. En la tabla 12-1 observamos que $\bar{x}_1 = 0.184$, $n_1 = 5$, $\bar{x}_4 = 1.334$ y $n_4 = 5$. A partir de los resultados anteriores del análisis de varianza, también sabemos que para los datos de la tabla 12-3, CM(del error) = 0.2723275. Ahora podemos evaluar el estadístico de prueba:

$$t = \frac{\bar{x}_1 - \bar{x}_4}{\sqrt{\text{CM}(\text{del error}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_4} \right)}} = \frac{0.184 - 1.334}{\sqrt{0.2723275 \cdot \left(\frac{1}{5} + \frac{1}{5} \right)}} = -3.484$$

El número de grados de libertad es $gl = N - k = 20 - 4 = 16$. Con un estadístico de prueba $t = -3.484$ y con $gl = 16$, el valor P de dos colas es 0.003065, pero ajustamos este valor P al multiplicarlo por 6 (el número de diferentes pares de muestras posibles) para obtener un valor P final de 0.01839. Como este valor P es pequeño (menor que 0.05), rechazamos la hipótesis nula. Parece que las muestras 1 y 4 tienen medias significativamente diferentes. [Si queremos obtener el valor crítico, usamos $gl = 16$ y un nivel de significancia ajustado de $0.05/6 = 0.00833$. (Este nivel de significancia ajustado no se aleja mucho del área de 0.01, que se incluye en la tabla A-3. Por lo tanto, los valores t críticos para la prueba de dos colas están cerca de -2.921 y 2.921). Los valores críticos reales son -3.008 y 3.008 . El estadístico de prueba de -3.484 se encuentra en la región crítica, de manera que nuevamente rechazamos la hipótesis nula de que $\mu_1 = \mu_4$].

En vez de continuar haciendo las pruebas de hipótesis separadas para los cinco pares restantes, observe la pantalla de SPSS, que incluye todos los resultados de la prueba de Bonferroni. (El tercer renglón de los resultados numéricos corresponde a los resultados obtenidos aquí). La pantalla indica que la media de la muestra 4 difiere significativamente de las otras tres medias muestrales. Con base en la prueba de Bonferroni, parece que los pesos de los álamos tratados con fertilizante y riego tienen una media diferente de las tres categorías de tratamiento restantes.

SPSS Bonferroni Results

Multiple Comparisons						
		Dependent Variable: WEIGHT				
		Bonferroni				
(I) SAMPLE	(J) SAMPLE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.1480	.33005	1.000	-1.1409	.8449
	3.00	.0200	.33005	1.000	-.9729	1.0129
	4.00	-1.1500*	.33005	.018	-2.1429	-.1571
2.00	1.00	.1480	.33005	1.000	-.8449	1.1409
	3.00	.1680	.33005	1.000	-.8249	1.1609
	4.00	-1.0020*	.33005	.047	-1.9949	-.0091
3.00	1.00	-.0200	.33005	1.000	-1.0129	.9729
	2.00	-.1680	.33005	1.000	-1.1609	.8249
	4.00	-1.1700*	.33005	.016	-2.1629	-.1771
4.00	1.00	1.1500*	.33005	.018	.1571	2.1429
	2.00	1.0020*	.33005	.047	.0091	1.9949
	3.00	1.1700*	.33005	.016	.1771	2.1629

*. The mean difference is significant at the .05 level.

Uso de la tecnología

STATDISK Anote los datos en las columnas de la ventana de datos. Seleccione **Analysis** de la barra del menú principal,

luego seleccione **One-Way Analysis of Variance** y proceda a elegir las columnas de los datos muestrales. Haga clic en **Evaluate** al finalizar.

MINITAB Primero ingrese los datos muestrales en las columnas C1, C2, C3, ... Despues seleccione **Stat**, **ANOVA**, **ONE-WAY(UNSTACKED)** e introduzca C1 C2 C3 ... en el recuadro identificado como "Responses" (en columnas separadas).

EXCEL Primero anote los datos en las columnas A, B, C, ... Despues seleccione **Tools** de la barra del menú principal, luego **Data Analysis**, seguido por **Anova: Single**

Factor. En el cuadro de diálogo, introduzca el rango que contiene los datos muestrales. (Por ejemplo, ingrese A1:C12 si el primer valor está en el renglón 1 de la columna A y el último dato está en el renglón 12 de la columna C).

TI-83/84 PLUS Primero registre los datos como listas en L1, L2, L3, ..., después presione **STAT**, seleccione **TESTS** y elija la opción **ANOVA**. Ingrese los rótulos de las columnas. Por ejemplo, si los datos están en las columnas L1, L2 y L3, ingrese esas columnas para obtener **ANOVA (L1, L2, L3)** y presione la tecla **ENTER**.

12-2 DESTREZAS Y CONCEPTOS BÁSICOS**Conocimientos estadísticos y pensamiento crítico**

- 1. ANOVA.** ¿Qué es el *análisis de varianza de un factor* y para qué se utiliza?
- 2. Tratamiento.** En el contexto del análisis de varianza, ¿qué es un tratamiento? En este contexto, ¿la palabra "tratamiento" tiene el mismo significado que en el lenguaje cotidiano?

- 3. Entre/dentro.** ¿Qué es la varianza entre muestras y qué es la varianza dentro de muestras?
- 4. Comparación de áreas de estudio.** Un estudiante del College of Newport aplica una prueba de razonamiento abstracto a estudiantes de inglés, matemáticas y ciencias de su universidad, elegidos al azar. Luego, utiliza el análisis de varianza con los datos muestrales y concluye que no todas las medias son iguales. ¿El estudiante puede concluir que, en Estados Unidos, los estudiantes de inglés, matemáticas y ciencias tienen puntuaciones medias de razonamiento abstracto que no son iguales? ¿Por qué?

Los ejercicios 5 y 6 se basan en el conjunto de datos 7 del apéndice B.

- 5. Pesos de álamos.** En el problema del capítulo señalamos que el peso de 1.34 kg parece ser un valor extremo. Si eliminamos ese valor, los resultados de STATDISK son los que se presentan abajo.
- ¿Cuál es la hipótesis nula?
 - ¿Cuál es la hipótesis alternativa?
 - Identifique el valor del estadístico de prueba.
 - Identifique el valor crítico para un nivel de significancia de 0.05.
 - Identifique el valor P .
 - Con base en los resultados anteriores, ¿qué concluye acerca de la igualdad de las medias poblacionales?
 - Compare estos resultados con los que se obtuvieron al incluir el peso de 1.34 kg. ¿El aparente valor extremo de 1.34 kg tiene un gran efecto en los resultados? ¿Altera la conclusión?

STATDISK

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Treatment:	3	5.216	1.739	8.448	3.2874	0.0016
Error:	15	3.087	0.206			
Total:	18	8.303	0.461			

TI-83/84 Plus

```
One-way ANOVA
F=6.141270387
P=.0055664528
Factor
df=3
SS=3.346455
↓ MS=1.115485
```

```
One-way ANOVA
↑ MS=1.115485
Error
df=16
SS=2.9062
MS=.1816375
SxP=.426189512
```

- 6. Pesos de álamos.** El problema del capítulo utiliza los pesos de álamos para el año 1 y el terreno 1. (Véase el conjunto de datos 7 del apéndice B). Si utilizamos los pesos del año 2 y del terreno 1 con la calculadora TI-83/84 Plus, obtenemos los resultados del análisis de varianza que se presentan al margen. Suponga que deseamos utilizar un nivel de significancia de 0.05 para probar la hipótesis nula de que los cuatro tratamientos dan por resultado pesos con la misma media poblacional.

- ¿Cuál es la hipótesis nula?
- ¿Cuál es la hipótesis alternativa?
- Identifique el valor del estadístico de prueba.
- Identifique el valor crítico para un nivel de significancia de 0.05.
- Identifique el valor P .
- Con base en los resultados anteriores, ¿qué concluye acerca de la igualdad de las medias poblacionales?

- 7. Pérdida de peso con diferentes dietas.** En una prueba de los programas Atkins, Zone, Weight Watchers y Ornish para perder peso, 160 sujetos siguieron los programas, de tal manera que cada dieta fue adoptada por 40 sujetos. Se pesó a los individuos un año después de seguir la dieta, y abajo se presentan los resultados del ANOVA, realizado con Excel (según datos de “Comparison of the Atkins, Ornish, Weight Watchers and Zone Diets for Weight Loss and Heart Disease Risk Reduction”, de Dansinger *et al.*, *Journal of the American Medical Association*, vol. 293, núm. 1). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la pérdida media de peso es igual en todas las dietas. Dado que las cantidades medias de pérdida de peso después de un año son -2.1 lb, -3.2 lb, -3.0 lb y -3.3 lb para las cuatro dietas, ¿parece que las dietas son efectivas?

Excel

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	35.99984	3	11.99995	0.35206	0.787709	2.662569
Within Groups	5317.256	156	34.08497			
Total	5353.256	159				

- 8. Pruebas de inflamabilidad de tela en diferentes laboratorios.** Se realizaron pruebas de inflamabilidad en ropa infantil para dormir. Se utilizó la prueba Vertical Semi-restrained, en la que se quemaron pedazos de tela en condiciones controladas. Una vez apagado el fuego, se midió y registró la longitud de la porción quemada. Las mismas muestras de tela se probaron en cinco laboratorios diferentes. A continuación se presentan los resultados del análisis de varianza realizado con Minitab. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los cinco laboratorios producen las mismas puntuaciones medias de prueba.

Minitab

Source	DF	SS	MS	F	P
Lab	4	2.987	0.747	4.53	0.003
Error	50	8.233	0.165		
Total	54	11.219			

- 9. Pesos medios de dulces M&M.** El conjunto de datos 13 del apéndice B incluye los pesos (en gramos) de dulces M&M sencillos, clasificados de acuerdo con su color (rojo, anaranjado, amarillo, café, azul y verde). Se utilizó STATDISK para hacer un análisis de varianza con esos datos; los resultados se presentan abajo. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el peso medio de los dulces M&M es igual en las seis poblaciones de diferente color. Si Mars Inc. tiene la intención de que las poblaciones de diferentes colores tengan el mismo peso medio, ¿estos resultados sugieren que la empresa tiene un problema que requiere una acción correctiva?

STATDISK

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Treatment:	5	0.006	0.001	0.443	2.3113	0.8173
Error:	94	0.259	0.003			
Total:	99	0.266	0.003			

TI-83/84 Plus

```
One-way ANOVA
F=38.03789731
P=1.3340195E-6
Factor
df=2
SS=6.914444444
↓ MS=3.457222222
```

- 10. Energía solar en diferentes climas.** Un alumno del autor vive en una casa que tiene un sistema eléctrico solar. A la misma hora, cada día, reunió las lecturas de voltaje de un medidor conectado al sistema y realizó un análisis de varianza con las lecturas obtenidas en tres tipos de días diferentes: soleados, nublados y lluviosos. Los resultados de la calculadora TI-83/84 Plus aparecen al margen. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la lectura media de voltaje es la misma en los tres tipos distintos de días. ¿Existe evidencia suficiente para sustentar una aseveración de medias poblacionales diferentes? Podríamos esperar que un sistema solar proporcione más energía eléctrica los días soleados que los días nublados o lluviosos. ¿Podemos concluir que los días soleados dan como resultado mayores cantidades de energía eléctrica?

```
One-way ANOVA
↑ MS=3.457222222
Error
df=15
SS=1.363333333
MS=.0908888889
SxP=.301477841
■
```

En los ejercicios 11 y 12, utilice los datos muestrales listados acerca de experimentos de choques de automóviles realizados por la National Transportation Safety Administration. Se adquirieron autos nuevos y se hicieron chocar contra una barrera fija a 35 mi/h; las mediciones se registraron con respecto al maniquí en el asiento del conductor. Los automóviles subcompactos son el Ford Escort, Honda Civic, Hyundai Accent, Nissan Sentra y Saturn SL4. Los automóviles compactos son Chevrolet Cavalier, Dodge Neon, Mazda 626 DX, Pontiac Sunfire y Subaru Legacy. Los automóviles medianos son Chevrolet Camaro, Dodge Intrepid, Ford Mustang, Honda Accord y Volvo S70. Los automóviles grandes son Audi A8, Cadillac Deville, Ford Crown Victoria, Oldsmobile Aurora y Pontiac Bonneville.

- 11. Traumatismo craneal en un choque de automóvil.** A continuación se presentan los datos de traumatismo craneal (de acuerdo con el *Head Injury Criterion*, HIC). Utilice un nivel de significancia de 0.05 para probar la hipótesis nula de que las diferentes categorías de peso tienen la misma media. ¿Sugieren los datos que los automóviles grandes son más seguros?

Subcompacto:	681	428	917	898	420
Compacto:	643	655	442	514	525
Mediano:	469	727	525	454	259
Grande:	384	656	602	687	360

- 12. Desaceleración del pecho en un choque de automóvil.** A continuación se presentan datos de desaceleración del pecho (*g*). Utilice un nivel de significancia de 0.05 para probar la hipótesis nula de que las distintas categorías de peso tienen la misma media. ¿Sugieren los datos que los automóviles más grandes son más seguros?

Subcompacto:	55	47	59	49	42
Compacto:	57	57	46	54	51
Mediano:	45	53	49	51	46
Grande:	44	45	39	58	44

- 13. Ejercicio y estrés.** Se realizó un estudio para investigar los efectos del ejercicio sobre el estrés. En la siguiente tabla se listan las lecturas de la presión sanguínea sistólica (en mmHg) de sujetos, antes de iniciar 25 minutos de ejercicio aeróbico en bicicleta y antes de generarles estrés por medio de una prueba de aritmética y otra de expresión verbal (según datos de “Sympathoadrenergic Mechanisms in Reduced Hemodynamic Stress Responses after Exercise”, de Kim Brownley *et al.*, *Medicine and Science in Sports and Exercise*, vol. 35, núm. 6). Utilice un nivel de significancia de 0.05 para probar la aseveración de que los diferentes grupos de sujetos tienen la misma presión sanguínea media. Con base en los resultados, ¿se puede considerar que los grupos provienen de la misma población?

Mujer/afroam.	Hombre/afroam.	Mujer/caucásica	Hombre/caucásico
117.00	115.67	119.67	124.33
130.67	120.67	106.00	111.00
102.67	133.00	108.33	99.67
93.67	120.33	107.33	128.33
96.33	124.67	117.00	102.00
92.00	118.33	113.33	127.33

- 14. Arqueología: Anchura de cráneos de distintas épocas.** Se obtuvieron muestras de anchuras de cráneos de hombres egipcios de distintas épocas, y abajo se presentan las mediciones (según datos de *Ancient Races of the Thebaid*, de Thomson y Randall-Maciver). Los cambios en la forma de la cabeza a lo largo del tiempo sugieren que hubo mestizaje con poblaciones inmigrantes. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las distintas épocas no tienen la misma media. ¿Qué concluye?

4000 a. C.	1850 a. C.	150 d. C.
131	129	128
138	134	138
125	136	136
129	137	139
132	137	141
135	129	142
132	136	137
134	138	145
138	134	137

En los ejercicios 15 y 16, utilice el conjunto de datos del apéndice B.

- 15. Conjunto de datos del apéndice B: Pesos de monedas de un centavo.** Remítase al conjunto de datos 14 del apéndice B y utilice los pesos de los peniques o centavos “con cabeza de indio”, de los peniques “de trigo”, de los peniques anteriores a 1983 y de los posteriores a 1983. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el peso medio de los peniques es el mismo en las cuatro categorías diferentes. Con base en los resultados, ¿parecería que una máquina contadora de monedas puede manejar los pesos de los peniques de la misma forma?
- 16. Conjunto de datos del apéndice B: Distancias de *home runs*.** Remítase al conjunto de datos 17 del apéndice B. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los *home runs* anotados por Barry Bonds, Mark McGwire y Sammy Sosa tienen distancias medias que no son iguales. ¿Explican las distancias de los *home runs* el hecho de que hasta hoy, Barry Bonds tiene el mayor número de *home runs* en una temporada, mientras que Mark McGwire posee el segundo número más grande de *home runs*? (En años recientes se ha afirmado que los jugadores de béisbol usaban esteroides para aumentar su fuerza, lo que los ayudó a batear las pelotas más lejos).

12-2 MÁS ALLÁ DE LO BÁSICO

- 17. Pruebas equivalentes.** En este ejercicio usted comprobará que cuando tiene dos conjuntos de datos muestrales, la prueba *t* agrupada para muestras independientes y el método ANOVA de esta sección son equivalentes. Remítase a los pesos de los álamos en la tabla 12-1, pero utilice únicamente los datos de los árboles que no recibieron tratamiento y de los que sólo recibieron fertilizante.
- Utilice un nivel de significancia de 0.05 y el método de la sección 9-3 para probar la aseveración de que las dos muestras provienen de poblaciones con la misma media. (Suponga que ambas poblaciones tienen la misma varianza).
 - Utilice un nivel de significancia de 0.05 y el método ANOVA de esta sección para probar la aseveración planteada en el inciso *a*).
 - Verifique que los cuadrados del estadístico de prueba *t* y el valor crítico del inciso *a*) son iguales al estadístico de prueba *F* y al valor crítico del inciso *b*).
- 18. Cálculo de componentes del ANOVA.** El ejercicio 9 se basa en los pesos de los 100 dulces M&M incluidos en el conjunto de datos 13 del apéndice B. Si se utilizan los pesos de todos los dulces M&M de la bolsa llena, $SC(\text{del tratamiento}) = 0.00644$, $gl(\text{del tratamiento}) = 5$, $SC(\text{del error}) = 1.47320$ y $gl(\text{del error}) = 459$.
- ¿Cuántos dulces M&M hay en la bolsa llena?
 - Calcule los valores del $CM(\text{del tratamiento})$ y del $CM(\text{del error})$.
 - Calcule el valor del estadístico de prueba *F* y del valor *P*.
 - Compare los resultados con los obtenidos en el ejercicio 9.

- 19. Uso de la prueba de Tukey.** En esta sección se incluyó una pantalla de resultados de la prueba de Bonferroni con los datos de la tabla 12-1. A continuación se muestran los resultados generados por SPSS de la prueba de Tukey. Compare los resultados de la prueba de Tukey con los de la prueba de Bonferroni.

SPSS

Multiple Comparisons						
				95% Confidence Interval		
(I) SAMPLE	(J) SAMPLE	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
1.00	2.00	.1480	.33005	.969	-1.0923	.7963
	3.00	.0200	.33005	1.000	-.9243	.9643
	4.00	-1.1500*	.33005	.015	-2.0943	-.2057
2.00	1.00	.1480	.33005	.969	-.7963	1.0923
	3.00	.1680	.33005	.956	-.7763	1.1123
	4.00	-1.0020*	.33005	.036	-1.9463	-.0577
3.00	1.00	-.0200	.33005	1.000	-.9643	.9243
	2.00	-.1680	.33005	.956	-1.1123	.7763
	4.00	-1.1700*	.33005	.013	-2.1143	-.2257
4.00	1.00	1.1500*	.33005	.015	.2057	2.0943
	2.00	1.0020*	.33005	.036	.0577	1.9463
	3.00	1.1700*	.33005	.013	.2257	2.1143

*The mean difference is significant at the .05 level.

- 20. Uso de la prueba de Bonferroni.** A continuación se presentan los resultados parciales de la prueba de Bonferroni con los datos muestrales utilizados en el ejercicio 14. Suponga que se utiliza un nivel de significancia de 0.05.
- ¿Qué nos indica la pantalla de resultados?
 - Utilice el procedimiento de la prueba de Bonferroni para calcular el valor de significancia para la prueba de una diferencia significativa entre la anchura media de los cráneos de 1850 a. C. y de 150 d. C. Identifique el estadístico de prueba y el valor *P* o los valores críticos. ¿Qué indican los resultados?

				95% Confidence Interval		
(I) SAMPLE	(J) SAMPLE	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
1.00	2.00	-1.7778	1.95105	1.000	-6.7991	3.2435
	3.00	-5.4444*	1.95105	.030	-10.4657	-.4231

12-3 ANOVA de dos factores

Concepto clave El procedimiento de análisis de varianza que estudiamos en la sección 12-2 se conoce como análisis de varianza de *un factor* (o análisis de varianza de una entrada), porque los datos están categorizados en grupos de acuerdo con un *solo factor* (o tratamiento). En esta sección estudiaremos el método del *análisis de varianza de dos factores*, que se utiliza con datos separados en categorías formadas de acuerdo con *dos factores*. El método de esta sección requiere que primero hagamos una prueba de *interacción* entre los dos factores. Después, hacemos una prueba para determinar si el factor de renglón tiene algún efecto, y también para determinar si el factor de columna tiene algún efecto.

La tabla 12-4 es un ejemplo de datos clasificados de acuerdo con dos factores. Un factor es la variable de renglón del terreno (terreno 1 y terreno 2), en tanto que el segundo factor es la variable de columna de tratamiento (ninguno, fertilizante, riego, fertilizante y riego). A las subcategorías de la tabla 12-4 se les conoce como *celdas*, de manera que la tabla 12-4 tiene ocho celdas, con cinco valores cada una.

Cuando analizamos los datos muestrales de la tabla 12-4 estudiamos el análisis de varianza de un solo factor, por lo que parecería razonable hacer simplemente un ANOVA de una entrada para el factor del terreno y otro ANOVA de una entrada para el factor de tratamiento. Por desgracia, si realizamos dos ANOVA de un factor por separado, perdemos información e ignoramos totalmente una característica muy importante: el efecto de la interacción entre los dos factores.

Definición

Existe una **interacción** entre dos factores si el efecto de uno de los factores cambia en las diferentes categorías del otro factor.

Como ejemplo de una *interacción* entre dos factores, considere el apareamiento de alimentos. La crema de maní y la mermelada interactúan bien, pero la salsa de tomate y el helado interactúan de tal forma que el resultado es un sabor desagradable. Los médicos deben tener cuidado de evitar la prescripción de fármacos que tienen interacciones que producen efectos adversos. Se descubrió que el fármaco

Tabla 12-4 Pesos de álamos (en kg)				
	Sin tratamiento	Fertilizante	Riego	Fertilizante y riego
Terreno 1 (fértil, húmedo)	0.15	1.34	0.23	2.03
	0.02	0.14	0.04	0.27
	0.16	0.02	0.34	0.92
	0.37	0.08	0.16	1.07
	0.22	0.08	0.05	2.38
Terreno 2 (arenoso, seco)	0.60	1.16	0.65	0.22
	1.11	0.93	0.08	2.13
	0.07	0.30	0.62	2.33
	0.07	0.59	0.01	1.74
	0.44	0.17	0.03	0.12



Encuestas y psicólogos

Los resultados de las encuestas se pueden ver muy afectados por la redacción de las preguntas. Las personas interpretan de forma diferente una frase como “durante los últimos años”. Durante los últimos años (en realidad desde 1980), los investigadores de encuestas y los psicólogos han trabajado en conjunto para mejorar las encuestas disminuyendo el sesgo e incrementando la exactitud. En un caso, los psicólogos estudiaron el hallazgo de que del 10 al 15 por ciento de los encuestados afirmaron haber votado en la última elección, cuando en realidad no lo hicieron. Experimentaron con teorías de problemas de memoria, el deseo de ser considerado responsable y la tendencia que manifiestan quienes generalmente votan de decir que votaron en la elección más reciente, aun cuando no lo hayan hecho. Se encontró que sólo esta última explicación formaba parte del problema.

antimicótico Nizoral (ketoconazole) interactuaba con el fármaco antihistamínico Seldane (terfenadine) de tal manera que el Seldane no se metabolizaba de manera adecuada, provocando arritmias cardíacas en algunos pacientes. Posteriormente el Seldane fue retirado del mercado.

Exploración de los datos Exploraremos los datos de la tabla 12-4 calculando la media de cada celda y construyendo una gráfica. En la tabla 12-5 se listan las medias de las celdas individuales. Esas medias van de 0.164 kg hasta 1.334 kg, de manera que parecen variar considerablemente. La figura 12-3 ilustra gráficas de esas medias e indica que las medias del terreno 2 son mayores que las medias del terreno 1 para tres de las cuatro categorías de tratamiento. Puesto que los segmentos lineales del terreno 2 parecen ser aproximadamente *paralelos* a los segmentos lineales correspondientes del terreno 1, aparentemente los pesos por terreno se comportan de la misma forma para las distintas categorías del tratamiento, de manera que no hay un efecto de interacción. En general, si una gráfica como la de la figura 12-3 da por resultado segmentos lineales que son aproximadamente *paralelos*, entonces tenemos evidencia de que *no hay una interacción* entre las variables de renglón y de columna. Si los segmentos lineales del terreno 2 no fueran paralelos a los segmentos lineales del terreno 1, tendríamos evidencia de una interacción entre el terreno y el tratamiento. Estas observaciones basadas en la tabla 12-5 y en la figura 12-3 son muy subjetivas, por lo que procederemos con el método más objetivo del análisis de varianza de dos factores.

Tabla 12-5 Medias (en kg) de las celdas de la tabla 12-4

	Tratamiento			
	Ninguno	Fertilizante	Riego	Fertilizante y riego
Terreno 1 (fértil, húmedo)	0.184	0.332	0.164	1.334
Terreno 2 (arenoso, seco)	0.458	0.630	0.278	1.308

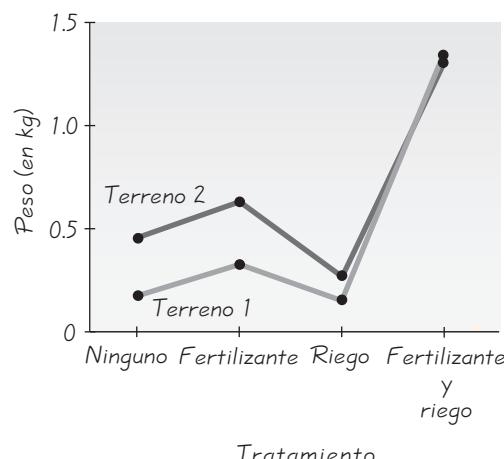


Figura 12-3 Gráfica de interacción de medias (en kg) de las celdas de la tabla 12-4

Cuando aplicamos el ANOVA de dos factores para los datos de la tabla 12-4, consideraremos tres posibles efectos sobre el peso de los álamos: **1.** los efectos de una interacción entre el terreno y el tratamiento; **2.** los efectos del terreno; **3.** los efectos del tratamiento. Los cálculos son bastante complicados, *de manera que supondremos que se utilizó un programa de cómputo o una calculadora TI-83/84 Plus*. (Al final de esta sección se describen los procedimientos para el uso de herramientas tecnológicas). A continuación se presentan los resultados de Minitab para los datos de la tabla 12-4.

Minitab

Source	DF	SS	MS	F	P
Site	1	0.2722	0.27225	0.81	0.374
Treatment	3	7.5470	2.51567	7.50	0.001
Interaction	3	0.1716	0.05721	0.17	0.915
Error	32	10.7267	0.33521		
Total	39	18.7176			

Ahora presentaremos los requisitos y el procedimiento básico para el análisis de varianza (ANOVA) de dos factores. El procedimiento también se resume en la figura 12-3.

Requisitos

1. Para cada celda, los valores muestrales provienen de una población con una distribución que es aproximadamente normal.
2. Las poblaciones tienen la misma varianza σ^2 (o desviación estándar σ).
3. Las muestras son aleatorias simples. (Es decir, las muestras del mismo tamaño tienen la misma probabilidad de ser seleccionadas).
4. Las muestras son independientes entre sí. (Las muestras no están apareadas o asociadas de ninguna manera).
5. Los valores muestrales se categorizan en dos factores. (Ésta es la base del nombre del método: análisis de varianza de *dos factores*).
6. Todas las celdas tienen el mismo número de valores muestrales. (Este diseño se conoce como diseño *balanceado*).

Procedimiento del ANOVA de dos factores (véase la figura 12-4)

Paso 1: *Efecto de interacción:* En el análisis de varianza de dos factores, inicie probando la hipótesis nula de que no existe interacción entre los dos factores. Si utilizamos Minitab para los datos de la tabla 12-4, obtenemos el siguiente estadístico de prueba:

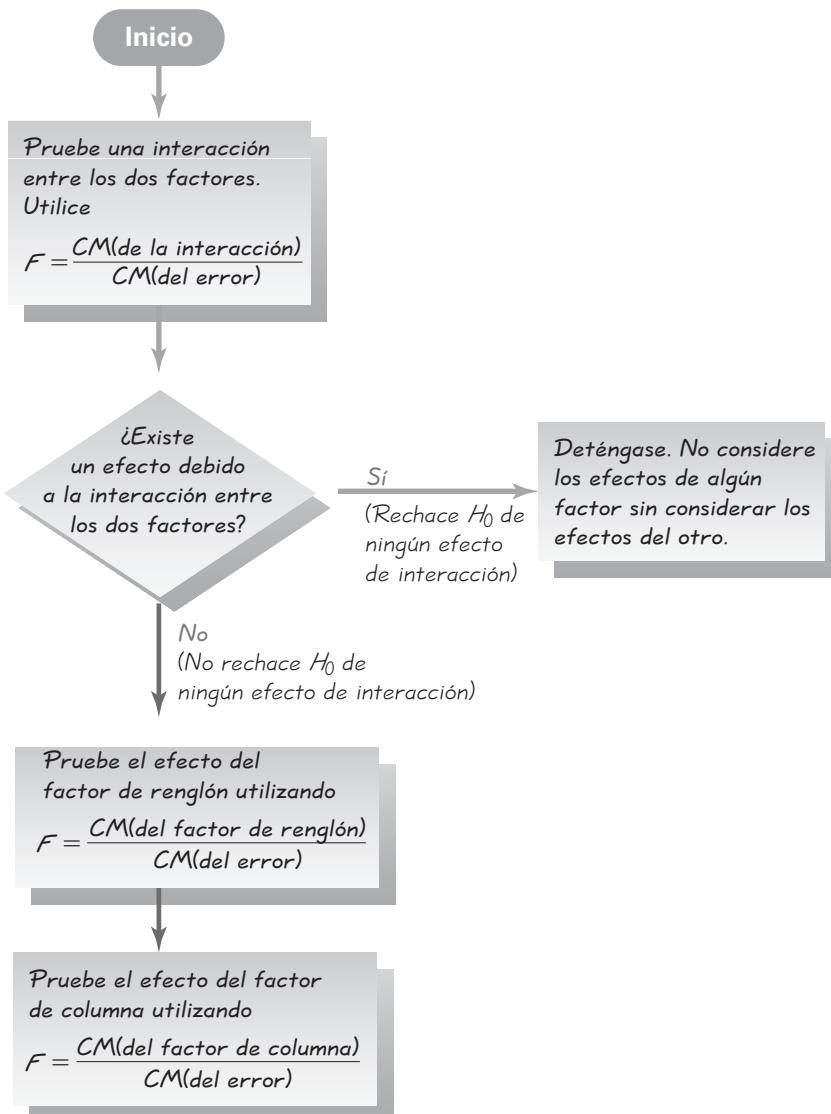
$$F = \frac{CM(\text{de la interacción})}{CM(\text{del error})} = \frac{0.05721}{0.33521} = 0.17$$

Interpretación: El valor P correspondiente aparece en los resultados de Minitab como 0.915, por lo que no rechazamos la hipótesis nula de ninguna interacción entre los dos factores. No parece que los pesos de los álamos estén afectados por una interacción entre el terreno y el tratamiento.

Paso 2: *Efectos renglón/columna:* Si rechazamos la hipótesis nula de ninguna interacción entre factores, entonces tenemos que detenernos aquí; no debemos proceder con las dos pruebas adicionales. (Si existe una interacción

Figura 12-4

Procedimiento del ANOVA de dos factores



entre los factores, no debemos considerar los efectos de alguno de los factores sin considerar los del otro).

Si no rechazamos la hipótesis nula de ninguna interacción entre los factores, entonces debemos proceder a probar las siguientes dos hipótesis:

H_0 : No existen efectos del factor de renglón (es decir, las medias de renglón son iguales).

H_0 : No existen efectos del factor de columna (es decir, las medias de columna son iguales).

En el paso 1 no rechazamos la hipótesis nula de ninguna interacción entre los factores, por lo que procedemos con las siguientes dos pruebas de hipótesis identificadas en el paso 2.

Para el factor de renglón del terreno obtenemos

$$F = \frac{CM(\text{del terreno})}{CM(\text{del error})} = \frac{0.27225}{0.33521} = 0.81$$

Interpretación: Este valor no es significativo, ya que el valor P correspondiente aparece en los resultados de Minitab como 0.374. No rechazamos la hipótesis nula de que no existen efectos por el terreno. Es decir, el terreno no parece tener un efecto sobre el peso de los álamos.

Para el factor de columna del tratamiento obtenemos

$$F = \frac{\text{CM}(\text{del tratamiento})}{\text{CM}(\text{del error})} = \frac{2.51567}{0.33521} = 7.50$$

Interpretación: Este valor es significativo, ya que el valor P correspondiente se indica como 0.001. Por lo tanto, rechazamos la hipótesis nula de ningún efecto del tratamiento. Parece que el tratamiento tiene un efecto sobre el peso de los álamos. Con base en los datos muestrales de la tabla 12-4, concluimos que los diferentes tratamientos parecen tener pesos medios diferentes, aunque al parecer los pesos tienen medias iguales en ambos terrenos. (Vea la figura 12-3 y observe que los segmentos lineales cambian drásticamente en las diferentes categorías de tratamiento, pero los segmentos lineales del terreno 1 y del terreno 2 no son muy diferentes entre sí).

Caso especial: Una observación por celda y ninguna interacción La tabla 12-4 contiene 5 observaciones por celda. Si nuestros datos muestrales consisten únicamente en una observación por celda, perdemos CM(de la interacción), SC(de la interacción) y gl(de la interacción), ya que estos valores están basados en varianzas muestrales calculadas para cada celda individual. Si existe sólo una observación por celda, no hay variación dentro de las celdas individuales y esas varianzas muestrales no pueden ser calculadas. Cuando tenemos una observación por celda procedemos de la siguiente manera: *si parece razonable suponer (con base en el conocimiento de las circunstancias) que no existe interacción entre los dos factores, haga dicha suposición y después proceda como antes a probar las siguientes dos hipótesis por separado:*

H_0 : No existen efectos del factor de renglón.

H_0 : No existen efectos del factor de columna.

Como ejemplo, suponga que tenemos únicamente el primer valor de cada celda de la tabla 12-4, lo que produce los datos de la tabla 12-6. En la tabla 12-6, las dos medias por renglón son 0.9375 y 0.6575. ¿Es esta diferencia significativa, lo que sugiere que existe un efecto debido al terreno? En la tabla 12-6, las cuatro medias de columna son 0.3750, 1.2500, 0.4400 y 1.1250. ¿Son significativas estas diferencias, lo que sugiere que existe un efecto debido al tratamiento? Suponiendo que los pesos no se ven afectados por alguna interacción entre el terreno y el tratamiento,

Tabla 12-6 Pesos (en kg) de álamos en el año 1

	Tratamiento			
	Ninguno	Fertilizante	Riego	Fertilizante y riego
Terreno 1 (fértil, húmedo)	0.15	1.34	0.23	2.03
Terreno 2 (arenoso, seco)	0.60	1.16	0.65	0.22

a continuación se presentan los resultados de Minitab. (Si creemos que existe una interacción, el método descrito aquí no se aplica).

Minitab

Source	DF	SS	MS	F	P
Site	1	0.15680	0.156800	0.28	0.634
Treatment	3	1.23665	0.412217	0.73	0.598
Error	3	1.68690	0.562300		
Total	7	3.08035			

Factor de renglón: Primero empleamos los resultados de la pantalla de Minitab para probar la hipótesis nula de ningún efecto del factor de renglón del terreno.

$$F = \frac{CM(\text{del terreno})}{CM(\text{del error})} = \frac{0.156800}{0.562300} = 0.28$$

Este estadístico de prueba no es significativo, debido a que el valor P correspondiente en la pantalla de Minitab es 0.634. No rechazamos la hipótesis nula; parece que los pesos de los álamos no se ven afectados por el terreno.

Factor de columna: Ahora utilizamos la pantalla de Minitab para probar la hipótesis nula de ningún efecto del factor de columna de la categoría de tratamiento. El estadístico de prueba es

$$F = \frac{CM(\text{de la edad})}{CM(\text{del error})} = \frac{0.412217}{0.562300} = 0.73$$

Este estadístico de prueba no es significativo, ya que el valor P correspondiente en la pantalla de Minitab es 0.598. No rechazamos la hipótesis nula, de manera que parece que el peso del álamo no se ve afectado por el tratamiento. Utilizando los datos de la tabla 12-6, concluimos que los pesos de los álamos no parecen verse afectados por el terreno ni por el tratamiento, pero cuando utilizamos 5 valores en cada celda (tabla 12-4), concluimos que los pesos parecen verse afectados por el tratamiento. Éste es el poder de las muestras grandes.

En esta sección explicamos brevemente una rama importante de la estadística. Pusimos énfasis en la interpretación de resultados de computadora y omitimos los cálculos y fórmulas manuales, que son bastante engorrosos.

12-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. ANOVA de dos factores.** ¿Qué es el *análisis de varianza de dos factores* y para qué se utiliza?
- 2. Interacción.** ¿Qué es una interacción y qué papel desempeña en el análisis de varianza de dos factores?
- 3. Interacción.** Cuando se realiza un análisis de varianza de dos factores, si hay una interacción entre los dos factores, ¿por qué no debemos continuar con pruebas separadas para los efectos de los factores de renglón y de columna?

Uso de la tecnología

STATDISK Haga clic en **Analysis** y seleccione **Two-Way Analysis of Variance**. Complete lo que se le pide en la ventana y luego haga clic en **Continue**. Proceda a ingresar o copiar los datos en la columna “**Values**” y luego haga clic en **Evaluate**.

MINITAB Primero ingrese todos los valores muestrales en la columna C1. Registre los números por renglón correspondientes en la columna C2. Ingrese los números de columna correspondientes en la columna C3. Seleccione **Stat** de la barra del menú principal, después **ANOVA** y luego **Two-Way**. En el cuadro de diálogo ingrese C1 para **Response**, C2 para **Row Factor** y C3 Para **Column Factor**. Haga clic en **OK**. **Sugerencia:** Evite confusiones y ponga rótulos a las columnas C1, C2 y C3 con nombres que tengan algún significado.

EXCEL Para tablas de dos factores con más de un dato por celda: los datos de la misma celda deben listarse en una columna, no en un renglón. Ingrese los rótulos correspondientes al conjunto de datos en la columna A y el renglón 1, como en este ejemplo, que corresponde a la tabla 12-4:

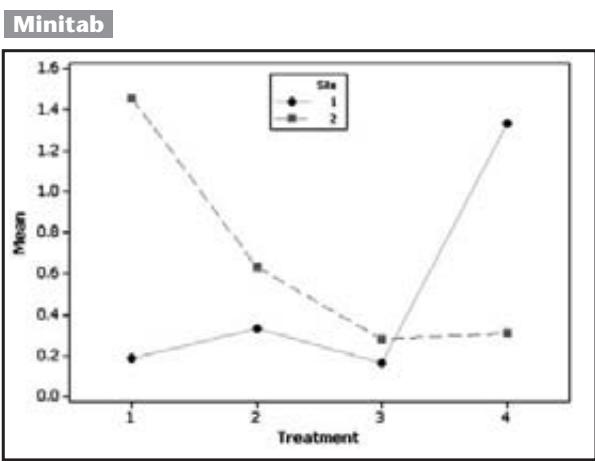
	A	B	C	D	E
1		Ninguno	Fert	Riego	Ferty Riego
2	Terreno 1	0.15	1.34	0.23	2.03
3	Terreno 1	0.02	0.14	0.04	0.27
:	:	:	:	:	:

Después de ingresar los datos muestrales y los rótulos, seleccione **Tools** de la barra del menú principal, luego **Data Analysis** y después **Anova: Two-Factor With Replication**. En el cuadro de diálogo ingrese el rango de entrada. Para los datos de la tabla 12-4, ingrese A1:E11. Para “rows per sample”, introduzca el número de valores en cada celda; ingrese 5 para los datos de la tabla 12-4. Haga clic en **OK**.

Para tablas de dos factores con exactamente un dato por celda, no se requieren los rótulos. Ingrese los datos muestrales como aparecen en la tabla. Seleccione **Tools**, luego **Data Analysis**, después **Anova: Two-Factor Without Replication**. En el cuadro de diálogo, introduzca el rango de entrada únicamente de los valores muestrales; no incluya rótulos en el rango de entrada. Haga clic en **OK**.

TI-83/84 PLUS El programa A1ANOVA de la calculadora TI-83/84 Plus puede descargarse del CD-ROM incluido con este libro. Seleccione el archivo del *software*. El programa debe descargarse a la calculadora; luego, los datos muestrales deben ingresarse como una matriz D con tres columnas. Presione **2nd** y la tecla x^{-1} . Muévase a la derecha hasta **EDIT**, muévase hacia abajo hasta **[D]**, luego presione **ENTER** y proceda a ingresar el número total de valores de datos, seguido por 3 (para las tres columnas). La primera columna de D lista todos los datos muestrales, la segunda columna lista el número de renglón correspondiente y la tercera columna lista el número de columna correspondiente. Después de ingresar todos los datos, los números de renglón y los números de columna en la matriz D, presione **PRGM**, seleccione **A1ANOVA** y presione **ENTER** dos veces; luego elija **RAN BLOCK DESI** (para diseño de bloque aleatorio) y presione **ENTER** dos veces. Seleccione **CONTINUE** y presione **ENTER**. En un momento aparecen los resultados. F(A) es el estadístico de prueba *F* para el factor de renglón, el cual será seguido por el valor *P* correspondiente. F(B) es el estadístico de prueba *F* para el factor de columna, el cual será seguido por el valor *P* correspondiente. (Es necesario presionar **ENTER** para poder ver el resto de los resultados). F(AB) es el estadístico de prueba *F* para el efecto de interacción, y va seguido por el valor *P* correspondiente.

- 4. Interacción.** A continuación aparece una gráfica de interacción generada con Minitab, similar en naturaleza a la figura 12-3. [Como en la figura 12-3, esta gráfica de interacción se basa en pesos de árboles, y los factores son el terreno (terreno 1 y terreno 2) y los tratamientos (ninguno, fertilizante, riego, fertilizante y riego)]. ¿Qué sugiere esta gráfica acerca de la interacción entre los dos factores?



Interpretación de la pantalla de resultados de una computadora: pesos de álamos. Los ejercicios 5 a 7 requieren de la siguiente pantalla de Minitab, que resulta de los pesos de álamos del segundo año que están listados en el conjunto de datos 7 del apéndice B. (La tabla 12-4 de esta sección incluye los pesos del primer año, pero la siguiente pantalla de Minitab se basa en los pesos del segundo año. El factor de renglón es el terreno (terreno 1 y terreno 2) y el factor de columna es el tratamiento (ninguno, fertilizante, riego, fertilizante y riego).

Minitab

Source	DF	SS	MS	F	P
Site	1	0.3258	0.325803	1.43	0.240
Treatment	3	1.7518	0.583949	2.57	0.072
Interaction	3	1.9060	0.635336	2.79	0.056
Error	32	7.2755	0.227360		
Total	39	11.2592			

5. **Efecto de interacción.** Remítase a la pantalla de Minitab y ponga a prueba la hipótesis nula de que los pesos de los álamos no se ven afectados por una interacción entre el terreno y el tratamiento. ¿Qué concluye?
6. **Efecto del terreno.** Remítase a la pantalla de Minitab y suponga que los pesos de los álamos no se ven afectados por una interacción entre el terreno y el tratamiento. ¿Existe evidencia suficiente para sustentar la aseveración de que el terreno tiene un efecto sobre el peso?
7. **Efecto del tratamiento.** Remítase a la pantalla de Minitab y suponga que los pesos de los álamos no se ven afectados por una interacción entre el terreno y el tratamiento. ¿Existe evidencia suficiente para sustentar la aseveración de que el tratamiento tiene un efecto sobre el peso?

Interpretación de una pantalla de resultados de computadora. En los ejercicios 8 a 10, utilice los resultados de Minitab obtenidos a partir de las puntuaciones listadas en la siguiente tabla. Los datos muestrales son calificaciones de la prueba SAT de las secciones verbal y de matemáticas (SAT-I) basadas en estadísticos reportados por el Consejo Universitario.

Verbal

Mujer	646	539	348	623	478	429	298	782	626	533
Hombre	562	525	512	576	570	480	571	555	519	596

Matemáticas

Mujer	484	489	436	396	545	504	574	352	365	350
Hombre	547	678	464	651	645	673	624	624	328	548

Minitab

Source	DF	SS	MS	F	P
Gender	1	52635	52635.0	5.03	0.031
Verb/Math	1	6027	6027.0	0.58	0.453
Interaction	1	31528	31528.2	3.01	0.091
Error	36	376748	10465.2		
Total	39	466938			

8. **Efecto de interacción.** Pruebe la hipótesis nula de que las calificaciones de la prueba SAT no se ven afectadas por una interacción entre el género y la prueba (verbal/de matemáticas). ¿Qué concluye?

9. Efecto del género. Suponga que las calificaciones de la prueba SAT no se ven afectadas por una interacción entre el género y el tipo de prueba (verbal/de matemáticas). ¿Existe suficiente evidencia para sustentar la aseveración de que el género tiene un efecto sobre las calificaciones de la prueba SAT?

10. Efecto del tipo de prueba del SAT. Suponga que las calificaciones de la prueba SAT no se ven afectadas por una interacción entre el género y el tipo de prueba (verbal/de matemáticas). ¿Existe evidencia suficiente para sustentar la aseveración de que el tipo de prueba (verbal/de matemáticas) tiene un efecto sobre las calificaciones del SAT?

Interpretación de una pantalla de resultados de computadora: una observación por celda. En los ejercicios 11 y 12, remítase a la pantalla de resultados de Minitab. Esta pantalla resulta de un estudio en el que se aplicó una prueba de audición a 24 sujetos, empleando cuatro listas diferentes de palabras. Los 24 sujetos tenían una audición normal y las pruebas se llevaron a cabo sin sonido de fondo. El principal objetivo fue determinar si las cuatro listas eran igualmente difíciles de comprender. En la tabla original de las puntuaciones de la prueba de audición, cada celda tiene un dato. Los datos originales provienen del informe “A Study of the Interlist Equivalency of the CID W-22 Word List Presented in Quiet and in Noise”, de Faith Loven, Universidad de Iowa. Los datos originales están disponibles en DASL (Data and Story Library) de Internet.

Minitab

Analysis of Variance for Hearing					
Source	DF	SS	MS	F	P
Subject	23	3231.6	140.5	3.87	0.000
List	3	920.5	306.8	8.45	0.000
Error	69	2506.5	36.3		
Total	95	6658.6			

11. Pruebas de audición: efecto de sujetos. Suponiendo que no existe un efecto sobre las puntuaciones de las pruebas de audición como resultado de una interacción entre el sujeto y las listas, ¿existe evidencia suficiente para sustentar la aseveración de que la selección del sujeto tiene un efecto en la puntuación de las pruebas de audición? Interprete el resultado explicando por qué tiene un sentido práctico.

12. Pruebas de audición: efecto de la lista de palabras. Suponiendo que no existe un efecto sobre las puntuaciones de las pruebas de audición como resultado de una interacción entre el sujeto y las listas, ¿existe evidencia suficiente para sustentar la aseveración de que la selección de la lista de palabras tiene un efecto en la puntuación de las pruebas de audición?

En los ejercicios 13 y 14 utilice un programa de cómputo o una calculadora TI-83/84 Plus para obtener los resultados del análisis de varianza de dos factores.

13. Pulso. La siguiente tabla lista pulsos del conjunto de datos 1 del apéndice B. ¿Se ven afectados los pulsos por una interacción entre el género y la edad? ¿Se ven afectados los pulsos por el género? ¿Se ven afectados los pulsos por la edad?

	Edad		
	Menor de 20	20–40	Mayor de 40
Hombre	96 64 68 60	64 88 72 64	68 72 60 88
Mujer	76 64 76 68	72 88 72 68	60 68 72 64

14. Tiempos de maratón. A continuación se presentan los tiempos (en segundos) de recorrido de la maratón de Nueva York, para corredores elegidos al azar que completaron la prueba. ¿Los tiempos de recorrido se ven afectados por una interacción entre el

género y el grupo de edad? ¿Los tiempos de recorrido se ven afectados por el género?
¿Los tiempos de recorrido se ven afectados por el grupo de edad?

Tiempos (en segundos) de corredores de la maratón de Nueva York

	Edad		
	21–29	30–39	40 y mayores
Hombres	13,615	14,677	14,528
	18,784	16,090	17,034
	14,256	14,086	14,935
	10,905	16,461	14,996
	12,077	20,808	22,146
Mujeres	16,401	15,357	17,260
	14,216	16,771	25,399
	15,402	15,036	18,647
	15,326	16,297	15,077
	12,047	17,636	25,898

12-3 MÁS ALLÁ DE LO BÁSICO

15. **Transformaciones de datos.** Suponga que se utiliza un ANOVA de dos factores para analizar datos muestrales consistentes en más de un dato por celda. ¿De qué manera se ven afectados los resultados del ANOVA en cada uno de los siguientes casos?
- Se añade la misma constante a cada valor muestral.
 - Cada valor muestral se multiplica por la misma constante distinta de cero.
 - Se traspone el formato de la tabla, de manera que se intercambien los factores de renglón y de columna.
 - Se cambia el primer valor muestral de la primera celda, de manera que se convierte en un valor extremo.

Repaso

En la sección 9-3 presentamos un procedimiento para probar la igualdad entre *dos* medias poblacionales, pero en la sección 12-2 utilizamos el análisis de varianza de un factor para probar la igualdad de tres o más medias poblacionales. Este método requiere: **1.** poblaciones distribuidas normalmente, **2.** poblaciones con la misma desviación estándar (o varianza), y **3.** muestras aleatorias simples que sean independientes entre sí. Los métodos del análisis de varianza de un factor se utilizan cuando tenemos tres o más muestras obtenidas de poblaciones que están caracterizadas de acuerdo con un solo factor. Las siguientes son características clave del análisis de varianza de un factor:

- El estadístico de prueba F está basado en la razón de dos estimados diferentes de la varianza poblacional común σ^2 , como se indica a continuación:

$$F = \frac{\text{varianza entre muestras}}{\text{varianza dentro de muestras}} = \frac{CM(\text{del tratamiento})}{CM(\text{del error})}$$

- Los valores críticos de F se pueden encontrar en la tabla A-5, pero nos enfocamos en la interpretación de los valores P que están incluidos como parte de un resultado por computadora.

En la sección 12-3 consideramos el análisis de varianza de dos factores, con los datos categorizados de acuerdo con dos factores diferentes. Un factor se utiliza para ordenar los datos muestrales en renglones diferentes, mientras que el otro factor se emplea para columnas distintas. El procedimiento del análisis de varianza de dos factores está resumido la figura 12-4 y requiere que primero probemos si existe una interacción entre los dos factores. Si no

existe una interacción significativa, entonces procedemos a elaborar pruebas individuales de los efectos de cada uno de los dos factores. También consideramos el análisis de varianza de dos factores para el caso especial en el que sólo existe una observación por celda.

Dada la naturaleza compleja de los cálculos requeridos a lo largo de este capítulo, enfatizamos la interpretación de resultados por computadora.

Conocimientos estadísticos y pensamiento crítico

- 1. Análisis de varianza.** ¿Por qué se le llama análisis de “varianza” al método empleado para probar la igualdad de tres o más medias poblacionales, cuando las medias son los parámetros de interés?
- 2. ¿Cuál prueba?** Un investigador médico realiza una prueba clínica de un fármaco creado para reducir el colesterol LDL. Se administraron 10 mg del fármaco a 50 sujetos, a otros 50 sujetos se les administraron 20 mg y a otros 50 se les administró un placebo. ¿Cuál método se debe utilizar para probar la igualdad de las tres cantidades medias de disminución del colesterol?
- 3. ¿Cuál prueba?** Suponga que la prueba clínica del ejercicio 2 se modifica para excluir el tratamiento con 20 mg del fármaco. ¿Qué método debe usarse para probar la igualdad de las cantidades medias de disminución del colesterol en el grupo placebo y en el grupo de tratamiento con 10 mg?
- 4. ANOVA de dos factores.** Un genetista reúne datos que consisten en el color de ojos; los colores se categorizan de acuerdo con el género (hombre, mujer) y el grupo de edad (menor de 30, 30 y mayor). ¿Se pueden emplear los métodos de este capítulo para probar los efectos del género y del grupo de edad sobre el color de ojos? ¿Por qué?

Ejercicios de repaso

- 1. Interpretación de una pantalla de resultados de computadora: beber y conducir.** El Associated Insurance Institute financia estudios sobre los efectos de la bebida al conducir. En uno de estos estudios, se seleccionaron al azar tres grupos de hombres adultos para un experimento que pretendía medir los niveles de alcohol en la sangre, después de haber consumido cinco bebidas. Los miembros del grupo A se probaron después de una hora, los miembros del grupo B se probaron después de dos horas, y los miembros del grupo C se probaron después de cuatro horas. Los resultados se presentan en la tabla adjunta; también se incluyen los resultados de Minitab para estos datos. Con un nivel de significancia de 0.05, pruebe la aseveración de que los tres grupos tienen el mismo nivel medio.

	A	B	C
	0.11	0.08	0.04
	0.10	0.09	0.04
	0.09	0.07	0.05
	0.09	0.07	0.05
	0.10	0.06	0.06
			0.04
			0.05

Minitab

Source	DF	SS	MS	F	P
Factor	2	0.0076571	0.0038286	46.90	0.000
Error	14	0.0011429	0.0000816		
Total	16	0.0088000			

- 2. Puntuaciones de facilidad de lectura.** A continuación se muestran puntuaciones en la escala de Flesch sobre la facilidad de lectura de 12 páginas seleccionadas al azar de: *El oso y el dragón*, de Tom Clancy; *Harry Potter y la piedra filosofal*, de J. K. Rowling; y *La guerra y la paz*, de León Tolstoi. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los tres libros tienen la misma puntuación media de facilidad de lectura de Flesch. Con base en los resultados, ¿parece que los tres libros tienen el mismo nivel de lectura?

Clancy: 58.2 73.4 73.1 64.4 72.7 89.2 43.9 76.3 76.4 78.9 69.4 72.9
 Rowling: 85.3 84.3 79.5 82.5 80.2 84.6 79.2 70.9 78.6 86.2 74.0 83.7
 Tolstoi: 69.4 64.2 71.4 71.6 68.5 51.9 72.2 74.4 52.8 58.4 65.4 73.6

Interpretación de una pantalla de resultados de computadora. En los ejercicios 3 a 5, utilice la pantalla de resultados de Minitab que proviene de los valores listados en la tabla adjunta. Los datos muestrales son estimados de los estudiantes (en pies) de la longitud de su salón de clases. La longitud real del salón de clases es de 24 pies, 7.5 pulgadas.

	Área de estudios		
	Matemáticas	Negocios	Artes liberales
Mujer	28 25 30	35 25 20	40 21 30
Hombre	25 30 20	30 24 25	25 20 32

Minitab

Source	DF	SS	MS	F	P
Gender	1	29.389	29.3889	0.78	0.395
Major	2	10.111	5.0556	0.13	0.876
Interaction	2	14.111	7.0556	0.19	0.832
Error	12	453.333	37.7778		
Total	17	506.944			

- 3. Efecto de interacción.** Pruebe la hipótesis nula de que las longitudes estimadas no se ven afectadas por una interacción entre el género y el área de estudios.
- 4. Efecto del género.** Suponga que las longitudes estimadas no se ven afectadas por una interacción entre el género y el área de estudios. ¿Existe evidencia suficiente para sustentar la aseveración de que la longitud estimada se ve afectada por el género?
- 5. Efecto del área de estudios.** Suponga que las longitudes estimadas no se ven afectadas por una interacción entre el género y el área de estudios. ¿Existe evidencia suficiente para sustentar la aseveración de que la longitud estimada se ve afectada por el área de estudios?
- 6. Tabaquismo, temperatura corporal y género.** La tabla adjunta lista temperaturas obtenidas de sujetos elegidos al azar (según el conjunto de datos 2 del apéndice B). Las temperaturas están clasificadas de acuerdo con el género y el hábito del tabaquismo. Utilice un nivel de significancia de 0.05 para probar una interacción entre el género y el hábito del tabaquismo, para probar un efecto del género y para probar un efecto del hábito del tabaquismo. ¿Qué concluye?

	Fuma				No fuma				
	Hombre	98.4	98.4	99.4	98.6	98.0	98.0	98.8	97.0
Mujer	98.8	98.0	98.7	98.4	97.7	98.0	98.2	99.1	

- 7. Contaminación de automóviles.** La tabla adjunta lista las cantidades de gases de invernadero emitidos por diferentes automóviles en un año.

- Suponiendo que no existe un efecto de interacción, ¿existe evidencia suficiente para sustentar la aseveración de que las cantidades de gases de invernadero emitidas se ven afectadas por el tipo de transmisión (automática/manual)?
- Suponiendo que no existe un efecto de interacción, ¿existe evidencia suficiente para sustentar la aseveración de que las cantidades de gases de invernadero emitidas se ven afectadas por el número de cilindros?
- Con base en los resultados de los incisos a) y b), ¿podemos concluir que la emisión de gases de invernadero no se ve afectada por el tipo de transmisión o el número de cilindros? ¿Por qué?

Emisión de gases de invernadero (toneladas/año)

	4 cilindros	6 cilindros	8 cilindros
Automática	10	12	14
Manual	10	12	12

- Longevidad.** Consulte la siguiente tabla que lista el número de años (desde 1690) que los presidentes estadounidenses, los papas y los monarcas británicos vivieron después de que asumieron su respectivo cargo. Determine si los tiempos de supervivencia de los tres grupos difieren. (Tabla basada en datos de *Computer-Interactive Data Analysis*, de Lunn y McNeil, John Wiley & Sons).

Presidentes	Papas	Reyes y reinas	
Washington	10	Alejandro VIII	2
J. Adams	29	Inocencio XII	9
Jefferson	26	Clemente XI	21
Madison	28	Inocencio XIII	3
Monroe	15	Benedicto XIII	6
J. Q. Adams	23	Clemente XII	10
Jackson	17	Benedicto XIV	18
Van Buren	25	Clemente XIII	11
Harrison	0	Clemente XIV	6
Tyler	20	Pío VI	25
Polk	4	Pío VII	23
Taylor	1	León XII	6
Fillmore	24	Pío VIII	2
Pierce	16	Gregorio XVI	15
Buchanan	12	Pío IX	32
Lincoln	4	León XIII	25
A. Johnson	10	Pío X	11
Grant	17	Benedicto XV	8
Hayes	16	Pío XI	17
Garfield	0	Pío XII	19
Arthur	7	Juan XXIII	5

continúa

	Presidentes	Papas	Reyes y reinas
Cleveland	24	Paulo VI	15
Harrison	12	Juan Pablo I	0
McKinley	4	Juan Pablo II	26
T. Roosevelt	18		
Taft	21		
Wilson	11		
Harding	2		
Coolidge	9		
Hoover	36		
F. Roosevelt	12		
Truman	28		
Kennedy	3		
Eisenhower	16		
L. Johnson	9		
Nixon	25		
Reagan	23		

Ejercicios de repaso acumulativo

- 1. Longevidad.** Remítase al número de años que los presidentes estadounidenses, los papas y los monarcas británicos vivieron después de asumir el cargo. Los datos se presentan en la tabla del ejercicio de repaso 8.
- Calcule la media para cada uno de los tres grupos.
 - Calcule la desviación estándar para cada uno de los tres grupos.
 - Pruebe la aseveración de que existe una diferencia entre la media de los presidentes y la media de los monarcas británicos.
 - Considere la longevidad de los presidentes y determine si los datos parecen provenir de una población con una distribución normal. Explique por qué la distribución parece o no ser normal.
 - Considere la longevidad de los presidentes y construya un estimado de un intervalo de confianza del 95% para la media poblacional.
- 2. Tratamiento M&M.** La tabla que se presenta a continuación incluye 60 calificaciones del SAT, separadas en categorías de acuerdo con el color de los dulces M&M utilizados como tratamiento. Las calificaciones del SAT están basadas en datos del Consejo Universitario, y el elemento del color de los dulces M&M está basado en un capricho del autor.
- Calcule la media de las 20 calificaciones del SAT en cada una de las tres categorías. Al parecer, ¿las tres medias son aproximadamente iguales?
 - Calcule la mediana de las 20 calificaciones del SAT en cada una de las tres categorías. Al parecer, ¿las tres medianas son aproximadamente iguales?
 - Calcule la desviación estándar de las 20 calificaciones del SAT en cada una de las tres categorías. Al parecer, ¿las tres desviaciones estándar son aproximadamente iguales?
 - Pruebe la hipótesis nula de que no existe una diferencia entre la calificación media del SAT de los sujetos tratados con dulces M&M rojos y la calificación media del SAT de sujetos tratados con dulces M&M verdes.

- e. Construya un estimado del intervalo de confianza del 95% para la puntuación media del SAT de la población de sujetos que recibieron el tratamiento de dulces M&M rojos.
- f. Pruebe la hipótesis nula de que las tres poblaciones (tratamientos con dulces M&M rojos, verdes y azules) tienen la misma calificación media en el SAT.

Rojo	1130	621	813	996	1030	1257	898	743	921	1179
	1092	855	896	858	1095	1133	896	1190	908	699
Verde	996	630	583	828	1121	993	1025	907	1111	1147
	780	916	793	1188	499	1180	1229	1450	1071	1153
Azul	706	1068	1013	892	1370	1590	939	1004	821	915
	866	848	1408	793	1097	1244	996	1131	1039	1159

3. **Genes azules.** Algunas parejas tienen características genéticas que causan que una cuarta parte de sus descendientes tengan ojos azules. Se realiza un estudio con 100 parejas que se cree tienen esas características, y los resultados revelan que 19 de sus 100 descendientes tienen ojos azules. Suponiendo que una cuarta parte de todos los descendientes tienen ojos azules, estime la probabilidad de que, de 100 descendientes, 19 o menos tengan ojos azules. Con base en esa probabilidad, ¿parece que la proporción de una cuarta parte es incorrecta? ¿Por qué?
4. **Pesos de bebés: cálculo de probabilidades.** En Estados Unidos los pesos de los recién nacidos se distribuyen de manera normal, con una media de 7.54 lb y una desviación estándar de 1.09 lb (según datos de “Birth Weight and Prenatal Mortality”, de Wilcox, Skjaerven, Buekens y Kiely, *Journal of the American Medical Association*, vol. 273, núm. 9).
- a. Si se selecciona al azar a un bebé recién nacido, ¿cuál es la probabilidad de que pese más de 8.0 lb?
 - b. Si se seleccionan al azar 16 bebés recién nacidos, ¿cuál es la probabilidad de que su peso medio sea mayor de 8.0 lb?
 - c. ¿Cuál es la probabilidad de que cada uno de los siguientes tres bebés nazca con un peso mayor de 7.54 lb?

Actividades de cooperación en equipo

1. **Actividad fuera de clase** El *World Almanac and Book of Facts* incluye una sección llamada “Personalidades notables”, con apartados correspondientes a arquitectos, artistas, líderes de negocios, caricaturistas, científicos sociales, líderes militares, filósofos, líderes políticos, científicos, escritores, compositores y animadores, entre otros. Diseñe y realice un estudio observacional que inicie con la selección de muestras de grupos selectos, seguida por una comparación de los tiempos de vida de personas de distintas categorías. ¿Algunos grupos en particular parecen tener tiempos de vida diferentes en comparación con los otros grupos? ¿Puede usted explicar estas diferencias?
2. **Actividad en clase** Comience pidiendo a cada estudiante en la clase que estime la longitud del salón de

clases. Especifique que la longitud es la distancia entre el pizarrón y la pared opuesta. (Véase la sección de ejercicios de repaso 3-5). En el mismo papel, cada estudiante debe anotar también su género (hombre/mujer) y área de estudios. Después formen grupos de tres o cuatro miembros y utilicen los datos de toda la clase para plantear estas preguntas:

- ¿Existe una diferencia significativa entre el estimado medio de los hombres y el estimado medio de las mujeres?
- ¿Existe evidencia suficiente para rechazar la igualdad de los estimados medios en las diferentes áreas de estudio? Describa cómo se clasificaron las áreas de estudio.

- ¿La interacción entre el género y el área de estudios tiene algún efecto sobre la longitud estimada?
- ¿Parece que el género tiene un efecto sobre la longitud estimada?
- ¿Parece que el área de estudios tiene un efecto sobre la longitud estimada?

3. Actividad fuera de clase Formen grupos de tres o cuatro estudiantes. Cada grupo debe encuestar a otros estudiantes de la misma universidad y pedirles que identifiquen su área de estudios y género. También podría incluir otros factores, tales como el empleo (ninguno, de medio tiempo, de tiempo completo) y la edad (menos de 21, 21-30, más de 30). Para cada sujeto encuestado, determine la exactitud de la hora de su reloj de pulso. Primero

ponga su reloj a la hora correcta por medio de una fuente exacta y confiable del tipo “Cuando escuche el tono, la hora es . . .”. Registre una hora positiva para los relojes que están adelantados. Registre una hora negativa para los relojes que están atrasados. Utilice los datos muestrales para plantear preguntas como éstas:

- ¿Parece que el género tiene algún efecto sobre la exactitud de los relojes de pulso?
- ¿El área de estudios tiene algún efecto sobre la exactitud de los relojes de pulso?
- ¿La interacción entre el género y el área de estudios tiene algún efecto sobre la exactitud de los relojes de pulso?

Proyecto tecnológico

Remítase a los datos muestrales de la tabla 12-4 (página 655). Los resultados en la sección 12-3 mostraron que los datos llevaron a la conclusión de que no hay interacción entre el terreno y el tratamiento, que no hay un efecto del terreno, pero que sí existe un efecto del tratamiento. (*Sugerencia: Gráficas de interacción similares a las de la figura 12-3 le serán útiles para lo siguiente.*)

- a. Modifique los valores del terreno 1 y del cuarto tratamiento (fertilizante y riego), de manera que haya una interacción entre el terreno y el tratamiento.
- b. Modifique los valores del terreno 1 y del cuarto tratamiento (fertilizante y riego), de manera que no

haya una interacción entre el terreno y el tratamiento, que no haya un efecto del terreno y que no haya un efecto del tratamiento.

- c. Trate de modificar los valores del terreno 1 y del cuarto tratamiento (fertilizante y riego), de manera que no haya interacción, pero que sí haya un efecto del terreno y no haya un efecto del tratamiento. Si no se puede modificar esa celda para lograr los resultados deseados, modifique otras celdas según sea necesario.

De los datos a la decisión

Pensamiento crítico: ¿Debe usted aprobar este fármaco?

Los fármacos deben ser sometidos a pruebas exhaustivas antes de ser aprobados para su uso general. Además de probar sus reacciones adversas, también debe probarse su eficacia, y el análisis de este tipo de resultados de pruebas suelen incluir métodos estadísticos. Considere la creación del Xynamine, un nuevo fármaco diseñado para disminuir la frecuencia cardiaca. Para obtener resultados más consistentes, que no incluyan una variable confusa del género, el fármaco se prueba únicamente en hombres. Abajo se incluyen los pulsos de un grupo placebo, de un grupo de hombres tratados con Xynamine en dosis de 10 mg y de un grupo de hombres tratados con Xynamine en dosis de 20 mg. El director del proyecto del fármaco realiza la investigación y encuentra que en hombres adultos el pulso se distribuye normalmente, con una media de alrededor de 70 latidos por minuto y una desviación estándar de aproximadamente 11 latidos por minuto. El resumen de su informe afirma que el fármaco

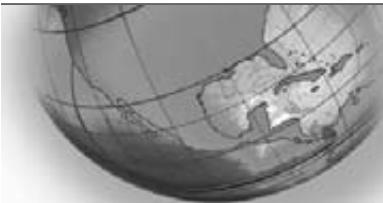
es eficaz, con base en esta evidencia: el grupo placebo tiene un pulso medio de 68.9, que se acerca al valor de 70 latidos por minuto de los hombres adultos en general, pero el grupo tratado con dosis de 10 mg de Xynamine tiene una media más baja de 66.2, y el grupo tratado con dosis de 20 mg de Xynamine tiene la media más baja de 65.2.

Análisis de los resultados

Analice los datos utilizando los métodos de este capítulo. Con base en los resultados, ¿parece que existe evidencia suficiente para sustentar la aseveración de que el fármaco reduce la frecuencia cardiaca? ¿Existen algunos problemas graves en el diseño de este experimento? Dado que únicamente se incluyeron hombres en el experimento, ¿los resultados también son aplicables a las mujeres? El director del proyecto comparó los pulsos postratamiento con el pulso medio de hombres adultos. ¿Existe alguna forma mejor de medir la eficacia del fármaco para disminuir el pulso? ¿Cómo calificaría la validez general del experimento? Con base en los resultados disponibles, ¿debe aprobarse el fármaco? Escriba un breve reporte que resuma sus hallazgos.

Grupo placebo	Grupo de tratamiento con 10 mg	Grupo de tratamiento con 20 mg
77	67	72
61	48	94
66	79	57
63	67	63
81	57	69
75	71	59
66	66	64
79	85	82
66	75	34
75	77	76
48	57	59
70	45	53





Proyecto de Internet

Análisis de varianza

Visite el sitio de Internet de este libro en

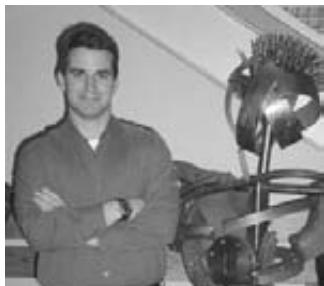
<http://www.pearsoneducacion.net/triola>

Siga el vínculo del Proyecto de Internet de este capítulo. El proyecto incluye antecedentes para experimentos en áreas tan variadas como el de-

sempeño atlético, la etiquetación de productos de consumo y la biología del cuerpo humano. En cada caso, los datos asociados se podrán agrupar de forma ideal para la aplicación de las técnicas este capítulo. Usted Fórmulará las hipótesis apropiadas, después realizará y resumirá pruebas ANOVA.

La estadística en el trabajo

“Si no tuviera conocimientos de estadística, no sería capaz de comprender plenamente los datos que genera mi empresa ... ni de ayudar a proteger a nuestros empleados y clientes”.



Jeffrey Foy

Jeffrey Foy es toxicólogo y trabaja para la Cabot Corporation, una empresa de productos químicos.

Jeffrey Foy también es responsable de evaluar los peligros de las sustancias químicas que produce la Cabot Corporation. Su trabajo consiste en entender de qué manera los productos de la empresa pueden afectar a los seres humanos o al ambiente, y en tomar decisiones sobre las mejores formas de proteger ambos.

¿Qué hace usted en su trabajo?

Mis responsabilidades incluyen la organización y evaluación de estudios toxicológicos, elaborar hojas de cálculo sobre la seguridad de los materiales y ayudar a nuestros grupos de investigación y desarrollo a producir materiales que sean seguros tanto para las personas como para el ambiente y a comprender qué daños potenciales pueden producir esos materiales.

¿Qué conceptos de estadística utiliza?

El principal concepto que utilizo es la prueba de hipótesis (pruebas de probabilidad).

¿Cómo utiliza la estadística en el trabajo?

Utilizo la estadística todos los días. Los métodos estadísticos se han utilizado y se siguen utilizando de dos formas en mi trabajo. En primer lugar, la estadística se utiliza para determinar la forma en que diseño mis experimentos. En segundo lugar, la estadística se usa para determinar si los datos generados son significativos o para saber si son lo suficientemente buenos para utilizarlos.

Los estudios en que participo pueden costar desde \$1000 hasta \$500,000 o más, y si no determino de manera apropiada la forma en que voy a evaluar los datos, esto podría costar a la empresa una gran cantidad de tiempo y dinero. Si el experimento se realiza adecuadamente, entonces procedemos a analizar los datos. Los datos de los estudios que realizamos se utilizan para evaluar cual-

quier efecto potencial que muchos productos pueden tener en la salud de nuestros empleados, clientes o en el ambiente. Los resultados se utilizan para determinar la manera en que se pueden vender o manejar los productos químicos. Cuando se realizan experimentos en un laboratorio de pruebas o en una compañía farmacéutica, se busca determinar si los materiales tienen algún efecto, ya sea deseable (como un fármaco que cura una enfermedad) o indeseable (por ejemplo, si un fármaco es tóxico). La estadística tiene un papel muy importante en nuestra evaluación de la importancia de los efectos.

Por favor, describa un ejemplo específico que ilustre cómo el uso de la estadística tuvo éxito en mejorar un producto o servicio.

Hace poco tiempo realizamos un estudio toxicológico que costó alrededor de \$300,000. Los datos del estudio se utilizarían para determinar si cierto producto químico causaba algún efecto en los sujetos estudiados. Después de que se llevó a cabo el estudio, se descubrieron fallas en los datos y en los estadísticos utilizados. Se requirieron dos años más para revisar los datos adecuadamente y finalizar la evaluación de salud. Si se hubieran elegido los métodos y los términos apropiados, tal vez no se hubiera requerido de tiempo y dinero adicionales. La comprensión de los datos y la evaluación estadística correcta ayudó a prevenir un fracaso y una posible repetición del estudio.



Estadística no paramétrica

13



13-1 Panorama general

13-2 Prueba del signo

**13-3 Prueba de rangos con signo de Wilcoxon
para datos apareados**

**13-4 Prueba de la suma de rangos de Wilcoxon
para dos muestras independientes**

13-5 Prueba de Kruskal-Wallis

13-6 Correlación de rangos

13-7 Prueba de rachas para detectar aleatoriedad



¿Los estudiantes clasifican a las universidades de la misma manera que el *U.S. News and World Report*?

Cada año, la revista *U.S. News and World Report* publica una clasificación de universidades con base en estadísticos tales como las tasas de admisión, las tasas de graduación, el tamaño de los grupos, la razón entre profesores y estudiantes, los salarios de los profesores y las calificaciones de los administradores otorgadas por sus compañeros. Los economistas Christopher Avery, Mark Glickman, Caroline Minter Hoxby y Andrew Metrick usaron un método alternativo para analizar la selección de universidades de 3240 estudiantes del último año de preparatoria con alto rendimiento escolar. Examinaron las universidades que ofrecen admisión junto con las universidades que los estudiantes eligen. La tabla 13-1 lista el orden de una pequeña muestra de universidades, así como también cierto acuerdo entre el orden de preferencia de los estudiantes y las calificaciones de la revista, aunque también indica cierto desacuerdo. Por ejemplo, de las ocho universidades consideradas, Harvard ocupó el primer lugar tanto para los estudiantes como para la revista *U.S. News and World Report*. Sin embargo, de las ocho universidades incluidas, la Universidad de Pensilvania fue considerada en séptimo lugar por los estudiantes pero en tercer lugar por la revista.

Consideremos el tema de una correlación entre la clasificación de los estudiantes y la clasificación de la revis-

ta. El concepto de correlación se estudió en la sección 10-2, donde el coeficiente de correlación lineal r se utilizó para medir la asociación entre dos variables. Los métodos de la sección 10-2 requieren datos apareados, y los datos de la tabla 13-1 están apareados. Sin embargo, existe una diferencia muy importante: los métodos de la sección 10-2 tienen requisitos como las distribuciones normales, y los rangos como los que aparecen en la tabla 13-1 no satisfacen estos requisitos. Los métodos de la sección 10-2 no se pueden utilizar con los datos muestrales de la tabla 13-1. En este capítulo se presentan varios métodos que se utilizan con datos que no satisfacen el requisito de una distribución normal. En particular, varios métodos de esta sección pueden emplearse con datos muestrales en el formato de rangos, como los de la tabla 13-1. En la sección 13-6 se estudiará un método para poner a prueba una correlación con datos apareados que no tienen el formato de rangos. Entonces, seremos capaces de analizar el grado de acuerdo y desacuerdo entre las clasificaciones de los estudiantes y de la revista, como aparecen en la tabla 13-1. Así, probaremos si existe una correlación entre las preferencias de los estudiantes y la clasificación de la revista, y podremos contestar la siguiente pregunta importante: ¿Los estudiantes coinciden con la revista?

Tabla 13-1 Universidades clasificadas por estudiantes y por *U.S. News and World Report*

Universidad	Clasificación según la preferencia de los estudiantes	Clasificación según la revista <i>U.S. News and World Report</i>
Harvard	1	1
Yale	2	2
Cal. Inst. of Tech.	3	5
M.I.T.	4	4
Brown	5	7
Columbia	6	6
U. of Penn.	7	3
Notre Dame	8	8

13-1 Panorama general

Los métodos de estadística inferencial presentados en los capítulos 7, 8, 9, 10 y 12 se llaman *métodos paramétricos* porque se basan en el muestreo de una población con parámetros específicos, como la media μ , la desviación estándar σ o la proporción p . Por lo regular, estos métodos paramétricos deben cumplir con algunas condiciones bastante estrictas, como el requisito de que los datos muestrales provengan de una población distribuida normalmente. Este capítulo describe métodos no paramétricos, los cuales no tienen esos estrictos requisitos.

Definiciones

Las **pruebas paramétricas** tienen requisitos acerca de la naturaleza o forma de las poblaciones implicadas; las **pruebas no paramétricas** no requieren que las muestras provengan de poblaciones con distribuciones normales o con cualquier otro tipo particular de distribución. En consecuencia, las pruebas de hipótesis no paramétricas suelen llamarse **pruebas de distribución libre**.

Aunque el término *no paramétrica* sugiere que la prueba no está basada en un parámetro, existen algunas pruebas no paramétricas que sí dependen de un parámetro como la mediana. Sin embargo, las pruebas no paramétricas no requieren de una distribución particular, por lo que algunas veces se les conoce como pruebas de *distribución libre*. Aunque *distribución libre* es una descripción más precisa, por lo regular se utiliza el término *no paramétrica*. A continuación se enuncian las principales ventajas y desventajas de los métodos no paramétricos.

Ventajas de los métodos no paramétricos

1. Los métodos no paramétricos pueden aplicarse a una amplia variedad de situaciones puesto que no tienen los requisitos más estrictos de los métodos paramétricos correspondientes. En particular, los métodos no paramétricos no requieren de poblaciones distribuidas normalmente.
2. A diferencia de los métodos paramétricos, los métodos no paramétricos a menudo pueden aplicarse a datos categóricos, como el género de quienes responden una encuesta.
3. Los métodos no paramétricos, por lo regular, implican cálculos más sencillos que los métodos paramétricos correspondientes y, por lo tanto, son más fáciles de comprender y aplicar. (Sin embargo, como la tecnología ha simplificado los cálculos, es probable que la facilidad de los cálculos no sea un factor tan importante).

Desventajas de los métodos no paramétricos

1. Los métodos no paramétricos tienden a desperdiciar información porque los datos numéricos exactos suelen reducirse a una forma cualitativa. Por ejemplo, en la prueba del signo no paramétrica (descrita en la sección 13-2), las pérdidas de peso de las personas sometidas a una dieta se registran

simplemente como signos negativos; las magnitudes reales de las pérdidas de peso se ignoran.

- Las pruebas no paramétricas no son tan eficientes como las pruebas paramétricas, de manera que con una prueba no paramétrica generalmente necesitamos evidencia más fuerte (como una muestra más grande o diferencias mayores) para rechazar una hipótesis nula.

Cuando se satisfacen los requisitos de distribuciones poblacionales, las pruebas no paramétricas generalmente son menos eficaces que sus contrapartes paramétricas, pero la reducción en la eficiencia puede compensarse con un tamaño muestral más grande. Por ejemplo, en la sección 13-6 se presentará un concepto llamado *correlación de rangos*, que tiene una tasa de eficiencia de 0.91 cuando se compara con la correlación lineal presentada en el capítulo 10. Esto significa que, si todo permanece igual, la correlación de rangos no paramétrica requiere 100 observaciones muestrales para obtener los mismos resultados que 91 observaciones muestrales analizadas por medio de la correlación lineal paramétrica, suponiendo que se satisfacen los requisitos más estrictos para la aplicación del método paramétrico. La tabla 13-2 lista los métodos no paramétricos cubiertos en este capítulo, junto con el método paramétrico correspondiente y la tasa de **eficiencia**. La tabla 13-2 indica que varias pruebas no paramétricas tienen tasas de eficiencia por encima de 0.90, por lo que la eficiencia más baja tal vez no sea un factor esencial para elegir entre los métodos paramétricos y no paramétricos. Sin embargo, puesto que las pruebas paramétricas tienen tasas de eficiencia más altas que sus contrapartes no paramétricas, generalmente es mejor utilizar las pruebas paramétricas cuando sus supuestos requeridos se satisfacen.

Rangos

Las secciones 13-3 a 13-6 utilizan métodos basados en rangos, que describiremos a continuación.

Definición

Los datos están *ordenados* cuando se acomodan de acuerdo con algún criterio, por ejemplo, del más pequeño al más grande o del mejor al peor. Un **rango** es un número asignado a un elemento muestral individual de acuerdo con su lugar en la lista ordenada. Al primer elemento se le asigna un rango de 1, al segundo elemento se le asigna un rango de 2 y así sucesivamente.

Tabla 13-2 Eficiencia: Comparación de pruebas paramétricas y no paramétricas

Aplicación	Prueba paramétrica	Prueba no paramétrica	Tasa de eficiencia de una prueba no paramétrica con población normal
Datos muestrales apareados	Prueba <i>t</i> o prueba <i>z</i>	Prueba del signo Prueba de rangos con signo de Wilcoxon	0.63 0.95
Dos muestras independientes	Prueba <i>t</i> o prueba <i>z</i>	Prueba de la suma de rangos de Wilcoxon	0.95
Varias muestras independientes	Análisis de varianza (prueba <i>F</i>)	Prueba de Kruskal-Wallis	0.95
Correlación	Correlación lineal	Prueba de correlación de rangos ordenados	0.91
Aleatoriedad	Prueba no paramétrica	Prueba de rachas	Sin bases para comparar

Manejo de rangos empatados: Si ocurre un empate en los rangos, el procedimiento habitual es calcular la media de los rangos implicados y luego asignar este rango medio a cada uno de los elementos empatados, como en el siguiente ejemplo.

EJEMPLO Los números 4, 5, 5, 5, 10, 11, 12 y 12 tienen rangos dados de 1, 3, 3, 3, 5, 6, 7.5 y 7.5, respectivamente. Vea la siguiente tabla y observe el procedimiento para manejar empates.

Datos ordenados	Rango preliminar	Rango
4	1	1
5	2	3
5	3	3
5	4	3
10	5	5
11	6	6
12	7	7.5
12	8	7.5

13-2 Prueba del signo

Concepto clave El objetivo principal de esta sección es entender el procedimiento de la *prueba del signo*, el cual implica convertir valores de datos en signos positivos y negativos, y luego hacer una prueba para ver si hay una cantidad desproporcionadamente mayor de uno u otro signo.

Definición

La **prueba del signo** es una prueba no paramétrica (de distribución libre) que utiliza signos positivos y negativos para probar diferentes aseveraciones, incluyendo:

1. Aseveraciones que implican datos muestrales apareados
2. Aseveraciones que implican datos nominales
3. Aseveraciones acerca de la mediana de una sola población

Concepto básico de la prueba del signo

La idea básica que subyace en la prueba del signo es el análisis de las frecuencias de los signos positivos y negativos para determinar si son significativamente diferentes. Por ejemplo, suponga que probamos un tratamiento diseñado para incrementar la probabilidad de que un bebé sea niña. Si se trata a 100 mujeres y 51 de ellas tienen niñas, el sentido común sugiere que no existe evidencia suficiente para afirmar que el tratamiento es efectivo, puesto que 51 niñas entre 100 bebés no son significativas. Pero ¿qué sucede con 52 niñas y 48 niños? ¿O con 90 niñas y 10 niños? La prueba del signo nos permite determinar cuándo este tipo de resultados son significativos.

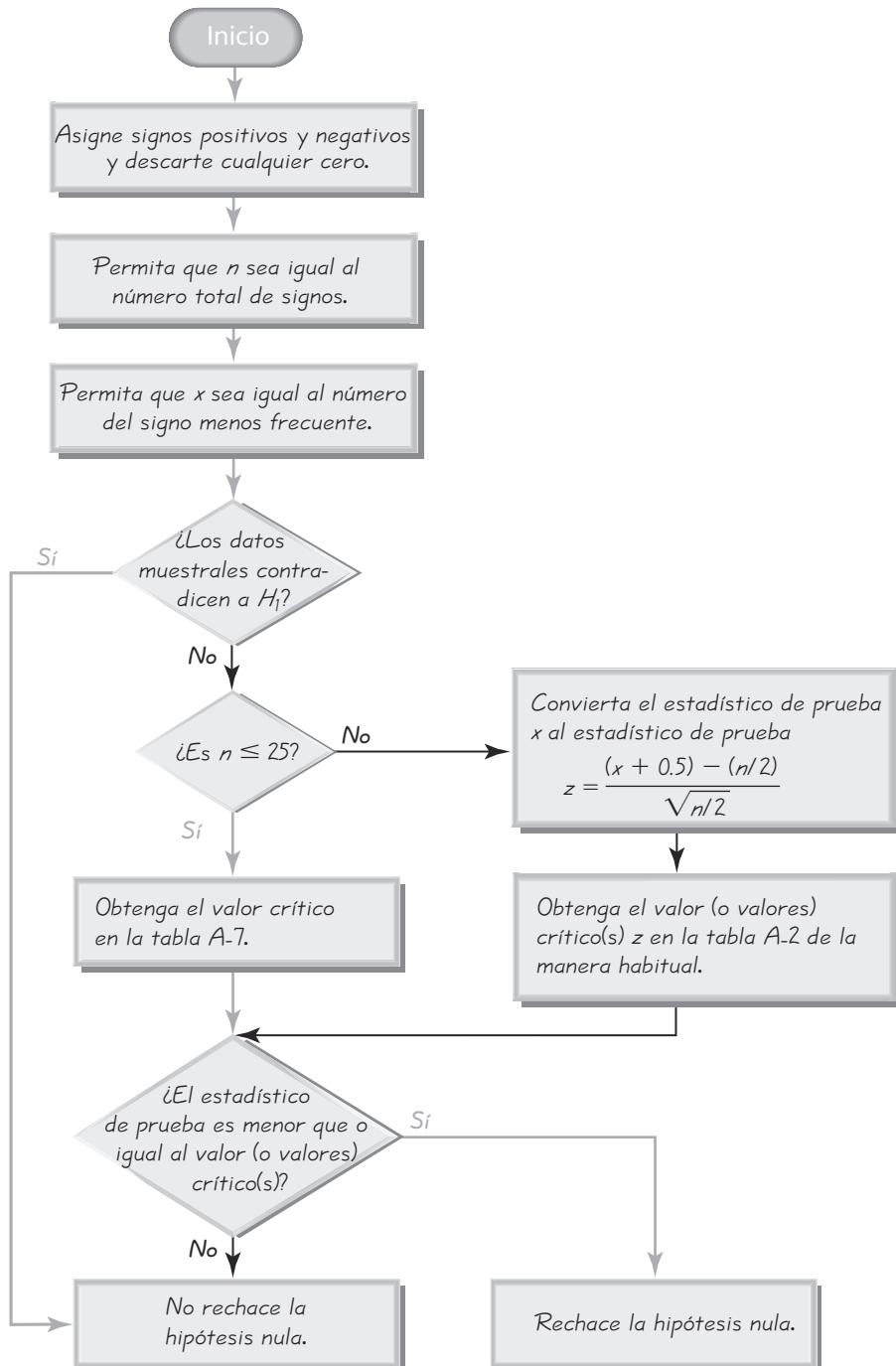


Figura 13-1

Procedimiento de la prueba del signo

Por razones de consistencia y simplicidad, utilizaremos un estadístico de prueba con base en el número de veces que ocurre el signo *menos frecuente*. En el cuadro adjunto se resumen los supuestos relevantes, la notación, el estadístico de prueba y los valores críticos. La figura 13-1 resume el procedimiento de la prueba del signo, que se ilustrará con los ejemplos que siguen.



Asistencia a clases y calificaciones

En un estudio de 424 estudiantes de licenciatura de la Universidad de Michigan, se encontró que los estudiantes con los peores registros de asistencia tendían a obtener las calificaciones más bajas. (¿Quién se sorprende?). Aquellos que estuvieron ausentes menos del 10% del tiempo tendieron a recibir calificaciones de B o superiores. El estudio también reveló que los estudiantes que se sientan al frente en el salón de clases tienden a obtener calificaciones significativamente mejores.

Prueba del signo

Requisitos

1. Los datos muestrales se seleccionaron aleatoriamente.
2. No existe el requisito de que los datos muestrales provengan de una población con una distribución particular, como una distribución normal.

Notación

x = el número de veces que ocurre el signo *menos frecuente*

n = el número total de signos positivos y negativos combinados

Estadístico de prueba

Para $n \leq 25$: x (el número de veces que ocurre el signo menos frecuente)

$$\text{Para } n > 25: z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\sqrt{\frac{n}{2}}}$$

Valores críticos

1. Para $n \leq 25$, los valores críticos x se encuentran en la tabla A-7.
2. Para $n > 25$, los valores críticos z se encuentran en la tabla A-2.

Cuidado: Cuando se aplica la prueba del signo en una prueba de una cola, necesitamos ser muy cuidadosos para evitar obtener la conclusión equivocada cuando un signo ocurre significativamente con más frecuencia que el otro, pero los datos muestrales contradicen la hipótesis alternativa. Por ejemplo, suponga que estamos probando la aseveración de que una técnica de selección del género favorece a los niños, pero obtenemos una muestra de 10 niños y 90 niñas. Con una proporción muestral de niños igual a 0.10, los datos contradicen la hipótesis alternativa $H_1: p > 0.5$. No hay forma de sustentar la aseveración de que $p > 0.5$ con ninguna proporción muestral menor que 0.5, por lo que no rechazamos la hipótesis nula y no procedemos con la prueba del signo. La figura 13-1 resume el procedimiento para la prueba del signo e incluye esta revisión: ¿Contradicen los datos muestrales a H_1 ? Si los datos muestrales van en el sentido opuesto de H_1 , no rechace la hipótesis nula. *Siempre es importante reflexionar acerca de los datos y evitar confiar a ciegas en cálculos o resultados de computadora.*

Aseveraciones que implican datos apareados

Cuando se utiliza la prueba del signo con datos que están ordenados en pares, convertimos los datos en bruto a datos con signos positivos y negativos como sigue:

1. Restamos cada valor de la segunda variable del valor correspondiente de la primera variable.
2. Registramos sólo el *signo* de la diferencia encontrada en el paso 1. *Excluimos los empates:* es decir, excluimos todos los datos apareados en los que ambos valores son iguales.

Éste es el concepto clave que subyace en la aplicación de la prueba del signo:

Si dos conjuntos de datos tienen medianas iguales, el número de signos positivos debe ser aproximadamente igual al número de signos negativos.

EJEMPLO ¿El tipo de semilla afecta el crecimiento del maíz? En 1908 William Gosset publicó el artículo “The Probable Error of a Mean” bajo el seudónimo de “Student” (*Biometrika*, vol. 6, núm. 1). Él incluyó los datos que se listan en la tabla 13-3 para dos tipos diferentes de semillas de maíz (normales y secadas en horno), que se utilizaron en parcelas de tierra *adyacentes*. Los valores corresponden a las cosechas de cabezas de maíz (o mazorcas) en libras por acre. Utilice la prueba del signo con un nivel de significancia de 0.05 para probar la aseveración de que no hay diferencia entre las cosechas de las semillas normales y las de las semillas secadas en horno.

SOLUCIÓN

REQUISITO ✓ El único requisito es que los datos muestrales se obtengan al azar. No hay un requisito sobre la distribución de la población, como el hecho de que los datos muestrales provengan de una población distribuida normalmente. Con base en el diseño de este experimento, suponemos que los datos muestrales son aleatorios. ✓

La siguiente es la idea básica: si no existe diferencia entre las cosechas de las semillas normales y las cosechas de las semillas secadas en horno, el número de signos positivos y negativos debe ser aproximadamente igual. En la tabla 13-3 tenemos 7 signos negativos y 4 signos positivos. ¿Son aproximadamente iguales los números de signos positivos y negativos, o son significativamente diferentes? Seguimos los mismos pasos básicos de prueba de hipótesis, tal como se describieron en la figura 8-9, y aplicamos el procedimiento de la prueba del signo que se resume la figura 13-1.

Pasos 1, 2 y 3: La hipótesis nula es la aseveración de que no hay diferencia entre las cosechas de las semillas normales y las cosechas de las semillas secadas en horno, y la hipótesis alternativa es la aseveración de que existe una diferencia.

H_0 : No existe diferencia (la mediana de las diferencias es igual a 0).

H_1 : Existe una diferencia (la mediana de las diferencias no es igual a 0).

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Utilizamos la prueba no paramétrica del signo.

continúa

Tabla 13-3 Cosechas de maíz de diferentes semillas

Normales	1903	1935	1910	2496	2108	1961	2060	1444	1612	1316	1511
Secadas en horno	2009	1915	2011	2463	2180	1925	2122	1482	1542	1443	1535
Signo de la diferencia	-	+	-	+	-	+	-	-	+	-	-

- Paso 6: El estadístico de prueba x es el número de veces que se presenta el signo menos frecuente. La tabla 13-3 incluye diferencias con 7 signos negativos y 4 signos positivos. (Si hubiera cualquier diferencia igual a 0, la descartaríamos). Permitimos que x sea igual al menor de 7 y 4, así que $x = 4$. Además, $n = 11$ (el número total de signos positivos y negativos combinados). Nuestra prueba es de dos colas con $\alpha = 0.05$. Nos remitimos a la tabla A-7, donde se encuentra el valor crítico de 1 para $n = 11$ y $\alpha = 0.05$ en dos colas. (Véase la figura 13-1).
- Paso 7: Con un estadístico de prueba de $x = 4$ y un valor crítico de 1, no rechazamos la hipótesis nula de no diferencia. [Véase la nota 2 incluida en la tabla A-7: “La hipótesis nula se rechaza si el número del signo menos frecuente (x) es menor que o igual al valor en la tabla”. Puesto que $x = 4$ no es menor que o igual al valor crítico de 1, no rechazamos la hipótesis nula].
- Paso 8: No hay suficiente evidencia para sustentar el rechazo de la aseveración de que la mediana de las diferencias es igual a 0; esto es, no existe suficiente evidencia para justificar el rechazo de la aseveración de que no existe una diferencia entre las cosechas de las semillas normales y las cosechas de las semillas secadas en horno. Ésta es la misma conclusión a la que se llegaría utilizando la prueba paramétrica t con datos apareados de la sección 9-4, pero los resultados de la prueba del signo no siempre coinciden con los resultados de la prueba paramétrica.

Aseveraciones que implican datos nominales

En el capítulo 1 definimos los datos nominales como aquellos que consisten sólo en nombres, etiquetas o categorías. La naturaleza de los datos nominales limita los cálculos posibles, pero podemos identificar la *proporción* de datos muestrales que pertenecen a una categoría en particular y podemos probar aseveraciones acerca de la proporción poblacional p correspondiente. El siguiente ejemplo utiliza datos nominales que consisten en el género (niñas/niños). La prueba del signo se utiliza representando a las niñas con signos positivos (+) y a los niños con signos negativos (-). (Créanme, estos signos se eligieron arbitrariamente). También observe el procedimiento para manejar casos en los que $n > 25$.

EJEMPLO Selección del género El Genetics and IVF Institute realizó un ensayo clínico de sus métodos de selección del género. Para cuando se escribía este libro, los resultados incluían a 325 bebés nacidos de padres que utilizaron el método XSORT para aumentar la probabilidad de concebir una niña, y 295 de esos bebés fueron niñas. Utilice la prueba del signo con un nivel de significancia de 0.05 y pruebe la aseveración de que este método de selección del género no tiene ningún efecto.

SOLUCIÓN

REQUISITO ✓ El único requisito es que los datos muestrales se seleccionen al azar. Con base en el diseño de este experimento, podemos suponer que los datos muestrales son aleatorios. Ahora podemos proceder con la prueba del signo. ✓

Permita que p denote la proporción poblacional de niñas. La aseveración de ningún efecto implica que las proporciones de niñas y niños sean iguales a

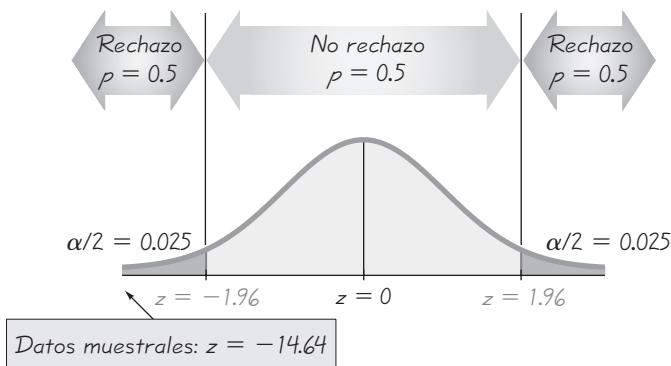


Figura 13-2 Prueba del efecto de un método de selección del género

0.5, de manera que $p = 0.5$. Por lo tanto, las hipótesis nula y alternativa pueden establecerse de la siguiente manera:

$$\begin{aligned} H_0: p &= 0.5 && \text{(la proporción de niñas es igual a 0.5)} \\ H_1: p &\neq 0.5 \end{aligned}$$

Si denotamos a las niñas con signo positivo (+) y a los niños con signo negativo (-), tenemos 295 signos positivos y 30 signos negativos. Ahora consulte el procedimiento de prueba del signo que se resume en la figura 13-1. El estadístico de prueba x es el menor de 295 y 30, así que $x = 30$. Esta prueba implica dos colas puesto que un número desproporcionadamente alto o bajo de niñas nos llevará a rechazar la aseveración de $p = 0.5$. Los datos muestrales no contradicen la hipótesis alternativa puesto que 295 y 30 no son precisamente iguales. (Esto es, los datos muestrales son consistentes con la hipótesis alternativa de una diferencia). Continuando con el procedimiento de la figura 13-1, notamos que el valor de $n = 325$ es superior a 25, por lo que el estadístico de prueba x se convierte (utilizando una corrección por continuidad) al estadístico de prueba z como sigue:

$$\begin{aligned} z &= \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\sqrt{n}} \\ &= \frac{(30 + 0.5) - \left(\frac{325}{2}\right)}{\sqrt{325}} = -14.64 \end{aligned}$$

Con $\alpha = 0.05$ en una prueba de dos colas, los valores críticos son $z = \pm 1.96$. El estadístico de prueba $z = -14.64$ es menor que -1.96 (véase la figura 13-2), por lo que rechazamos la hipótesis nula de que la proporción de niñas es igual a 0.5. Tenemos evidencia muestral suficiente para justificar el rechazo de la aseveración de que el método de selección del género no tiene efecto alguno (ya que las proporciones de niñas y niños son iguales a 0.5). Parece que este método afecta el género de los bebés.

Aseveraciones acerca de la mediana de una sola población

El siguiente ejemplo ilustra el procedimiento para utilizar la prueba del signo al probar una aseveración acerca de la mediana de una sola población. Vea cómo los signos positivos y negativos se basan en el valor aseverado de la mediana.

EJEMPLO Temperaturas corporales El conjunto de datos 2 del apéndice B incluye temperaturas corporales medidas en adultos. Utilice las 106 temperaturas listadas para las 12:00 a.m. del día 2 con la prueba del signo para probar la aseveración de que la mediana es menor que 98.6°F. El conjunto de datos tiene 106 sujetos, 68 sujetos con temperaturas por debajo de 98.6°F, 23 sujetos con temperaturas por arriba de 98.6°F y 15 sujetos con temperaturas iguales a 98.6°F.

SOLUCIÓN

REQUISITO ✓ El único requisito es que los datos muestrales se seleccionen al azar y, con base en el diseño de este experimento, suponemos que los datos muestrales son aleatorios. Ahora podemos proceder con la prueba del signo. ✓

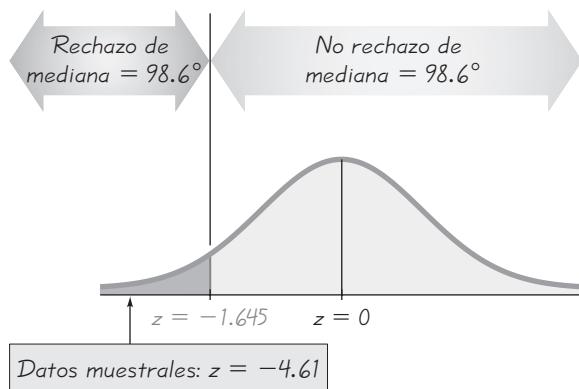
La aseveración de que la mediana es menor que 98.6°F es la hipótesis alternativa, mientras que la hipótesis nula es la aseveración de que la mediana es igual a 98.6°F.

$$\begin{aligned} H_0: \text{ La mediana es igual a } 98.6^{\circ}\text{F.} & \quad (\text{mediana} = 98.6^{\circ}\text{F}) \\ H_1: \text{ La mediana es menor que } 98.6^{\circ}\text{F.} & \quad (\text{mediana} < 98.6^{\circ}\text{F}) \end{aligned}$$

Siguiendo el procedimiento descrito en la figura 13-1, descartamos los 15 ceros, utilizamos el signo negativo (–) para denotar cada temperatura que está por debajo de 98.6°F y utilizamos el signo positivo (+) para denotar cada temperatura que está por encima de 98.6°F. Por lo tanto, tenemos 68 signos negativos y 23 signos positivos, entonces $n = 91$ y $x = 23$ (el número del signo menos frecuente). Los datos muestrales no contradicen la hipótesis alternativa, puesto que la mayoría de las 91 temperaturas están por debajo de 98.6°F. (Si los datos muestrales indicaran un conflicto con la hipótesis alternativa, podríamos terminar inmediatamente la prueba concluyendo que no rechazamos la hipótesis nula). El valor de n excede a 25, por lo que convertimos el estadístico de prueba x al estadístico de prueba z :

$$\begin{aligned} z &= \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\sqrt{\frac{n}{2}}} \\ &= \frac{(23 + 0.5) - \left(\frac{91}{2}\right)}{\sqrt{\frac{91}{2}}} = -4.61 \end{aligned}$$

En esta prueba de una cola con $\alpha = 0.05$, utilizamos la tabla A-2 para obtener el valor crítico z de -1.645 . En la figura 13-3 podemos ver que el estadístico de prueba $z = -4.61$ cae dentro de la región crítica; por lo tanto, rechazamos la hipótesis nula. Con base en la evidencia muestral disponible, sustentamos la aseveración de que la mediana de la temperatura corporal de adultos saludables es menor que 98.6°F.

**Figura 13-3**

Prueba de la aseveración de que la mediana es menor que 98.6°F

En esta prueba del signo para la aseveración de que la mediana está por debajo de 98.6°F, obtenemos un estadístico de prueba de $z = -4.61$ con un valor P de 0.00000202, pero una prueba paramétrica de la aseveración de que $\mu < 98.6°F$ produce un estadístico de prueba $t = -6.611$, con un valor P de 0.00000000813. Puesto que el valor P de la prueba del signo no es tan bajo como el valor P de la prueba paramétrica, vemos que la prueba del signo no es tan sensible como la prueba paramétrica. Ambas pruebas nos llevan al rechazo de la hipótesis nula, pero la prueba del signo no considera que los datos muestrales sean tan extremos, en parte porque la prueba del signo utiliza sólo información acerca de la *dirección* de los datos, ignorando las *magnitudes* de los valores de los datos. En la siguiente sección se estudiará la prueba de rangos con signo de Wilcoxon, que supera con creces esta desventaja.

Fundamentos para el estadístico de prueba que se utiliza cuando $n > 25$:

Cuando se calculan valores críticos para la prueba del signo, utilizamos la tabla A-7 sólo para n hasta 25. Cuando $n > 25$, el estadístico de prueba z se basa en una aproximación normal a la distribución de probabilidad binomial con $p = q = 1/2$. Recuerde que en la sección 6-6 vimos que la aproximación normal a la distribución binomial es aceptable cuando $np \geq 5$ y $nq \geq 5$. Recuerde también que en la sección 5-4 vimos que $\mu = np$ y $\sigma = \sqrt{npq}$ para distribuciones de probabilidad binomial. Puesto que esta prueba del signo supone que $p = q = 1/2$, satisfacemos los requisitos de que $np \geq 5$ y $nq \geq 5$ siempre y cuando $n \geq 10$. Además, con el supuesto de que $p = q = 1/2$, obtenemos $\mu = np = n/2$ y $\sqrt{npq} = \sqrt{n/4} = \sqrt{n}/2$, por lo tanto

$$z = \frac{x - \mu}{\sigma}$$

se convierte en

$$z = \frac{x - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

Finalmente, reemplazamos x con $x + 0.5$ como una corrección por continuidad. Esto es, los valores de x son discretos, pero puesto que estamos utilizando una distribución de probabilidad continua, un valor discreto como 10 se representa realmente con el intervalo de 9.5 a 10.5. Como x representa el signo menos frecuente, actuamos conservadoramente interesándonos sólo por $x + 0.5$; así obtenemos el estadístico de prueba z , como en la ecuación y en la figura 13-1.

Uso de la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego seleccione **Sign Test**. Elija la opción **Given Number of Signs** si usted conoce el número de signos positivos y negativos, o seleccione **Given Pairs of Values** si hay datos apareados en la ventana de datos. Después de realizar las entradas requeridas en el cuadro de diálogo, los resultados en la pantalla incluirán el estadístico de prueba, el valor crítico y la conclusión.

MINITAB Primero debe crear una columna de valores que representen las diferencias entre los datos apareados o el número de signos positivos y negativos. (Para más detalles véase el manual *Minitab Student Laboratory Manual and Workbook*). Seleccione **Stat**, luego **Nonparametrics** y después **1-Sample Sign**. Haga clic en el botón para **Test Median**. Ingrese el valor de la mediana y seleccione el tipo de prueba, luego haga clic en **OK**. Minitab dará el valor P . Usted debe rechazar la hipótesis nula si el valor P es menor o igual que el nivel de significancia. De no ser así, no rechace la hipótesis nula.

EXCEL Excel no tiene una función predeterminada para la prueba del signo, pero usted puede utilizar la función **BINOM.DIST** de Excel para calcular el valor P para una prueba del signo. Haga clic en **fx** en la barra del menú principal, luego seleccione la categoría de función **Statistical** y después **BINOMDIST**. En el cuadro de diálogo, ingrese primero x , luego el número de ensayos n y luego una probabilidad de 0.5. Ingrese **TRUE** en el cuadro para “**cumulative**”. El valor resultante es la probabilidad de obtener x o menos éxitos entre n ensayos. *Duplicque este valor para pruebas de dos colas*. El resultado final es el valor P .

También se puede utilizar el complemento **DDXL** al seleccionar **Nonparametric Tests** y luego **Sign Test**.

TI-83/84 PLUS La calculadora TI-83/84 Plus no tiene una función predeterminada para la prueba del signo, pero se puede utilizar la función **binomcdf** para calcular el valor P para una prueba del signo. Oprima **2nd, VARS** (para obtener el menú **DISTR**); luego baje el cursor para seleccionar **binomcdf**. Complete la entrada de **binomcdf (n, p, x)** con n para el número total

de signos positivos y negativos, 0.5 para p y el número del signo menos frecuente para x . Ahora oprima **ENTER**; el resultado será la probabilidad de obtener x o menos éxitos entre n ensayos. *Duplicue este valor para pruebas de dos colas*. El resultado final es el valor P ; por lo tanto, rechace la hipótesis nula si el valor P es menor o igual que el nivel de significancia. De no ser así, no rechace la hipótesis nula. Por ejemplo, vea la siguiente pantalla de resultados sobre el primer ejemplo de esta sección. Con 7 signos negativos y 4 signos positivos, $n = 11$. La probabilidad se duplica para obtener un valor P de 0.548828125.

TI-83/84 Plus

```
binomcdf(11,.5,4
          ) .2744140625
Ans*2      .548828125
```

13-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Prueba no paramétrica.** ¿Por qué la prueba del signo se considera una prueba “no paramétrica” o una prueba “de distribución libre”?
- Prueba del signo.** ¿Por qué el procedimiento descrito en esta sección se conoce como prueba “del signo”?
- Procedimiento de la prueba del signo.** Se le asignó la tarea de probar la aseveración de que un método de selección del género tiene el efecto de aumentar la probabilidad de que un bebé sea niña, y los datos muestrales consisten en 20 niñas entre 80 bebés recién nacidos. Sin aplicar el procedimiento formal de la prueba del signo, ¿qué concluye usted acerca de la aseveración? ¿Por qué?
- Eficiencia de la prueba del signo.** Remítase a la tabla 13-2 e identifique la eficiencia de la prueba del signo. ¿Qué nos indica ese valor acerca de la prueba del signo?

En los ejercicios 5 a 8, suponga que los datos apareados dan por resultado el número dado de signos cuando el valor de la segunda variable se resta del correspondiente valor de la primera variable. Utilice la prueba del signo con un nivel de significancia de 0.05 y pruebe la hipótesis nula de ninguna diferencia.

- Signos positivos: 15; signos negativos: 4; empates: 1
- Signos positivos: 3; signos negativos: 12; empates: 2

7. Signos positivos: 30; signos negativos: 35; empates: 3

8. Signos positivos: 50; signos negativos: 40; empates: 4

En los ejercicios 9 a 18, utilice la prueba del signo.

9. **¿El viernes 13 es de mala suerte?** Investigadores reunieron datos sobre el número de admisiones hospitalarias resultantes de choques de vehículos, y a continuación se presentan los resultados de los viernes 6 de un mes y de los siguientes viernes 13 del mismo mes (según datos de “Is Friday the 13th Bad for Your Health?”, de Scanlon *et al.*, *BMJ*, vol. 307, tal como aparece en el recurso de datos en línea *Data and Story Line*). Utilice un nivel de significancia de 0.05 para probar la aseveración de que, cuando el día 13 del mes cae en viernes, el número de admisiones hospitalarias por choques de vehículos no se ve afectado.

Viernes 6: 9 6 11 11 3 5

Viernes 13: 13 12 14 10 4 12

10. **Prueba de semillas de maíz.** En 1908 William Gosset publicó al artículo “The Probable Error of a Mean”, bajo el seudónimo de “Student” (*Biometrika*, vol. 6, núm. 1). Él incluyó la lista que parece abajo, acerca de las cosechas de dos tipos diferentes de semillas (normales y secadas en horno), que se utilizaron en parcelas de tierra adyacentes. Los valores listados son las cosechas de paja en cwt por acre, donde cwt representa 100 libras. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que no hay diferencia entre las cosechas de los dos tipos de semillas. ¿Parece que alguna de las semillas es mejor?

Normales	19.25	22.75	23	23	22.5	19.75	24.5	15.5	18	14.25	17
Secadas en horno	25	24	24	28	22.5	19.5	22.25	16	17.25	15.75	17.25

11. **Prueba para la diferencia entre estaturas de hombres reportadas y medidas.** Como parte de la National Health and Nutrition Examination Survey, realizada por el Departamento de Salud de Estados Unidos, se obtuvieron las estaturas reportadas y medidas de varones entre 12 y 16 años de edad. Abajo se listan los resultados muestrales. ¿Existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas reportadas y las estaturas medidas de varones de 12 a 16 años de edad? Utilice un nivel de significancia de 0.05.

Estaturas reportadas	68	71	63	70	71	60	65	64	54	63	66	72
Estaturas medidas	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67.0	55.6	74.2	65.0	70.8

12. **Estatura de ganadores y de segundos lugares.** A continuación se listan las estaturas de candidatos que ganaron las elecciones presidenciales y las estaturas de los candidatos que obtuvieron el segundo número más alto de votos del electorado. Los datos aparecen en orden cronológico, de manera que las estaturas de las dos listas están apareadas. Para los candidatos que ganaron más de una vez, sólo se incluyen las estaturas de la primera elección, y no se incluyen elecciones previas a 1900. Una creencia generalizada asegura que los candidatos ganadores tienden a ser más altos que los candidatos perdedores correspondientes. Utilice un nivel de significancia de 0.05 para probar esa aseveración. Al parecer, ¿la estatura es un factor importante para ganar la presidencia?

Ganador de la presidencia	Segundo lugar											
71 74.5 74 73 69.5 71.5 75 72	73	74	68	69.5	72	71	72	71.5				
70.5 69 74 70 71 72 70 67	70	68	71	72	70	72	72	72				

13. **Prueba para temperatura corporal media de 98.6°F.** En una clase de estadística, se pide a una estudiante del curso propedéutico de medicina que desarrolle un proyecto.

Inspirada por las temperaturas corporales del conjunto de datos 2 del apéndice B, ella planea reunir sus propios datos muestrales para probar la aseveración de que la temperatura corporal media es menor que 98.6°F. Por restricciones de tiempo, se da cuenta de que sólo podrá reunir datos de 12 personas. Después de planear con cuidado un procedimiento para obtener una muestra aleatoria de 12 adultos saludables, ella mide sus temperaturas corporales y obtiene los resultados que se listan abajo. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que estas temperaturas corporales provienen de una población con una mediana que es menor que 98.6°F.

97.6 97.5 98.6 98.2 98.0 99.0 98.5 98.1 98.4 97.9 97.9 97.7

- 14. Prueba para la mediana del peso de monedas de 25 centavos.** A continuación se listan los pesos (en gramos) de monedas de 25 centavos, acuñadas después de 1964, seleccionadas al azar. Se supone que el peso de las monedas tiene una mediana de 5.670 g. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la mediana es igual a 5.670 g. Al parecer, ¿las monedas están acuñadas según las especificaciones?

5.7027 5.7495 5.7050 5.5941 5.7247 5.6114 5.6160 5.5999 5.7790 5.6841

- 15. Datos nominales: Selección del género de niños.** El Genetics and IVF Institute realizó un ensayo clínico del método YSORT, diseñado para incrementar la probabilidad de concebir un niño. Por el tiempo en que se escribió este libro, habían nacido 51 bebés de padres que utilizaron el método YSORT, y 39 de ellos fueron niños. Utilice los datos muestrales con un nivel de significancia de 0.01 para probar la aseveración de que, con este método, la probabilidad de que un bebé sea niño es mayor que 0.5. ¿Parece que el método funciona?

- 16. Datos nominales: Choques de automóviles.** En un estudio de 11,000 choques de automóviles, se descubrió que 5720 de ellos ocurrieron a una distancia no mayor de 5 millas de casa del conductor (según datos de Progressive Insurance). Utilice un nivel de significancia de 0.01 para probar la aseveración de que más del 50% de los choques de automóviles ocurren a una distancia no mayor de 5 millas de casa del conductor. ¿Los resultados son cuestionables porque se basan en una encuesta financiada por una compañía de seguros?

- 17. Datos nominales: Viaje a través de Internet.** De 734 usuarios de Internet elegidos al azar, se descubrió que 360 de ellos usan Internet para planear viajes (según datos de una encuesta Gallup). Utilice un nivel de significancia de 0.01 para probar la aseveración de que, de los usuarios de Internet, menos del 50% utiliza este medio para planear viajes. ¿Los resultados son importantes para los agentes de viajes?

- 18. Posposición de la muerte.** Una creencia interesante y generalizada asegura que las personas pueden posponer temporalmente su muerte para estar presentes en una festividad o un suceso importante como un cumpleaños. En un estudio de este fenómeno, se registraron 6062 muertes la semana previa al Día de Acción de Gracias, y 5938 muertes la semana posterior a esa festividad (según datos de “Holidays, Birthdays, and Postponement of Cancer Death”, de Young y Hade, *Journal of the American Medical Association*, vol. 292, núm. 24). Si la gente puede posponer su muerte para después del Día de Acción de Gracias, entonces la proporción de muertes la semana anterior debe ser menor que 0.5. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la proporción de muertes la semana anterior al Día de Acción de Gracias es menor que 0.5. Con base en el resultado, ¿parece haber alguna indicación de que las personas pueden posponer temporalmente su muerte para estar presentes el Día de Acción de Gracias?

13-2 MÁS ALLÁ DE LO BÁSICO

- 19. Procedimientos para manejar empates.** En el procedimiento de la prueba del signo descrito en esta sección, excluimos los empates (representados por 0 en vez de un signo de + o de -). Un segundo método consiste en tratar a la mitad de los ceros como

signos positivos y a la otra mitad como negativos. (Si el número de ceros es impar, excluya uno para que puedan dividirse por igual). Con un tercer método, en pruebas de dos colas, haga la mitad de los ceros positivos y la mitad negativos. En pruebas de una cola haga todos los ceros positivos o negativos; cualquier signo sustenta la hipótesis nula. Suponga que cuando se utiliza la prueba del signo para una aseveración de que el valor de la mediana es menor que 100, obtenemos 60 valores por debajo de 100, 40 valores por encima de 100, y 21 valores iguales a 100. Identifique el estadístico de prueba y la conclusión con las tres formas diferentes de manejar empates (con diferencias de 0). Suponga un nivel de significancia de 0.05 en los tres casos.

- 20. Cálculo de valores críticos.** La tabla A-7 lista valores críticos para alternativas limitadas de α . Utilice la tabla de A-1 para añadir una nueva columna en la tabla de A-7 (bajando hasta $n = 15$) que represente un nivel de significancia de 0.03 en una cola o de 0.06 en dos colas. Para cualquier n particular, utilice $p = 0.5$, puesto que la prueba del signo requiere el supuesto de que $P(\text{signo positivo}) = P(\text{signo negativo}) = 0.5$. La probabilidad de x o menos signos del mismo tipo es la suma de las probabilidades de los valores hasta x , inclusive.
- 21. Error de aproximación normal.** La compañía Compulife.com contrató a 18 mujeres entre los últimos 54 empleados nuevos. De los solicitantes, alrededor de la mitad son mujeres y la otra mitad hombres; todos los aspirantes están calificados. Utilizando un nivel de significancia de 0.01 con la prueba del signo, ¿existe suficiente evidencia para acusar de favoritismo? ¿La conclusión cambia si se utiliza la distribución binomial en vez de la distribución normal aproximada?

Prueba de rangos con signo

13-3 de Wilcoxon para datos apareados

Concepto clave Esta sección se ocupa de la *prueba de rangos con signo de Wilcoxon*, que se usa con datos muestrales apareados. Esta prueba se utiliza con la hipótesis nula de que la población de diferencias de los datos apareados tiene una mediana igual a cero.

La prueba del signo (sección 13-2) también se puede usar con datos apareados, pero la prueba del signo sólo utiliza los signos de las diferencias. Al utilizar los rangos, la prueba de rangos con signo (o de rangos signados) de Wilcoxon toma en cuenta las magnitudes de las diferencias. Puesto que la prueba de rangos con signo de Wilcoxon incorpora y utiliza más información que la prueba del signo, tiende a arrojar conclusiones que reflejan mejor la verdadera naturaleza de los datos.

Definición

La **prueba de rangos con signo de Wilcoxon** es una prueba no paramétrica que utiliza rangos ordenados de datos muestrales que consisten en datos apareados. Se usa para probar la hipótesis nula de que la población de diferencias tiene una mediana de cero, de manera que las hipótesis nula y alternativa son las siguientes:

H_0 : Los datos apareados tienen diferencias que provienen de una población con una mediana igual a cero.

H_1 : Los datos apareados tienen diferencias que provienen de una población con una mediana diferente de cero.

(La prueba de rangos con signo de Wilcoxon también puede usarse para probar la aseveración de que una muestra proviene de una población con una mediana específica. Véase el ejercicio 13 para esta aplicación).



Brecha de género en las pruebas de fármacos

Un estudio de la relación entre los ataques cardíacos y las dosis administradas de aspirina incluyó a 22,000 médicos varones. Este estudio, como muchos otros, excluyó a las mujeres. La General Accounting Office criticó hace poco a los institutos nacionales de salud por no incluir a ambos sexos en muchos estudios, ya que los resultados de pruebas médicas en hombres no necesariamente se aplican a las mujeres. Por ejemplo, los corazones de las mujeres son diferentes de los de los hombres en muchos aspectos importantes. Cuando saquemos conclusiones con base en resultados muestrales, debemos ser cuidadosos al generalizar las inferencias a una población más grande que aquella de la cual se obtuvo la muestra.

Prueba de rangos con signo de Wilcoxon

Requisitos

1. Los datos consisten en datos apareados que se seleccionaron aleatoriamente.
2. La población de las diferencias (calculadas a partir de los pares de datos) tiene una distribución que es aproximadamente *simétrica*, lo que quiere decir que la mitad izquierda de su histograma es aproximadamente una imagen de espejo de la mitad derecha. (*No* existe el requisito de que los datos tengan una distribución normal).

Notación

El procedimiento para calcular la suma de rangos T se incluye después de este recuadro.

$T =$ la más pequeña de las siguientes dos sumas:

1. La suma de los valores absolutos de los rangos negativos de las diferencias d que no sean cero.
2. La suma de los rangos positivos de las diferencias d que no sean cero.

Estadístico de prueba

Si $n \leq 30$, el estadístico de prueba es T .

$$T = \frac{n(n + 1)}{4}$$

Si $n > 30$, el estadístico de prueba es $z = \frac{\sqrt{\frac{n(n + 1)(2n + 1)}{24}}}{\sqrt{\frac{n(n + 1)(2n + 1)}{24}}}$

Valores críticos

1. Si $n \leq 30$, el valor crítico T se encuentra en la tabla A-8.
2. Si $n > 30$, los valores críticos z se encuentran en la tabla A-2.

Procedimiento de la prueba de rangos con signo de Wilcoxon

- Paso 1: Para cada par de datos, calcule la diferencia d restando el segundo valor del primero. Mantenga los signos, pero descarte cualquier par para el que $d = 0$.
- Paso 2: *Ignore los signos de las diferencias*, luego acomode las diferencias de la menor a la mayor y reemplácelas por el valor del rango correspondiente (como se describe en la sección 13-1). Cuando las diferencias tengan el mismo valor numérico, asígneles la media de los rangos implicados en el empate.
- Paso 3: Agregue a cada rango el signo de la diferencia de la que provino. Esto es, inserte aquellos signos que se ignoraron en el paso 2.
- Paso 4: Calcule la suma de los valores absolutos de los rangos negativos. También calcule la suma de los rangos positivos.
- Paso 5: Permita que T sea *la más pequeña* de las dos sumas calculadas en el paso 4. Podría utilizarse cualquier suma, pero para simplificar el procedimiento seleccionamos arbitrariamente la más pequeña de las dos sumas. (Véase la notación para T en el recuadro anterior).
- Paso 6: Permita que n sea el número de pares de datos para los que la diferencia d no es 0.
- Paso 7: Determine el estadístico de prueba y los valores críticos con base en el tamaño muestral, como se indica en el recuadro anterior.

Paso 8: Cuando plantea la conclusión, rechace la hipótesis nula si los datos muestrales le llevan a un estadístico de prueba que se ubica en la región crítica, esto es, cuando el estadístico de prueba sea menor o igual que el valor (o los valores) crítico(s). De otra forma, no rechace la hipótesis nula.

EJEMPLO ¿El tipo de semilla afecta el crecimiento del maíz? En 1908 William Gosset publicó el artículo “The Probable Error of a Mean” bajo el seudónimo de “Student” (*Biometrika*, vol. 6, núm. 1). Él incluyó los datos que se listan en la tabla 13-4 para dos tipos diferentes de semillas de maíz (normales y secadas en horno), que se utilizaron en parcelas de tierra *adyacentes*. Los valores corresponden a las cosechas de cabezas de maíz (o mazorcas) en libras por acre. Utilice la prueba de rangos con signos de Wilcoxon, con un nivel de significancia de 0.05, para probar la aseveración de que no hay diferencia entre las cosechas de las semillas normales y de las semillas secadas en horno.

SOLUCIÓN

REQUISITOS ✓ Debemos tener datos apareados seleccionados al azar. Los datos están apareados y, dado el diseño de este experimento, es razonable suponer que los datos apareados fueron elegidos al azar. Además, el siguiente histograma generado por Minitab indica que la distribución de las diferencias es aproximadamente simétrica, tal como se requiere. (Es decir, el lado izquierdo de la gráfica es aproximadamente una imagen de espejo del lado derecho. A simple vista no parecen ser simétricos, pero con sólo 11 valores, la diferencia entre el lado izquierdo y el lado derecho no es demasiado pronunciada). ✓

continúa

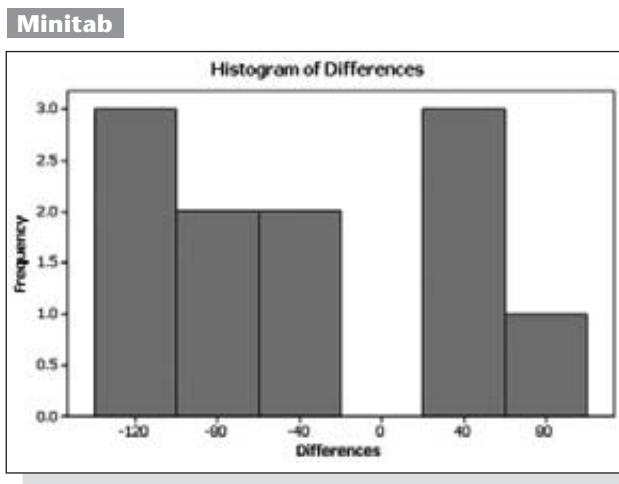


Tabla 13-4 Cosechas de maíz de diferentes semillas

Normales	1903	1935	1910	2496	2108	1961	2060	1444	1612	1316	1511
Secadas en horno	2009	1915	2011	2463	2180	1925	2122	1482	1542	1443	1535
Diferencias <i>d</i>	-106	20	-101	33	-72	36	-62	-38	70	-127	-24
Rangos de diferencias	10	1	9	3	8	4	6	5	7	11	2
Rangos con signo	-10	1	-9	3	-8	4	-6	-5	7	-11	-2

Las hipótesis nula y alternativa son como sigue:

H_0 : Las cosechas de las semillas normales y de las semillas secadas en horno son tales que la mediana de la población de las diferencias es igual a cero.

H_1 : La mediana de la población de diferencias no es igual a cero.

El nivel de significancia es $\alpha = 0.05$. Estamos utilizando el procedimiento de la prueba de rangos con signo de Wilcoxon, por lo que el estadístico de prueba se calcula aplicando el procedimiento de ocho pasos presentado con anterioridad en esta sección.

Paso 1: En la tabla 13-4 el renglón de diferencias se obtiene calculando esta diferencia para cada par de datos:

$$d = \text{cosecha de las semillas normales} - \text{cosecha de las semillas secadas en horno}$$

Paso 2: Ignorando sus signos, ordenamos los rangos de las diferencias absolutas de la menor a la mayor. (Si existiera algún empate en los rangos tendríamos que manejarlos asignando la media de los rangos implicados a cada uno de los valores empatados. Además, tendríamos que descartar cualquier diferencia de 0).

Paso 3: El renglón inferior de la tabla 13-4 se crea insertando a cada rango el signo de la diferencia correspondiente. Si en realidad no existe diferencia entre las cosechas de los dos tipos de semillas (como en la hipótesis nula), esperamos que la suma de los rangos positivos sea aproximadamente igual a la suma de los valores absolutos de los rangos negativos.

Paso 4: Ahora calculamos la suma de los valores absolutos de los rangos negativos y también calculamos la suma de los rangos positivos.

$$\begin{aligned} \text{Suma de los valores absolutos de los rangos negativos: } & 51 \quad (\text{de } 10 + 9 \\ & + 8 + 6 + 5 + 11 + 2) \end{aligned}$$

$$\begin{aligned} \text{Suma de los rangos positivos: } & 15 \quad (\text{de } 1 + 3 + 4 + 7) \end{aligned}$$

Paso 5: Permitiendo que T sea la menor de las dos sumas calculadas en el paso 4, encontramos que $T = 15$.

Paso 6: Permitiendo que n sea el número de pares de datos para los que la diferencia d no es 0, tenemos $n = 11$.

Paso 7: Puesto que $n = 11$, tenemos que $n \leq 30$, por lo cual utilizamos un estadístico de prueba de $T = 15$ (y no calculamos un estadístico de prueba z). Además, puesto que $n \leq 30$, utilizamos la tabla de A-8 para encontrar el valor crítico de 11 (utilizando $n = 11$ y $\alpha = 0.05$ en dos colas).

Paso 8: El estadístico de prueba $T = 15$ no es menor o igual que el valor crítico de 11, por lo que no rechazamos la hipótesis nula. Parece que no hay una diferencia entre las cosechas de las semillas normales y de las semillas secadas en horno.

Si utilizamos la prueba del signo con el ejemplo anterior, llegaremos a la misma conclusión. Aunque la prueba de signo y la prueba de rangos con signo de Wilcoxon coinciden en este caso en particular, existen otros casos en los que no concuerdan.

Fundamentos: En este ejemplo los rangos sin signo de 1 hasta 11 tienen un total de 66, de manera que si no existen diferencias significativas, cada uno de los dos totales de rangos con signo debe ser de alrededor de $66 \div 2$, o 33. Esto es, los

rangos negativos y los rangos positivos deberían repartirse como 33-33 o algo cercano, tal como 31-35. La tabla de valores críticos indica que a un nivel de significancia de 0.05, con 11 pares de datos, un reparto de 11-55 representa una desviación significativa de la hipótesis nula, y cualquier reparto que esté más separado (como 10-56 o 2-64) también representará una desviación significativa de la hipótesis nula. Por el contrario, repartos como 12-52 no representan desviaciones significativas de un reparto de 33-33, y no justificarán el rechazo de la hipótesis nula. La prueba de rangos con signo de Wilcoxon está basada en el total del rango más bajo, por lo que en vez de analizar los dos números que constituyen el reparto, consideramos sólo el número más bajo.

La suma $1 + 2 + 3 + \dots + n$ de todos los rangos es igual a $n(n + 1)/2$, y si ésta es una suma de rangos a dividirse por igual entre dos categorías (positivo y negativo), cada uno de los dos totales debería estar cerca de $n(n + 1)/4$, que es la mitad de $n(n + 1)/2$. El reconocimiento de este principio nos ayuda a entender el estadístico de prueba que se usa cuando $n > 30$. El denominador en esa expresión representa una desviación estándar de T y se basa en el principio de que

$$1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$$

La prueba de rangos con signo de Wilcoxon puede utilizarse sólo para datos apareados. La siguiente sección describirá una prueba de suma de rangos que puede aplicarse a dos conjuntos de datos independientes que no están asociados en pares.

Uso de la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego seleccione **Wilcoxon Tests**. Ahora elija **Signed-Ranks Test** y proceda a seleccionar las columnas de datos. Haga clic en **Evaluate**.

MINITAB Ingrese los datos apareados en las columnas C1 y C2. Haga clic en **Editor**, luego en **Enable Command Editor** e ingrese el comando **LET C3 = C1 - C2**. Oprima la tecla **Enter**. Seleccione las opciones **Stat**, **Nonparametrics** y **1-Sample Wilcoxon**. Ingrese C3 para la variable y haga clic en el botón para **Test Median**. La pantalla de Minitab incluirá el valor P . Vea la pantalla de resultados de Minitab para el ejemplo de esta sección. El valor P de

0.120 es mayor que el nivel de significancia de 0.05, por lo que no se rechaza la hipótesis nula.

EXCEL Excel no está programado para la prueba de rangos con signo de Wilcoxon, pero se puede emplear el complemento DDXL al seleccionar **Non-parametric Tests** y luego **Paired Wilcoxon**.

TI-83/84 PLUS La calculadora TI-83/84 Plus no está programada para la prueba de rangos con signo de Wilcoxon, pero se puede usar el programa SRTEST (de Michael Lloyd), el cual se puede descargar del sitio <http://www.pearsonedicion.net/triola>. Primero descargue e instale el programa. (También descargue el programa ZZRANK, necesario para el programa SRTEST). Luego, elabore una lista de las

diferencias entre los valores de los datos apareados. (El primer conjunto de valores se puede insertar en la lista L1, el segundo conjunto de valores en la lista L2, y luego las diferencias se pueden almacenar en la lista L3 ingresando $L1 - L2 \rightarrow L3$, donde la tecla **STO** se utiliza para la flecha). Presione la tecla **PRGM** y seleccione **SRTEST**. Presione la tecla **ENTER** dos veces. Cuando aparezca el indicador de **DATA=**, introduzca la lista que contiene las diferencias. Presione **ENTER** para ver la suma de los rangos positivos y la suma de los rangos negativos. Presione **ENTER** nuevamente para ver la media y la desviación estándar, y presione **ENTER** una vez más para ver la puntuación z . Si $n \leq 30$, busque de valor crítico T en la tabla A-8, pero si $n > 30$, obtenga los valores críticos z en la tabla A-2.

Minitab

Wilcoxon Signed Rank Test: Differences					
Test of median = 0.000000 versus median not = 0.000000					
Differences	N	for Wilcoxon		P	Estimated Median
		N	Test Statistic		
11	11	15.0	0.120	-34.50	

13-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Prueba de rangos con signo de Wilcoxon.** ¿Por qué se utiliza la prueba de rangos con signo de Wilcoxon para datos apareados en vez de los métodos presentados en la sección 9-4?
- Prueba de rangos con signo de Wilcoxon y prueba del signo.** Tanto la prueba de rangos con signo de Wilcoxon como la prueba del signo se pueden utilizar con datos muestrales apareados. ¿Ambas podrían conducirnos a la misma conclusión? ¿Ambas siempre conducen a la misma conclusión?
- Prueba de rangos con signo de Wilcoxon.** Si tenemos datos muestrales consistentes en datos apareados, podemos comparar los valores utilizando la prueba del signo o la prueba de rangos con signo de Wilcoxon. ¿Qué importante ventaja tiene la prueba de rangos con signo de Wilcoxon con respecto a la prueba del signo?
- Prueba de rangos con signo de Wilcoxon.** Un investigador de mercados diseña un experimento que incluye una encuesta de parejas casadas en centros comerciales elegidos al azar. Después de separar a los hombres y a las mujeres, se pregunta a cada uno cuánto gastó en sus compras durante la última hora. Los resultados se analizan con la prueba de rangos con signo de Wilcoxon. ¿Las conclusiones se aplican a la población de parejas casadas de Estados Unidos? ¿Por qué?

Uso de la prueba de rangos con signo de Wilcoxon. En los ejercicios 5 y 6, remítase a los datos muestrales apareados indicados y utilice la prueba de rangos con signo de Wilcoxon para probar la aseveración de que los datos apareados tienen diferencias que provienen de una población con una mediana igual a cero. Utilice un nivel de significancia de 0.05.

5.

x	60	55	89	92	78	84	93	87
y	35	27	47	44	39	48	51	54

6.

x	90	93	112	97	102	115	148	152	121
y	88	91	115	95	103	116	150	147	119

Uso de la prueba de rangos con signo de Wilcoxon. En los ejercicios 7 a 10, remítase a los datos muestrales para los ejercicios de la sección 13-2. En vez de la prueba del signo, utilice la prueba de rangos con signo de Wilcoxon para probar la aseveración de que los datos apareados tienen diferencias que provienen de una población con una mediana igual a cero. Use un nivel de significancia de 0.05.

7. Ejercicio 9

8. Ejercicio 10

9. Ejercicio 11

10. Ejercicio 12

Conjuntos de datos del apéndice B. En los ejercicios 11 y 12, utilice la prueba de rangos con signo de Wilcoxon con los datos del apéndice B.

- Prueba de la diferencia entre temperaturas pronosticadas y reales.** Remítase al conjunto de datos 8 del apéndice B y utilice las temperaturas máximas reales y el pronóstico de temperaturas máximas de tres días. ¿Parece existir una diferencia?
- Prueba para la diferencia entre los tiempos en que se presenta consumo de alcohol y de tabaco.** Remítase al conjunto de datos 5 del apéndice B. Utilice sólo aquellas películas que presentan algún consumo de tabaco o alcohol (es decir, ignore aquellas películas con tiempos de cero para consumo de tabaco y para consumo del alcohol). ¿Parece existir una diferencia?

13-3 MÁS ALLÁ DE LO BÁSICO

- 13. Uso de la prueba de rangos con signo de Wilcoxon para aseveraciones acerca de una mediana.** La prueba de rangos con signo de Wilcoxon puede utilizarse para probar la aseveración de que una muestra proviene de una población con una mediana específica. El procedimiento que se utiliza es el mismo que se describió en esta sección, excepto que las diferencias (paso 1) se obtienen restando el valor de la mediana hipotética de cada valor. Utilice los datos muestrales que consisten en las 106 temperaturas corporales listadas para las 12 a.m. del día 2 en el conjunto de datos 2 del apéndice B. Con un nivel de significancia de 0.05, pruebe la aseveración de que los adultos saludables tienen una mediana de temperatura corporal igual a 98.6°F.

Prueba de la suma de rangos de Wilcoxon para dos muestras

13-4 independientes

Concepto clave Esta sección describe la *prueba de la suma de rangos de Wilcoxon*, que utiliza los rangos de los valores de dos conjuntos independientes de datos muestrales para probar la hipótesis nula de que las dos poblaciones tienen medianas iguales.

La prueba de rangos con signo de Wilcoxon (sección 13-3) implica datos apareados, pero la prueba de suma de rangos de Wilcoxon de esta sección implica dos muestras independientes que no están relacionadas ni asociadas o apareadas. (Para evitar confusiones entre la prueba de suma de rangos de Wilcoxon para muestras independientes y la prueba de rangos con signo de Wilcoxon para datos apareados, considere el uso de las siglas ISR del Impuesto Sobre la Renta como recurso mnemotécnico para recordarnos “independiente: suma de rangos”).

Definición

La **prueba de la suma de rangos de Wilcoxon** es una prueba no paramétrica que utiliza rangos de datos muestrales de dos poblaciones independientes. Se utiliza para probar la hipótesis nula de que las dos muestras independientes provienen de poblaciones con medianas iguales. La hipótesis alternativa es la aseveración de que las dos poblaciones tienen medianas diferentes.

H_0 : Las dos muestras provienen de poblaciones con medianas iguales.

H_1 : Las dos muestras provienen de poblaciones con medianas diferentes.

Concepto básico: La prueba de la suma de rangos de Wilcoxon es equivalente a la **prueba U de Mann-Whitney** (véase el ejercicio 13), que se incluye en algunos otros libros de texto y programas de cómputo (como el Minitab). La idea fundamental que subyace en la prueba de la suma de rangos de Wilcoxon es la siguiente: si dos muestras se obtienen de poblaciones idénticas y los valores individuales se acomodan en rangos como un conjunto combinado de valores, entonces el rango alto y el bajo deberían caer de manera uniforme entre las dos muestras. Si los rangos bajos se encuentran predominantemente en una muestra y los rangos altos se encuentran predominantemente en la otra muestra, sospechamos que las dos poblaciones tienen medianas diferentes.

Prueba de la suma de rangos de Wilcoxon

Requisitos

1. Hay dos muestras independientes de datos seleccionados al azar.
2. Cada una de las dos muestras tiene más de 10 valores. (Para muestras con 10 valores o menos, en libros de referencia están disponibles tablas especiales, como las *CRC Standard Probability and Statistics Tables and Formulae*, publicadas por CRC Press).
3. No existe el requisito de que las dos poblaciones tengan una distribución normal o cualquier otra distribución particular.

Notación

n_1 = tamaño de la muestra 1

n_2 = tamaño de la muestra 2

R_1 = suma de rangos de la muestra 1, que se calcula utilizando el procedimiento que se describe a continuación

R_2 = suma de rangos de la muestra 2, que se calcula utilizando el procedimiento que se describe a continuación

R = lo mismo que R_1 (suma de rangos de la muestra 1)

μ_R = media de los valores muestrales R que se espera cuando las dos poblaciones tienen medianas iguales

σ_R = desviación estándar de los valores muestrales R que se espera cuando las dos poblaciones tienen medianas iguales

Estadístico de prueba

$$z = \frac{R - \mu_R}{\sigma_R}$$

donde

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

n_1 = tamaño de la muestra a partir de la cual se calcula la suma de rangos R

n_2 = tamaño de la otra muestra

R = suma de rangos de la muestra con tamaño n_1

Valores críticos: Los valores críticos pueden encontrarse en la tabla A-2 (puesto que el estadístico de prueba está basado en la distribución normal).

Procedimiento para calcular el valor del estadístico de prueba

1. Combine temporalmente las dos muestras en una muestra grande, entonces reemplace cada valor muestral por su rango. (El valor más bajo toma un rango de 1, el siguiente valor más bajo toma un rango de 2, etcétera. Si los valores están empacados, asigneles la media de los rangos implicados en el empate. Consulte la sección 13-1 para ver una descripción de los rangos y el procedimiento para manejar empates).
2. Calcule la suma de los rangos de las dos muestras.

3. Calcule el valor del estadístico de prueba z como se indicó en el recuadro anterior, donde cualquier muestra puede utilizarse como la “muestra 1”. (Si ambos tamaños muestrales son mayores que 10, entonces la distribución muestral de R es aproximadamente normal, con media μ_R y desviación estándar σ_R , y el estadístico de prueba es como se mostró en el recuadro anterior).

Note que, a diferencia de las pruebas de hipótesis correspondientes en la sección 9-3, la prueba de la suma de rangos de Wilcoxon *no* requiere poblaciones distribuidas normalmente. Además, la prueba de la suma de rangos de Wilcoxon puede utilizarse con datos en el nivel de medición ordinal, como los datos que consisten en rangos. En contraste, los métodos paramétricos de la sección 9-3 no se pueden utilizar con datos en el nivel de medición ordinal. En la tabla 13-2 notamos que la prueba de la suma de rangos de Wilcoxon tiene una tasa de eficiencia de 0.95 en comparación con la prueba paramétrica t o con la prueba z . Puesto que esta prueba tiene una alta tasa de eficiencia y requiere cálculos más fáciles, suele preferirse sobre las pruebas paramétricas descritas en la sección 9-3, aun cuando se satisfaga el requisito de normalidad.

EJEMPLO IMC de hombres y mujeres Remítase al conjunto de datos 1 del apéndice B y use únicamente los primeros 13 valores muestrales del índice de masa corporal (IMC) de los hombres y los primeros 12 valores muestrales del IMC de las mujeres. En la tabla 13-5 se incluyen los valores muestrales del IMC. (Sólo se emplea parte de los valores muestrales disponibles para que los cálculos en este ejemplo sean más fáciles). Utilice un nivel de significancia de 0.05 para probar la aseveración de que la mediana del IMC de los hombres es igual a la mediana del IMC de las mujeres.

SOLUCIÓN

REQUISITOS ✓ La prueba de suma de rangos de Wilcoxon requiere de dos muestras independientes y aleatorias, cada una con más de 10 valores. Los datos muestrales son independientes y aleatorios, y los tamaños muestrales son 13 y 12. Los requisitos se satisfacen, así que procedemos con la prueba. ✓

Las hipótesis nula y alternativa son las siguientes:

H_0 : Los hombres y las mujeres tienen valores del IMC con medianas iguales.

H_1 : Los hombres y las mujeres tienen valores del IMC con medianas que no son iguales.

Acomode en rangos las 25 mediciones combinadas del IMC, comenzando con un rango de 1 (asignado al valor más bajo de 17.7). Los empates en los rangos se manejan como se describió en la sección 13-1: calcule la media de los rangos implicados y asigne este rango medio a cada uno de los valores empatados. (Los valores segundo y tercero son ambos de 19.6, por lo que se asigna el rango de 2.5 a cada uno de estos valores. Los valores decimoprimero y decimosegundo son ambos de 23.8, por lo que se asigna el rango de 11.5 a cada uno de estos valores. Los valores decimoquinto y decimosexto son ambos de 25.2, y se asigna un rango de 15.5 a cada uno de ellos). En la tabla 13-5, los rangos correspondientes a los valores muestrales individuales se presentan entre paréntesis. R denota la suma de los rangos para la muestra que elegimos como muestra 1. Si elegimos los valores del IMC de los hombres, obtenemos

$$R = 11.5 + 9 + 14 + \cdots + 15.5 = 187$$

Puesto que existen 13 valores para los hombres, tenemos $n_1 = 13$. Además,
continúa

Tabla 13-5	
Mediciones del IMC	
Hombres	Mujeres
23.8 (11.5)	19.6 (2.5)
23.2 (9)	23.8 (11.5)
24.6 (14)	19.6 (2.5)
26.2 (17)	29.1 (22)
23.5 (10)	25.2 (15.5)
24.5 (13)	21.4 (5)
21.5 (6)	22.0 (7)
31.4 (24)	27.5 (19)
26.4 (18)	33.5 (25)
22.7 (8)	20.6 (4)
27.8 (20)	29.9 (23)
28.1 (21)	17.7 (1)
25.2 (15.5)	
$n_1 = 13$	$n_2 = 12$
$R_1 = 187$	$R_2 = 138$

$n_2 = 12$, ya que existen 12 valores para las mujeres. Ahora podemos determinar los valores de μ_R , σ_R y el estadístico de prueba z .

$$\begin{aligned}\mu_R &= \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{13(13 + 12 + 1)}{2} = 169 \\ \sigma_R &= \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{(13)(12)(13 + 12 + 1)}{12}} = 18.385 \\ z &= \frac{R - \mu_R}{\sigma_R} = \frac{187 - 169}{18.385} = 0.98\end{aligned}$$

La prueba es de dos colas, puesto que un valor positivo grande de z indicaría que los rangos más altos se encuentran de forma desproporcionada en la muestra 1, y un valor negativo grande de z indicaría que la muestra 1 tuvo una parte desproporcionada de los rangos más bajos. En cualquier caso, tendríamos una fuerte evidencia en contra de la aseveración de que las dos muestras provienen de poblaciones con medianas iguales.

La significancia del estadístico de prueba z puede tratarse de la misma forma que en los capítulos anteriores. Ahora estamos probando (con $\alpha = 0.05$) la hipótesis de que las dos poblaciones tienen medianas iguales, de manera que tenemos una prueba de dos colas con valores críticos z de 1.96 y -1.96 . El estadístico de prueba de $z = 0.98$ *no* cae dentro de la región crítica, por lo que no rechazamos la hipótesis nula de que los valores del IMC de hombres y mujeres tienen medianas iguales. Parece que los valores del IMC de hombres y mujeres son básicamente iguales.

Podemos verificar que, si intercambiamos los dos conjuntos de valores muestrales y consideramos que la muestra de los valores del IMC de las mujeres es la primera, $R = 138$, $\mu_R = 156$, $\sigma_R = 18.385$ y $z = -0.98$, así que la conclusión es exactamente la misma.

EJEMPLO IMC de hombres y mujeres El ejemplo anterior utilizó sólo 13 de los 40 valores muestrales del IMC de los hombres, listados en el conjunto de datos 1 del apéndice B, y sólo 12 de los 40 valores muestrales del IMC de las mujeres. ¿Cambian los resultados si se emplean los 40 valores muestrales de los hombres y los 40 valores muestrales de las mujeres? Utilice la prueba de suma de rangos de Wilcoxon.

SOLUCIÓN

REQUISITOS ✓ Como en el ejemplo anterior, los datos muestrales son independientes y aleatorios. Además, los tamaños de ambas muestras son mayores que 10. Los requisitos se satisfacen, así que continuamos con la prueba. ✓

Las hipótesis nula y alternativa son iguales a las del ejemplo anterior. En vez de calcular manualmente las sumas de los rangos, nos remitimos a la pantalla de Minitab que se incluye aquí. En esta pantalla de Minitab, “ETA1” y “ETA2” denotan la mediana de la primera muestra y la mediana de la segunda muestra, respectivamente. Los componentes fundamentales de la pantalla de Minitab son los siguientes: la suma de rangos para los valores del IMC de los hombres es $W = 1727.5$, el valor P es 0.3032 (o 0.3031 después de un ajuste por los empates). Puesto que el valor P es mayor que el nivel de significancia de 0.05, no rechazamos la hipótesis nula. No existe evidencia suficiente para justificar el rechazo de la aseveración de que los hombres y las mujeres tienen valores de IMC con medianas iguales. Se sabe que los hombres suelen ser más altos y más pesados que las mujeres, aunque los datos utilizados aquí sugieren que los valores del IMC son aproximadamente iguales para los hombres y para las mujeres.

Minitab

	N	Median
Men	40	26.200
Women	40	23.900

Point estimate for ETA1-ETA2 is 1.250
 95.1 Percent CI for ETA1-ETA2 is (-1.300,3.200)
 $W = 1727.5$
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.3032
 The test is significant at 0.3031 (adjusted for ties)

Uso de la tecnología

STATDISK Ingrese los datos en las columnas de la ventana de datos de Statdisk. Seleccione **Analysis** de la barra del menú principal, luego seleccione **Wilcoxon Tests**, seguido por la opción **Rank-Sum Test**. Seleccione las columnas de los datos y luego haga clic en **Evaluate** para obtener una pantalla que incluye las sumas de rangos, el tamaño muestral, el estadístico de prueba, el valor crítico y la conclusión.



MINITAB Primero ingrese los dos conjuntos de datos muestrales en las columnas C1 y C2. Luego seleccione las opciones **Stat**, **Nonparametrics** y **Mann-Whitney**, y proceda a ingresar C1 para la primera muestra y C2 para la segunda muestra. El nivel de confianza de 95.0 corresponde a un nivel de significancia de $\alpha = 0.05$, y el cuadro “alternate: not equal” se refiere a la hipótesis alternativa, donde “not equal” corresponde a una prueba de hipótesis de dos colas. Minitab da el valor P y la conclusión.

EXCEL Excel no está programado para la prueba de la suma de rangos de Wilcoxon, pero el complemento DDXL se puede utilizar al seleccionar **Nonparametric Tests** y luego **Mann-Whitney Rank Sum**.

TI-83/84 PLUS La calculadora TI-83/84 Plus no está programada para la prueba de la suma de rangos de

Wilcoxon, pero se puede utilizar el programa RSTEST. Este programa (diseñado por Michael Lloyd) se puede descargar del sitio <http://www.pearson-educacion.net/Triola>. Primero descargue e instale el programa. (También descargue el programa ZZRANK, necesario para el programa RSTEST). Luego, registre los dos conjuntos de datos muestrales en forma de lista en L1 y L2. Presione la tecla **PRGM** y seleccione **RSTEST**. Presione la tecla **ENTER** dos veces. Cuando aparezca el indicador de GROUP A=, ingrese **L1** y presione la tecla **ENTER**. Cuando aparezca el indicador de GROUP B=, ingrese **L2** y presione la tecla **ENTER**. La segunda suma de rangos se presentará como el valor de R . También aparecerán la media y la desviación estándar basadas en ese valor de R . Presione **ENTER** nuevamente para obtener la puntuación z basada en la segunda suma de rangos.

13-4 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Prueba de suma de rangos de Wilcoxon.** ¿Cuál es la principal diferencia entre la prueba de rangos con signo de Wilcoxon y la prueba de la suma de rangos de Wilcoxon?
- Prueba de suma de rangos de Wilcoxon.** El estadístico de prueba de la prueba de la suma de rangos de Wilcoxon se basa en la suma de rangos R , cuya distribución es aproximadamente normal. ¿La prueba de la suma de rangos de Wilcoxon es una prueba paramétrica porque requiere de una distribución normal?
- Prueba de suma de rangos de Wilcoxon.** La prueba de la suma de rangos de Wilcoxon y los métodos de prueba de hipótesis descritos en la sección 9-3 se aplican a dos muestras independientes. ¿Qué ventaja tiene la prueba de suma de rangos de Wilcoxon sobre la prueba descrita en la sección 9-3?
- Eficiencia.** Remítase a la tabla 13-2 e identifique la eficiencia de la prueba de la suma de rangos de Wilcoxon. ¿Qué nos indica ese valor acerca de la prueba?

Identificación de las sumas de rangos. En los ejercicios 5 y 6, utilice un nivel de significancia de 0.05 con los métodos de esta sección para identificar las sumas de rangos R_1 y R_2 , μ_R , σ_R , el estadístico de prueba z , los valores críticos z y luego establezca la conclusión sobre una aseveración de medianas iguales.

- | | | | | | | | | | | | | |
|-----------|--------------------------|---|----|----|----|----|----|----|----|----|----|----|
| 5. | Valores de la muestra 1: | 2 | 7 | 10 | 16 | 20 | 22 | 23 | 26 | 27 | 30 | 33 |
| | Valores de la muestra 2: | 3 | 4 | 11 | 14 | 28 | 35 | 40 | 46 | 47 | 52 | 53 |
| 6. | Valores de la muestra 1: | 8 | 15 | 27 | 39 | 45 | 62 | 68 | 72 | 77 | 80 | 87 |
| | Valores de la muestra 2: | 3 | 5 | 9 | 11 | 14 | 21 | 33 | 44 | 61 | 70 | 85 |

Uso de la prueba de la suma de rangos de Wilcoxon. En los ejercicios 7 a 12, utilice la prueba de la suma de rangos de Wilcoxon.

7. ¿Los trastornos psiquiátricos severos están relacionados con factores biológicos?

Un estudio utilizó tomografía computarizada (TC) por rayos X para reunir datos de volúmenes cerebrales de un grupo de pacientes con trastorno obsesivo-compulsivo y un grupo de control de personas saludables. La lista adjunta presenta los resultados muestrales (en mililitros) para volúmenes del hemisferio derecho (según datos de "Neuroanatomical Abnormalities in Obsessive-Compulsive Disorder Detected with Quantitative X-Ray Computed Tomography", de Luxenberg *et al.*, *American Journal of Psychiatry*, vol. 145, núm. 9). Utilice un nivel de significancia de 0.01 y pruebe la aseveración de que los pacientes obsesivo-compulsivos y las personas saludables tienen la misma mediana de volúmenes cerebrales. Con base en este resultado, ¿podemos concluir que el trastorno obsesivo-compulsivo tiene una base biológica?

Pacientes obsesivo-compulsivos				Grupo de control			
0.308	0.210	0.304	0.344	0.519	0.476	0.413	0.429
0.407	0.455	0.287	0.288	0.501	0.402	0.349	0.594
0.463	0.334	0.340	0.305	0.334	0.483	0.460	0.445

- 8. Prueba del efecto de anclaje.** Se dio un lapso de 5 segundos a estudiantes de estadística, seleccionados al azar, para estimar el valor de un producto de números; los resultados se incluyen en la tabla adjunta. (Vea las Actividades de cooperación en equipos al final del capítulo 3). ¿Parece que las muestras son significativamente diferentes?

Estimados de estudiantes a quienes se pidió calcular $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$

1560	169	5635	25	842	40,320	5000	500	1110	10,000
200	1252	4000	2040	175	856	42,200	49,654	560	800

Estimados de estudiantes a quienes se pidió calcular $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

100,000	2000	42,000	1500	52,836	2050	428	372	300	225	64,582
23,410	500	1200	400	49,000	4000	1876	3600	354	750	640

- 9. Prueba de hipótesis de la diferencia de la antigüedad de automóviles y taxis.**

Cuando el autor visitó Dublín en Irlanda, registró la antigüedad de automóviles y taxis seleccionados al azar. A continuación se listan las antigüedades (en años). Utilice un nivel de significancia de 0.05 para probar la aseveración de que existe una diferencia entre la mediana de la antigüedad de un automóvil de Dublín y la mediana de la antigüedad de un taxi de Dublín. Podríamos esperar que los taxis fueran más nuevos, pero, ¿qué sugieren los resultados?

Automóviles	Taxis
4 0 8 11 14 3 4 4 3 5	8 8 0 3 8 4 3 3 6 11
8 3 3 7 4 6 6 1 8 2 15	7 7 6 9 5 10 8 4 3 4
11 4 1 6 1 8	

- 10. Pulos.** Remítase al conjunto de datos 1 del apéndice B, para los pulsos de hombres y mujeres. Utilice únicamente las primeras 13 pulsos de los hombres y únicamente los primeros 12 pulsos de las mujeres para probar la aseveración de que las dos muestras de pulsos provienen de poblaciones con la misma mediana. Utilice un nivel de significancia de 0.05.
- 11. Conjunto de datos grande del apéndice B: Pulos.** Repita el ejercicio 10 utilizando los 40 pulsos de los hombres y los 40 pulsos de las mujeres.
- 12. Conjunto de datos grande del apéndice B: Pesos de peniques.** Remítase al conjunto de datos 14 del apéndice B y utilice los pesos de los “peniques de trigo” y los pesos de los peniques acuñados antes de 1983. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las dos muestras provienen de poblaciones con la misma mediana.

13-4 MÁS ALLÁ DE LO BÁSICO

- 13. Uso de la prueba U de Mann-Whitney.** La prueba U de Mann-Whitney es equivalente a la prueba de la suma de rangos de Wilcoxon para muestras independientes, ya que ambas se aplican a las mismas situaciones y siempre llevan a las mismas conclusiones. En la prueba U de Mann-Whitney calculamos

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

donde

$$U = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R$$

Utilice las mediciones del IMC de la tabla 13-5 de esta sección y calcule el estadístico de prueba z para la prueba U de Mann-Whitney y compárela con el estadístico de prueba z que se calculó utilizando la prueba de la suma de rangos de Wilcoxon.

- 12. Cálculo de valores críticos.** Suponga que tenemos dos tratamientos (A y B) que producen resultados cuantitativos, y que tenemos sólo dos observaciones del tratamiento A y dos observaciones del tratamiento B. No podemos utilizar el estadístico de prueba descrito en esta sección puesto que ambos tamaños muestrales no exceden de 10.

	1	2	3	4	Suma de rangos del tratamiento A
A	A	B	B		3

- Complete la tabla adjunta listando los cinco renglones que corresponden a los otros cinco casos y registre las sumas de rangos correspondientes del tratamiento A.
- Haga una lista de los valores posibles de R , junto con sus probabilidades correspondientes. [Suponga que los renglones de la tabla del inciso a) son igualmente probables].
- ¿Es posible, con un nivel de significancia de 0.10, rechazar la hipótesis nula de que no existe diferencia entre los tratamientos A y B? Explique.

13-5 Prueba de Kruskal-Wallis

Concepto clave En esta sección se describe la *prueba de Kruskal-Wallis*, que utiliza rangos de datos de tres o más muestras independientes para probar la hipótesis nula de que las muestras provienen de poblaciones con medianas iguales.

En la sección 12-2 utilizamos el análisis de varianza de un factor (ANOVA) para probar la hipótesis nula de que tres o más poblaciones tienen la misma media, pero el ANOVA requiere que todas las poblaciones implicadas tengan distribuciones normales. La prueba de Kruskal-Wallis no requiere distribuciones normales.

Definición

La **prueba de Kruskal-Wallis** (también llamada la **prueba H**) es una prueba no paramétrica que utiliza rangos de datos muestrales de tres o más poblaciones independientes. Se utiliza para probar la hipótesis nula de que las muestras independientes provienen de poblaciones con medianas iguales; la hipótesis alternativa es la aseveración de que las poblaciones tienen medianas que no son iguales.

H_0 : Las muestras provienen de poblaciones con medianas iguales.

H_1 : Las muestras provienen de poblaciones con medianas que no son iguales.

Para aplicar la prueba de Kruskal-Wallis, calculamos el *estadístico de prueba H* , el cual tiene una distribución que puede aproximarse por medio la distribución chi cuadrada, siempre y cuando cada muestra tenga al menos cinco observaciones. Cuando utilizamos la distribución chi cuadrada en este contexto, el número de grados de libertad es $k - 1$, donde k es el número de muestras. (Para una revisión rápida de las características clave de la distribución chi cuadrada, véase la sección 7-5).

Prueba de Kruskal-Wallis

Requisitos

1. Tenemos al menos tres muestras independientes, las cuales se seleccionan al azar.
2. Cada muestra tiene al menos cinco observaciones. (Si las muestras tienen menos de cinco observaciones, remítase a tablas especiales de valores críticos, como las *CRC Standard Probability and Statistics Tables and Formulae*, publicadas por CRC Press).
3. No existe el requisito de que las poblaciones tengan una distribución normal o alguna otra distribución particular.

Notación

N = número total de observaciones en todas las muestras combinadas

k = número de muestras

R_1 = suma de los rangos de la muestra 1, que se calcula utilizando el procedimiento que se describe a continuación

n_1 = número de observaciones de la muestra 1

Para la muestra 2, la suma de los rangos es R_2 y el número de observaciones es n_2 , y se utiliza una notación similar para las otras muestras.

Estadístico de prueba

$$H = \frac{12}{N(N+1)} \left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k} \right) - 3(N+1)$$

Valores críticos

1. La prueba es de *cola derecha*.
2. gl = $k - 1$. (Puesto que el estadístico de prueba H puede aproximarse por medio de una distribución chi cuadrada, utilice la tabla A-4 con $k - 1$ grados de libertad, donde k es el número de muestras diferentes).

Procedimiento para calcular el valor del estadístico de prueba H

1. Combine temporalmente todas las muestras en una muestra grande y asigne un rango a cada valor muestral. (Ordene los valores del menor al mayor, y en caso de empates, asigne a cada observación la media de los rangos implicados).
2. En cada muestra, calcule la suma de los rangos y calcule el tamaño muestral.
3. Calcule H utilizando los resultados del paso 2, con la notación y el estadístico de prueba descritos en el recuadro anterior.

El estadístico de prueba H es básicamente una medida de la varianza de las sumas de rangos R_1, R_2, \dots, R_k . Si los rangos están distribuidos de forma equitativa entre los grupos muestrales, entonces H debe ser un número relativamente pequeño. Si las muestras son muy diferentes, entonces los rangos serán excesivamente bajos en algunos grupos y altos en otros, con el efecto neto de que H será grande. En consecuencia, sólo los valores grandes de H nos llevan al rechazo de la hipótesis nula de que las muestras provienen de poblaciones idénticas. *La prueba de Kruskal-Wallis es, por lo tanto, una prueba de cola derecha.*

EJEMPLO Efectos de tratamientos en los pesos de álamos La tabla 13-6 lista los pesos (en kg) de álamos que recibieron tratamientos diferentes. En la sección 12-2 utilizamos el análisis de varianza para probar la hipótesis nula de que las cuatro muestras de pesos provienen de poblaciones con la misma media. Ahora usaremos la prueba de Kruskal-Wallis para probar la hipótesis nula de que las cuatro muestras provienen de poblaciones con medianas iguales.

SOLUCIÓN

REQUISITO ✓ La prueba de Kruskal-Wallis requiere de tres o más muestras aleatorias e independientes, con al menos 5 valores cada una. Las cuatro muestras son independientes y aleatorias, y cada una tiene 5 valores. Una vez satisfechos todos los requisitos, procedemos con la prueba. ✓

Las hipótesis nula y alternativa son las siguientes:

H_0 : Las poblaciones de los pesos de álamos con los cuatro tratamientos tienen medianas iguales.

H_1 : Las medianas de las cuatro poblaciones no son todas iguales.

continúa

Tabla 13-6 Pesos (en kg) de álamos

Tratamiento			
Ninguno	Fertilizante	Riego	Fertilizante y riego
0.15 (8)	1.34 (18)	0.23 (12)	2.03 (19)
0.02 (1.5)	0.14 (7)	0.04 (3)	0.27 (13)
0.16 (9.5)	0.02 (1.5)	0.34 (14)	0.92 (16)
0.37 (15)	0.08 (5.5)	0.16 (9.5)	1.07 (17)
0.22 (11)	0.08 (5.5)	0.05 (4)	2.38 (20)
$n_1 = 5$	$n_2 = 5$	$n_3 = 5$	$n_4 = 5$
$R_1 = 45$	$R_2 = 37.5$	$R_3 = 42.5$	$R_4 = 85$

Para determinar el valor del estadístico de prueba H , primero tenemos que ordenar en rangos todos los datos. Comenzamos con los valores más bajos de 0.02 y 0.02. Como hay un empate entre los valores correspondientes a los rangos 1 y 2, asignamos el rango medio de 1.5 a cada uno de los valores empata-dos. En la tabla 13-6 los rangos aparecen entre paréntesis, junto al peso original del árbol. Después calculamos el tamaño muestral n y la suma de rangos R para cada muestra, los cuales se presentan en la parte inferior de la tabla 13-6. Puesto que el número total de observaciones es 20, tenemos $N = 20$. Ahora podemos evaluar el estadístico de prueba de la siguiente manera:

$$\begin{aligned} H &= \frac{12}{N(N+1)} \left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \dots + \frac{R_k^2}{n_k} \right) - 3(N+1) \\ &= \frac{12}{20(20+1)} \left(\frac{45^2}{5} + \frac{37.5^2}{5} + \frac{42.5^2}{5} + \frac{85^2}{5} \right) - 3(20+1) \\ &= 8.214 \end{aligned}$$

Puesto que cada muestra tiene al menos cinco observaciones, la distribución de H es aproximadamente una distribución chi cuadrada con $k - 1$ grados de libertad. El número de muestras es $k = 4$, entonces tenemos $4 - 1 = 3$ grados de libertad. Remítase a la tabla A-4 para encontrar el valor crítico de 7.815, que corresponde a 3 grados de libertad y a un nivel de significancia de 0.05 (con una área de 0.05 en la cola derecha).

El estadístico de prueba $H = 8.214$ está en la región crítica acotada por 7.815; por lo tanto, rechazamos la hipótesis nula de medianas iguales. (En la sección 12-2, rechazamos la hipótesis nula de medias iguales).

INTERPRETACIÓN Existe suficiente evidencia para rechazar la aseveración de que la población de los pesos de álamos con los cuatro tratamientos tienen medianas iguales. Parece que al menos una de las medianas difiere de las demás.

Fundamentos: El estadístico de prueba H de la prueba de Kruskal-Wallis es la versión con rangos del estadístico de prueba F utilizado en el análisis de varianza que se estudió en el capítulo 12. Cuando tratamos con rangos R en vez de valores x originales, muchos componentes están predeterminados. Por ejemplo, la suma de todos los rangos puede expresarse como $N(N+1)/2$, donde N es el número total de valores en todas las muestras combinadas. La expresión

$$H = \frac{12}{N(N+1)} \sum n_i (\bar{R}_i - \bar{\bar{R}})^2$$

donde

$$\bar{R}_i = \frac{R_i}{n_i} \quad \text{y} \quad \bar{\bar{R}} = \frac{\sum R_i}{\sum n_i}$$

combina varianzas ponderadas de rangos para producir el estadístico de prueba H que se dio aquí. Esta expresión de H es equivalente en términos algebraicos a la expresión de H que se dio antes como estadístico de prueba.

Uso de la tecnología

STATDISK Ingrese los datos en las columnas de la ventana de datos. Seleccione **Analysis** de la barra del menú principal, luego seleccione **Kruskal-Wallis Test** y proceda a seleccionar las columnas de datos. STATDISK desplegará en la pantalla la suma de los rangos para cada muestra, el estadístico de prueba H , el valor crítico y la

conclusión. Véase la pantalla adjunta de STATDISK para el ejemplo de esta sección.

MINITAB Remítase al manual *Minitab Student Laboratory Manual and Workbook* para ver el procedimiento que se requiere para utilizar las opciones **Stat**, **Nonparametrics** y **Kruskal-Wallis**. La idea básica es hacer una lista de todos los datos muestrales en una gran columna, con otra columna que identifique la muestra para los valores correspondientes. Para los datos de la tabla 13-6, ingrese los 20 valores en la columna C1 de Minitab. En la columna C2, ingrese cinco números 1 seguidos de cinco números 2, de cinco números 3 y de cinco números 4. Ahora seleccione **Stat**, **Nonparametrics** y **Kruskal-Wallis**. En el cuadro de diálogo, ingrese C1 para la respuesta, C2 para el factor y haga clic en **OK**. La pantalla de Minitab incluye el estadístico de prueba H y el valor P .

EXCEL Excel no está programado para la prueba de Kruskal-Wallis, pero se puede usar el complemento DDXL seleccionando **Nonparametric Tests** y **Kruskal-Wallis**. Los datos muestrales deben aparecer en una columna y los nombres de las muestras deben aparecer en otra columna (Factor).

TI-83/84 PLUS La calculadora TI-83/84 Plus no está programada para la prueba de Kruskal-Wallis, pero se puede usar el programa KWTEST. Este programa (desarrollado por Michael Lloyd) se descarga del sitio <http://www.pearson-educacion.net/Triola>. Primero descargue e instale el programa. (También descargue el programa ZZRANK, necesario para el programa KWTEST). Despues, registre la lista de datos muestrales en columnas separadas de la matriz [A]. Presione la tecla **PRGM**, seleccione **KWTEST** y luego presione la tecla **ENTER**. La calculadora dará el valor del estadístico de prueba y el número de grados de libertad. (*Nota:* Si las muestras tienen tamaños diferentes y uno de los valores de los datos es cero, añada una constante adecuada a todos los valores muestrales para evitar la presencia de ceros).

STATDISK

Total Num Values:	20
Rank Sum 1:	45.0000
Rank Sum 2:	37.5000
Rank Sum 3:	42.5000
Rank Sum 4:	85.0000
Test Statistic, H:	8.2143
Critical H:	7.8147
P-value:	0.0418
Reject the Null Hypothesis	
Data provides evidence that the samples come from different populations	

13-5 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Prueba de Kruskal-Wallis.** ¿Cuál es la principal ventaja de la prueba de Kruskal-Wallis sobre la prueba del análisis de varianza de un factor?
- Prueba de Kruskal-Wallis.** Si se utiliza la prueba de Kruskal-Wallis y el análisis de varianza de un factor con tres conjuntos de muestras independientes, ¿ambas pruebas siempre conducirán a la misma conclusión?
- Eficiencia.** Remítase a la tabla 13-2 e identifique la eficiencia de la prueba de Kruskal-Wallis. ¿Qué nos indica ese valor acerca de la prueba?

- 4. Requisitos.** Se seleccionan 50 familias al azar y se aplican pruebas de CI a la madre, al padre y al primogénito. ¿Se puede utilizar la prueba de Kruskal-Wallis para probar la aseveración de que las tres poblaciones de madres, padres y primogénitos tienen puntuaciones de CI con la misma mediana? ¿Por qué?

Uso de la prueba de Kruskal-Wallis. En los ejercicios 5 a 10, utilice la prueba de Kruskal-Wallis.

- 5. ¿Afecta el peso de un automóvil las heridas en la cabeza producidas en un choque?** Se obtuvieron datos de experimentos de choques realizados por la National Transportation Safety Administration. Se compraron automóviles nuevos, se impactaron contra una barrera fija a 35 mi/h y se registraron las mediciones en un maniquí en el asiento del conductor. Utilice los datos muestrales listados abajo para probar las diferencias en las mediciones de heridas en la cabeza (de acuerdo con el *Head Injury Criterion*, HIC) en cuatro categorías de peso. ¿Existe evidencia suficiente para concluir que las mediciones de heridas en la cabeza para las cuatro categorías de peso de automóviles no son las mismas? ¿Sugieren los datos que los automóviles más pesados son más seguros en un choque?

Subcompacto:	681	428	917	898	420
Compacto:	643	655	442	514	525
Mediano:	469	727	525	454	259
Grande:	384	656	602	687	360

- 6. ¿Afecta el peso de un automóvil las heridas en el pecho producidas en un choque?** Se obtuvieron datos de experimentos de choques realizados por la National Transportation Safety Administration. Se compraron automóviles nuevos y se impactaron contra una barrera fija a 35 mi/h; abajo se presentan los datos de desaceleración del pecho (g). Utilice los datos para probar la hipótesis nula de que las diferentes categorías de peso tienen medianas que no son iguales. ¿Sugieren los datos que los automóviles más pesados son más seguros en un choque?

Subcompacto:	55	47	59	49	42
Compacto:	57	57	46	54	51
Mediano:	45	53	49	51	46
Grande:	44	45	39	58	44

- 7. ¿La energía solar es la misma todos los días?** Un alumno del autor vive en una casa con sistema eléctrico solar. A la misma hora de cada día, reúne lecturas de voltaje con un medidor conectado al sistema y los resultados se listan en la tabla al margen. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las lecturas de voltaje tienen la misma mediana para los tres diferentes tipos de día. Podríamos esperar que un sistema solar proporcione más energía eléctrica en días soleados que en días nublados o lluviosos. ¿Podemos concluir que los días soleados dan como resultado mayores cantidades de energía eléctrica?

- 8. Prueba de diferencias de amplitud craneana en distintas épocas.** Los valores adjuntos son amplitudes máximas medidas de cráneos de hombres egipcios de diferentes épocas (según datos de *Ancient Races of the Thebaid*, de Thomson y Randall-MacIver). Los cambios en la forma de la cabeza a través del tiempo sugieren que ocurrió mestizaje con poblaciones inmigrantes. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las tres muestras provienen de poblaciones idénticas. ¿Sugieren los datos un mestizaje de culturas?

- 9. Ejercicio y estrés.** Se realizó un estudio para investigar los efectos del ejercicio sobre el estrés. La siguiente tabla lista lecturas de presión sanguínea sistólica (en mmHg) de sujetos, antes de iniciar 25 minutos de ejercicio aeróbico en bicicleta y antes de generarles estrés por medio de una prueba de aritmética y otra de expresión verbal (según datos de “Sympathoadrenergic Mechanisms in Reduced Hemodynamic Stress Responses after Exercise”, de Kim Brownley *et al.*, *Medicine and Science in Sports and*

Datos del ejercicio 7

Soleado	Nublado	Lluvioso
13.5	12.7	12.1
13.0	12.5	12.2
13.2	12.6	12.3
13.9	12.7	11.9
13.8	13.0	11.6
14.0	13.0	12.2

Datos del ejercicio 8

4000 a.C.	1850 a.C.	150 d.C.
131	129	128
138	134	138
125	136	136
129	137	139
132	137	141
135	129	142
132	136	137
134	138	145
138	134	137

Exercise, vol. 35, núm. 6). Utilice un nivel de significancia de 0.05 para probar la aseveración de que los diferentes grupos de sujetos tienen la misma mediana de la presión sanguínea. Con base en los resultados, ¿se puede considerar que los grupos son muestras de la misma población?

Mujer/afroam.	Hombre/afroam.	Mujer/caucásica	Hombre/caucásico
117.00	115.67	119.67	124.33
130.67	120.67	106.00	111.00
102.67	133.00	108.33	99.67
93.67	120.33	107.33	128.33
96.33	124.67	117.00	102.00
92.00	118.33	113.33	127.33

- 10. Prueba de laboratorio de inflamabilidad de ropa de dormir para niños.** Se realizaron pruebas de inflamabilidad en ropa de dormir para niños. Se utilizó la prueba Vertical Semirestrained, en la cual se quemaron piezas de tela en condiciones controladas. Después de apagar el fuego, se midió y registró la longitud de la porción quemada. Al margen se presentan los resultados para la misma tela probada en laboratorios diferentes. Puesto que se utilizó la misma tela, los diferentes laboratorios deberían haber obtenido los mismos resultados. ¿Fue así?

En los ejercicios 11 y 12, use la prueba de Kruskal-Wallis con el conjunto de datos del apéndice B.

- 11. Conjunto de datos del apéndice B: Pesos de peniques.** Remítase al conjunto de datos 14 del apéndice B y use los pesos de los “peniques indios”, los “peniques de tri-glo”, los peniques acuñados antes de 1983 y los acuñados después de 1983. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la mediana del peso de los peniques es la misma para las cuatro categorías diferentes. Con base en los resultados, ¿parece que la máquina que acuña las monedas puede tratar los pesos de los peniques de la misma forma?
- 12. Conjunto de datos del apéndice B: ¿Pesan lo mismo los dulces M&M de todos los colores?** Remítase al conjunto de datos 13 del apéndice B. Con un nivel de significancia de 0.05, pruebe la aseveración de que los pesos correspondientes a la mediana de los dulces M&M son iguales para cada una de las seis poblaciones de colores diferentes. Si la intención de Mars Inc. es hacer los dulces para que las diferentes poblaciones de color sean las mismas, ¿sugieren sus resultados que la compañía tiene un problema que requiere de una acción correctiva?

Datos del ejercicio 10

Laboratorio

1	2	3	4	5
2.9	2.7	3.3	3.3	4.1
3.1	3.4	3.3	3.2	4.1
3.1	3.6	3.5	3.4	3.7
3.7	3.2	3.5	2.7	4.2
3.1	4.0	2.8	2.7	3.1
4.2	4.1	2.8	3.3	3.5
3.7	3.8	3.2	2.9	2.8
3.9	3.8	2.8	3.2	
3.1	4.3	3.8	2.9	
3.0	3.4	3.5		
2.9	3.3			

13-5 MÁS ALLÁ DE LO BÁSICO

- 13. Corrección del estadístico de prueba H por empates.** Al emplear la prueba de Kruskal-Wallis, existe un factor de corrección que debe aplicarse siempre que existan muchos empates: divida H entre

$$1 - \frac{\sum T}{N^3 - N}$$

Para cada grupo de observaciones empata das en el conjunto combinado de datos muestrales, calcule $T = t^3 - t$, donde t es el número de observaciones que están empata das en el grupo individual. Calcule t para cada grupo de valores empata dos, luego calcule el valor de T para cada grupo, y después sume los valores T para obtener $\sum T$. El número total de observaciones en todas las muestras combinadas es N . Utilice este procedimiento para calcular el valor corregido de H para el ejercicio 7. ¿El valor corregido de H difiere sustancialmente del valor calculado en el ejercicio 7?

- 14. Pruebas equivalentes.** Demuestre que para el caso de dos muestras, la prueba de Kruskal-Wallis es equivalente a la prueba de la suma de rangos de Wilcoxon. Esto

puede hacerse demostrando que para el caso de dos muestras, el estadístico de prueba H es igual al cuadrado del estadístico de prueba z que se utiliza en la prueba de la suma de rangos de Wilcoxon. Además, note que con 1 grado de libertad, los valores críticos de χ^2 corresponden al cuadrado de la puntuación crítica z .

13-6 Correlación de rangos

Concepto clave En esta sección se describe el método no paramétrico de la correlación de rangos, el cual se utiliza con datos apareados para probar una asociación entre dos variables. En el capítulo 10 utilizamos datos muestrales apareados para calcular valores del coeficiente de correlación lineal r , pero en esta sección utilizaremos *rangos* como base para calcular el coeficiente de correlación de rangos r_s .

Definición

La **prueba de correlación de rangos** (o **prueba de correlación de rangos de Spearman**) es una prueba no paramétrica que utiliza rangos de datos muestrales consistentes en datos apareados. Se utiliza para probar una asociación entre dos variables, por lo que las hipótesis nula y alternativa son las siguientes (donde ρ_s denota el coeficiente de correlación de rangos de la población completa):

$$H_0: \rho_s = 0 \text{ (No existe correlación entre las dos variables).}$$

$$H_1: \rho_s \neq 0 \text{ (Existe una correlación entre las dos variables).}$$

Ventajas: La correlación de rangos tiene las siguientes ventajas sobre los métodos paramétricos analizados en el capítulo 10:

1. El método no paramétrico de correlación de rangos puede utilizarse en una variedad más amplia de circunstancias que el método paramétrico de correlación lineal. Con la correlación de rangos, podemos analizar datos apareados que sean rangos o puedan convertirse en rangos. Por ejemplo, si dos jueces califican a 30 gimnastas, podemos utilizar la correlación de rangos, pero no la correlación lineal. A diferencia de los métodos paramétricos del capítulo 10, el método de correlación de rangos *no* requiere una distribución normal de cualquier población.
2. La correlación de rangos puede utilizarse para detectar algunas relaciones (no todas) que no son lineales. (Más adelante se dará un ejemplo en esta sección).

Desventaja: Una desventaja de la correlación de rangos es su tasa de eficiencia de 0.91, como se describe en la sección 13-1. Esta tasa de eficiencia indica que, con todas las demás circunstancias iguales, el método no paramétrico de correlación de rangos requiere de 100 pares de datos muestrales para tener los mismos resultados que sólo 91 pares de observaciones muestrales analizadas a través del método paramétrico, suponiendo que los requisitos más estrictos del método paramétrico se satisfacen.

Utilizamos la notación r_s para el coeficiente de correlación de rangos para no confundirlo con el coeficiente de correlación lineal r . El subíndice s no tiene nada que ver con la desviación estándar; se usa en honor de Charles Spearman (1863-1945), quien originó el método de correlación de rangos. De hecho, r_s suele llamarse **coeficiente de correlación de rangos de Spearman**. El procedimiento de la correlación de rangos se resume en la figura 13-4.

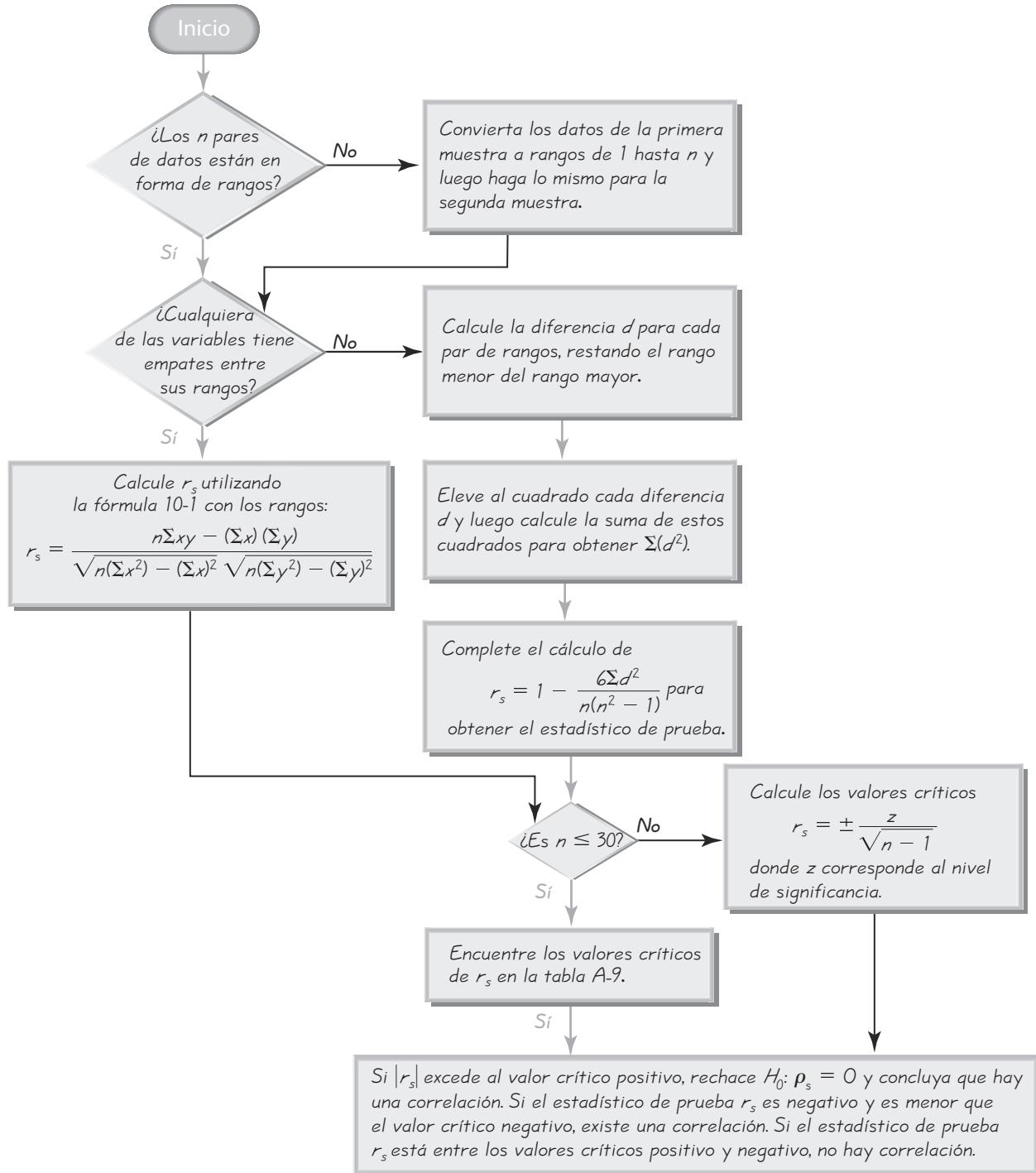


Figura 13-4 Procedimiento de correlación de rangos para probar $H_0: \rho_s = 0$

Correlación de rangos

Requisitos

1. Los datos muestrales apareados se seleccionaron aleatoriamente.
2. A diferencia de los métodos paramétricos de la sección 10-2, *no* existe el requisito de que los datos muestrales apareados tengan una distribución normal bivariada (como se describió en la sección 10-2). *No* existe el requisito de una distribución normal para cualquier población.

Notación

r_s = coeficiente de correlación de rangos para datos muestrales apareados (r_s es un estadístico muestral)

ρ_s = coeficiente de correlación de rangos para todos los datos poblacionales (ρ_s es un parámetro poblacional)

n = número de pares de datos muestrales

d = diferencia entre los rangos de los dos valores dentro de un par

Estadístico de prueba

Sin empates: Despues de convertir los datos de cada muestra a rangos, si no existen empates entre los rangos para la primera variable y no existen empates entre los rangos para la segunda variable, el valor exacto del estadístico de prueba puede calcularse utilizando esta fórmula:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Empates: Despues de convertir los datos de cada muestra a rangos, si cualquier variable tiene empates entre sus rangos, el valor exacto del estadístico de prueba r_s puede calcularse utilizando la fórmula 10-1 con los rangos:

$$r_s = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

Valores críticos

1. Si $n \leq 30$, los valores críticos se encuentran en la tabla A-9.
2. Si $n > 30$, los valores críticos de r_s se calculan utilizando la fórmula 13-1.

Fórmula 13-1 $r_s = \frac{\pm z}{\sqrt{n-1}}$ (valores críticos cuando $n > 30$)

donde el valor de z corresponde al nivel de significancia. (Por ejemplo, si $\alpha = 0.05$, $z = 1.96$).



EJEMPLO Clasificación de universidades por estudiantes y el U.S. News and World Report El problema del capítulo incluye las clasificaciones de universidades por parte de estudiantes y de la revista *U.S. News and World Report*. Esta clasificación está incluida en la tabla 13-7, que también incluye las diferencias d y los cuadrados de las diferencias d^2 . Calcule el valor del coeficiente de correlación de rangos y utilícelo para determinar si existe una correlación entre la clasificación de los estudiantes y la clasificación de la revista. Use un nivel de significancia de 0.05.

Tabla 13-7 Universidades clasificadas por estudiantes y por *U.S. News and World Report*

Universidad	Clasificación según la preferencia de los estudiantes	Clasificación según la revista <i>U.S. News and World Report</i>	Diferencia d	d^2
Harvard	1	1	0	0
Yale	2	2	0	0
Cal. Inst. of Tech.	3	5	2	4
M.I.T.	4	4	0	0
Brown	5	7	2	4
Columbia	6	6	0	0
U. de Penn.	7	3	4	16
Notre Dame	8	8	0	0
			Total:	24



Vínculo directo entre el tabaquismo y el cáncer

Cuando encontramos una correlación estadística entre dos variables, debemos ser sumamente cuidadosos para evitar el error de concluir que existe un vínculo de causa y efecto. La industria tabacalera ha enfatizado una y otra vez que la correlación no implica causalidad. Sin embargo, el doctor David Sidransky, de la John Hopkins University, asegura: "Tenemos pruebas moleculares tan fuertes que podemos estudiar un caso de cáncer individual y, potencialmente, con base en los patrones de cambio genético, determinar si fumar cigarrillos fue la causa de ese cáncer". A partir de sus hallazgos, también afirma que "el fumador tuvo una incidencia mucho más alta de mutación, lo que se confirmó con el patrón tan claro de mutaciones... así que prácticamente encontramos la pistola humeante". Aunque los métodos estadísticos no pueden probar que fumar causa cáncer, con evidencia física como la que presenta el doctor Sidransky es posible establecer demostraciones como ésta.

SOLUCIÓN

REQUISITO ✓ El único requisito es que los datos muestrales apareados sean elegidos al azar. Las universidades incluidas se eligieron al azar entre aquellas que estaban disponibles, de manera que procedemos con la prueba. ✓

El coeficiente de correlación lineal r (sección 10-2) no debe utilizarse puesto que requiere de distribuciones normales, y los datos consisten en rangos que no están distribuidos normalmente. En vez de ello, utilizamos el coeficiente de correlación de rangos para probar una relación entre los rangos de los estudiantes y de la revista.

Las hipótesis nula y alternativa son las siguientes:

$$H_0: \rho_s = 0$$

$$H_1: \rho_s \neq 0$$

Siguiendo el procedimiento de la figura a 13-4, los datos están en forma de rangos y ninguna de las dos variables tiene empates entre los rangos, por lo que el valor exacto del estadístico de prueba puede calcularse como se indica abajo. Utilizamos $n = 8$ (para 8 pares de datos) y $\sum d^2 = 24$ (como se indica en la tabla de 13-7) para obtener

$$\begin{aligned} r_s &= 1 - \frac{6\sum d^2}{n(n^2 - 1)} = 1 - \frac{6(24)}{8(8^2 - 1)} \\ &= 1 - \frac{144}{504} = 0.714 \end{aligned}$$

Ahora nos remitimos a la tabla de A-9 para determinar que los valores críticos son ± 0.738 (con $\alpha = 0.05$ y $n = 8$). Puesto que el estadístico de prueba $r_s = 0.714$ no excede al valor crítico de 0.738, no rechazamos la hipótesis nula. No existe suficiente evidencia para sustentar una aseveración de correlación entre la clasificación de los estudiantes y la clasificación de la revista. Parece que en lo que se refiere a la clasificación de universidades, los estudiantes y la revista no coinciden. (Si coincidieran, habría una correlación significativa, pero no la hay).

EJEMPLO Caso de muestra grande Suponga que el ejemplo anterior se amplía para incluir un total de 40 universidades y que se encuentra que el estadístico de prueba r_s es 0.300. Si el nivel de significancia es $\alpha = 0.05$, ¿qué concluye usted acerca de la correlación?

SOLUCIÓN Puesto que existen 40 pares de datos, tenemos $n = 40$. Como n excede a 30, calculamos los valores críticos con la fórmula 13-1, en vez de emplear la tabla A-9. Con $\alpha = 0.05$ en dos colas, permitimos que $z = 1.96$ para obtener

$$r_s = \frac{\pm 1.96}{\sqrt{40 - 1}} = \pm 0.314$$

El estadístico de prueba $r_s = 0.300$ no excede al valor crítico de 0.314, por lo tanto, no rechazamos la hipótesis nula. No existe suficiente evidencia para sustentar la aseveración de una correlación entre los estudiantes y la revista.

El siguiente ejemplo tiene la intención de ilustrar el principio de que la correlación de rangos algunas veces se puede utilizar para detectar relaciones que no son lineales.

EJEMPLO Detección de un patrón no lineal Se utiliza una máquina de *pinball Raiders of the Lost Ark* (modelo L-7) para medir el aprendizaje que resulta de repetir funciones manuales. Los sujetos se seleccionaron para que fueran similares en características importantes de edad, género, inteligencia, educación, etcétera. La tabla 13-8 lista los números de juegos que se realizaron y las últimas puntuaciones (en millones) de sujetos seleccionados al azar del grupo con características similares. Esperamos que haya una asociación entre el número de juegos realizados y la puntuación del *pinball*. ¿Existe suficiente evidencia para sustentar la aseveración de que existe una asociación como ésta?

SOLUCIÓN Probaremos la hipótesis nula de inexistencia de correlación de rangos ($\rho_s = 0$).

$$\begin{aligned} H_0: \rho_s &= 0 && \text{(sin correlación)} \\ H_1: \rho_s &\neq 0 && \text{(correlación)} \end{aligned}$$

Remítase a la figura 13-4, la cual seguimos para esta solución. Las puntuaciones originales no son rangos, así que las convertimos a rangos y colocamos los resultados entre paréntesis en la tabla 13-8. (En la sección 13-1 se describe el procedimiento para convertir puntuaciones en rangos). No hay empates entre

Tabla 13-8 Puntuaciones de *pinball* (rangos entre paréntesis)

Número de juegos	9 (2)	13 (4)	21 (5)	6 (1)	52 (7)	78 (8)	33 (6)	11 (3)	120 (9)
Puntuación	22 (2)	62 (4)	70 (6)	10 (1)	68 (5)	73 (8)	72 (7)	58 (3)	75 (9)
d	0	0	1	0	2	0	1	0	0
d^2	0	0	1	0	4	0	1	0	0

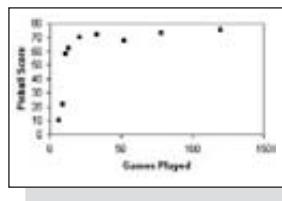
los rangos del número de juegos realizados, ni hay empates entre los rangos de las puntuaciones, de manera que procedemos a calcular las diferencias d y luego las elevamos al cuadrado. La suma de los valores de d^2 es 6. Ahora calculamos

$$\begin{aligned} r_s &= 1 - \frac{6\sum d^2}{n(n^2 - 1)} = 1 - \frac{6(6)}{9(9^2 - 1)} \\ &= 1 - \frac{36}{720} = 0.950 \end{aligned}$$

Continuando con la figura 13-4, tenemos $n = 9$, por lo que respondemos sí cuando se pregunta si $n \leq 30$. Utilizamos la tabla A-9 para obtener los valores críticos de ± 0.700 . Finalmente, el estadístico muestral de 0.950 excede a 0.700, por lo que concluimos que existe una correlación significativa. Los números más altos de juegos parecen estar asociados con puntuaciones más altas. Los sujetos parecen aprender mejor el juego al jugar más.

En el ejemplo anterior, si calculamos el coeficiente de correlación lineal r (utilizando la fórmula 9.1) para los datos originales, obtenemos $r = 0.586$, lo que nos lleva a la conclusión de que no existe evidencia suficiente para sustentar la aseveración de una correlación lineal significativa al nivel 0.05 de significancia. Si examinamos el diagrama de dispersión de Excel, podemos ver que el patrón de puntos no es un patrón de línea recta. Este último ejemplo ilustra una ventaja del método no paramétrico sobre el método paramétrico: con la correlación de rangos, algunas veces podemos detectar relaciones que no son lineales.

Excel



Uso de la tecnología

STATDISK Anote los datos muestrales en las columnas de la ventana de datos. Seleccione **Analysis** de la barra del menú principal, luego elija **Rank Correlation**. Seleccione las dos columnas de datos que se incluirán y luego haga clic en **Evaluate**. Los resultados de STATDISK incluyen el valor exacto del estadístico de prueba r_s , el valor crítico y la conclusión.

MINITAB Ingrese los datos apareados en las columnas C1 y C2. Si los datos todavía no son rangos, utilice las opciones **Manip** y **Rank** de Minitab para convertir los datos a rangos, después seleccione **Stat**, seguido por **Basic Statistics**, y luego **Correlation**. Minitab mostrará en la pantalla el valor exacto del estadístico de prueba r_s . Aunque Minitab identifica esto como el coeficiente de correlación de Pearson descrito en la sección 10-2, en realidad se trata del coeficiente de correlación de Spearman descrito en esta sección (puesto que está basado en rangos).

EXCEL Excel no tiene una función que calcule el coeficiente de correlación de rangos a partir de valores muestrales originales, pero el valor exacto del estadístico de prueba r_s se puede calcular como sigue. Primero reemplace cada uno de los valores muestrales originales por su rango correspondiente. Ingrese estos rangos en las columnas A y B. Haga clic en el botón de función **fx** localizado en la barra del menú principal. Seleccione la categoría de función **Statistical** y el

nombre de la función **CORREL**, luego haga clic en **OK**. En el cuadro de diálogo, ingrese el rango de la celda de los valores para x , como A1:A10. También ingrese el rango de la celda de los valores para y , como B1:B10. Excel mostrará en la pantalla el valor exacto del coeficiente de correlación de rango r_s . También es posible emplear el complemento DDXL al seleccionar **Nonparametric Tests** y luego **Spearman Rank Test**.

TI-83/84 PLUS Sí utilizamos una calculadora TI-83/84 Plus o cualquier otra con estadísticos para 2 variables, es posible calcular el valor exacto de r_s como sigue: 1. reemplace cada valor muestral por su rango correspondiente, luego 2. calcule el valor del coeficiente de correlación lineal r con los mismos procedimientos utilizados en la sección 10-2. Ingrese los rangos apareados en las listas L1 y L2, después oprima **STAT** y seleccione **TESTS**. El uso de la opción **LinRegTTest** dará como resultado varios valores, incluyendo el valor exacto del coeficiente de correlación de rangos r_s .



13-6 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Correlación de rango.** ¿Qué ventaja importante tiene el método de correlación de rangos de esta sección sobre el método de correlación lineal de la sección 10-2?
- Correlación de rango.** Si dos jueces clasifican a 25 gimnastas desde 1 hasta 25 sin empates, y sus clasificaciones coinciden perfectamente, ¿cuál es el valor del coeficiente de correlación de rangos?
- Notación.** Representamos el estadístico de prueba del coeficiente de correlación de rangos con la notación r_s , y el parámetro poblacional correspondiente se representa por ρ_s . ¿Por qué se utiliza el subíndice s ? El subíndice s representa la misma desviación estándar s que se estudió en la sección 3-3?
- Eficiencia.** Remítase a la tabla 13-2 e identifique la eficiencia de la prueba de correlación de rangos. ¿Qué nos indica ese valor acerca de la prueba?

En los ejercicios 5 y 6, dibuje un diagrama de dispersión, calcule el valor de r_s y determine si al parecer existe una correlación entre x y y .

- | | | | | | |
|--------|---|---|---|---|---|
| 5. x | 2 | 4 | 1 | 5 | 3 |
| y | 2 | 4 | 1 | 5 | 3 |
-
- | | | | | | |
|--------|---|---|---|---|---|
| 6. x | 1 | 2 | 3 | 4 | 5 |
| y | 5 | 4 | 3 | 2 | 1 |

Cálculo de valores críticos. En los ejercicios 7 y 8, calcule el valor (o valores) crítico(s) r_s utilizando la tabla A-9 o la fórmula 13-1, según sea más apropiado. Suponga que la hipótesis nula es $\rho_s = 0$, de manera que la prueba es de dos colas. Además, n denota el número de pares de datos.

7. a. $n = 12, \alpha = 0.05$
- b. $n = 20, \alpha = 0.01$
- c. $n = 60, \alpha = 0.05$
- d. $n = 80, \alpha = 0.01$
8. a. $n = 14, \alpha = 0.05$
- b. $n = 23, \alpha = 0.01$
- c. $n = 120, \alpha = 0.01$
- d. $n = 75, \alpha = 0.05$

Prueba para correlación de rangos. En los ejercicios 9 a 16, utilice el coeficiente de correlación de rangos para probar una correlación entre las dos variables. Utilice un nivel de significancia de $\alpha = 0.05$.

- Bolsa de Valores y ventas de automóviles.** Para una serie reciente de 10 años, se obtuvieron los valores máximos anuales del Promedio Industrial Dow Jones (DJIA) y los números de automóviles (en miles) correspondientes que se vendieron en Estados Unidos. La siguiente tabla lista los rangos de cada conjunto de valores. Pruebe una correlación entre el DJIA y el número de automóviles vendidos.

DJIA máximo	1	2	3	4	5	6	7	8	10	9
Ventas de automóviles	2	3	5	10	7	6	4	1	8	9

- 10. Manchas solares y puntos del Súper Bowl.** Para una serie reciente de 10 años, se obtuvieron el número de manchas solares y los puntos totales anotados en el Súper Bowl. La siguiente tabla lista los rangos de cada conjunto de valores. Pruebe una correlación entre el número de manchas solares y los puntos anotados en el Súper Bowl. ¿El resultado coincide con lo que se esperaría?

Número de manchas solares	10	8	5	4	2	1	3	6	7	9
Puntos del Súper Bowl	8	9	3	10	4	7	6	5	1	2

- 11. Correlación entre salario y estrés.** La siguiente tabla lista rangos de salario y rangos de estrés de empleos seleccionados al azar (según datos de *The Jobs Rated Almanac*). ¿Parece que el salario se incrementa a medida que se incrementa el estrés?

Empleo	Rango de salario	Rango de estrés
Corredor de bolsa	2	2
Zoólogo	6	7
Ingeniero eléctrico	3	6
Director de escuela	5	4
Gerente de hotel	7	5
Funcionario bancario	10	8
Inspector de seguridad laboral	9	9
Economista doméstico	8	10
Psicólogo	4	3
Piloto de aerolínea	1	1

- 12. Correlación entre salario y demanda física.** El ejercicio 11 incluye rangos aparentados de salario y nivel de estrés para 10 empleos seleccionados al azar. Las demandas físicas de los empleos también se ordenaron por rangos; los rangos de salario y demanda física se presentan abajo (según datos de *The Jobs Rated Almanac*). ¿Parece existir una relación entre el salario de un empleo y sus demandas físicas?

Salario	2	6	3	5	7	10	9	8	4	1
Demanda física	5	2	3	8	10	9	1	7	6	4

- 13. Grillos y temperatura.** Se estudió la relación entre la temperatura y el número de veces que un grillo chirría en un minuto. Abajo se listan los números de chirridos por minuto y las temperaturas correspondientes en grados Fahrenheit (según datos de *The Song of Insects*, de George W. Pierce, Harvard University Press). ¿Existe evidencia suficiente para concluir que existe una relación entre el número de chirridos por minuto y la temperatura?

Chirridos en un minuto	882	1188	1104	864	1200	1032	960	900	
Temperatura (en °F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6	

- 14. Muertes por vehículos automotores y asesinatos.** A continuación se lista el número de muertes en vehículos automotores (en cientos) y el número de asesinatos (en cientos) en Estados Unidos para varios años diferentes. Pruebe una correlación entre esas dos variables.

Muertes en vehículos automotores	435	410	418	425	434	436	434	435	413	430
Asesinatos	247	238	245	233	216	197	182	170	155	156

- 15. Audiencia y ventas de canciones.** En la siguiente tabla se lista el número de impresiones de audiencia (en cientos de millones) de canciones y el número de álbumes vendidos correspondientes (en cientos de miles). El número de impresiones de audiencia es

un conteo del número de veces que la gente ha escuchado la canción. La tabla se basa en datos de *USA Today*. ¿Parece que las ventas del álbum se ven muy afectadas por el número de impresiones de audiencia?

Impresiones de audiencia	28	13	14	24	20	18	14	24	17
Álbumes vendidos	19	7	7	20	6	4	5	25	12

- 16. Presupuestos y ganancias netas de películas.** A continuación se listan los presupuestos (en millones de dólares) y las ganancias brutas (en millones de dólares) de películas seleccionadas al azar (según datos de la Motion Picture Association of America). ¿Al parecer existe una correlación significativa entre el dinero gastado para hacer la película y la cantidad que se recupera en las salas de cine? Además de la cantidad del presupuesto, identifique otro factor importante que tal vez afecte la cantidad que gana una película.

Presupuesto	62	90	50	35	200	100	90
Ganancia bruta	65	64	48	57	601	146	47

Conjuntos de datos del apéndice B: En los ejercicios 17 y 18, utilice los conjuntos de datos del apéndice B para hacer una prueba de correlación de rango con un nivel de significancia de 0.05.

- 17. Conjunto de datos del apéndice B: Nocividad de los cigarrillos.** Remítase al conjunto de datos 3 del apéndice B.

- a. Utilice los datos apareados referentes a cantidades de alquitrán y nicotina. Con base en el resultado, ¿parece existir una correlación significativa entre el alquitrán y la nicotina de los cigarrillos? Si es así, ¿pueden los investigadores reducir sus gastos de laboratorio midiendo sólo una de estas dos variables?
- b. Utilice los datos apareados referentes a monóxido de carbono y nicotina. Con base en el resultado, ¿parece existir una correlación significativa entre el monóxido de carbono y la nicotina de los cigarrillos? Si es así, ¿pueden los investigadores reducir sus gastos de laboratorio midiendo sólo una de estas dos variables?
- c. Suponga que los investigadores quieren desarrollar un método para predecir la cantidad de nicotina y sólo quieren medir algún otro elemento. Entre alquitrán y monóxido de carbono, ¿cuál es la mejor opción? ¿Por qué?

- 18. Conjunto de datos del apéndice B: Pronósticos del clima.** Remítase al conjunto de datos 8 del apéndice B.

- a. Utilice las temperaturas máximas pronosticadas para cinco días y las temperaturas máximas reales. ¿Existe una correlación? ¿Una correlación significativa implica que las temperaturas del pronóstico de cinco días son exactas?
- b. Utilice las temperaturas máximas pronosticadas para un día y las temperaturas máximas reales. ¿Existe una correlación? ¿Una correlación significativa implica que las temperaturas de pronóstico para un día son exactas?
- c. ¿Cómo esperaría usted obtener una correlación más alta con las temperaturas máximas reales: con las temperaturas máximas del pronóstico para cinco días o con las temperaturas máximas del pronóstico para un día? ¿Los resultados de los incisos a) y b) son como usted esperaba? Si existe una correlación muy alta entre las temperaturas de pronóstico y las temperaturas reales, ¿se deduce que las temperaturas pronosticadas son exactas?

13-6 MÁS ALLÁ DE LO BÁSICO

- 19. Curva de aprendizaje.** Un psicólogo diseña una prueba de aprendizaje. Los sujetos reciben letras que deben memorizar, y sus tiempos de estudio (en segundos) se aparean con el número de letras que logran recordar. Los resultados se listan en la siguiente tabla. ¿Hay una correlación entre el tiempo de estudio del número de letras que logran

recordar? Además, dibuje un diagrama de dispersión y calcule los resultados obtenidos utilizando el coeficiente de correlación lineal descrito en la sección 10-2. Compare los resultados.

Tiempo	1	25	31	33	36	38	48	55	95	165	300
Letras	1	2	20	36	51	72	74	75	77	78	79

20. **Efecto de empates en r_s .** Remítase al conjunto de datos 5 del apéndice B para los tiempos (en segundos) de consumo de tabaco y consumo de alcohol presentados en películas de dibujos animados para niños. Calcule el valor del estadístico de prueba r_s utilizando cada una de las dos fórmulas dadas en esta sección (fórmulas 13-1 y 10-1). ¿Existe una diferencia sustancial entre los dos resultados? ¿Cuál resultado es mejor? ¿La conclusión se ve afectada por la fórmula utilizada?

13-7 Prueba de rachas para detectar aleatoriedad

Concepto clave En esta sección estudiaremos la prueba de rachas para detectar aleatoriedad, la cual puede utilizarse para determinar si los datos muestrales en una secuencia están en un orden aleatorio. Esta prueba se basa en datos muestrales que tienen dos características y analiza rachas de esas características para determinar si las rachas parecen ser el resultado de algún proceso aleatorio, o si las rachas sugieren que el orden de los datos no es aleatorio.

Definiciones

Una **racha** es una secuencia de datos que tienen la misma característica; la secuencia es precedida y seguida por datos con una característica diferente o por ningún dato en absoluto.

La **prueba de rachas** utiliza el número de rachas en una secuencia de datos muestrales para probar la aleatoriedad del orden de los datos.

Principio fundamental de la prueba de rachas

El principio fundamental de la prueba de rachas puede establecerse brevemente como sigue:

Rechace la aleatoriedad si el número de rachas es muy bajo o muy alto.

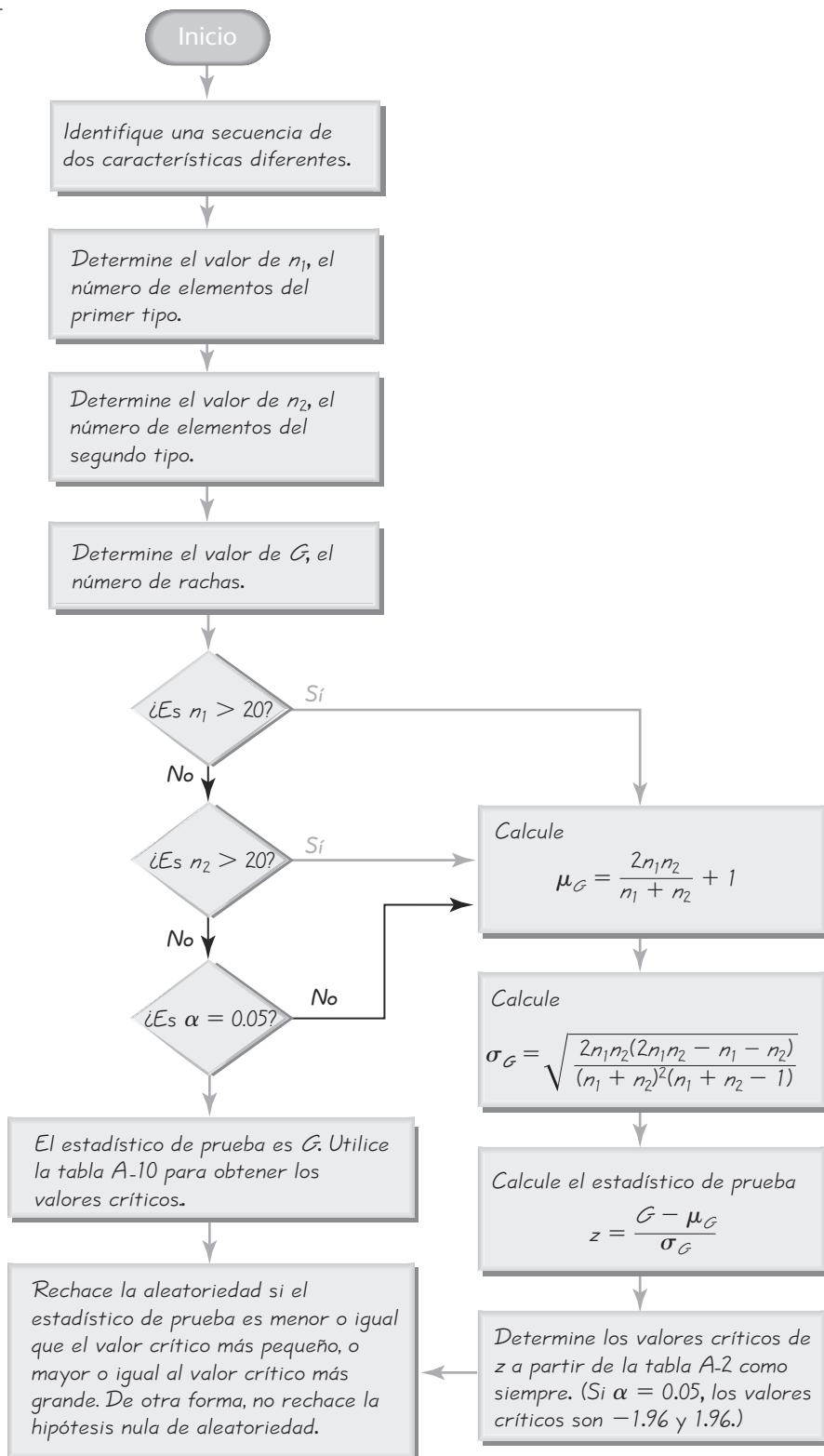
- Ejemplo: La secuencia de género MMMMMHHHHH no es aleatoria puesto que tiene sólo dos rachas, es decir, el número de rachas es muy *bajo*.
- Ejemplo: La secuencia de género MHMHMHMHMH no es aleatoria puesto que existen 10 rachas, lo cual se considera un número muy *alto*.

Los criterios exactos para determinar si un número de rachas es muy alto o muy bajo se encuentran en el recuadro de la página 723, que resume los elementos clave de la prueba de rachas para detectar aleatoriedad. Además, el procedimiento de la prueba de rachas para detectar aleatoriedad también se resume en la figura 13-5.

Es importante señalar que la prueba de rachas para detectar aleatoriedad se basa en el *orden* en el que se presentan los datos; *no* está basada en la *frecuencia* de los datos. Por ejemplo, una secuencia de 3 hombres y 20 mujeres podría parecer aleatoria, pero la prueba de rachas no se ocupa del problema de si 3 hombres y 20 mujeres constituyen una muestra *sesgada* (con un número desproporcionadamente mayor de mujeres).

Figura 13-5

Prueba de rachas para detectar aleatoriedad



Prueba de rachas para detectar aleatoriedad

Requisitos

1. Los datos muestrales están acomodados de acuerdo con algún esquema de orden, por ejemplo, el orden en el que se obtuvieron los valores muestrales.
2. Cada valor de los datos se puede categorizar en una de *dos* categorías separadas (como hombre/mujer).

Notación

n_1 = número de elementos en la secuencia que tienen una característica particular.
(La característica elegida para n_1 es arbitraria).

n_2 = número de elementos en la secuencia que tienen la otra característica

G = número de rachas

Estadístico de prueba

Para muestras pequeñas y $\alpha = 0.05$: Si $n_1 \leq 20$ y $n_2 \leq 20$ y el nivel de significancia es $\alpha = 0.05$, el estadístico de prueba es el número de rachas G . Los valores críticos se encuentran en la tabla A-10. He aquí el criterio de decisión:

Rechace la aleatoriedad si el número de rachas G es

- menor o igual al valor crítico más pequeño encontrado en la tabla A-10.
- mayor o igual al valor crítico más grande encontrado en la tabla A-10.

Para muestras grandes o $\alpha \neq 0.05$: If $n_1 > 20$ o $n_2 > 20$ o $\alpha \neq 0.05$, utilice el estadístico de prueba y los valores críticos siguientes.

$$\text{Estadístico de prueba: } z = \frac{G - \mu_G}{\sigma_G}$$

donde

$$\mu_G = \frac{2n_1 n_2}{n_1 + n_2} + 1$$

$$\text{y } \sigma_G = \sqrt{\frac{(2n_1 n_2)(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

Valores críticos de z : Utilice la tabla A-2.



Rachas de suerte en los deportes

Existe la creencia de que los atletas suelen tener “rachas de suerte”, es decir, períodos breves de éxito extraordinario. El psicólogo Amos Tversky, de la Universidad de Stanford, y otros investigadores utilizaron la estadística para analizar los miles de tiros de los 76 de Filadelfia en una temporada completa y la mitad de otra. Encontraron que el número de “rachas de suerte” no difería de lo que se esperaría en pruebas aleatorias, donde el resultado de cada prueba es independiente de cualquier resultado previo. Es decir, la probabilidad de una anotación no depende de las anotaciones o fallas previas.

EJEMPLO Muestra pequeña: Género de osos A continuación se listan los géneros de los primeros 10 osos del conjunto de datos 6 del apéndice B. Utilice un nivel de significancia de 0.05 para probar la aleatoriedad de la secuencia de géneros

M M M M H H M M H H

SOLUCIÓN

REQUISITOS ✓ Los datos están acomodados en orden, y cada valor está clasificado en una de dos categorías separadas (macho/hembra). Los requisitos se satisfacen y podemos proceder con la prueba de rachas para detectar aleatoriedad. ✓

continúa

Seguiremos el procedimiento que se resume en la figura 13-5. Se ha identificado la secuencia de dos características (macho y hembra). Ahora debemos calcular los valores de n_1 , n_2 y el número de rachas G . La secuencia se muestra abajo con espacios que se usan para identificar mejor las rachas separadas.

$\overbrace{M M M M}$	$\overbrace{H H}$	$\overbrace{M M}$	$\overbrace{H H}$
1 ^a racha	2 ^a racha	3 ^a racha	4 ^a racha

Podemos ver que hay 6 machos y 4 hembras, y el número de rachas es 4. Por lo tanto, tenemos

$$n_1 = \text{número total de machos} = 6$$

$$n_2 = \text{número total de hembras} = 4$$

$$G = \text{número de rachas} = 4$$

Puesto que $n_1 \leq 20$ y $n_2 \leq 20$ y $\alpha = 0.05$, el estadístico de prueba es $G = 4$, y nos remitimos a la tabla A-10 para encontrar los valores críticos de 2 y 9. Puesto que $G = 4$ no es menor o igual que el valor crítico de 2, ni tampoco es mayor o igual que el valor crítico de 9, *no rechazamos la aleatoriedad*. No existe evidencia suficiente para rechazar la aleatoriedad en la secuencia de géneros. Parece que la secuencia de géneros es aleatoria.

EJEMPLO Muestra grande: Iluvia en Boston los lunes Remítase a las cantidades de lluvia en Boston que se listan en el conjunto de datos 10 del apéndice B. ¿Existe suficiente evidencia para sustentar la aseveración de que la lluvia de los lunes no es aleatoria? Utilice un nivel de significancia de 0.05.

SOLUCIÓN

REQUISITOS Permita que S (para seco) represente los lunes sin lluvia (indicados por valores de 0.00), y permita que L represente los lunes con algo de lluvia (cualquier valor mayor que 0.00). A continuación se listan las cantidades de lluvia para los 52 lunes consecutivos. Vemos que los datos están acomodados en orden, y que cada dato está clasificado en una de dos categorías separadas. Los requisitos se satisfacen, así que procedemos con la prueba de rachas para determinar aleatoriedad.

S S S S L S L S S L S S L S S S L S S L L L S S S S
L S L S L L L S L S S S L S S S L S L S S S L

Las hipótesis nula y alternativa son las siguientes:

$$H_0: \text{La secuencia es aleatoria.}$$

$$H_1: \text{La secuencia no es aleatoria.}$$

El estadístico de prueba se obtiene buscando primero el número de “eses”, el número de “eles” y el número de rachas. Es fácil examinar la secuencia para encontrar que

$$n_1 = \text{número de S} = 33$$

$$n_2 = \text{número de L} = 19$$

$$G = \text{número de rachas} = 30$$

Al seguir el procedimiento de la figura 13-5, contestamos sí a la pregunta “¿Es $n_1 > 20$?”. Por lo tanto, necesitamos evaluar el estadístico de prueba z dado

en el recuadro que resume los elementos clave de la prueba de rachas para detectar aleatoriedad. Primero debemos evaluar μ_G y σ_G . Tenemos

$$\begin{aligned}\mu_G &= \frac{2n_1n_2}{n_1 + n_2} + 1 = \frac{2(33)(19)}{33 + 19} + 1 = 25.115 \\ \sigma_G &= \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \\ &= \sqrt{\frac{(2)(33)(19)[2(33)(19) - 33 - 19]}{(33 + 19)^2(33 + 19 - 1)}} = 3.306\end{aligned}$$

Ahora podemos calcular el estadístico de prueba:

$$z = \frac{G - \mu_G}{\sigma_G} = \frac{30 - 25.115}{3.306} = 1.48$$

Puesto que el nivel de significancia es $\alpha = 0.05$ y tenemos una prueba de dos colas, los valores críticos son $z = -1.96$ y $z = 1.96$. El estadístico de prueba de $z = 1.48$ no cae dentro de la región crítica, por lo que no rechazamos la hipótesis nula de aleatoriedad. La secuencia dada parece ser aleatoria.

Datos numéricos: Aleatoriedad por arriba o por debajo de la media o de la mediana

En cada uno de los ejemplos anteriores, los datos se ajustan claramente dentro de dos categorías, pero también podemos probar la aleatoriedad por la forma como los datos numéricos fluctúan por encima o por debajo de una media o mediana. Para probar la aleatoriedad por arriba o por debajo de la mediana, por ejemplo, utilice los datos muestrales para calcular el valor de la mediana, luego reemplace cada valor individual con la letra A si éste se encuentra por *arriba* de la mediana, y reemplácelo con D si se encuentra por *debajo* de la mediana. Excluya cualquier valor que sea igual a la mediana. Es útil escribir las A y las D directamente arriba de los números que éstas representan ya que esto hace más sencilla la revisión y además reduce la posibilidad de tener un número equivocado de letras. Después

Uso de la tecnología

STATDISK Primero determine los valores n_1 , n_2 y el número de rachas G . Selec-

cione **Analysis** de la barra del menú principal, luego seleccione **Runs Test** y proceda a ingresar los datos requeridos en el cuadro de diálogo. La pantalla de STATDISK incluirá el estadístico de prueba (G o z , según sea lo apropiado), los valores críticos y la conclusión.

MINITAB Minitab efectuará una prueba de rachas únicamente con una secuencia de datos numéricos, pero vea el manual *Minitab Student Laboratory Manual and Workbook* para evitar esta restricción. Ingrese los datos numéricos en la columna C1, luego seleccione **Stat**, **Nonparametrics** y **Runs**

Test. En el cuadro de diálogo, ingrese C1 para la variable, luego elija probar la aleatoriedad por arriba o por debajo de la media, o ingrese un valor a utilizar. Haga clic en **OK**. Los resultados del Minitab incluyen el número de rachas y el valor P (“la prueba es significativa a...”).

EXCEL Excel no está programado para la prueba de rachas para detectar aleatoriedad.

TI-83/84 PLUS La calculadora TI-83/84 Plus no está programada para la prueba de rachas para detectar aleatoriedad.



de encontrar la secuencia de las letras A y B, procedemos a aplicar la prueba de rachas tal como se describió. Los economistas utilizan la prueba de rachas para detectar aleatoriedad por arriba y por debajo de la media cuando tratan de identificar tendencias o ciclos. Una tendencia económica a la alza contendría una predominancia de D al principio y de A al final, de manera que el número de rachas sería pequeño. En una tendencia a la baja las A dominarían al principio y las D al final, con un número bajo de rachas. Un patrón cíclico produciría una secuencia que cambia sistemáticamente, de manera que el número de rachas tendería a ser grande. (Véase el ejercicio 14).

13-7 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- 1. Prueba de rachas.** Un encuestador realiza una encuesta llenando una página de respuestas de cada sujeto que entrevista. Luego mezcla las páginas antes de enviarlas para su análisis. ¿Se puede emplear la prueba de rachas de aleatoriedad para determinar si los sujetos encuestados fueron seleccionados en una secuencia aleatoria?
- 2. Prueba de rachas.** ¿Qué es una racha?
- 3. Prueba de rachas.** Si una prueba de rachas no lleva a la conclusión de que una muestra en particular es aleatoria, ¿podemos concluir que los datos se seleccionaron de una forma adecuada para fines estadísticos?
- 4. Muestra sesgada.** Después de utilizar la prueba de rachas para determinar aleatoriedad, se descubre que una secuencia de 20 hombres y 80 mujeres parece ser aleatoria. ¿Esto implica que la muestra es representativa de la población de adultos estadounidenses?

Identificación de rachas y cálculo de valores críticos. En los ejercicios 5 a 8, utilice la secuencia dada para determinar los valores de n_1 , n_2 , el número de rachas G y los valores críticos de la tabla A-10; además, utilice esos resultados para determinar si la secuencia parece ser aleatoria.

5. S N N S N N S N N S N N
6. T T T T F F F F T T T T
7. H H H M M H M M H M M H
8. X Y X X Y Y X X X Y Y Y X X X X Y Y Y Y

Uso de la prueba de rachas para detectar aleatoriedad. En los ejercicios 9 a 18, utilice la prueba de rachas de esta sección para determinar si la secuencia dada es aleatoria. Utilice un nivel de significancia de $\alpha = 0.05$. (Todos los datos están listados en orden por renglón).

9. **Géneros de osos.** El primer ejemplo de esta sección utilizó los géneros de los primeros 10 osos del conjunto de datos 6 del apéndice B. Realice una prueba de rachas para detectar aleatoriedad utilizando los géneros de los primeros 20 osos del conjunto de datos 6. A continuación se listan los géneros.

M M M M H H M M H H M M H M H M M H M M

10. **Ganadores olímpicos de medallas de oro.** Abajo se listan los ganadores de la medalla de oro en la carrera de 400 metros planos para hombres, donde E representa a Estados Unidos y O a cualquier otro país. Al parecer, ¿los países de los ganadores están en una secuencia aleatoria?

E	E	E	E	O	E	O	O	E	E	E	O	O
E	E	E	E	E	O	O	E	E	E	E	E	E

- 11. Prueba de la aleatoriedad de las victorias en las series mundiales de béisbol.** Pruebe la aseveración de que la secuencia de triunfos en las series mundiales de los equipos de la American League y la National League es aleatoria. Abajo se dan los resultados de los equipos de las ligas Americana y Nacional representados por A y N, respectivamente. ¿Qué sugieren los resultados acerca de las habilidades de las dos ligas?

N	A	A	A	N	N	A	A	N	N	N	N	A	A	A	N	A
N	A	N	A	A	A	N	A	N	A	A	A	N	A	N	A	

- 12. Prueba de aleatoriedad de ganadores de elección presidencial.** Para una secuencia reciente de elecciones presidenciales, el partido político del ganador se indica con D para demócrata y R para republicano. ¿Parece que se eligieron candidatos demócratas y republicanos en una secuencia que es aleatoria?

R	R	D	R	D	R	R	R	R	D	D	R	R	R	D	D	D
D	D	R	R	D	D	R	R	D	R	R	R	D	D	R	R	

- 13. Salas de cine: Prueba de aleatoriedad por arriba y por debajo de la mediana.** Las tendencias en los negocios y la economía a menudo se analizan con la prueba de rachas. A continuación se incluye el número de salas de cine, listadas en orden por renglón para cada año, comenzando en 1987 (según datos de la National Association of Theater Owners). Primero calcule la mediana, luego reemplace cada valor por A si está por arriba de la mediana y por D si está por debajo de la mediana. Despues aplique la prueba de rachas a la secuencia resultante de A y D. ¿Qué sugiere el resultado sobre la tendencia en el número de salas de cine?

20,595	21,632	21,907	22,904	23,740	24,344	24,789	25,830	26,995
28,905	31,050	33,418	36,448	35,567	34,490	35,170	35,361	

- 14. Bolsa de Valores: Prueba de aleatoriedad por arriba y por debajo de la mediana.** A continuación se listan los puntajes máximos anuales del promedio industrial Dow Jones para una secuencia de años recientes. Primero calcule la mediana de los valores, luego reemplace cada valor por A si está por encima de la mediana y por D si está por debajo de la mediana. Despues aplique la prueba de rachas a la secuencia resultante de A y D. ¿Qué sugiere el resultado acerca del mercado bursátil como una opción para invertir?

969	995	943	985	969	842	951	1036	1052	892
882	1015	1000	908	898	1000	1024	1071	1287	1287
1553	1956	2722	2184	2791	3000	3169	3413	3794	3978
5216	6561	8259	9374	11,497	11,723	11,338	10,635	10,454	10,895

- 15. Muestra grande: Géneros de osos.** En el primer ejemplo de esta sección se utilizaron los géneros de los primeros 10 osos del conjunto de datos 6 del apéndice B. Realice una prueba de rachas para detectar aleatoriedad utilizando los géneros de todos los osos del conjunto de datos 6.

- 16. Muestra grande: Lluvia sabatina.** En el segundo ejemplo de esta sección se utilizaron las cantidades de lluvia de los lunes, listadas en el conjunto de datos 10 del apéndice B. Utilice las cantidades de lluvia de los sábados para hacer una prueba de la aleatoriedad de los días con precipitación. Considere que un día es seco si la cantidad de precipitación es 0.00. Considere que un día es lluvioso si la cantidad de precipitación es cualquier valor mayor que 0.00.

- 17. Muestra grande: Prueba de aleatoriedad de dígitos impares y pares en el valor π .** Un artículo del *New York Times* acerca del cálculo de lugares decimales de π señaló que “los matemáticos están bastante seguros de que los dígitos de π son indistinguibles

de cualquier secuencia aleatoria". A continuación se presentan los primeros 100 lugares decimales de π . Pruebe la aleatoriedad de dígitos impares (I) y pares (P).

1 4 1 5 9 2 6 5 3 5 8 9 7 9 3 2 3 8 4 6 2 6 4 3 3 8 3 2 7 9 5 0 2 8 8 4 1 9 7 1
6 9 3 9 9 3 7 5 1 0 5 8 2 0 9 7 4 9 4 4 5 9 2 3 0 7 8 1 6 4 0 6 2 8 6 2 0 8 9 9
8 6 2 8 0 3 4 8 2 5 3 4 2 1 1 7 0 6 7 9

- 18. Muestra grande: Prueba de aleatoriedad de victorias en las series mundiales de béisbol.** Pruebe la aseveración de que la secuencia de victorias en series mundiales de los equipos de la American League y de la National League es aleatoria. A continuación se presentan resultados recientes, con los equipos de las ligas Americana y Nacional representados por A y N, respectivamente.

A N A N N N A A A A N A A A A N A N N A A N N A A A A N
N A A A A A N A N A N A N A A A A A A A A N N A N A N N A A N
N N A N A N A A A A N N A A N N N N A A A A N A N A N A A A
N A N A A A N A N A

13-7 MÁS ALLÁ DE LO BÁSICO

- 19. Cálculo de números críticos de rachas.** Al utilizar los elementos A, A, B y B, ¿cuál es el número mínimo posible de rachas que pueden acomodarse? ¿Cuál es el número máximo de rachas? Ahora remítase a la tabla A-10 para encontrar los valores críticos G para $n_1 = n_2 = 2$. ¿Qué concluye usted acerca de este caso?
- 20. Cálculo de valores críticos.**
- Utilice todos los elementos A, A, A, B, B, B, B, B, y haga una lista de las 84 secuencias diferentes posibles.
 - Calcule el número de rachas para cada una de las 84 secuencias.
 - Utilice los resultados de los incisos *a*) y *b*) para calcular sus propios valores críticos de G .
 - Compare sus resultados con los de la tabla A-10.

Repasso

En este capítulo examinamos seis pruebas no paramétricas diferentes para analizar datos muestrales. Las pruebas no paramétricas también se denominan pruebas de distribución libre porque no requieren que las poblaciones tengan una distribución particular; por ejemplo, una distribución normal. Sin embargo, las pruebas no paramétricas no son tan eficientes como las pruebas paramétricas, de manera que generalmente necesitamos una evidencia más fuerte antes de rechazar la hipótesis nula.

La tabla 13-9 lista las pruebas no paramétricas presentadas en este capítulo, junto con sus funciones. La tabla lista además las pruebas paramétricas correspondientes.

Conocimientos estadísticos y pensamiento crítico

- Prueba no paramétrica.** ¿Qué es una prueba *no paramétrica*?
- Prueba de distribución libre.** ¿Cuál es la diferencia entre una prueba no paramétrica y una prueba de distribución libre?
- Rango.** Muchas pruebas no paramétricas se basan en rangos. ¿Qué es un rango?

Tabla 13-9 Resumen de pruebas no paramétricas

Prueba no paramétrica	Función	Prueba paramétrica
Prueba del signo (sección 13-2)	Prueba del valor aseverado del promedio con una muestra	Prueba z o prueba t (secciones 8-4, 8-5)
	Prueba de las diferencias entre datos apareados	Prueba t (sección 9-4)
	Prueba del valor aseverado de una proporción	Prueba z (sección 8-3)
Prueba de rangos con signo de Wilcoxon (sección 13-3)	Prueba de las diferencias entre datos apareados	Prueba t (sección 9-4)
Prueba de la suma de rangos de Wilcoxon (sección 13-4)	Prueba de la diferencia entre dos muestras independientes	Prueba t o prueba z (sección 9-3)
Prueba de Kruskal-Wallis (sección 13-5)	Prueba si más de dos poblaciones independientes tienen la misma mediana	Análisis de varianza (sección 12-2)
Correlación de rangos (sección 13-6)	Prueba de la relación entre dos variables	Correlación lineal (sección 10-2)
Prueba de rachas (sección 13-7)	Prueba de la aleatoriedad de datos muestrales	(No hay prueba paramétrica)

- 4. Eficiencia.** Las pruebas no paramétricas no suelen ser tan *eficientes* como las pruebas paramétricas correspondientes, suponiendo que los requisitos necesarios se satisfacen. ¿Qué mide la eficiencia?

Ejercicios de repaso

Uso de pruebas no paramétricas. En los ejercicios 1 a 8, utilice un nivel de significancia de 0.05 con la prueba indicada. Si no se especifica una prueba en particular, utilice la prueba no paramétrica adecuada de este capítulo.

- 1. Medición de la inteligencia en niños.** La siguiente tabla lista datos apareados de los tiempos (en segundos) obtenidos de una muestra aleatoria de niños a los que se les dieron bloques junto con las instrucciones de construir una torre lo más alta posible (según datos de “Tower Building”, de Johnson y Courtney, *Child Development*, vol. 3). Este procedimiento se utiliza para medir la inteligencia en niños. Utilice la prueba del signo y un nivel de significancia de 0.05 para probar la aseveración de que no hay diferencia entre los tiempos del primero y del segundo ensayo.

Niño	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Primer ensayo	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
Segundo ensayo	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15

- 2. Medición de la inteligencia en niños.** Repita el ejercicio 1 utilizando la prueba de rangos con signo de Wilcoxon.
- 3. Clasificación de escuelas de negocios y de leyes.** La revista *U.S. News and World Report* clasificó escuelas de negocios y de leyes; en la siguiente tabla se muestran las clasificaciones basadas en esos resultados. ¿Existe una correlación entre la clasificación de las escuelas de negocios y la clasificación de las escuelas de leyes?

Universidad	Hvd	Yale	Sfd	NYU	Chi	UPenn	BU	Ohio	Gtwn	USC
Negocios	1	6	2	5	4	3	10	8	9	7
Leyes	2	1	3	4	5	6	9	10	7	8

- 4. Clasificación de escuelas de medicina.** La revista *U.S. News and World Report* clasificó escuelas de investigación médica y presentó una lista con las calificaciones correspondientes “promedio” en la prueba MCAT. La siguiente tabla se basa en esos resultados. ¿Existe una correlación entre la clasificación y la calificación promedio en la prueba MCAT?

Universidad	Hvd	NYU	Yale	Chi	USC	Ohio	BU	Utah
Clasificación	1	4	2	3	5	6	7	8
MCAT	11.3	11.0	11.4	10.3	10.8	10.5	9.5	9.6

- 5. Prueba de discriminación por género.** La compañía de contabilidad Axipon asegura que contrata empleados sin discriminación por género. De los 40 empleados recién contratados, 15 son mujeres. Aproximadamente la mitad de los solicitantes son hombres y la mitad mujeres, todos bien calificados. ¿Existe evidencia suficiente para acusarla de discriminar en favor de los hombres?

- 6. ¿Las personas que beben cerveza y las que beben licor tienen diferentes niveles de CAS?** Los datos muestrales de la siguiente lista indican los niveles de CAS (concentración de alcohol en sangre) en el momento del arresto de personas seleccionadas al azar por los delitos de conducir en estado de ebriedad (CEE) o conducir bajo la influencia del alcohol (CIA). Los datos están clasificados según el tipo de bebida consumida (según datos del Departamento de Justicia de Estados Unidos). Pruebe la aseveración de que los bebedores de cerveza y los bebedores de licor tienen los mismos niveles de CAS. Con base en estos resultados, ¿parece que ambos grupos son igualmente peligrosos, o existe un grupo más peligroso que el otro?

Cerveza				Licor				
0.129	0.146	0.148	0.152		0.220	0.225	0.185	0.182
0.154	0.155	0.187	0.212		0.253	0.241	0.227	0.205
0.203	0.190	0.164	0.165		0.247	0.224	0.226	0.234
					0.190	0.257		

- 7. ¿Es aleatoria la lotería?** A continuación se presentan los primeros dígitos seleccionados de 40 tomas consecutivas de la urna del juego de lotería Win 4 del estado de Nueva York. ¿Los dígitos impares y pares parecen haberse tomado en una secuencia aleatoria?

9 7 0 7 5 5 1 9 0 0 8 7 6 0 1 6 7 2 4 7
 5 5 5 2 0 4 4 9 9 0 5 3 3 1 9 2 5 6 8 2

- 8. ¿El peso de un automóvil afecta las heridas en la pierna durante un choque?** Se obtuvieron datos de experimentos de choques de automóviles realizados por la National Transportation Safety Administration. Se adquirieron automóviles nuevos, se hicieron chocar contra una barrera fija a 35 mi/h y se registraron las mediciones con un maniquí en el asiento del conductor. Utilice los datos muestrales listados abajo para probar las diferencias en las mediciones de carga (en libras) del fémur izquierdo entre las cuatro categorías de peso. ¿Existe evidencia suficiente para concluir que las mediciones de heridas en la pierna para las cuatro categorías de peso de automóviles no son las mismas? ¿Sugieren los datos que los automóviles más pesados son más seguros en un choque?

Subcompacto:	595	1063	885	519	422
Compacto:	1051	1193	946	984	584
Mediano:	629	1686	880	181	645
Grande:	1085	971	996	804	1376

Ejercicios de repaso acumulativo

En los ejercicios 1 a 8, use los datos de la siguiente tabla. Los valores se basan en un diagrama de dispersión que está incluido en “An SAT Coaching Program That Works”, de Jack Kaplan, *Chance*, vol. 15, núm. 1. Los valores son calificaciones de estudiantes en la prueba de matemáticas SAT, antes y después de tomar el curso de preparación para la prueba SAT.

Antes	460	470	490	490	510	510	600	620	610
Después	480	510	500	610	590	630	630	660	690

- Cálculo de estadísticos.** Calcule la media, la mediana, el rango, la desviación estándar y la varianza de las calificaciones previas al curso.
- Gráfica.** Construya un diagrama de dispersión. ¿Al parecer existe una correlación entre las calificaciones antes y después del curso?
- Correlación lineal.** Haga una prueba de correlación lineal entre las calificaciones antes y después del curso. Si hay una correlación lineal, ¿eso significa que el curso de preparación es eficaz? ¿Por qué?
- Correlación de rango.** Use la correlación de rango para probar una correlación entre las calificaciones antes y después del curso.
- Datos apareados.** Utilice una prueba *t* para probar la aseveración de que la diferencia media entre las calificaciones antes y después del curso es igual a cero.
- Prueba del signo.** Utilice la prueba del signo para probar la aseveración de que no existe diferencia entre las calificaciones antes y después del curso.
- Prueba de rangos con signo de Wilcoxon.** Utilice la prueba de rangos con signo de Wilcoxon para probar la aseveración de que no existe diferencia entre las calificaciones antes y después del curso.
- Eficiencia de prueba.** ¿Cuál de los ejercicios anteriores produce los resultados más útiles para determinar si el curso de preparación es eficaz? ¿Al parecer el curso de preparación es eficaz? ¿Por qué?

Actividades de cooperación en equipo

- Actividad en clase** Utilice el orden de los asientos en su clase y aplique la prueba de rachas para determinar si los estudiantes se acomodan aleatoriamente de acuerdo con el género. Después de registrar el orden de asientos, se puede hacer el análisis en subgrupos de tres o cuatro estudiantes.
- Actividad en clase** Formen grupos de 8 a 12 personas. Para cada miembro del grupo, *midan* su estatura y *midan* la envergadura de sus brazos. Para medir la envergadura de los brazos, el sujeto debe ponerse de pie con los brazos extendidos, como las alas de un avión. Es fácil marcar la altura y la envergadura de los brazos en un pizarrón y luego medir las distancias ahí. Divida las siguientes tareas en subgrupos de tres o cuatro personas.
 - Utilice la correlación de rangos con los datos muestrales apareados para determinar si existe una correlación entre la estatura y la envergadura de los brazos.
 - Utilice la prueba del signo para probar la diferencia entre las dos variables.
 - Utilice la prueba de rangos con signo de Wilcoxon para probar la diferencia entre las dos variables.
- Actividad en clase** Realice la actividad 2 utilizando el pulso en vez de la envergadura de los brazos. Mida los pulsos contando el número de latidos cardíacos en un minuto.
- Actividad fuera de clase** Formen grupos de tres o cuatro estudiantes. Investiguen la relación entre dos variables reuniendo sus propios datos muestrales apareados y utili-

zando los métodos de la sección 13-6 para determinar si existe una correlación significativa. Temas sugeridos:

- ¿Existe una relación entre el sabor y el costo de marcas diferentes de galletas con chispas de chocolate (o bebidas de cola)? (El sabor puede medirse en alguna escala numérica, como es de 1 a 10).
- ¿Existe una relación entre los salarios de los jugadores profesionales de béisbol (o básquetbol, o fútbol) y sus logros por temporada?
- Tasas versus pesos: ¿Existe una relación entre las tasas de consumo de combustible de los automóviles y los pesos de los automóviles?
- ¿Existe una relación entre la longitud de los pies de los hombres (o de las mujeres) y su estatura?
- ¿Existe una relación entre el promedio de las calificaciones de los estudiantes y la cantidad de tiempo que ven televisión?
- ¿Existe una relación entre la estatura de los padres (o de las madres) y las estaturas de sus hijos (o hijas) primogénitos?

- 5. Actividad fuera de clase** Vea el proyecto “De los datos a la decisión”, relacionado con el análisis del sorteo de 1970 utilizado para reclutar hombres en el ejército estadounidense. Puesto que los resultados de 1970 generaron preocupación por la aleatoriedad al seleccionar números prioritarios, diseñe un nuevo procedimiento para generar los 366 números prioritarios. Utilice su procedimiento para generar los 366 números y pruebe sus resultados utilizando las técnicas que se sugieren en los incisos *a*, *b* y *c*) del proyecto “De los datos a la decisión”. ¿Cómo se comparan sus resultados con los obtenidos en 1970? ¿Su procedimiento de selección aleatoria parece ser mejor que el que se usó en 1970? Escriba

un reporte que describa con claridad el proceso que diseñó. También incluya su análisis y conclusiones.

- 6. Actividad fuera de clase** Formen grupos de tres o cuatro personas. Encuesten a estudiantes, pidiéndoles que identifiquen su área de estudios y su género. Para cada sujeto entrevistado, determine la exactitud de la hora en su reloj. Primero ajuste su propio reloj a la hora correcta utilizando una fuente precisa y confiable (del tipo “Al escuchar el tono, la hora es...”). Para los relojes que estén adelantados, registre tiempos positivos; para los relojes que estén atrasados, registre tiempos negativos. Utilice los datos muestrales para responder estas preguntas:
- Al parecer, ¿ambos géneros tienen los mismos errores?
 - Al parecer, ¿las diferentes áreas de estudio tienen los mismos errores?

- 7. Actividad en clase** Formen grupos de 8 a 12 personas. Para cada miembro del grupo, midan su estatura y la altura de su ombligo, que es la altura desde el piso hasta el ombligo. Utilice el coeficiente de correlación de rangos para determinar si existe una correlación entre la estatura y la altura del ombligo.

- 8. Actividad en clase** Formen grupos de tres o cuatro personas. El apéndice B incluye muchos conjuntos de datos que todavía no pueden resolverse por medio de los métodos de este capítulo. Revise el apéndice B y busque las variables de interés, luego investigue empleando los métodos apropiados de estadística no paramétrica. Enumere sus conclusiones y trate de identificar aplicaciones prácticas.

Proyecto tecnológico

En el pasado ha habido intentos por identificar o establecer contacto con vida inteligente extraterrestre, los cuales incluyen esfuerzos para enviar mensajes de radio llevando información acerca de nosotros, los terrícolas. El doctor Frank Drake, de la Universidad Cornell, elaboró un mensaje de radio de este tipo, que podría ser transmitido en series de pulsos y silencios. Se puede pensar en los pulsos y silencios como ceros y unos. A continuación se lista un mensaje consistente en 77 ceros y unos. Si factorizamos 77 en los números primos 7 y 11, y luego hacemos una cuadrícula de 11×7 y ponemos un punto en aquellas posiciones correspondientes a un pulso o 1, podemos obtener una imagen sencilla de algo.

Suponga que la secuencia de 77 unos y ceros se envía como un mensaje de radio que es interceptado por vida extraterrestre con una inteligencia suficiente para estudiar este libro. Si el mensaje de radio se prueba utilizando los métodos de este capítulo, ¿la secuencia aparecerá como un “ruido aleatorio” o se identificará como un patrón que no es aleatorio? Construya la imagen representada por los dígitos e identifíquela.

```
0 0 1 1 1 0 0 0 0 1 1 1 0 0 0 0 0 1 0 0 0
1 1 1 1 1 1 0 0 1 1 1 0 0 0 0 1 1 1 0 0
0 0 1 1 1 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0
1 0 0 0 0 1 0 1 0 0 0 0 1 0
```

De los datos a la decisión

Pensamiento crítico: ¿Fue aleatorio el sorteo?

En 1970 se realizó un sorteo para determinar quién sería reclutado en el ejército estadounidense. Las 366 fechas del año se colocaron en cápsulas individuales. Primero, las 31 cápsulas de enero se ubicaron en una caja; luego se añadieron las 29 cápsulas de febrero y se mezclaron los dos meses. Despues, se agregaron las 31 cápsulas de marzo y se mezclaron los tres meses. Este procedimiento continuó hasta que se incluyeron todos los meses. La primera cápsula seleccionada fue el 14 de septiembre, por lo tanto los hombres que nacieron en esa fecha fueron reclutados primero. La lista adjunta muestra las 366 fechas en el orden de su selección.

Análisis de los resultados

- Utilice la prueba de rachas para probar la aleatoriedad de la secuencia por arriba y por debajo de la mediana de 183.5.
- Utilice la prueba de Kruskal-Wallis para probar la aseveración de que los 12 meses tuvieron números prioritarios obtenidos de la misma población.
- Calcule las 12 medias mensuales. Luego registre esas 12 medias en una gráfica. (En el eje horizontal liste los 12 meses, y el eje vertical debe ir de 100 hasta 260). Observe cualquier patrón que sugiera que los números prioritarios originales no fueron seleccionados aleatoriamente.
- Con base en los resultados de los incisos a), b) y c), decida si este sorteo en particular fue justo. Escriba una aseveración sustentando su postura de que fue

justo, o explicando por qué cree que no fue justo. Si usted decide que este sorteo no fue justo, describa un procedimiento para seleccionar los números que habrían sido justos.



Ene:	305	159	251	215	101	224	306	199	194	325	329	221	318	238	017	121
	235	140	058	280	186	337	118	059	052	092	355	077	349	164	211	
Feb:	086	144	297	210	214	347	091	181	338	216	150	068	152	004	089	212
	189	292	025	302	363	290	057	236	179	365	205	299	285			
Mar:	108	029	267	275	293	139	122	213	317	323	136	300	259	354	169	166
	033	332	200	239	334	265	256	258	343	170	268	223	362	217	030	
Abr:	032	271	083	081	269	253	147	312	219	218	014	346	124	231	273	148
	260	090	336	345	062	316	252	002	351	340	074	262	191	208		
May:	330	298	040	276	364	155	035	321	197	065	037	133	295	178	130	055
	112	278	075	183	250	326	319	031	361	357	296	308	226	103	313	
Jun:	249	228	301	020	028	110	085	366	335	206	134	272	069	356	180	274
	073	341	104	360	060	247	109	358	137	022	064	222	353	209		
Jul:	093	350	115	279	188	327	050	013	277	284	248	015	042	331	322	120
	098	190	227	187	027	153	172	023	067	303	289	088	270	287	193	
Ago:	111	045	261	145	054	114	168	048	106	021	324	142	307	198	102	044
	154	141	311	344	291	339	116	036	286	245	352	167	061	333	011	
Sep:	225	161	049	232	082	006	008	184	263	071	158	242	175	001	113	207
	255	246	177	063	204	160	119	195	149	018	233	257	151	315		
Oct:	359	125	244	202	024	087	234	283	342	220	237	072	138	294	171	254
	288	005	241	192	243	117	201	196	176	007	264	094	229	038	079	
Nov:	019	034	348	266	310	076	051	097	080	282	046	066	126	127	131	107
	143	146	203	185	156	009	182	230	132	309	047	281	099	174		
Dic:	129	328	157	165	056	010	012	105	043	041	039	314	163	026	320	096
	304	128	240	135	070	053	162	095	084	173	078	123	016	003	100	



Proyecto de Internet

Pruebas no paramétricas

En este capítulo se estudiaron métodos de prueba de hipótesis de carácter no paramétrico o de distribución libre. Los métodos no paramétricos le permiten probar hipótesis sin hacer suposiciones al respecto de la distribución poblacional

subyacente que se está muestreando. Para continuar su trabajo con este importante tipo de métodos estadísticos de prueba, visite el sitio de Internet de *Estadística Elemental*:

<http://www.pearsoneducacion.net/triola>

La estadística en el trabajo

“Quienes solicitan empleo deben tener conocimientos fundamentales de estadística y sobre sus implicaciones en el mundo de los negocios”.



Angela Gillespi

Analista de tráfico, Lycos.com

Como analista de tráfico para Lycos, Inc., Angela realiza reportes sobre mediciones amplias y menores de tráfico. Ella hace un seguimiento de los cambios en las tendencias y los patrones de comportamiento del uso del sitio de Internet, con el fin de mejorarlo incrementando su alcance y “adhesividad” (también llamada *stickiness*, la cantidad de tiempo que las personas permanecen conectadas a un sitio Web en particular).

¿Cuál es su trabajo en Lycos?

Realizo reportes de tráfico de las actividades en nuestro sitio cada semana, los cuales son revisados por nuestros equipos de trabajo de producción y gerentes. Ellos ven qué aumenta, qué disminuye y toman decisiones referentes al gasto de los recursos.

Mis reportes analizan básicamente las tendencias en los sitios y permiten hacer proyecciones acerca de dónde estaremos dentro de un año o en cualquier otro período.

¿Qué conceptos de estadística utiliza?

Análisis de regresión y valores de R cuadrada.

¿De qué manera utiliza la estadística en su trabajo?

Para determinar qué es lo que funciona y lo que no funciona para nuestros usuarios. Para determinar la eficacia de las campañas de publicidad y para crear proyectos de crecimiento futuro.

Por favor, describa un ejemplo específico e ilustre la manera en que la aplicación de la estadística permitió mejorar un producto o un servicio.

Al final de nuestro último año fiscal, nuestro director ejecutivo, Bob Davis, presentó a la compañía una meta promedio diaria de visitantes del sitio que debía lograrse al final del siguiente año fiscal. Utilizando los datos de visitantes de dos años anteriores, hice una proyección que mostró

en donde estaríamos al final del siguiente año fiscal si las cosas permanecían estables. El uso de un valor de R cuadrada le dio a estas gráficas el impulso que necesitaban para ser eficaces. Actualicé las gráficas cada semana y las presenté al equipo de gerencia de producción. Los datos les ayudaron a comprender qué reajustes debían hacer a sus productos, y cada semana se acercaron más y más a sus metas. Cuando Bob presentó por primera vez la meta de visitantes, todos pensamos que se había vuelto loco, pero estoy feliz de decir que al final del siguiente año fiscal alcanzaremos nuestra meta o al menos un 98% de ella. Sin la representación que realicé, la gerencia de producción no hubiera sabido en dónde enfocar su energía y sus recursos. Puesto que forman un equipo eficiente, hemos alcanzado nuestra meta inalcanzable.

¿El uso que usted hace de la probabilidad y la estadística está aumentando, disminuyendo o permanece estable?

Conforme Lycos se vuelve más compleja, ellos (la gerencia) esperan reportes cada vez más elaborados, de manera que está en aumento.

¿Considera que los solicitantes de empleo que tienen algunos estudios de estadística son evaluados de forma más favorable?

Definitivamente, y no sólo en lo que respecta a los reportes de Lycos, sino también en marketing y finanzas.



Control estadístico de procesos

14



14-1 Panorama general

14-2 Gráficas de control para la variación y la media

14-3 Gráficas de control para atributos

¿La producción de altímetros para aviones es peligrosa para quienes vuelan?

La Altigauge Manufacturing Company produce altímetros para aviones, los cuales proporcionan a los pilotos lecturas de su altitud con respecto al nivel del mar. La exactitud de los altímetros es importante, ya que los pilotos confían en ellos para mantenerse a altitudes con espacio vertical seguro sobre montañas, torres y rascacielos, así como para continuar con una separación vertical pertinente en relación con otras aeronaves. La exactitud de los altímetros es especialmente importante cuando los pilotos van a aterrizar sin poder ver el suelo. En el pasado, pilotos y pasajeros han muerto en accidentes causados por lecturas incorrectas de altímetros, que provocaron que el piloto creyera que se encontraban a salvo en el aire, cuando en realidad la nave volaba demasiado bajo.

Puesto que los altímetros de aviones son esenciales para la seguridad de la aviación, su exactitud se controla cuidadosamente mediante normas gubernamentales. Según la Norma 43 de la Federal Aviation Administration (apéndice E), un altímetro debe proporcionar lecturas

con un error de no más de 20 pies al probarse para una altitud de 1000 pies.

En la Altigauge Manufacturing Company se seleccionan al azar cuatro altímetros diariamente, durante 20 días hábiles consecutivos; en la tabla 14-1 se listan los errores (en pies) cuando se prueban en una cámara de presión que simula una altitud de 1000 pies. Por ejemplo, el día 1 las lecturas reales de los cuatro altímetros seleccionados fueron 1002 pies, 992 pies, 1005 pies y 1011 pies, de manera que los errores correspondientes (en pies) son 2, -8, 5 y 11.

En este capítulo evaluaremos este proceso de fabricación de altímetros, analizando el comportamiento de los errores al paso del tiempo. Estudiaremos la forma en que se utilizan los métodos estadísticos para supervisar un proceso de fabricación con el fin de identificar y corregir cualquier problema grave. Además de ayudar a que las empresas permanezcan en operación, los métodos estadísticos pueden afectar de forma positiva nuestra seguridad de forma muy significativa.

Tabla 14-1 Errores de altímetros de aviones (en pies)

Día	Error					Media	Mediana	Rango	Desviación estándar
1	2	-8	5	11	2.50	3.5	19		7.94
2	-5	2	6	8	2.75	4.0	13		5.74
3	6	7	-1	-8	1.00	2.5	15		6.98
4	-5	5	-5	6	0.25	0.0	11		6.08
5	9	3	-2	-2	2.00	0.5	11		5.23
6	16	-10	-1	-8	-0.75	-4.5	26		11.81
7	13	-8	-7	2	0.00	-2.5	21		9.76
8	-5	-4	2	8	0.25	-1.0	13		6.02
9	7	13	-2	-13	1.25	2.5	26		11.32
10	15	7	19	1	10.50	11.0	18		8.06
11	12	12	10	9	10.75	11.0	3		1.50
12	11	9	11	20	12.75	11.0	11		4.92
13	18	15	23	28	21.00	20.5	13		5.72
14	6	32	4	10	13.00	8.0	28		12.91
15	16	-13	-9	19	3.25	3.5	32		16.58
16	8	17	0	13	9.50	10.5	17		7.33
17	13	3	6	13	8.75	9.5	10		5.06
18	38	-5	-5	5	8.25	0.0	43		20.39
19	18	12	25	-6	12.25	15.0	31		13.28
20	-27	23	7	36	9.75	15.0	63		27.22

14-1 Panorama general

En el capítulo 2 señalamos que al describir, explorar y comparar conjuntos de datos, es sumamente importante tomar en cuenta las características centrales, la variación, la distribución, los valores extremos y los aspectos que cambian con el paso del tiempo. El principal objetivo de este capítulo es estudiar el último aspecto: las características cambiantes de los datos a lo largo del tiempo. Cuando se investigan características tales como el centro y la variación, es importante saber si se trata de una población estable o de una que está cambiando conforme transcurre el tiempo.

Actualmente existe una fuerte tendencia a tratar de mejorar la calidad de los bienes y servicios, y un número creciente de empresas utilizan los métodos que se presentan en este capítulo. La evidencia de la creciente importancia de la calidad se encuentra en la publicidad y en el gran número de libros y artículos que se enfocan en el tema de la calidad. En muchos casos, quienes solicitan empleo (*¿usted?*) poseen una ventaja definitiva cuando pueden decir a los empleadores que estudiaron estadística y métodos de control de calidad. Este capítulo presentará algunas de las herramientas básicas que se utilizan comúnmente para controlar la calidad.

Minitab, Excel y otros paquetes estadísticos de cómputo incluyen programas para generar automáticamente el tipo de gráficas que se estudian en este capítulo: incluiremos diversos ejemplos de esas representaciones gráficas. Las gráficas de control son un buen ejemplo de los maravillosos recursos gráficos que nos permiten *ver* y *comprender* algunas propiedades de los datos que, de otra forma, serían muy difíciles o imposibles de entender. El mundo necesita más personas capaces de construir e interpretar gráficas importantes, tales como las gráficas de control descritas en este capítulo.

14-2 Gráficas de control para la variación y la media

Concepto clave El principal objetivo de esta sección es la construcción de gráficas de rachas, gráficas R y gráficas \bar{x} para controlar características importantes de datos a lo largo del tiempo. Usaremos este tipo de gráficas para determinar si un proceso es estadísticamente estable (o si está bajo control estadístico).

La siguiente definición describe formalmente el tipo de datos que manejaremos en este capítulo.



Definición

Los **datos de proceso** son datos ordenados de acuerdo con alguna secuencia de tiempo. Son mediciones de una característica de bienes o servicios que resultan de alguna combinación de equipo, personas, materiales, métodos y condiciones.

Por ejemplo, la tabla 14-1 incluye datos de proceso que consisten en el error medido (en pies) de las lecturas de altímetros durante 20 días consecutivos de producción. Cada día se seleccionaron cuatro altímetros al azar y se probaron. Puesto que los

datos en la tabla 14-1 están ordenados de acuerdo con el momento en que se seleccionaron, se trata de datos de proceso. Es muy importante reconocer este punto:

Las características importantes de datos de proceso pueden cambiar a lo largo del tiempo.

Al producir altímetros, el fabricante empleará personal competente y bien capacitado, además de máquinas correctamente calibradas, pero si el personal es reemplazado o si las máquinas se estropean con el uso, los altímetros podrían empezar a resultar defectuosos. Existen compañías que han ido a la bancarrota por haber permitido, involuntariamente, que el proceso de fabricación se deteriorara al no ejercer un control constante.

Gráficas de rachas

Existen varios métodos que pueden emplearse para controlar un proceso y así asegurar que las características deseadas importantes no cambien; el análisis de una *gráfica de rachas* es un método de este tipo.

Definición

Una **gráfica de rachas** es una gráfica secuencial de valores de datos *individuales* a lo largo del tiempo. Un eje (generalmente el eje vertical) se utiliza para los valores de los datos, y el otro eje (generalmente el eje horizontal) se emplea para la secuencia de tiempo.

PROBLEMA
14
DEL CAPÍTULO

EJEMPLO Fabricación de altímetros para aviones Trate los 80 errores de los altímetros de la tabla 14-1 como una secuencia de mediciones consecutivas, construya una gráfica de rachas utilizando el eje vertical para los errores y el eje horizontal para identificar el orden de los datos muestrales.

SOLUCIÓN La figura 14-1 es la gráfica de rachas generada por Minitab para los datos de la tabla 14-1. La escala vertical se diseñó para ajustarse a los errores

continúa

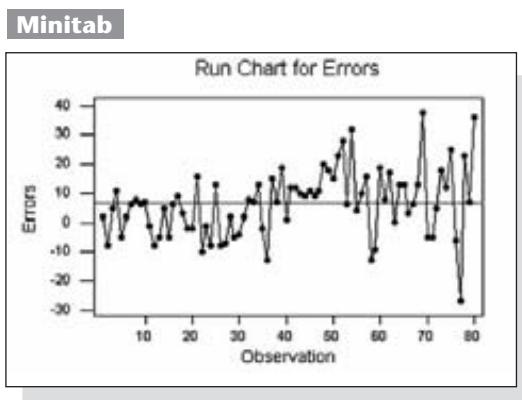


Figura 14-1 Gráfica de rachas de los errores individuales de altímetros de la tabla 14-1



El efecto Flynn: tendencia a la alza en puntuaciones de CI

Una gráfica de rachas o gráfica de control de las puntuaciones de CI revelaría que éstas exhiben una tendencia a incrementarse, ya que las puntuaciones de CI han estado aumentando de forma estable desde que empezaron a utilizarse hace casi 70 años. Esta tendencia es mundial y es igual en los distintos tipos de pruebas de inteligencia, incluso en aquellas que se basan casi por completo en el razonamiento abstracto y no verbal, con mínima influencia de la cultura. A esta tendencia al incremento se le ha llamado *efecto Flynn*, ya que el científico político James R. Flynn descubrió esta tendencia en sus estudios con reclutas del ejército de Estados Unidos. La cantidad del incremento es muy sustancial: con base en la puntuación media del CI de 100, se estima que el CI medio en 1920 era de cerca de 77. Por lo tanto, el estudiante común actual es brillante, si se le compara con sus bisabuelos. Hasta ahora no hay una explicación aceptable para el efecto Flynn.

de los altímetros, que van desde -27 pies hasta 38 pies, que son los valores mínimo y máximo de la tabla 14-1. La escala horizontal está diseñada para incluir los 80 valores ordenados en secuencia. El primer punto representa el primer valor de 2 pies, el segundo punto representa el segundo valor de -8 pies y así sucesivamente.

En la figura 14-1 la escala horizontal identifica el número de muestra, de manera que el número 20 indica el vigésimo artículo. La escala vertical representa el error del altímetro (en pies). Ahora examine la figura 14-1 y trate de identificar cualquier *patrón* que resalte a la vista. La figura 14-1 revela este problema: conforme el tiempo avanza de izquierda a derecha, las alturas de los puntos parecen mostrar un patrón de variación creciente. Observe cómo los puntos a la izquierda fluctúan mucho menos que los puntos de la derecha. Las normas de la Federal Aviation Administration exigen errores menores de 20 pies (o que estén entre 20 pies y -20 pies), de tal manera que los altímetros representados por puntos a la izquierda son aceptables, mientras que varios de los puntos de la derecha corresponden a altímetros que no cumplen con las especificaciones requeridas. Parece que el proceso de fabricación empezó bien, pero se deterioró con el paso del tiempo. Si se deja así, este proceso de fabricación provocará que la empresa tenga que cerrar.

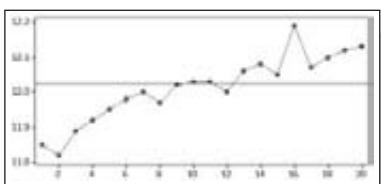
Interpretación de las gráficas de rachas Los datos de un proceso se pueden tratar como si provinieran de una población con una media, desviación estándar, distribución y otras características constantes únicamente cuando el proceso es *estadísticamente estable*.

Definición

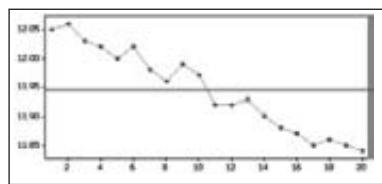
Un proceso es **estadísticamente estable** (o está **bajo control estadístico**) si sólo varía de forma natural, sin patrones, ciclos o puntos fuera de lo común.

La figura 14-2 consiste en gráficas de rachas que ilustran los patrones típicos que indican formas en las cuales el proceso de llenado de latas de bebida de cola de 12 onzas puede no ser estadísticamente estable.

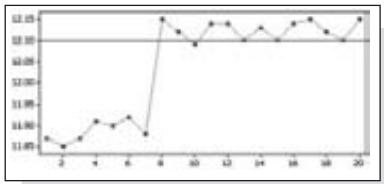
- **Figura 14-2a):** Hay una evidente *tendencia creciente*, que corresponde a valores que se incrementan con el paso del tiempo. Si el proceso de llenado continúa con este tipo de patrón, las latas se llenarían con más y más bebida de cola hasta que el líquido empezara a derramarse y los empleados terminarían nadando en bebida de cola.
- **Figura 14-2b):** Hay una evidente *tendencia descendente* que corresponde a valores que disminuyen de manera estable. Las latas se llenarían con menos y menos bebida de cola hasta que estarían casi vacías. Un proceso como éstos requeriría de una revisión completa de las latas, con la finalidad de llenarlas con suficiente cantidad para distribuirlas a los consumidores.
- **Figura 14-2c):** Hay un *cambio hacia arriba*. Una gráfica de rachas como ésta podría resultar de un ajuste en el proceso de llenado, provocando que los valores posteriores sean más altos.

Minitab

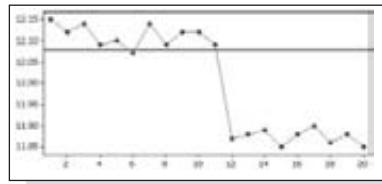
a)



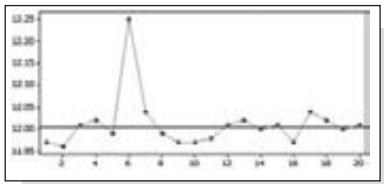
b)



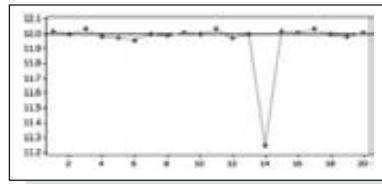
c)



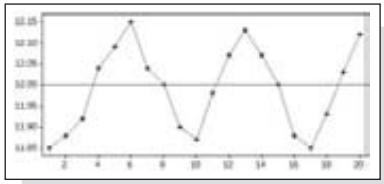
d)



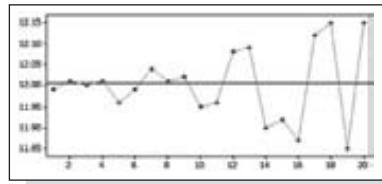
e)



f)



g)



h)

Figura 14-2

Procesos que no son estadísticamente estables

- **Figura 14-2d):** Existe un *cambio hacia abajo*. Los primeros valores son relativamente estables, pero después algo sucede, ya que los últimos valores son relativamente estables, aunque en un nivel mucho más bajo.
- **Figura 14-2e):** El proceso es estable, salvo por un *valor excepcionalmente alto*. La causa de un valor tan fuera de lo común debe investigarse. Tal vez las latas se atascaron temporalmente y una lata en particular se llenó dos veces en lugar de una.
- **Figura 14-2f):** Existe un *valor excepcionalmente bajo*.

- **Figura 14-2g):** Hay un *patrón cíclico* (o ciclo repetitivo). Como es evidente, este patrón no es aleatorio y, por lo tanto, revela un proceso estadísticamente inestable. Quizá se hayan hecho reajustes periódicos a la maquinaria con el fin de buscar continuamente algún valor deseado, pero sin éxito.
- **Figura 14-2h):** La *variación aumenta con el paso del tiempo*. Éste es un problema común en el control de calidad. El efecto neto es que los productos varían más y más hasta que casi todos son defectuosos. Por ejemplo, algunas latas de bebida de cola se derramarán, desperdimando el producto, y otras no se llenarán por completo y no podrán distribuirse a los consumidores.

Una meta común de muchos métodos diferentes de control de calidad es la siguiente: *reducir la variación* de un producto o servicio. Por ejemplo, la Ford se preocupó por la variación cuando se dio cuenta de que sus transmisiones requerían significativamente más reparaciones por garantía que el mismo tipo de transmisiones fabricadas por Mazda en Japón. Un estudio reveló que las transmisiones de Mazda tenían mucho menos variación en las cajas de velocidades; es decir, las medidas cruciales en las cajas de velocidades variaban mucho menos en las transmisiones Mazda. Aun cuando las transmisiones Ford estaban construidas dentro de los límites permitidos, las transmisiones Mazda eran más confiables gracias a su menor variación. La variación en un proceso puede resultar por dos causas.

Definiciones

La **variación aleatoria** se debe al azar; es el tipo de variación inherente a cualquier proceso que no es capaz de producir un bien o servicio exactamente de la misma forma cada vez.

La **variación assignable** resulta de causas identificables (como maquinaria defectuosa, empleados sin capacitación adecuada, etcétera).

Más adelante en este capítulo, consideraremos formas de distinguir entre la variación assignable y la variación aleatoria.

La gráfica de rachas es una herramienta para supervisar la estabilidad de un proceso. Ahora estudiaremos las *gráficas de control*, que también son sumamente útiles para los mismos propósitos.

Gráfica de control para supervisar la variación: Gráfica *R*

En el artículo “The State of Statistical Process Control as We Proceed into the 21st Century” (Stoumbos, Ryan y Woodall, *Journal of the American Statistical Association*, vol. 95, núm. 451), los autores afirman que “las gráficas de control son una de las herramientas más importantes y más utilizadas en estadística. Sus aplicaciones han pasado de los procesos de fabricación a la ingeniería, las ciencias ambientales, la biología, la genética, la epidemiología, la medicina, las finanzas e incluso al cumplimiento de la ley y los deportes”. Iniciamos con la definición de una gráfica de control.

Definición

Una **gráfica de control** de una característica de proceso (como la media o la variación) consiste en valores graficados en secuencia a lo largo del tiempo e incluye una **línea central** así como un **límite de control inferior** (LCI) y un **límite de control superior** (LCS). La línea central representa un valor central de las mediciones características, mientras que los límites de control son las fronteras utilizadas para separar e identificar cualesquiera puntos considerados *fueras de lo común*.

Supondremos que desconocemos la desviación estándar poblacional σ , mientras consideramos únicamente dos de diversos tipos de *gráficas de control*: 1. las gráficas R (o gráficas de rangos), que se utilizan para hacer un seguimiento de la variación, y 2. las gráficas \bar{x} que se emplean para verificar medias. Al utilizar gráficas de control para supervisar un proceso, es común que se consideren las gráficas R y las gráficas \bar{x} en conjunto, ya que un proceso estadísticamente inestable puede ser el resultado de un aumento en la variación, de cambios en las medias o de ambos.

Una **gráfica R** (o **gráfica de rangos**) es una gráfica de los rangos muestrales, en vez de valores muestrales individuales, y se aplica para verificar la *variación* en un proceso. (Podría parecer más sensato utilizar desviaciones estándar, pero las gráficas de rangos se emplean con mayor frecuencia en la práctica. Esto es una consecuencia de los tiempos en que no se disponía de calculadoras ni de computadoras. Véase el ejercicio 17, en el que se incluye una gráfica de control basada en desviaciones estándar). Además de graficar los valores de los rangos, incluimos una línea central localizada en \bar{R} , que denota la media de todos los rangos muestrales, así como otra línea para el límite de control inferior y una tercera línea para el límite de control superior. A continuación se presenta un resumen de la notación de los componentes de la gráfica R .

Notación

Se sabe que los datos de proceso consisten en una secuencia de muestras, todas del mismo tamaño n , y la distribución de los datos de proceso es esencialmente normal.

n = tamaño de cada muestra o *subgrupo*

\bar{R} = media de los rangos muestrales (es decir, la suma de los rangos muestrales dividida entre el número de muestras)

Verificación de un proceso de variación: Gráfica de control para R

Puntos graficados: rangos muestrales

Línea central: \bar{R}

Límite de control superior (LCS): $D_4\bar{R}$ (donde D_4 se encuentra en la tabla 14-2)

Límite de control inferior (LCI): $D_3\bar{R}$ (donde D_3 se encuentra en la tabla 14-2)

Los valores D_4 y D_3 fueron calculados por expertos en control de calidad y sirven para simplificar los cálculos. Los límites de control superior e inferior de $D_4\bar{R}$ y $D_3\bar{R}$ son valores casi equivalentes a los límites de un intervalo de confianza del 99.7%. Por lo tanto, es muy poco probable que los valores de un proceso estadísticamente estable caigan más allá de estos límites. Si un valor cae fuera de esos límites, es muy probable que el proceso no sea estadísticamente estable.



Variación assignable costosa

El Mars Climate Orbiter fue lanzado por la NASA y enviado a Marte, pero se destruyó cuando voló demasiado cerca de su planeta de destino. La pérdida se calculó en \$125 millones. Se descubrió que la causa de la colisión fue la confusión en el empleo de las unidades utilizadas para realizar cálculos. Los datos de la aceleración estaban en las unidades inglesas de libras de fuerza, pero el Jet Propulsion Laboratory supuso que las unidades utilizadas eran "newtons" métricos en vez de libras. Quienes dirigían la nave espacial dieron consecutivamente cantidades erróneas de la fuerza para ajustar la posición de la nave. Los errores causados por la discrepancia fueron muy pequeños al principio, pero el error acumulado a lo largo de los meses de travesía de la nave espacial fue la causa de su fracaso.

En 1962 la nave espacial que transportaba al satélite Mariner I fue destruida por controladores en Tierra, cuando se salió de curso debido a la falta de un signo menos en un programa de cómputo.



¡Cuidado con los ajustes excesivos!

La empresa Nashua Corp. tuvo problemas con su máquina para recubrimiento de papel y consideró gastar millones de dólares para reemplazarla. La máquina estaba funcionando bien y con un proceso estable, pero las muestras se empezaron a tomar con mucha frecuencia; en consecuencia, con base en esos resultados, se le hicieron ajustes. Estos ajustes excesivos, denominados *alteración indebida*, causaron desviaciones de la distribución que hasta entonces había sido buena. El efecto fue un incremento en los defectos. Cuando el experto en estadística y control, W. Edwards Deming, estudió el proceso, recomendó que no se le hicieran ajustes, a menos que hubiera una señal de que el proceso había cambiado o se había vuelto inestable. La compañía funcionó mejor sin ajustes que con la alteración realizada.

Tabla 14-2 Constantes de una gráfica de control

<i>n</i> : Número de observaciones en subgrupo	\bar{x}		<i>s</i>		<i>R</i>	
	<i>A</i> ₂	<i>A</i> ₃	<i>B</i> ₃	<i>B</i> ₄	<i>D</i> ₃	<i>D</i> ₄
2	1.880	2.659	0.000	3.267	0.000	3.267
3	1.023	1.954	0.000	2.568	0.000	2.574
4	0.729	1.628	0.000	2.266	0.000	2.282
5	0.577	1.427	0.000	2.089	0.000	2.114
6	0.483	1.287	0.030	1.970	0.000	2.004
7	0.419	1.182	0.118	1.882	0.076	1.924
8	0.373	1.099	0.185	1.815	0.136	1.864
9	0.337	1.032	0.239	1.761	0.184	1.816
10	0.308	0.975	0.284	1.716	0.223	1.777
11	0.285	0.927	0.321	1.679	0.256	1.744
12	0.266	0.886	0.354	1.646	0.283	1.717
13	0.249	0.850	0.382	1.618	0.307	1.693
14	0.235	0.817	0.406	1.594	0.328	1.672
15	0.223	0.789	0.428	1.572	0.347	1.653
16	0.212	0.763	0.448	1.552	0.363	1.637
17	0.203	0.739	0.466	1.534	0.378	1.622
18	0.194	0.718	0.482	1.518	0.391	1.608
19	0.187	0.698	0.497	1.503	0.403	1.597
20	0.180	0.680	0.510	1.490	0.415	1.585
21	0.173	0.663	0.523	1.477	0.425	1.575
22	0.167	0.647	0.534	1.466	0.434	1.566
23	0.162	0.633	0.545	1.455	0.443	1.557
24	0.157	0.619	0.555	1.445	0.451	1.548
25	0.153	0.606	0.565	1.435	0.459	1.541

Fuente: Adaptado del *ASTM Manual on the Presentation of Data and Control Chart Analysis*, © 1976 ASTM, pp. 134-136. Reproducido bajo permiso de American Society of Testing and Materials.

PROBLEMA
14
DEL CAPÍTULO

EJEMPLO Fabricación de altímetros para aviones Remítase a los errores de los altímetros de la tabla 14-1. Empleando muestras de tamaño *n* = 4, reunidas cada día de fabricación, construya una gráfica de control para *R*.

SOLUCIÓN Iniciamos con el cálculo del valor de \bar{R} , la media de los rangos muestrales.

$$\bar{R} = \frac{19 + 13 + \dots + 63}{20} = 21.2$$

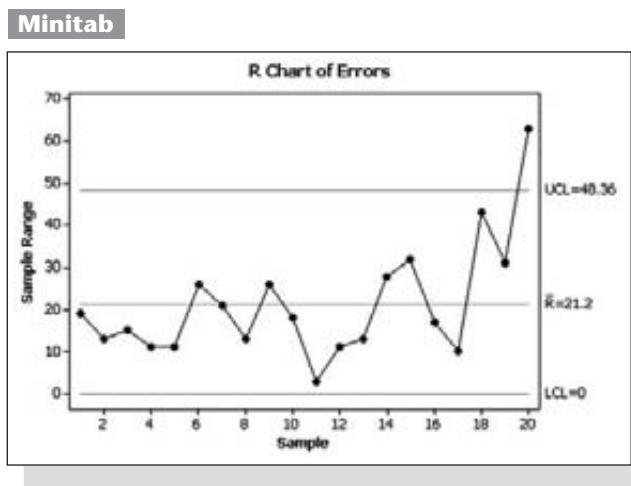
Por lo tanto, la línea central de la gráfica *R* está ubicada en $\bar{R} = 21.2$. Para calcular los límites de control superior e inferior, debemos obtener los valores de

D_3 y D_4 . Si nos remitimos a la tabla 14-2, para $n = 4$, obtenemos $D_3 = 0.000$ y $D_4 = 2.282$, de manera que los límites de control son los siguientes:

$$\text{Límite de control superior: } D_4 \bar{R} = (2.282)(21.2) = 48.4$$

$$\text{Límite de control inferior: } D_3 \bar{R} = (0.000)(21.2) = 0.0$$

Con un valor de línea central de $\bar{R} = 21.2$ y límites de control de 48.4 y 0.0, ahora procedemos a graficar los rangos muestrales. Los resultados se presentan en la pantalla de Minitab.



Interpretación de las gráficas de control

Al interpretar las gráficas de control, el siguiente punto es extremadamente importante:

Los límites de control superior e inferior de una gráfica de control están basados en el comportamiento *real* del proceso, no en el comportamiento *deseado*. Los límites de control superior e inferior se desvinculan totalmente de cualesquiera *especificaciones del proceso estipuladas por el fabricante*.

Al investigar la calidad de algún proceso, hay comúnmente dos preguntas importantes que deben plantearse:

1. Con base en el comportamiento actual del proceso, ¿podemos concluir que el proceso está bajo control estadístico?
2. ¿Cumplen con las especificaciones del diseño los bienes y servicios del proceso?

Los métodos de este capítulo se dirigen a responder la primera pregunta, pero no la segunda. Es decir, nos enfocamos en el comportamiento del proceso con el objetivo de determinar si éste se encuentra bajo control estadístico. El hecho de que el proceso dé como resultado bienes y servicios que cumplen con algunas especificaciones establecidas es otro aspecto que no se cubre con los métodos este capítulo. Por ejemplo, la anterior gráfica R de Minitab incluye límites de control superior e inferior de 48.36 y 0, los cuales resultan de los valores muestrales que se incluyen en la tabla 14-1. Las normas gubernamentales consideran aceptables los errores de

los altímetros entre –20 pies y 20 pies, pero tales especificaciones deseadas (o requeridas) no se incluyen en la gráfica de control de R .

Además, debemos comprender con claridad los criterios específicos para determinar si un proceso está bajo control estadístico (es decir, si es estadísticamente estable). Hasta ahora, hemos señalado que un proceso no es estadísticamente estable si su patrón se asemeja a cualquiera de los presentados en la figura 14-2. Este criterio se incluye con algunos otros de la siguiente lista.

Criterios para determinar cuando un proceso no es estadísticamente estable (es decir, cuando está fuera de control estadístico)

1. Hay un patrón, una tendencia o un ciclo que evidentemente no es aleatorio (tales como los que se incluyen en la figura 14-2).
2. Hay un punto que está fuera de la región entre los límites superior e inferior. (Es decir, existe un punto por encima del límite de control superior o por debajo del límite de control inferior).
3. *Regla de la racha de 8:* Existen ocho puntos consecutivos, todos por encima o por debajo de la línea central. (En un proceso estadísticamente estable existe una probabilidad de 0.5 de que un punto esté por encima o por debajo de la línea central, de manera que es muy poco probable que ocho puntos consecutivos aparezcan por encima o por debajo de la línea central).

Únicamente utilizaremos los tres criterios listados antes para establecer una falta de control, pero algunas empresas emplean criterios adicionales como éstos:

- Existen seis puntos consecutivos, todos crecientes o decrecientes.
- Hay 14 puntos consecutivos alternantes que se incrementan o disminuyen (tales como incremento, decremento, incremento, decremento y así sucesivamente).
- Dos de cada tres puntos consecutivos están más allá de los límites de control que se encuentran a dos desviaciones estándar de la línea central.
- Cuatro de cada cinco puntos consecutivos están más allá de los límites de control que están a una desviación estándar de la línea central.



EJEMPLO Control estadístico de procesos Examine la gráfica R del ejemplo anterior, que aparece en la pantalla de Minitab, y determine si la variación del proceso está bajo control estadístico.

SOLUCIÓN Podemos interpretar gráficas de control de R aplicando los tres criterios para establecer una falta de control que listamos anteriormente. Si aplicamos los tres criterios a la gráfica R de la pantalla de resultados de Minitab, concluimos que la variación de este proceso está fuera de control estadístico. No hay ocho puntos consecutivos por encima o por debajo de la línea central, de manera que no se viola la tercera condición, pero las primeras dos condiciones no se cumplen.

1. Existe un patrón, tendencia o ciclo que evidentemente no es aleatorio: de izquierda a derecha existe un patrón de tendencia creciente, como en la figura 14-2a).
2. Hay un punto (el punto a la extrema derecha) que está por arriba del límite de control superior.

INTERPRETACIÓN Concluimos que la variación (no necesariamente la media) del proceso está fuera de control estadístico. Como parece que la variación se incrementa con el tiempo, debe tomarse una medida correctiva de inmediato para fijar la *variación* entre los errores de los altímetros.

Gráfica de control para hacer un seguimiento de medias: la gráfica \bar{x}

Una **gráfica \bar{x}** es una gráfica de las medias muestrales y se utiliza para llevar un control del *centro* en un proceso. Además de graficar las medias muestrales, incluimos una línea central localizada en $\bar{\bar{x}}$, que denota la media de todas las medias muestrales (igual a la media de todos los valores muestrales combinados), así como otra línea para el límite de control inferior y una tercera para el límite de control superior. Utilizando el método común en los negocios y en la industria, la línea central y los límites de control están basados en rangos y no en desviaciones estándar. Véase el ejercicio 18, que incluye una gráfica \bar{x} basada en desviaciones estándar.

Seguimiento de la media del proceso: gráfica de control de \bar{x}

Puntos graficados: medias muestrales

Línea central: $\bar{\bar{x}} =$ media de todas las medias muestrales

Límite de control superior (LCS): $\bar{\bar{x}} + A_2 \bar{R}$ (donde A_2 se encuentra en la tabla 14-2)

Límite de control inferior (LCI): $\bar{\bar{x}} - A_2 \bar{R}$ (donde A_2 se encuentra en la tabla 14-2)



EJEMPLO Fabricación de altímetros para aviones Remítase a los errores de los altímetros en la tabla 14-1. Empleando las muestras de tamaño $n = 4$, reunidas cada día laboral, construya una gráfica de control de \bar{x} . Con base únicamente en la gráfica de control de \bar{x} determine si la media del proceso está bajo control estadístico.

SOLUCIÓN Antes de graficar los 20 puntos correspondientes a los 20 valores de \bar{x} , primero debemos calcular el valor de la línea central y los valores de los límites de control. Obtenemos

$$\bar{\bar{x}} = \frac{2.50 + 2.75 + \cdots + 9.75}{20} = 6.45$$

$$\bar{R} = \frac{19 + 13 + \cdots + 63}{20} = 21.2$$

Si nos remitimos a la tabla 14-2, encontramos que para $n = 4$, $A_2 = 0.729$. Conociendo los valores de $\bar{\bar{x}}$, A_2 , y \bar{R} , ahora podemos evaluar los límites de control.

$$\text{Límite de control superior: } \bar{\bar{x}} + A_2 \bar{R} = 6.45 + (0.729)(21.2) = 21.9$$

$$\text{Límite de control inferior: } \bar{\bar{x}} - A_2 \bar{R} = 6.45 - (0.729)(21.2) = -9.0$$

INTERPRETACIÓN La gráfica de control de \bar{x} resultante sería como la que aparece en la pantalla de Excel. El examen de la gráfica de control indica que la media del proceso está fuera de control estadístico, porque al menos uno de los tres criterios para establecer una falta de control no se satisface. Específicamente,

continúa

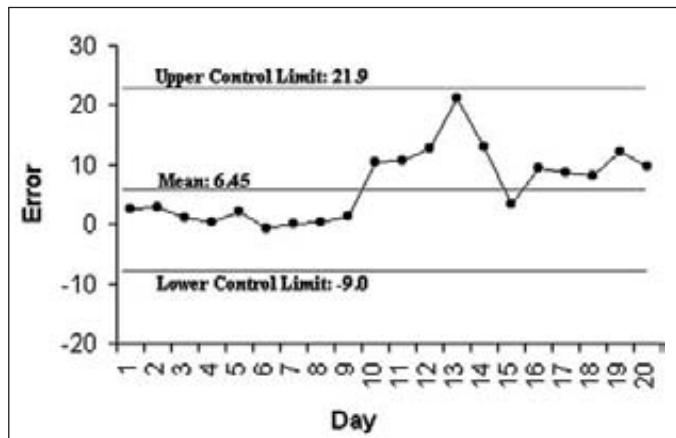


Detección de soborno con gráficas de control

Las gráficas de control se utilizaron para enviar a prisión a una persona que sobornaba a jugadores de jai alai de Florida para que perdieran. (Véase “Using Control Charts to Corroborate Bribery in Jai Alai”, de Charnes y Gitlow, *The American Statistician*, vol. 49, núm. 4). El auditor de una cancha de jai alai notó que cantidades anormalmente grandes de dinero se jugaban en ciertos tipos de apuestas, y que algunos participantes no ganaban tanto como se esperaba cuando se realizaban tales apuestas. Se utilizaron gráficas R y \bar{x} en la Corte como evidencia de patrones sumamente inusuales de apuestas. El examen de las gráficas de control señala claramente puntos que se encuentran muy lejos de límite de control superior, lo que indica que el proceso de apuestas estaba fuera de control estadístico. El especialista en estadística fue capaz de identificar una fecha en la cual la variación assignable parecía detenerse y los fiscales supieron que se trataba de la fecha del arresto del acusado.

el tercer criterio no se satisface porque existen ocho (o más) puntos consecutivos que están por debajo de la línea central. Además, parece existir un patrón de tendencia creciente. Nuevamente, se requieren acciones para corregir el proceso de producción.

Excel



Uso de la tecnología

STATDISK Véase el *Student Laboratory Manual and Workbook* de STATDISK que complementa este libro.

MINITAB **Gráfica de rachas:** Para construir la gráfica de rachas, como la de la figura 14-1, inicie anotando todos los datos muestrales en la columna C1. Seleccione la opción **Stat**, luego **Quality Tools** y después **Run Chart**. En los recuadros indicados, registre C1 en la única columna de variable,

para el tamaño del subgrupo y después haga clic en **OK**.

Gráfica R: Primero anote en la columna C1 los valores muestrales individuales de manera secuencial. Después, seleccione las opciones **Stat**, **Control Charts**, **Variables Chart for Subgroups** y **R**. Ingrese C1 en el recuadro de datos, indique el tamaño muestral en el recuadro del tamaño del subgrupo y haga clic en **R Options** y luego en **estimate**. Seleccione **R bar**. (La selección del estimado R bar hace que la variación de la distribución poblacional se estime con los rangos muestrales en vez de las desviaciones estándar muestrales, que es la que se aplica si no se especifica otra cosa). Haga clic en **OK** dos veces.

Gráfica \bar{x} Primero ingrese los valores muestrales individuales de manera secuencial en la columna C1. Después, seleccione las opciones **Stat**, **Control Charts**, **Variables Charts for Subgroups** y **Xbar**. Ingrese C1 en el recuadro de datos y anote el tamaño de cada

muestra en el recuadro de “subgroup sizes”. Haga clic en **Xbar Options**, luego seleccione **estimate** y elija la opción **Rbar**. Haga clic en **OK** dos veces.

EXCEL Para utilizar el complemento Data Desk XL, haga clic en **DDXL** y seleccione **Process Control**. Proceda a seleccionar el tipo de gráfica que desea. (Primero debe registrar los datos en la columna A y los códigos identificadores de muestra ingresados en la columna B. Por ejemplo, para los datos de la tabla 14-1, ingrese un 1 en la columna B, adyacente a cada valor del día 1, un 2 para cada valor del día 2 y así sucesivamente).

Para utilizar el elemento de construcción de gráficas de Excel en vez de Data Desk XL, vea lo siguiente:

Gráfica de rachas: Anote todos los datos muestrales en la columna A. En la barra del menú principal, haga clic en el ícono **Chart Wizard**, que aparece como una gráfica de barras. Para el tipo de gráfica, seleccione **Line**.

continúa



Para el subtipo de gráfica seleccione la primera gráfica del segundo renglón, luego haga clic en **Next**. Continúe haciendo clic en **Next** y luego en **Finish**. La gráfica se puede editar para incluir rótulos, borrar líneas, etcétera.

Gráfica R Paso 1: Ingrese los datos muestrales en renglones y columnas correspondientes al conjunto de datos. Por ejemplo, ingrese los datos de la tabla 14-1 en cuatro columnas (A, B, C, D) y 20 renglones, como aparecen en la tabla.

Paso 2: Después, cree una columna para los valores de rango con el siguiente procedimiento. Coloque el cursor en la primera celda vacía a la derecha del bloque de datos muestrales, después introduzca esta expresión en el recuadro de la fórmula: = AVERAGE(A1:D1), donde el rango A1:D1 debe modificarse para describir el primer renglón de su conjunto de datos. Después de presionar la tecla **Enter**, debe aparecer el rango para el primer renglón. Utilice el ratón para deslizar

la esquina inferior derecha de esta celda, de manera que la columna completa se llene con los rangos de los diferentes renglones.

Paso 3: Ahora, produzca una gráfica siguiendo el mismo procedimiento descrito para las gráficas de rachas, pero asegúrese de remitirse a la columna de *rangos* cuando ingrese el rango de entrada. Puede insertar la línea central y los límites superior e inferior requeridos editando la gráfica. Haga clic sobre la línea al final de la pantalla, después haga clic y deslice para colocar la línea correctamente.

Gráfica \bar{x} : **Paso 1:** Ingrese los datos muestrales en renglones y columnas correspondientes al conjunto de datos. Por ejemplo, introduzca los datos de la tabla 14-1 en 4 columnas (A, B, C, D) y 20 renglones, tal como aparece en la tabla.

Paso 2: Después, cree una columna para las medias muestrales utilizando el siguiente procedimiento. Coloque el cursor en la primera celda vacía a la derecha del bloque de

datos muestrales, después introduzca esta expresión en el recuadro de la fórmula: = AVERAGE(A1:D1), donde el rango A1:D1 debe modificarse para describir el primer renglón de su conjunto de datos. Después de presionar la tecla **Enter**, debe aparecer la media del primer renglón. Utilice el ratón para deslizar la esquina inferior derecha de esta celda, de manera que la columna completa se llene con las medias de los distintos renglones.

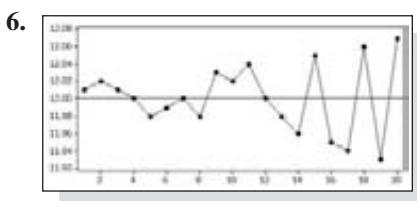
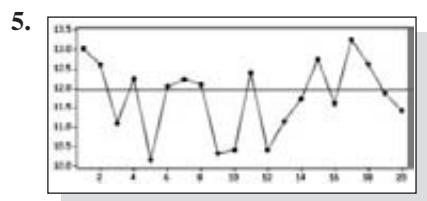
Paso 3: Ahora produzca una gráfica siguiendo el mismo procedimiento descrito para la gráfica de rachas, pero asegúrese de remitirse a la columna de *medias* cuando ingrese el rango de entrada. Puede insertar la línea central y los límites de control superior e inferior requeridos editando la gráfica. Haga clic en la línea de la parte inferior de la pantalla, después haga clic y deslice para colocar la línea correctamente. Esto no es sencillo.

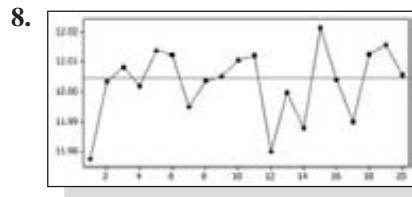
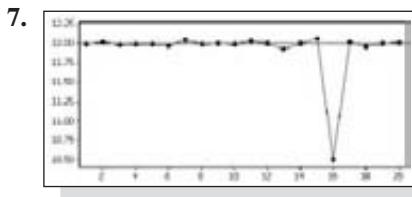
14-2 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

1. **Datos de proceso.** ¿Qué son los *datos de proceso*?
2. **Control estadístico.** ¿Qué significa el hecho de que un proceso esté fuera de control estadístico?
3. **Gráficas de control.** ¿Qué es una gráfica de control? ¿Qué es una gráfica \bar{x} ? ¿Qué diferencia hay entre una gráfica *R* y una gráfica \bar{x} ?
4. **Variación.** ¿Qué diferencia hay entre una variación aleatoria y una variación assignable?

Interpretación de gráficas de rachas. En los ejercicios 5 a 8, examine la gráfica de rachas de un proceso de llenado de latas de 12 oz de bebida de cola y haga lo siguiente:
 a) Determine si el proceso está bajo control estadístico; b) si el proceso no está bajo control estadístico, identifique las razones de ello; c) sin tomar en cuenta el control estadístico, ¿parece que el proceso se comporta como debe hacerlo?





Construcción de gráficas de control para latas de aluminio. Los ejercicios 9 y 10 están basados en las cargas axiales (en libras) de latas del aluminio que tienen un grosor de 0.0109 pulgadas, tal como se listan en el conjunto de datos 15 del apéndice B. La carga axial de una lata es el peso máximo soportado por sus costados y es importante tener una carga axial lo suficientemente alta para que la lata no se aplaste cuando se coloque la tapa superior en su lugar. Los datos provienen de un proceso de fabricación real de acuerdo con un estudiante que utilizó una edición anterior de este libro.

- 9. Gráfica R.** Durante cada día de producción, se seleccionaron al azar siete latas de aluminio con un grosor de 0.0109 pulgadas y se midieron sus cargas axiales. A continuación se presentan los rangos de los diferentes días, pero también se pueden encontrar a partir de los valores en el conjunto de datos 15 del apéndice B. Construya una gráfica R y determine si la variación del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una variación estadísticamente estable.

78	77	31	50	33	38	84	21	38	77	26	78	78
17	83	66	72	79	61	74	64	51	26	41	31	

- 10. Gráfica \bar{x} .** Durante cada día de producción, se seleccionaron al azar siete latas de aluminio con un grosor de 0.0109 pulgadas y se midieron sus cargas axiales. Las medias de los distintos días se presentan abajo, pero también se pueden encontrar a partir de los valores en el conjunto de datos 15 del apéndice B. Construya una gráfica \bar{x} y determine si la media del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una variación estadísticamente estable.

252.7	247.9	270.3	267.0	281.6	269.9	257.7	272.9	273.7	259.1	275.6	262.4	256.0
277.6	264.3	260.1	254.7	278.1	259.7	269.4	266.6	270.9	281.0	271.4	277.3	

Control de la acuñación de monedas de 25 centavos. En los ejercicios 11 a 13, utilice la siguiente información: la Casa de Moneda de Estados Unidos tiene la meta de acuñar monedas de 25 centavos con un peso de 5.670 g, pero cualquier peso entre 5.443 g y 5.897 g se considera aceptable. Se pone en servicio una nueva máquina acuñadora de monedas y se registran los pesos de una moneda seleccionada aleatoriamente cada 12 minutos, durante 20 horas consecutivas. Los resultados se listan en la tabla de la siguiente página.

- 11. Acuñación de monedas: Gráfica de rachas.** Construya una gráfica de rachas para los 100 valores. ¿Parece haber un patrón que sugiera que el proceso no está bajo control estadístico? ¿Cuáles son las implicaciones prácticas de la gráfica de rachas?
- 12. Acuñación de monedas: Gráfica R.** Construya una gráfica R y determine si la variación del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una variación estadísticamente estable.
- 13. Acuñación de monedas: Gráfica \bar{x} .** Construya una gráfica \bar{x} y determine si la media del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una media estadísticamente estable. ¿Necesita este proceso una acción correctiva?

Pesos (en gramos de monedas acuñadas)

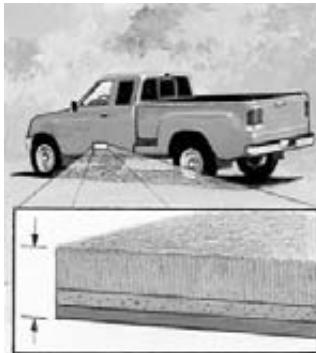
Hora	Peso (g)					\bar{x}	s	Rango
1	5.639	5.636	5.679	5.637	5.691	5.6564	0.0265	0.055
2	5.655	5.641	5.626	5.668	5.679	5.6538	0.0211	0.053
3	5.682	5.704	5.725	5.661	5.721	5.6986	0.0270	0.064
4	5.675	5.648	5.622	5.669	5.585	5.6398	0.0370	0.090
5	5.690	5.636	5.715	5.694	5.709	5.6888	0.0313	0.079
6	5.641	5.571	5.600	5.665	5.676	5.6306	0.0443	0.105
7	5.503	5.601	5.706	5.624	5.620	5.6108	0.0725	0.203
8	5.669	5.589	5.606	5.685	5.556	5.6210	0.0545	0.129
9	5.668	5.749	5.762	5.778	5.672	5.7258	0.0520	0.110
10	5.693	5.690	5.666	5.563	5.668	5.6560	0.0534	0.130
11	5.449	5.464	5.732	5.619	5.673	5.5874	0.1261	0.283
12	5.763	5.704	5.656	5.778	5.703	5.7208	0.0496	0.122
13	5.679	5.810	5.608	5.635	5.577	5.6618	0.0909	0.233
14	5.389	5.916	5.985	5.580	5.935	5.7610	0.2625	0.596
15	5.747	6.188	5.615	5.622	5.510	5.7364	0.2661	0.678
16	5.768	5.153	5.528	5.700	6.131	5.6560	0.3569	0.978
17	5.688	5.481	6.058	5.940	5.059	5.6452	0.3968	0.999
18	6.065	6.282	6.097	5.948	5.624	6.0032	0.2435	0.658
19	5.463	5.876	5.905	5.801	5.847	5.7784	0.1804	0.442
20	5.682	5.475	6.144	6.260	6.760	6.0642	0.5055	1.285

Conjunto de datos del apéndice B: construcción de gráficas de control para la lluvia en Boston. En los ejercicios 14 a 16, remítase a las cantidades diarias de lluvia en Boston en un año que aparecen en el conjunto de datos 10 del apéndice B. Omita el último dato para el miércoles, de manera que cada día de la semana tenga exactamente 52 valores.

14. **Lluvia en Boston: construcción de una gráfica de rachas.** Utilice únicamente las 52 cantidades de lluvia de los lunes y construya una gráfica de rachas. ¿Parece que el proceso está bajo control estadístico?
15. **Lluvia en Boston: construcción de una gráfica R.** Utilice las 52 muestras, con siete valores cada una para construir una gráfica R y determine si la variación del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una variación estadísticamente estable.
16. **Lluvia en Boston: construcción de una gráfica \bar{x} .** Utilice las 52 muestras, con siete valores cada una para construir una gráfica \bar{x} y determine si la media del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una media estadísticamente estable.

14-2 MÁS ALLÁ DE LO BÁSICO

17. **Construcción de una gráfica s .** En esta sección describimos las gráficas de control de R y \bar{x} basadas en rangos. Las gráficas de control para llevar control de la variación y el centro (media) también pueden estar basadas en desviaciones estándar. Una *gráfica s* para llevar control de la variación se construye graficando desviaciones estándar muestrales, con una línea central en \bar{s} (la media de las desviaciones estándar muestrales) y los límites de control en $B_4\bar{s}$ y $B_3\bar{s}$, donde B_4 y B_3 se obtienen de la tabla 14-2. Construya una gráfica s para los datos de la tabla 14-1. Compare el resultado con la gráfica R dada en esta sección.



Control de calidad en Perstorp

Perstorp Components, Inc. utiliza una computadora que genera automáticamente gráficas de control para verificar el grosor del aislamiento para el piso que fabrica para las camionetas Ford Ranger y Jeep Grand Cherokee. El costo de la computadora de \$20,000 se pagó con los ahorros de \$40,000 del primer año de operaciones, que se habían empleado para generar gráficas de control manuales que aseguraban que el grosor del aislamiento cumpliera con las especificaciones establecidas entre 2.912 y 2.988 mm. Por medio de gráficas de control y de otros métodos de control de calidad, Perstorp redujo sus mermas en más de dos tercios.

- 18. Construcción de una gráfica \bar{x} basada en desviaciones estándar.** Una gráfica \bar{x} basada en desviaciones estándar (y no en rangos) se construye graficando las medias muestrales con una línea central en $\bar{\bar{x}}$ y los límites de control en $\bar{\bar{x}} + A_3\bar{s}$ y $\bar{\bar{x}} - A_3\bar{s}$, donde A_3 se obtiene de la tabla 14-2 y \bar{s} es la media de las desviaciones estándar muestrales. Utilice los datos de la tabla 14-1 para construir una gráfica \bar{x} basada en desviaciones estándar. Compare el resultado con la gráfica \bar{x} basada en rangos muestrales (presentada en esta sección).

14-3 Gráficas de control para atributos

Concepto clave En esta sección se presenta un método para construir una gráfica de control con el fin de llevar control de la proporción p de algún *atributo* y así saber si el servicio o artículo manufacturado está defectuoso o es discordante. (Un bien o servicio es discordante si no cumple con ciertas especificaciones o requisitos; en ocasiones los bienes discordantes se desechan, se reparan o se denominan “de segunda” y se venden a precios más bajos). La gráfica de control se interpreta usando los mismos tres criterios de la sección 14-2 para determinar si el proceso es estadísticamente estable.

En la sección 14-2 estudiamos gráficas de control para datos *cuantitativos*, pero en esta sección se describe la construcción de gráficas de control para datos *cualitativos*. Como en la sección 14-2, seleccionamos muestras de tamaño n a intervalos de tiempo regulares y dibujamos los puntos en una gráfica secuencial con una línea central y límites de control. (Hay formas de manejar muestras con tamaños diferentes, pero no las estudiaremos aquí).

Definición

La **gráfica de control de p** (o **gráfica p**) es una gráfica de que se dibuja en secuencia en función del paso del tiempo y que incluye una línea central, un límite de control inferior (LCI) y un límite de control superior (LCS).

La notación y los valores de la gráfica de control son los siguientes (donde el atributo de “defectuoso” se puede reemplazar por cualquier otro atributo relevante).

Notación

$$\begin{aligned}\bar{p} &= \text{estimado agrupado de la proporción de artículos defectuosos en el proceso} \\ &= \frac{\text{número total de defectos encontrados en todos los artículos muestreados}}{\text{número total de artículos muestreados}}\end{aligned}$$

$$\begin{aligned}\bar{q} &= \text{estimado agrupado de la proporción de artículos del proceso que } \text{no} \text{ son defectuosos} \\ &= 1 - \bar{p}\end{aligned}$$

$$n = \text{tamaño de cada muestra (no el número de muestras)}$$

Gráfica de control de p

Línea central: \bar{p}

$$\text{Límite de control superior: } \bar{p} + 3\sqrt{\frac{\bar{p}\bar{q}}{n}}$$

$$\text{Límite de control inferior: } \bar{p} - 3\sqrt{\frac{\bar{p}\bar{q}}{n}}$$

(Si el cálculo del límite de control inferior da como resultado un valor negativo, utilice 0 en su lugar. Si el cálculo del límite de control superior excede a 1, utilice 1 en su lugar).



Utilizamos \bar{p} para la línea central porque es el mejor estimado de la proporción de defectos del proceso. Las expresiones de los límites de control corresponden a límites de un intervalo de confianza del 99.7%, como se describió en la sección 7-2.

**PROBLEMA
14
DEL CAPÍTULO**

EJEMPLO Altímetros defectuosos para aviones El problema del capítulo describe el proceso de fabricación de altímetros para aviones. La sección 14-2 incluye ejemplos de gráficas de control para verificar errores en las lecturas de los altímetros. Se considera que un altímetro está defectuoso si no es posible calibrarlo o corregirlo para que proporcione lecturas exactas (que no difieran más de 20 pies de la altitud verdadera). La Altigauge Manufacturing Company produce altímetros en lotes de 100, y cada altímetro se prueba para determinar si está o no defectuoso. A continuación se lista el número de altímetros defectuosos en lotes sucesivos de 100. Construya una gráfica de control para la proporción p de altímetros defectuosos y determine si el proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios de control se aplica.

Defectos: 2 0 1 3 1 2 2 4 3 5 3 7

SOLUCIÓN La línea central de la gráfica de control se localiza en el valor de \bar{p} :

$$\begin{aligned}\bar{p} &= \frac{\text{número total de defectos de todas las muestras combinadas}}{\text{número total de altímetros muestreados}} \\ &= \frac{2 + 0 + 1 + \dots + 7}{12 \cdot 100} = \frac{33}{1200} = 0.0275\end{aligned}$$

Puesto que $\bar{p} = 0.0275$, se infiere que $\bar{q} = 1 - \bar{p} = 0.9725$. Al utilizar $\bar{p} = 0.0275$, $\bar{q} = 0.9725$ y $n = 100$, calculamos los límites de control de la siguiente manera:

Límite de control superior:

$$\bar{p} + 3\sqrt{\frac{\bar{p}\bar{q}}{n}} = 0.0275 + 3\sqrt{\frac{(0.0275)(0.9725)}{100}} = 0.0766$$

Límite de control inferior:

$$\bar{p} - 3\sqrt{\frac{\bar{p}\bar{q}}{n}} = 0.0275 - 3\sqrt{\frac{(0.0275)(0.9725)}{100}} = -0.0216$$

Seis Sigma en la industria

Seis Sigma es el término utilizado en la industria para describir un proceso que da por resultado una tasa de no más de 3.4 defectos en un millón. La referencia a Seis Sigma sugiere seis desviaciones estándar a partir del centro de una distribución normal, pero el supuesto de un proceso perfectamente estable se reemplaza por el supuesto de un proceso que cambia ligeramente, de manera que la tasa de defectos no es mayor que 3 o 4 por millón.

Los programas Seis Sigma, iniciados alrededor de 1985 en Motorola, ahora intentan mejorar la calidad e incrementar las ganancias al reducir la variación de los procesos. Motorola ahorró más de \$940 millones en tres años. Allied Signal reportó ahorros de \$1,500 millones. GE, Polaroid, Ford, Honeywell, Sony y Texas Instruments son otras compañías grandes que han adoptado la meta Seis Sigma.

continúa

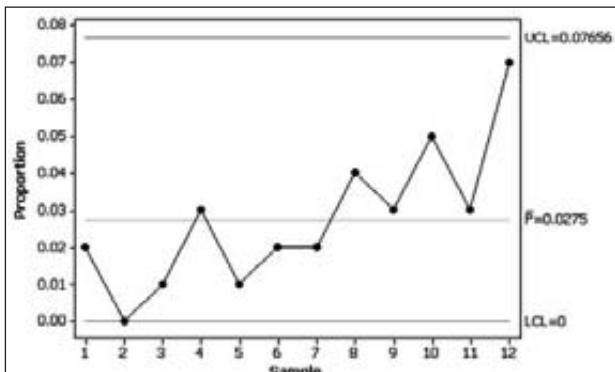
LA ESTADÍSTICA EN LAS NOTICIAS

El alto costo de la baja calidad

La Federal Drug Administration (FDA) recientemente llegó a un acuerdo en el que una compañía farmacéutica, la Schering-Plough Corporation, pagaría la cantidad récord de \$500 millones por no lograr corregir problemas en la producción de fármacos. Según un artículo del *New York Times*, de Melody Petersen, “algunos de los problemas se relacionan con la falta de controles que identifican medicamentos defectuosos, mientras otros provienen de equipos demasiado viejos. Tales problemas están relacionados con unos 200 medicamentos, incluyendo el Claritin, el fármaco para alergias que es el producto de mayor venta de Schering”.

Puesto que el límite de control inferior es menor que 0, usamos el 0 en su lugar. Una vez que calculamos los valores de la línea central y los límites de control, procedemos a graficar la proporción de altímetros defectuosos. La gráfica de control de p de Minitab se presenta en la siguiente pantalla.

Minitab



INTERPRETACIÓN Podemos interpretar la gráfica de control de p considerando los tres criterios para establecer una falta de control que se listan en la sección 14-2. Con esos criterios podemos concluir que este proceso está fuera de control estadístico por la siguiente razón: parece existir una tendencia creciente. La empresa debe tomar acciones inmediatas para corregir la proporción creciente de defectos.

Uso de la tecnología

MINITAB Ingrese los números de defectos (o artículos con algún atributo particular) en la columna C1. Seleccione la opción **Stat**, luego **Control Charts**, **Attributes Charts** y después **P**. Ingrese C1 en el recuadro identificado como variable e ingrese el tamaño de las muestras en el recuadro identificado como tamaño del subgrupo, después haga clic en **OK**.



EXCEL **Uso de DDXL:** Para utilizar el complemento DDXL, primero ingrese los números de defectos o éxitos en la columna A y registre los tamaños muestrales en la columna B. Para el ejemplo de esta sección, los primeros tres artículos se ingresarían en la hoja de cálculo de Excel como se indica abajo.

	A	B
1	2	100
2	0	100
3	1	100

Haga clic en **DDXL**, seleccione **Process control** y después **Summ Prop Control Chart** (para la gráfica de control del resumen de proporciones). Debe aparecer un cuadro de diálogo. Haga clic en el icono del lápiz de “Success Variable” e indique el rango de valores para la columna A, tal como A1:A12. Haga clic en el icono del lápiz de “Totals Variable” e indique

el rango de valores para la columna B, tal como B1:B12. Haga clic en **OK**. Ahora haga clic en la barra **Open Control Chart** y aparecerá la gráfica de control.

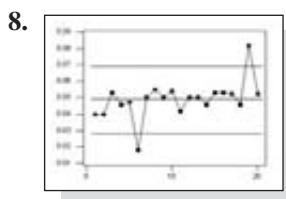
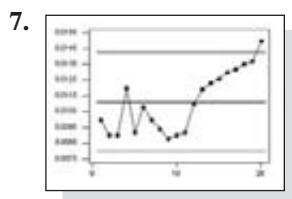
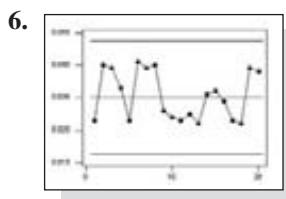
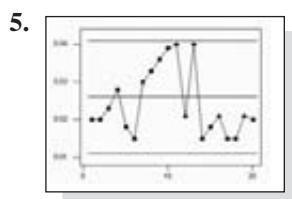
Uso del Chart Wizard de Excel: Anote las proporciones muestrales en la columna A. Haga clic en el ícono de **Chart Wizard**, que tiene la apariencia de una gráfica de barras. Para el tipo de gráfica, seleccione **Line**. Para el subtipo de gráfica, seleccione la primera gráfica del segundo renglón y luego haga clic en **Next**. Continúe haciendo clic en **Next** y luego en **Finish**. La gráfica puede editarse para incluir rótulos, borrar líneas, etcétera. Usted puede insertar la línea central y los límites de control inferior y superior requeridos, editando la gráfica. Haga clic sobre la línea de la parte inferior de la pantalla, luego haga clic y coloque la línea en la posición correcta.

14-3 DESTREZAS Y CONCEPTOS BÁSICOS

Conocimientos estadísticos y pensamiento crítico

- Gráfica p .** ¿Qué es una gráfica p y cuál es su finalidad?
- Especificaciones de contrato.** La empresa Paper Chase, que vende artículos para oficina, necesita un proveedor de bolígrafos para mantener un proceso de producción con una tasa de defectos menor al 2%. Si el proveedor usa una gráfica p para determinar que el proceso está bajo control estadístico, ¿esto indica que la tasa de defectos es menor al 2%?
- Límites de control.** Un proceso se analiza con los métodos de esta sección y se descubre que los límites de control superior e inferior son 0.250 y -0.50 , respectivamente. ¿Qué controles superior e inferior se usan para construir la gráfica de control?
- Interpretación de una gráfica de control.** Al llevar control del proceso de producción de altímetros, una empresa descubre que el proceso está fuera de control estadístico porque hay un patrón decreciente de defectos que no es aleatorio. ¿Debe corregirse el patrón decreciente? ¿Qué debe hacer la empresa?

Determinar si un proceso está bajo control. En los ejercicios 5 a 8, examine la gráfica de control de p dada y determine si el proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control se aplica.



Construcción de gráficas de control de p . En los ejercicios 9 a 12, utilice los datos de proceso para construir una gráfica de control de p . En cada caso, utilice los tres criterios para establecer una falta de control listados en la sección 14-2 y determine si el proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control se aplica.

- Gráfica p para tasa de nacimientos.** En cada uno de 10 años consecutivos recientes se seleccionaron al azar 10,000 personas y se registró el número de nacimientos generados por ellos. A continuación se listan los resultados (según datos del National Center for Health Statistics). ¿Cómo se podrían explicar los resultados?

Nacimientos: 155 152 148 147 145 146 145 144 141 139

- Gráfica p para tasa de divorcios.** En cada uno de 10 años consecutivos recientes se seleccionaron al azar 10,000 personas y se registró el número de divorcios protagonizados por ellos. A continuación se listan los resultados (según datos del National Center for Health Statistics). ¿Cómo se podrían explicar los resultados?

Divorcios: 48 46 46 44 43 43 42 41 42 40

- 11. Gráfica p para acumuladores defectuosos de automóvil.** Los acumuladores de automóvil defectuosos constituyen un problema porque pueden causar molestias a los conductores. Se considera que un acumulador está defectuoso si falla antes de que venza su garantía. Los defectos se identifican cuando los acumuladores son devueltos bajo el programa de garantía. La corporación Powerco Battery fabrica acumuladores de automóvil en lotes de 1000; a continuación se lista el número de defectos de 12 lotes consecutivos. ¿El proceso de fabricación requiere de corrección?

Defectos: 8 6 5 9 10 7 7 4 6 11 5 8

- 12. Encuesta.** Cuando la organización de encuestas Infopop realiza una encuesta telefónica, una llamada se considera un defecto si la persona no está disponible o se rehúsa a responder preguntas. Para una encuesta específica sobre preferencias de consumidores se llama a 200 personas diariamente; a continuación se lista el número de defectos. ¿El proceso de llamadas requiere de una acción correctiva?

Defectos: 92 83 85 87 98 108 96 115 121 125 112 127 109 131 130

14-3 MÁS ALLÁ DE LO BÁSICO

- 13. Gráfica p para la lluvia en Boston.** Remítase a las cantidades de lluvia en Boston en el conjunto de datos 10 del apéndice B. Para cada una de las 52 semanas, permita que la proporción muestral sea la proporción de días que llovió. (Borre el valor 53° del miércoles). Durante la primera semana, por ejemplo, la proporción muestral es $3/7 = 0.429$. ¿Los datos representan un proceso estadísticamente estable?
- 14. Construcción de una gráfica np .** Una variante de la gráfica de control de p es la gráfica np , en la cual se grafica el *número real* de defectos, en vez de las *proporciones* de defectos. La gráfica np tiene un valor de línea central $n\bar{p}$, y los límites de control tienen valores de $n\bar{p} + 3\sqrt{n\bar{p}\bar{q}}$ y $n\bar{p} - 3\sqrt{n\bar{p}\bar{q}}$. La gráfica p y la gráfica np difieren únicamente en la escala de valores empleada en el eje vertical. Construya una gráfica np para el ejemplo presentado en esta sección. Compare el resultado con la gráfica de control de p obtenida en esta sección.

Repaso

En el capítulo 2 identificamos características importantes de los datos: centro, variación, distribución, valores extremos y el patrón de cambio con el paso del tiempo. Este capítulo se enfoca en el patrón de cambio de los datos a lo largo del tiempo. Los datos de proceso se definieron como datos ordenados de acuerdo con alguna secuencia temporal; datos como éstos pueden analizarse con gráficas de rachas y gráficas de control. Las gráficas de control tienen una línea central, un límite de control superior y un límite de control inferior. Un proceso es estadísticamente estable (o está bajo control estadístico) sólo si tiene una variación natural sin patrones, ciclos o puntos poco comunes. Las decisiones sobre la estabilidad estadística están basadas en la forma en que el proceso se está comportando en realidad y no en la forma en que nos gustaría que se comportara de acuerdo con factores como las especificaciones de fabricación. Se describieron las siguientes gráficas:

- *Gráfica de rachas:* gráfica secuencial de datos *individuales* a lo largo del tiempo
- *Gráfica R:* gráfica de control que utiliza rangos en un intento de dar seguimiento a la *variación* en un proceso
- *Gráfica \bar{x} :* gráfica de control utilizada para determinar si la *media* del proceso está bajo control estadístico
- *Gráfica p :* gráfica de control utilizada para dar seguimiento a la proporción de algún *atributo* del proceso, tal como si los artículos son defectuosos.

Conocimientos estadísticos y pensamiento crítico

- Patrón temporal.** ¿Por qué es importante llevar control de un patrón de cambio de datos a lo largo del tiempo?
- Control estadístico de procesos.** El título de este capítulo es “Control estadístico de procesos”. ¿Qué significa esto?
- Pérdida de control de fabricación.** ¿Cuál sería una posible consecuencia adversa de una planta embotelladora de Pepsi que tiene un proceso sin verificación?
- Gráficas de control.** Al llevar control de los tiempos que tardan los técnicos en reparar computadoras, ¿por qué es importante usar una gráfica \bar{x} en conjunto con una gráfica R ?

Ejercicios de repaso

Construcción de gráficas de control para el consumo de electricidad. La siguiente tabla lista cantidades de consumo eléctrico (en kWh) de la casa del autor, tal como aparecen en el conjunto de datos 9 del apéndice B. Utilice la tabla para los ejercicios 1 a 3.

Consumo de electricidad (kWh)			
Año 1: primer semestre	3375	2661	2073
Año 1: segundo semestre	2579	2858	2296
Año 2: primer semestre	2812	2433	2266
Año 2: segundo semestre	3128	3286	2749
Año 3: primer semestre	3427	578	3792

- Gráfica de rachas.** Construya una gráfica de rachas para los 15 valores. ¿Parece haber un patrón que sugiera que el proceso no está bajo control estadístico?
- Gráfica R .** Utilice subgrupos de tamaño $n = 3$, correspondientes a los renglones de la tabla, construya una gráfica R y determine si la variación del proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una variación estadísticamente estable.
- Gráfica \bar{x}** Utilice subgrupos de tamaño $n = 3$, correspondientes a los renglones de la tabla, construya una gráfica \bar{x} y determine si la media del proceso está bajo control estadístico. ¿Parece que el proceso está bajo control estadístico? Si no es así, identifique cuál de los tres criterios para establecer una falta de control conduce al rechazo de una variación estadísticamente estable.
- Construcción de una gráfica de control para enfermedades infecciosas.** En cada uno de 13 años consecutivos recientes se seleccionaron al azar 100,000 adultos de 65 años de edad o mayores, y entre ellos se registró el número de muertes por enfermedades infecciosas; los resultados se presentan abajo (según datos de “Trends in Infectious Diseases Mortality in the United States”, de Pinner *et al.*, *Journal of the American Medical Association*, vol. 275, núm. 3). Construya una gráfica de control apropiada y determine si el proceso está bajo control estadístico. Si no es así, identifique cuáles criterios condujeron al rechazo de la estabilidad estadística.

Número de muertes: 270 264 250 278 302 334 348 347 377 357 362 351 343

- Gráfica de control de defectos.** La empresa Medassist Pharmaceutical fabrica tabletas de aspirina. Cada día se seleccionan al azar 100 tabletas y se ponen a prueba. Se considera que una tableta es defectuosa si presenta deformidades físicas evidentes o si el contenido de aspirina es menor que 490 mg o mayor que 510 mg. A continuación se

listan los números de defectos de días consecutivos. Construya una gráfica de control apropiada y determine si el proceso está bajo control estadístico. Si no es así, identifique cuáles criterios condujeron al rechazo de la estabilidad estadística.

Defectos: 4 2 2 3 5 2 9 12 1 11 3 2 12 14

Ejercicios de repaso acumulativo

- Gráfica de control para cinturones de seguridad defectuosos.** La Flint Accessory Corporation fabrica cinturones de seguridad para automóvil. Las especificaciones federales exigen que la banda tenga una fuerza de rompimiento de por lo menos 5000 libras. Durante cada semana de producción, se seleccionan al azar 200 cinturones y se prueba su fuerza de rompimiento; un cinturón se considera defectuoso si se rompe antes de alcanzar la fuerza de 5000 libras. A continuación se listan los números de defectos para una frecuencia de 10 semanas. Utilice una gráfica de control de p para verificar si el proceso está bajo control estadístico. Si no está bajo control, explique por qué.

6 4 12 3 7 2 3 5 4 2

- Intervalo de confianza para cinturones de seguridad defectuosos.** Remítase a los datos del ejercicio 1 y, con todos los datos de los 2000 cinturones de seguridad que se probaron, construya un intervalo de confianza del 95% para la proporción de defectos.
- Prueba de hipótesis para cinturones de seguridad defectuosos.** Remítase a los datos del ejercicio 1 y, con todos los datos de los 2000 cinturones de seguridad que se probaron, utilice un nivel de significancia de 0.05 para probar la aseveración de que la tasa de defectos es mayor que el 1%.
- Uso de la probabilidad en gráficas de control.** Al interpretar gráficas de control, uno de los tres criterios para establecer la falta de control es la presencia de ocho puntos consecutivos por arriba o por debajo de la línea central. En un proceso estadísticamente estable, hay una probabilidad de 0.5 de que un punto esté por arriba de la línea central y una probabilidad de 0.5 de que un punto esté por debajo de la línea central. Para resolver lo siguiente, suponga que los valores muestrales son independientes y que el proceso es estadísticamente estable.
 - Calcule la probabilidad de que, cuando se seleccionen ocho puntos al azar, se ubiquen por arriba de la línea central.
 - Calcule la probabilidad de que, cuando se seleccionen ocho puntos al azar, se ubiquen por debajo de la línea central.
 - Calcule la probabilidad de que, cuando se seleccionen ocho puntos al azar, se ubiquen todos por arriba o todos por debajo de la línea central.

Actividades de cooperación en equipo

- Actividad fuera de clase** Reúna su propio conjunto de datos de proceso y analícelos utilizando los métodos de esta sección. Sería ideal que pudiera reunir datos de un proceso real de fabricación, aunque esto tal vez resulte difícil. Si es así, considere realizar una simulación o remitirse a datos publicados, tales como los que se encuentran en un almanaque. He aquí algunas sugerencias:

- Lance cinco tiros libres de básquetbol (o lance cinco papeles arrugados en un bote de basura) y registre el número de canastas anotadas; después repita este

procedimiento 20 veces y utilice una gráfica p para probar la estabilidad estadística de la proporción de canastas anotadas.

- Puede medir su pulso contando el número de latidos de su corazón en un minuto. Mida su pulso cuatro veces cada hora, durante varias horas, después construya una gráfica de control apropiada. ¿Qué factores contribuyen a la variación aleatoria? ¿Y a la variación assignable?
- Busque periódicos de las 12 últimas semanas y registre el cierre del promedio industrial Dow Jones

(DJIA). Utilice gráficas de rachas y de control para explorar la estabilidad estadística del DJIA. Identifique al menos una consecuencia práctica de que este proceso sea estadísticamente estable e identifique al menos una consecuencia práctica de que este proceso esté fuera de control estadístico.

- Calcule la tasa de matrimonios por 10,000 habitantes de varios años. (Véase el *Information Please Almanac* o el *Statistical Abstract of the United States*). Suponga que se seleccionaron 10,000 personas cada año y que se entrevistaron para determinar si estaban casadas. Utilice en una gráfica p para probar la estabilidad estadística de la tasa de matrimonios. (Otras tasas posibles: muertes, muertes en accidentes, crímenes).

Imprima una copia de los resultados del programa de cómputo y escriba un reporte que resuma sus conclusiones.

2. Actividad en clase Si el profesor registra el número de ausencias en cada clase, grupos de tres o cuatro estudiantes pueden analizar las cifras para conocer su estabilidad estadística y hacer recomendaciones con base en las conclusiones.

3. Actividad fuera de clase Realice una investigación para identificar el *experimento del embudo de Deming*, después utilice un embudo y canicas para reunir datos respecto a las diferentes reglas con el fin de ajustar la ubicación del embudo. Construya gráficas de control apropiadas para las diferentes reglas del ajuste del embudo. ¿Qué ilustra el experimento del embudo? ¿Qué concluye?

Proyecto tecnológico

- a. Simule el siguiente proceso durante 20 días: cada día se fabrican 200 calculadoras, con una tasa del 5% de unidades defectuosas; la proporción de defectos se registra durante cada uno de los 20 días. Las calculadoras de un día se simulan generando aleatoriamente 200 números, donde cada número está entre 1 y 100. Considere que un resultado de 1, 2, 3, 4 o 5 es un defecto, mientras que uno del 6 al 100 es aceptable. Esto corresponde a una tasa de defectos del 5%. (Véase más adelante las instrucciones para el uso de herramientas tecnológicas).
- b. Construya una gráfica p para la proporción de calculadoras defectuosas y determine si el proceso está bajo control estadístico. Como sabemos que el proceso es en realidad estable, con $p = 0.05$, la conclusión de que no es estable sería un error tipo I; es decir, tendríamos una señal falsa positiva, lo que nos haría suponer que el proceso necesita ajustarse, cuando en realidad se debe dejar como está.
- c. El resultado del inciso a) es una simulación de 20 días. Ahora simule otros 10 días de fabricación de calculadoras, pero modifique estos últimos días de manera que la tasa de defectos sea del 10% en vez del 5%.
- d. Combine los datos generados en los incisos a) y c) para representar un total de 30 días de resultados muestrales. Construya una gráfica p para este conjunto de datos combinado. ¿Está el proceso fuera de control? Si concluimos que el proceso no está fuera de control, estaríamos cometiendo un error tipo II; es decir, pensaríamos que el proceso está bien cuando en realidad debería arreglarse o ajustarse para corregir el cambio a una tasa de defectos del 10%.

Instrucciones para el uso de la tecnología para el inciso a)

STATDISK Seleccione **Data**, **Uniform Generator** y proceda a generar 200 valores con un mínimo de 1 y un máximo de 100. Copie los valores en la ventana de datos y luego ordénelos con el botón **Data Tools**. Repita este procedimiento hasta obtener los 20 días simulados.

MINITAB Seleccione **Calc**, **Random Data** y luego **Integer**. Anote 200 en el recuadro para el número de renglones de datos, ingrese C1 como la columna para almacenar los datos, 1 para el valor mínimo y 100 para el valor máximo. Repita este procedimiento hasta obtener los resultados de 20 días simulados.

EXCEL Haga clic en el ícono **fx** en la barra del menú principal, luego seleccione la categoría de la función **Math & Trig**, seguida por **RANDBETWEEN**. En el cuadro de diálogo anote 1 para la parte más baja y 100 para la parte más alta. Debe aparecer un valor aleatorio en el primer renglón de la columna A. Utilice el ratón para hacer clic y deslizar la esquina inferior derecha de esa celda, después baje la celda para cubrir los primeros 200 renglones de la columna A. Cuando suelte el botón del ratón, la columna A debe contener 200 números aleatorios. También puede deslizar la esquina inferior derecha de la celda inferior moviendo el ratón a la derecha, de manera que obtenga 20 columnas de 200 números cada una. Las diferentes columnas representan los distintos días de fabricación.

TI-83/84 PLUS Presione la tecla **MATH**, seleccione **PRB**, luego seleccione el quinto elemento del menú, **randInt**, y proceda a ingresar 1, 100, 200; después presione la tecla **ENTER**. Presione **STO** y **L1** para almacenar los datos en la lista L1. Después de registrar el número de defectos, repita este procedimiento hasta obtener resultados para 20 días simulados.

De los datos a la decisión

Pensamiento crítico: ¿Están bajo control estadístico las cargas axiales?

¿Está procediendo como debe el proceso de fabricación de latas?

En los ejercicios 9 y 10 de la sección 14-2 se utilizaron datos de proceso de una compañía de Nueva York que fabrica latas de aluminio con un grosor de 0.0109 pulgadas para un distribuidor importante de bebidas. Remítase al conjunto de datos 15 del apéndice B y realice un análisis de los datos de proceso para las

latas que tienen 0.0111 pulgadas de grosor. Los valores en el conjunto de datos son las cargas axiales medidas de las latas, y las tapas superiores se colocan en su lugar con presiones que varían entre 158 y 165 libras.

superiores. También compare el comportamiento de las latas de 0.0111 pulgadas con el comportamiento de las latas de 0.0109 pulgadas y recomiende el grosor que debe utilizarse.



Gráficas de control

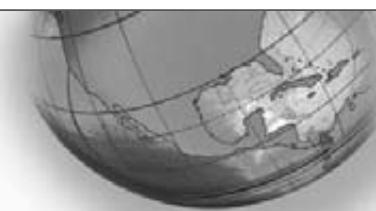
Este capítulo describe diferentes técnicas de graficación utilizadas para resumir y estudiar datos asociados con un proceso, junto con métodos para analizar la estabilidad de ese proceso. Con excepción de la gráfica de rachas, no se requieren datos individuales para construir una gráfica. Por ejemplo, la gráfica R se elabora a partir de rangos muestrales, mientras que la gráfica p se basa en proporciones muestrales. Éste es un punto importante, ya que los datos reunidos de fuentes terciarias suelen presentarse en

términos de estadísticos resumidos. Visite el sitio de Internet de *Estadística Elemental*:

<http://www.pearsoneducacion.net/Triola>

Localice el proyecto de Internet que se refiere a gráficas de control. Ahí será conducido a conjuntos de datos y fuentes de información que utilizará en la construcción de gráficas de control. Se le pedirá que interprete y discuta las tendencias en los procesos subyacentes a partir de las gráficas resultantes.

Proyecto de Internet



La estadística en el trabajo

“Una persona que sabe de estadística y puede explicarla a alguien que no sabe infunde cierto respeto”.



Dan O'Toole

Ejecutivo de cuenta: A. C. Nielsen

En su trabajo en el Advanced Analytics Group en A.C. Nielsen, Dan se encarga de idear soluciones estadísticas para ayudar a que clientes como Polaroid, Ocean Spray y Gillette comprendan cuáles de sus iniciativas de marketing producen mayores ganancias. Dan tiene una maestría en Economía y negocios del Bentley College.

¿Qué conceptos de estadística utiliza?

Trabajo con análisis tan sencillos como la correlación y las pruebas generales de significancia, hasta análisis complejos como la regresión múltiple, el análisis factorial, el análisis de correspondencia y el análisis de conglomerados.

¿Cómo utiliza la estadística en el trabajo?

Mi trabajo consiste en descubrir problemas de los clientes y después encontrar si podemos aplicar alguna de nuestras técnicas estadísticas a su problema específico. Si una técnica no ayuda a un cliente, es necesario saberlo. Un ejemplo de cómo utilizo la estadística es el siguiente: un cliente diría, yo vendo el producto “X”, ya sea jugo, pan o una cámara fotográfica. Ellos tal vez controlen el 20% del mercado y recurren a nosotros para ver si pueden incrementar sus ventas en el mercado bajando su precio. Mi trabajo consistiría en diseñar un estudio para analizar ese asunto. Para hacerlo debo diseñar un estudio que tome en cuenta todo lo que afecte las ventas de un producto. Empleando técnicas como la regresión, si soy capaz de crear un modelo con buena significancia, aislaré influencias específicas sobre las ventas y ofreceré recomendaciones. Tienen que incluirse aspectos como la distribución estacional, así como cualquier esfuerzo de comercialización que se haya presentado. Además, tengo que tomar en cuenta el precio de productos complementarios (la mantequilla es un complemento del pan, como la película lo es para la cámara) y también

productos competitivos. Por ejemplo, el pan compite con los “muffins” ingleses (como yo mismo he constatado).

¿Cree que las personas que buscan un empleo son mejor evaluadas si tienen estudios de estadística?

Claro que sí. Una persona que sabe de estadística y puede explicarla a alguien que no sabe infunde cierto respeto (porque eso significa que realmente sabe y que no sólo recita las páginas de un libro de texto). Casi cualquier empleo utiliza estadística (particularmente correlaciones y regresiones). En las empresas la gente a menudo dice frases como “Verifica si se correlacionan”.

¿El uso que usted hace de la probabilidad y de la estadística está aumentando, disminuyendo o permanece estable?

Definitivamente está aumentando. En este negocio de consultoría uno se enfrenta constantemente al reto de aprender una nueva técnica o de recurrir a una antigua técnica para mejorarla. Además, dado que constantemente se lanzan nuevos productos, nuestra comprensión de la estadística se debe incrementar para utilizar esas técnicas de manera eficiente.

¿Qué tan benéfico considera que es su conocimiento de estadística para cumplir con sus responsabilidades?

No es cuestión de beneficio, se trata de una necesidad. De hecho, encontramos que debemos conocerla tan bien que podemos explicarla a nuestros clientes en términos coloquiales.



Proyectos, procedimientos y perspectivas

15

15-1 Proyectos

Concepto clave Esta sección incluye sugerencias para un estudio que se puede utilizar como proyecto final para el curso de introducción a la estadística. Una ventaja fantástica de este curso es que trata con destrezas y conceptos que se pueden aplicar de inmediato al mundo real. Después de sólo un semestre divertido, los estudiantes son capaces de realizar sus propios estudios. Algunos de los temas sugeridos pueden tratarse realizando experimentos reales, mientras que otros implican estudios observacionales que requieren de la investigación de resultados que ya están disponibles. Por ejemplo, no se recomienda en absoluto probar la eficacia de las bolsas de aire haciendo chocar automóviles en la realidad, pero las pruebas de dilución del sabor de galletas con chispas de chocolate pueden ser un experimento sencillo y un tanto disfrutable. A continuación se presenta una sugerencia de formato, así como una lista de temas sugeridos.

Proyecto grupal o proyecto individual Los temas se pueden asignar de forma individual, pero los proyectos en grupo son particularmente efectivos puesto que ayudan a desarrollar las destrezas interpersonales que son tan necesarias en el ambiente laboral actual. Un estudio mostró que la “incapacidad para llevarse bien con los demás” es el motivo principal del despido de empleados, por lo que un proyecto grupal resultará muy útil si se quiere preparar a los estudiantes para sus ambientes de trabajo futuros.

Reporte oral Todos los miembros del grupo deben participar en una presentación en clase de 10 a 15 minutos de duración en un esfuerzo combinado para describir claramente los componentes esenciales del estudio. Los estudiantes por lo regular tienen cierta reticencia a hablar en público, por lo que un breve reporte oral resultará muy útil en el desarrollo de la confianza que ellos bien se merecen. El reporte oral es una actividad que ayuda a los estudiantes a estar mejor preparados para actividades profesionales futuras.

Reporte escrito El objetivo principal del proyecto no es producir un documento escrito equivalente a un trabajo final, pero se debe presentar un reporte escrito que incluya los siguientes componentes:

1. Lista de los datos reunidos, junto con la descripción de cómo se obtuvieron
2. Descripción del método de análisis
3. Gráficas y/o estadísticos relevantes, incluyendo pantallas de resultados de STATDISK, Minitab, Excel o la calculadora TI-83/84 Plus
4. Planteamiento de las conclusiones
5. Las razones por las que los resultados pudieran no ser correctos, junto con una descripción de las formas en las cuales el estudio podría mejorarse, contando con el tiempo y dinero suficientes.

Grupos grandes o clases en línea: carteles o PowerPoint Algunos grupos son demasiado grandes para trabajar proyectos individuales o proyectos grupales con tres, cuatro o cinco estudiantes por equipo. Algunas clases que se imparten en línea no ofrecen la posibilidad de reunirse como grupo. En estos casos se pueden presentar reportes de proyectos individuales o de grupos pequeños a través de carteles, similares a las sesiones de conferencias con carteles. Los carteles que resumen los elementos importantes de un proyecto se envían a los profesores para su evaluación. También se pueden usar presentaciones de PowerPoint.

Encuesta Una encuesta puede ser una excelente fuente de datos. Observe la encuesta que se presenta más adelante, la cual ofrece oportunidades para muchos proyectos interesantes que tratan preguntas como éstas:

1. Cuando las personas eligen dígitos “aleatoriamente” (como en la pregunta 2), ¿los resultados son realmente aleatorios?
2. Al parecer, ¿los cuatro últimos dígitos de los números del seguro social son aleatorios?
3. ¿Los hombres y las mujeres llevan consigo diferentes cantidades de cambio?
4. ¿Los hombres y las mujeres tienen números de tarjetas de crédito diferentes?
5. ¿Hay una diferencia en el pulso de las personas que hacen ejercicio y las que no lo practican?
6. ¿Hay una diferencia en el pulso de las personas que fuman y las que no lo hacen?
7. ¿Existe una relación entre la práctica de ejercicio y el tabaquismo?
8. ¿Existe una relación entre el color de ojos y la práctica de ejercicio?
9. ¿Hay una relación entre la práctica de ejercicio y el número de horas trabajadas por semana?
10. ¿Hay una correlación entre la estatura y el pulso?

Encuesta

1. Mujer Hombre
2. Elija al azar cuatro dígitos y anótelos aquí: _____
3. Color de ojos: _____
4. Anote su estatura en pulgadas: _____
5. ¿Cuál es el valor total de las monedas que lleva con usted? _____
6. ¿Cuántas llaves tiene en este momento? _____
7. ¿Cuántas tarjetas de crédito tiene en este momento? _____
8. Anote los cuatro últimos dígitos de su número del seguro social: _____
9. Registre su pulso contando el número de latidos por minuto:

10. ¿Se ejercita vigorosamente (corre, nada, practica ciclismo, tenis, básquetbol, etcétera) durante al menos 20 minutos por lo menos dos veces a la semana?
 Sí No
11. ¿Cuántas horas de créditos de clases está cursando este semestre? _____
12. ¿Trabaja actualmente? Sí No
Si su respuesta es afirmativa, ¿cuántas horas trabaja cada semana? _____
13. Durante los últimos 12 meses, ¿condujo algún automóvil que se viera involucrado en un choque? Sí No
14. ¿Fuma? Sí No
15. Zurdo Diestro Ambidiestro

Temas para proyectos Las anteriores preguntas de encuesta son una fuente de buenas ideas para proyectos. Consulte también las “Actividades de cooperación en equipo” que aparecen casi al final de cada capítulo. La siguiente lista ofrece sugerencias adicionales para proyectos.

1. Rehacer una gráfica de un periódico o revista con el fin de hacer una mejor descripción de los datos.
2. Reescribir un artículo de periódico acerca de una encuesta con el fin de informar mejor al lector.
3. Utilizar lanzamientos de monedas para obtener mejores resultados de una encuesta con preguntas sobre temas delicados.
4. Comparar la antigüedad de los automóviles de los estudiantes con la de los automóviles de los profesores y del personal administrativo de la universidad.
5. Comparar la proporción de automóviles extranjeros conducidos por estudiantes con la proporción de automóviles extranjeros conducidos por profesores.
6. Comparar la antigüedad de los automóviles en el estacionamiento de una tienda de descuento con la antigüedad de los automóviles en el estacionamiento de un almacén departamental de lujo.
7. ¿Los esposos son de mayor edad que sus esposas?

8. ¿Las diferencias en edad de los esposos y las esposas son las mismas para las parejas jóvenes que para las parejas de más edad?
9. Análisis de la antigüedad de los libros en la biblioteca de la universidad
10. ¿Cómo se compara la antigüedad de los libros de la biblioteca de su universidad con la de los libros de la biblioteca de una universidad vecina?
11. Comparación de la antigüedad de los libros de ciencias y los libros de inglés en la biblioteca de la universidad.
12. Estimación de la cantidad de horas que los estudiantes emplean para estudiar cada semana.
13. ¿Existe una relación entre las horas de estudio y las calificaciones obtenidas?
14. ¿Existe una relación entre las horas trabajadas y las calificaciones obtenidas?
15. Un estudio de estaturas *reportadas*, comparadas con estaturas *medidas*.
16. Un estudio de la exactitud de relojes.
17. ¿Existe una relación entre el sabor y el costo de marcas diferentes de galletas con chispas de chocolate?
18. ¿Hay una relación entre el sabor y el costo de marcas diferentes de mantequilla de maní?
19. ¿Existe una relación entre el sabor y el costo de marcas diferentes de bebidas de cola?
20. ¿Hay una relación entre los salarios de los jugadores profesionales de béisbol (o básquetbol, o fútbol) y sus logros en la temporada?
21. Tasas contra pesos: ¿Existe una relación entre las tasas de consumo de combustible de los automóviles y su peso? Si es así, ¿cuál es?
22. ¿Existe una relación entre la longitud de los pies de los hombres (o de las mujeres) y sus estaturas?
23. ¿Hay diferencias entre el sabor del agua común de la llave y el de las diferentes marcas de agua embotellada?
24. ¿Las tasas de mortalidad en accidentes automovilísticos se vieron afectadas por las leyes que exigen el uso de cinturones de seguridad?
25. ¿Las tasas de mortalidad en accidentes automovilísticos en Estados Unidos se vieron afectadas cuando se eliminó el límite de velocidad de 55 mi/h?
26. ¿Las tasas de mortalidad en accidentes automovilísticos se vieron afectadas por la presencia de las bolsas de aire?
27. ¿Existe alguna diferencia de sabor entre la Coca y la Pepsi?
28. ¿Hay una relación entre el promedio de calificaciones de un estudiante y la cantidad de televisión que ve? Si así es, ¿cuál es?
29. ¿Existe una relación entre el precio de venta de una casa y su área habitable (en metros cuadrados), el tamaño del terreno (en metros cuadrados), el número de habitaciones, el número de baños y el impuesto predial anual?
30. ¿Hay una relación entre la estatura de una persona y la altura de su ombligo?
31. ¿Existen evidencias que sustenten la teoría de que la proporción que guarda la estatura de una persona con la altura de su ombligo es la proporción áurea de alrededor de 1.6:1?
32. Una comparación del número de llaves que llevan consigo los hombres y las mujeres.

33. Una comparación del número de tarjetas de crédito que llevan consigo los hombres y las mujeres.
34. ¿Los homicidas son ahora más jóvenes que antes?
35. ¿Las personas que practican ejercicio vigoroso tienden a mostrar tasas de pulso más bajas que quienes no se ejercitan?
36. ¿Las personas que practican ejercicio vigoroso tienden a mostrar tiempos de reacción diferentes que quienes no se ejercitan?
37. ¿Las personas que fuman tienden a mostrar tasas de pulso más altas que quienes no lo hacen?
38. En el caso de las personas que no practican ejercicio, ¿cómo se ve afectado su pulso al subir un tramo de escaleras?
39. ¿Los estudiantes de estadística tienden a mostrar tasas de pulso diferentes que las personas que no estudian estadística?
40. Una comparación del promedio de calificaciones de estudiantes de estadística con el de los estudiantes que no estudian estadística.
41. ¿Las personas zurdas tienden a verse involucradas en más choques de automóviles que las diestras?
42. ¿Los hombres sufren más choques automovilísticos que las mujeres?
43. ¿Los conductores jóvenes sufren más choques automovilísticos que los conductores de mayor edad?
44. ¿Los conductores que reciben multas tienen mayores probabilidades de verse involucrados en choques?
45. ¿Los fumadores tienden a verse involucrados en más choques automovilísticos que quienes no fuman?
46. ¿Las personas con un pulso más alto tienden a verse involucradas en más o en menos choques de automóviles?
47. Una comparación de los tiempos de reacción medidos de las manos derecha e izquierda.
48. ¿Son iguales las proporciones de fumadores hombres y mujeres?
49. ¿Los estudiantes de estadística tienden a fumar más (o menos) que la población general?
50. ¿Es más probable que las personas fumen si sus padres fumaron?
51. Investigar la evidencia para sustentar o refutar la creencia de que fumar tiende a frenar el crecimiento.
52. ¿Un equipo deportivo tiene una ventaja al jugar en casa en comparación con las ocasiones en que es visitante?
53. Análisis de los tiempos de servicio (en segundos) de una sucursal bancaria
54. Una comparación de los tiempos de servicio de las ventanillas de dos bancos diferentes.
55. Análisis del tiempo que tardan los clientes de McDonald's en conseguir una mesa.
56. Análisis de los tiempos en que los clientes de McDonald's esperan en la fila.
57. Análisis de los tiempos que requieren los automóviles para llenar el tanque de combustible.
58. ¿La lotería estatal es una inversión sensata?

59. Comparación de juegos de casino: los dados contra la ruleta.
60. Comenzando con \$1, ¿es más fácil ganar un millón de dólares apostando a los dados en un casino o jugando en la lotería estatal?
61. Estrategias de apuestas audaces contra apuestas prudentes: Cuando se apuestan \$100, ¿es diferente apostar \$1 por juego que apostar los \$100 en un solo juego?
62. Diseño y análisis de resultados de una prueba de percepción extrasensorial.
63. Análisis de datos apareados consistentes en estaturas de padres (o madres) y las estaturas de su primer hijo (o hija).
64. Diferencias de género en preferencias de acompañantes para cenar entre las opciones de Brad Pitt, Tiger Woods, el presidente, Nicole Kidman, Cameron Diaz, Julia Roberts y el Papa.
65. Diferencias de género en preferencias de actividades entre las opciones de ir a cenar, ir al cine, ver la televisión, leer un libro, jugar golf, jugar tenis, nadar, asistir a un partido de béisbol y asistir a un partido de fútbol.
66. ¿Hay evidencias que sustenten la teoría de que los cereales con alto contenido de azúcar se colocan en los estantes al nivel visual de los niños?
67. ¿Hay evidencias que sustenten la aseveración de que la temperatura corporal media es menor que 98.6°F?
68. ¿Existe una relación entre fumar y beber café?
69. ¿Hay una relación entre las calificaciones del curso y el tiempo empleado en juegos de video?
70. ¿Hay evidencias que sustenten la teoría de que el martes 13 es de mala suerte?

15-2 Procedimientos

Concepto clave En esta sección se describe un procedimiento general para realizar un análisis estadístico de datos. Los datos se pueden reunir por medio de experimentos o estudios observacionales. Es absolutamente esencial hacer una crítica del método utilizado para reunir los datos, porque un método inadecuado de recolección de datos destruye su utilidad. Busque con cuidado sesgos en la forma en que se reúnen los datos, así como también sesgos por parte de la persona o grupo que recolecta los datos. Muchos de los procedimientos de este libro se basan en el supuesto de que se trabaja con una muestra aleatoria simple, lo que significa que cada muestra posible del mismo tamaño tiene la misma oportunidad de ser seleccionada. Si se selecciona una muestra de respuesta voluntaria, no servirá para hacer inferencias acerca de una población.

Exploración, comparación y descripción Después de reunir los datos, primero considere la exploración, descripción y comparación de los conjuntos de datos utilizando las herramientas básicas incluidas en los capítulos 2 y 3. Asegúrese de explorar lo siguiente:

1. *Centro*: Calcule la media y la mediana, que son medidas de tendencia central con valores representativos o de promedio, y que nos indican dónde se encuentra la parte media del conjunto de datos.
2. *Variación*: Calcule al rango y la desviación estándar, que son medidas de la cantidad en que los valores muestrales varían entre sí.

3. **Distribución:** Construya un histograma para ver la naturaleza o forma de la distribución de los datos. También construya una gráfica cuantílica normal para determinar si los datos provienen de una población con una distribución normal.
4. **Valores extremos:** Identifique cualquier valor muestral que se encuentre muy alejado de la gran mayoría de los demás valores muestrales. Si hay valores extremos, trate de determinar si se trata de errores que deben corregirse. Si los valores extremos son datos correctos, estudie sus efectos repitiendo el análisis sin incluirlos.
5. **Tiempo:** Determine si la proporción es estable o si sus características cambian con el tiempo.

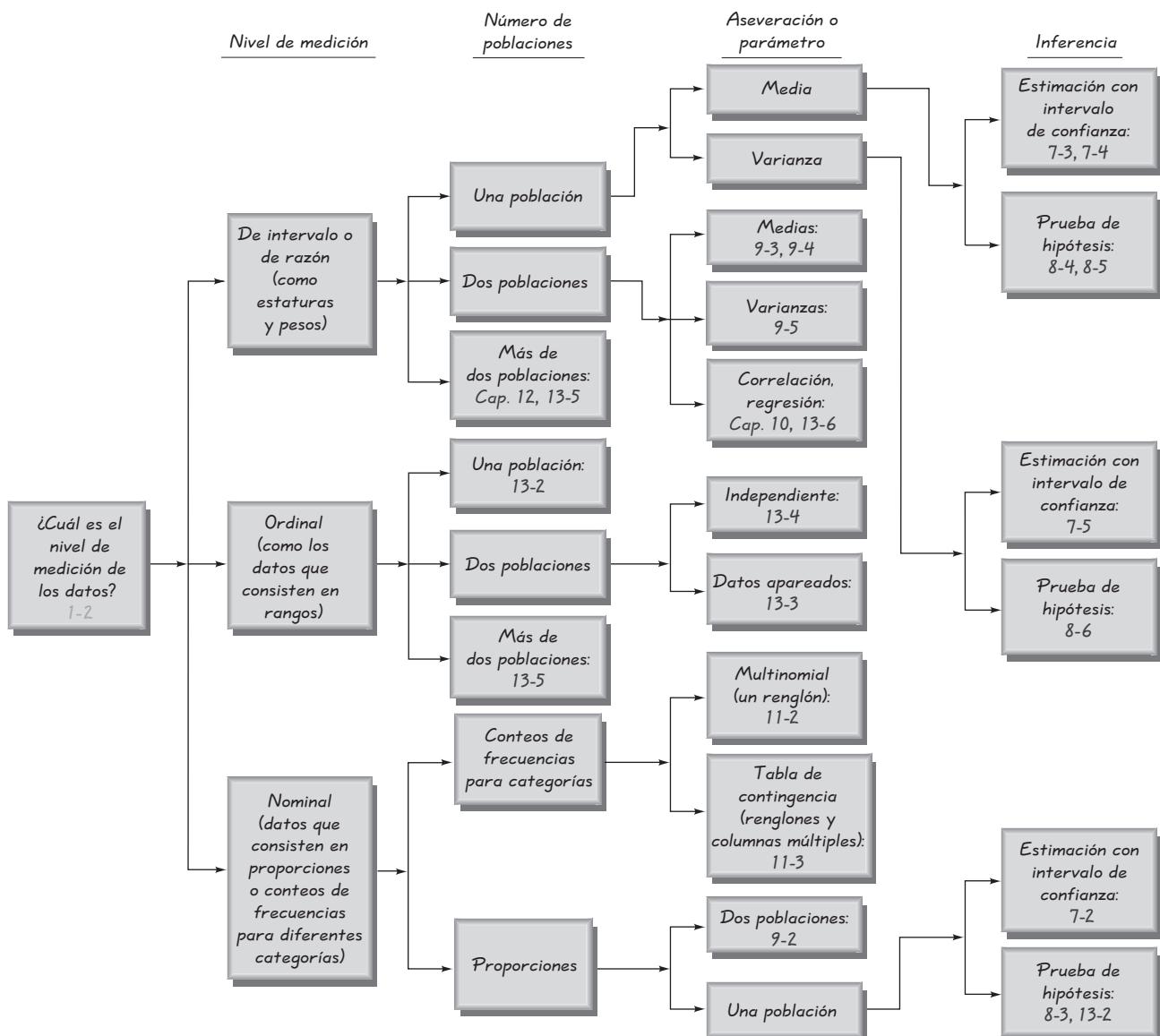


Figura 15-1 Selección del procedimiento adecuado

Inferencias: Estimación de parámetros y prueba de hipótesis Cuando se trata de utilizar datos muestrales para hacer inferencias acerca de una población, suele ser difícil elegir el procedimiento específico que se debe aplicar. Este libro incluye una amplia variedad de procedimientos que se aplican a muchas circunstancias diferentes. A continuación se incluyen algunas preguntas clave que deben responderse:

- ¿Cuál es el nivel de medición (nominal, ordinal, de intervalo o de razón) de los datos?
- ¿El estudio implica una, dos o más poblaciones?
- ¿Existe una aseveración a probar o un parámetro a estimar?
- ¿Cuál es el parámetro relevante (media, desviación estándar, proporción)?
- ¿Se conoce la desviación estándar poblacional? (La respuesta casi siempre es “no”).
- ¿Existe una razón para creer que la población está distribuida normalmente?
- ¿Cuál es la pregunta básica o el tema al que usted se quiere abocar?

En la figura 15-1 se listan los principales métodos incluidos en este libro, junto con un esquema para determinar cuáles de ellos se deben utilizar según el caso. Para utilizar la figura 15-1, inicie en el extremo izquierdo y comience por identificar el nivel de medición de los datos. Proceda siguiendo el flujo sugerido por el nivel de medición, el número de poblaciones y la aseveración o parámetro a considerar.

Nota: Esta figura se aplica a una población fija. Si los datos provienen de un proceso que puede cambiar con el tiempo, construya una gráfica de control (véase el capítulo 14) para determinar si el proceso es estadísticamente estable. Esta figura se aplica a datos de proceso sólo si el proceso es estadísticamente estable.

La figura 15-1 puede utilizarse con los métodos estadísticos presentados en este libro, pero quizás existan otros métodos que podrían ser más adecuados para un análisis estadístico en particular. Consulte a un especialista en estadística amistoso para que lo oriente acerca de otros métodos.

15-3 Perspectivas

Concepto clave Ningún curso de introducción a la estadística es capaz de transformar a alguien en un experto en la materia. El curso de introducción tiene un enfoque limitado y no incluye muchos temas importantes. Recuerde que puede recibir ayuda de expertos en el tema, y este curso de introducción a la estadística le servirá cuando consulte a uno de ellos.

La cobertura exitosa de un curso de introducción a la estadística produce beneficios que van más allá de la obtención de créditos para un título universitario. Usted también fortalecerá su calificación para conseguir empleo; estará más preparado para analizar de forma crítica informes de los medios de comunicación y de revistas científicas; comprenderá los conceptos básicos de la probabilidad y el azar. También sabrá que, para entender un conjunto de datos, es importante investigar medidas de tendencia central (como la media y la mediana), medidas de variación (como el rango y la desviación estándar), la naturaleza de la distribución (por medio de una distribución de frecuencias o de una gráfica), la presencia de valores extremos y si la población permanece estable o cambia con el tiempo. Además, conocerá y comprenderá la importancia de estimar parámetros poblacionales (como

una media, una desviación estándar y una proporción), así como la prueba de aseveraciones sobre parámetros poblacionales.

A lo largo de este libro hemos destacado la importancia de un buen muestreo. Hay que reconocer que una muestra mala suele ser muy difícil de reparar, incluso para los más expertos y utilizando las técnicas más elaboradas. Existen muchas encuestas por correo, de revistas y de respuesta telefónica que permiten que quienes responden se “autoseleccionen”. Los resultados de encuestas de este tipo generalmente son inútiles cuando se juzgan de acuerdo con los criterios de la metodología estadística sólida. Tenga esto en mente cuando encuentre encuestas de respuesta voluntaria (autoseleccionadas), para que no permita que afecten sus creencias y decisiones. Sin embargo, también debe reconocerse que muchas encuestas y entrevistas tienen muy buenos resultados, aun cuando los tamaños de las muestras puedan parecer relativamente pequeños. Aunque muchas personas se rehúsan a creerlo, una encuesta a nivel nacional de sólo 1200 votantes puede arrojar buenos resultados si el muestreo se planea y ejecuta cuidadosamente.

A través de este libro hemos destacado la importancia de la *interpretación* de los resultados. La conclusión final de “rechazar la hipótesis nula” básicamente no tiene ningún valor para todas aquellas personas que carecen de la visión y la sensatez como para tomar un curso de estadística. Las computadoras y las calculadoras son buenas para dar resultados, pero éstos por lo regular requieren de una interpretación cuidadosa que da paso a resultados que de otra forma carecerían de significado. Debemos reconocer que un resultado no es automáticamente válido y bueno sólo porque fue generado por computadora. Las computadoras no piensan y son capaces de proporcionar resultados bastante ridículos cuando se consideran en el contexto del mundo real. Siempre debemos aplicar la herramienta más importante e indispensable en toda la estadística: *¡el sentido común!*

El papel de la estadística en la educación Hubo una época en que una persona se consideraba educada simplemente si sabía leer. Sin embargo, ahora vivimos una época que demanda mucho más. En la actualidad un individuo educado es capaz de pensar de manera crítica y de aprender en vez de únicamente actuar. Una persona educada se comunica eficazmente de manera oral y escrita, y es capaz de relacionarse de forma adecuada con los demás, incluyendo a gente de diferentes culturas y a individuos que tienen poca educación. El curso de introducción a la estadística ofrece mucho más que la mera adquisición de habilidades técnicas. La cobertura exitosa de un curso de introducción a la estadística contribuye a que los estudiantes crezcan a nivel individual y profesional, de manera que puedan progresar mucho para convertirse en profesionistas productivos, ciudadanos responsables e individuos verdaderamente educados.

Apéndices

Apéndice A: Tablas

Apéndice B: Conjuntos de datos

Apéndice C: Glosario

Apéndice D: Bibliografía

Apéndice E: Soluciones de los ejercicios impares (y de todos los ejercicios de repaso y de los ejercicios de repaso acumulativo)

Apéndice A: Tablas

- | | |
|-------------------|---|
| Tabla A-1 | Probabilidades binomiales |
| Tabla A-2 | Distribución normal estándar |
| Tabla A-3 | Distribución <i>t</i> |
| Tabla A-4 | Distribución chi cuadrada (χ^2) |
| Tabla A-5 | Distribución <i>F</i> |
| Tabla A-6 | Valores críticos del coeficiente de correlación de Pearson <i>r</i> |
| Tabla A-7 | Valores críticos para la prueba del signo |
| Tabla A-8 | Valores críticos de <i>T</i> para la prueba de rangos con signo de Wilcoxon |
| Tabla A-9 | Valores críticos del coeficiente de correlación de rangos de Spearman r_s |
| Tabla A-10 | Valores críticos para el número de rachas <i>G</i> |

TABLA A-1 Probabilidades binomiales

n	x	p														x
		.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99		
2	0	.980	.902	.810	.640	.490	.360	.250	.160	.090	.040	.010	.002	0+	0	
	1	.020	.095	.180	.320	.420	.480	.500	.480	.420	.320	.180	.095	.020	1	
	2	0+	.002	.010	.040	.090	.160	.250	.360	.490	.640	.810	.902	.980	2	
3	0	.970	.857	.729	.512	.343	.216	.125	.064	.027	.008	.001	0+	0+	0	
	1	.029	.135	.243	.384	.441	.432	.375	.288	.189	.096	.027	.007	0+	1	
	2	0+	.007	.027	.096	.189	.288	.375	.432	.441	.384	.243	.135	.029	2	
	3	0+	0+	.001	.008	.027	.064	.125	.216	.343	.512	.729	.857	.970	3	
4	0	.961	.815	.656	.410	.240	.130	.062	.026	.008	.002	0+	0+	0+	0	
	1	.039	.171	.292	.410	.412	.346	.250	.154	.076	.026	.004	0+	0+	1	
	2	.001	.014	.049	.154	.265	.346	.375	.346	.265	.154	.049	.014	.001	2	
	3	0+	0+	.004	.026	.076	.154	.250	.346	.412	.410	.292	.171	.039	3	
	4	0+	0+	0+	.002	.008	.026	.062	.130	.240	.410	.656	.815	.961	4	
5	0	.951	.774	.590	.328	.168	.078	.031	.010	.002	0+	0+	0+	0+	0	
	1	.048	.204	.328	.410	.360	.259	.156	.077	.028	.006	0+	0+	0+	1	
	2	.001	.021	.073	.205	.309	.346	.312	.230	.132	.051	.008	.001	0+	2	
	3	0+	.001	.008	.051	.132	.230	.312	.346	.309	.205	.073	.021	.001	3	
	4	0+	0+	0+	.006	.028	.077	.156	.259	.360	.410	.328	.204	.048	4	
	5	0+	0+	0+	0+	.002	.010	.031	.078	.168	.328	.590	.774	.951	5	
6	0	.941	.735	.531	.262	.118	.047	.016	.004	.001	0+	0+	0+	0+	0	
	1	.057	.232	.354	.393	.303	.187	.094	.037	.010	.002	0+	0+	0+	1	
	2	.001	.031	.098	.246	.324	.311	.234	.138	.060	.015	.001	0+	0+	2	
	3	0+	.002	.015	.082	.185	.276	.312	.276	.185	.082	.015	.002	0+	3	
	4	0+	0+	0+	.001	.015	.060	.138	.234	.311	.324	.246	.098	.031	001	
	5	0+	0+	0+	0+	.002	.010	.037	.094	.187	.303	.393	.354	.232	.057	
	6	0+	0+	0+	0+	.001	.004	.016	.047	.118	.262	.531	.735	.941	6	
7	0	.932	.698	.478	.210	.082	.028	.008	.002	0+	0+	0+	0+	0+	0	
	1	.066	.257	.372	.367	.247	.131	.055	.017	.004	0+	0+	0+	0+	1	
	2	.002	.041	.124	.275	.318	.261	.164	.077	.025	.004	0+	0+	0+	2	
	3	0+	.004	.023	.115	.227	.290	.273	.194	.097	.029	.003	0+	0+	3	
	4	0+	0+	0+	.003	.029	.097	.194	.273	.290	.227	.115	.023	.004	0+	
	5	0+	0+	0+	0+	.004	.025	.077	.164	.261	.318	.275	.124	.041	.002	
	6	0+	0+	0+	0+	.004	.017	.055	.131	.247	.367	.372	.257	.066	6	
8	0	.923	.663	.430	.168	.058	.017	.004	.001	0+	0+	0+	0+	0+	0	
	1	.075	.279	.383	.336	.198	.090	.031	.008	.001	0+	0+	0+	0+	1	
	2	.003	.051	.149	.294	.296	.209	.109	.041	.010	.001	0+	0+	0+	2	
	3	0+	.005	.033	.147	.254	.279	.219	.124	.047	.009	0+	0+	0+	3	
	4	0+	0+	0+	.005	.046	.136	.232	.273	.232	.136	.046	.005	0+	0+	
	5	0+	0+	0+	0+	.009	.047	.124	.219	.279	.254	.147	.033	.005	0+	
	6	0+	0+	0+	0+	.001	.010	.041	.109	.209	.296	.294	.149	.051	.003	
7	0+	0+	0+	0+	.001	.008	.031	.090	.198	.336	.383	.279	.075	075	7	
	8	0+	0+	0+	0+	.001	.004	.017	.058	.168	.430	.663	.923	0.923	8	

NOTA: 0+ representa una probabilidad menor que 0.0005.

(continúa)

TABLA A-1 Probabilidades binomiales (*continuación*)

n	x	p													x
		.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	
9	0	.914	.630	.387	.134	.040	.010	.002	0+	0+	0+	0+	0+	0+	0
	1	.083	.299	.387	.302	.156	.060	.018	.004	0+	0+	0+	0+	0+	1
	2	.003	.063	.172	.302	.267	.161	.070	.021	.004	0+	0+	0+	0+	2
	3	0+	.008	.045	.176	.267	.251	.164	.074	.021	.003	0+	0+	0+	3
	4	0+	.001	.007	.066	.172	.251	.246	.167	.074	.017	.001	0+	0+	4
	5	0+	0+	.001	.017	.074	.167	.246	.251	.172	.066	.007	.001	0+	5
	6	0+	0+	0+	.003	.021	.074	.164	.251	.267	.176	.045	.008	0+	6
	7	0+	0+	0+	0+	.004	.021	.070	.161	.267	.302	.172	.063	.003	7
	8	0+	0+	0+	0+	0+	.004	.018	.060	.156	.302	.387	.299	.083	8
	9	0+	0+	0+	0+	0+	0+	.002	.010	.040	.134	.387	.630	.914	9
10	0	.904	.599	.349	.107	.028	.006	.001	0+	0+	0+	0+	0+	0+	0
	1	.091	.315	.387	.268	.121	.040	.010	.002	0+	0+	0+	0+	0+	1
	2	.004	.075	.194	.302	.233	.121	.044	.011	.001	0+	0+	0+	0+	2
	3	0+	.010	.057	.201	.267	.215	.117	.042	.009	.001	0+	0+	0+	3
	4	0+	.001	.011	.088	.200	.251	.205	.111	.037	.006	0+	0+	0+	4
	5	0+	0+	.001	.026	.103	.201	.246	.201	.103	.026	.001	0+	0+	5
	6	0+	0+	0+	.006	.037	.111	.205	.251	.200	.088	.011	.001	0+	6
	7	0+	0+	0+	0+	.001	.009	.042	.117	.215	.267	.201	.057	.010	0+
	8	0+	0+	0+	0+	0+	.001	.011	.044	.121	.233	.302	.194	.075	004
	9	0+	0+	0+	0+	0+	0+	.002	.010	.040	.121	.268	.387	.315	.091
	10	0+	0+	0+	0+	0+	0+	0+	.001	.006	.028	.107	.349	.599	.904
11	0	.895	.569	.314	.086	.020	.004	0+	0+	0+	0+	0+	0+	0+	0
	1	.099	.329	.384	.236	.093	.027	.005	.001	0+	0+	0+	0+	0+	1
	2	.005	.087	.213	.295	.200	.089	.027	.005	.001	0+	0+	0+	0+	2
	3	0+	.014	.071	.221	.257	.177	.081	.023	.004	0+	0+	0+	0+	3
	4	0+	.001	.016	.111	.220	.236	.161	.070	.017	.002	0+	0+	0+	4
	5	0+	0+	.002	.039	.132	.221	.226	.147	.057	.010	0+	0+	0+	5
	6	0+	0+	0+	.010	.057	.147	.226	.221	.132	.039	.002	0+	0+	6
	7	0+	0+	0+	0+	.002	.017	.070	.161	.236	.220	.111	.016	.001	0+
	8	0+	0+	0+	0+	0+	.004	.023	.081	.177	.257	.221	.071	.014	0+
	9	0+	0+	0+	0+	0+	0+	.001	.005	.027	.089	.200	.295	.213	.087
	10	0+	0+	0+	0+	0+	0+	0+	.001	.005	.027	.093	.236	.384	.329
	11	0+	0+	0+	0+	0+	0+	0+	0+	.004	.020	.086	.314	.569	.895
12	0	.886	.540	.282	.069	.014	.002	0+	0+	0+	0+	0+	0+	0+	0
	1	.107	.341	.377	.206	.071	.017	.003	0+	0+	0+	0+	0+	0+	1
	2	.006	.099	.230	.283	.168	.064	.016	.002	0+	0+	0+	0+	0+	2
	3	0+	.017	.085	.236	.240	.142	.054	.012	.001	0+	0+	0+	0+	3
	4	0+	.002	.021	.133	.231	.213	.121	.042	.008	.001	0+	0+	0+	4
	5	0+	0+	.004	.053	.158	.227	.193	.101	.029	.003	0+	0+	0+	5
	6	0+	0+	0+	.016	.079	.177	.226	.177	.079	.016	0+	0+	0+	6
	7	0+	0+	0+	0+	.003	.029	.101	.193	.227	.158	.053	.004	0+	0+
	8	0+	0+	0+	0+	0+	.001	.008	.042	.121	.213	.231	.133	.021	.002
	9	0+	0+	0+	0+	0+	0+	.001	.012	.054	.142	.240	.236	.085	.017
	10	0+	0+	0+	0+	0+	0+	0+	.002	.016	.064	.168	.283	.230	.099
	11	0+	0+	0+	0+	0+	0+	0+	0+	.003	.017	.071	.206	.377	.341
	12	0+	0+	0+	0+	0+	0+	0+	0+	0+	.002	.014	.069	.282	.540

NOTA: 0+ representa una probabilidad positiva menor que 0.0005.

(continúa)

TABLA A-1 Probabilidades binomiales (*continuación*)

n	x	p														x
		.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99		
13	0	.878	.513	.254	.055	.010	.001	0+	0+	0+	0+	0+	0+	0+	0+	0
	1	.115	.351	.367	.179	.054	.011	.002	0+	0+	0+	0+	0+	0+	0+	1
	2	.007	.111	.245	.268	.139	.045	.010	.001	0+	0+	0+	0+	0+	0+	2
	3	0+	.021	.100	.246	.218	.111	.035	.006	.001	0+	0+	0+	0+	0+	3
	4	0+	.003	.028	.154	.234	.184	.087	.024	.003	0+	0+	0+	0+	0+	4
	5	0+	0+	.006	.069	.180	.221	.157	.066	.014	.001	0+	0+	0+	0+	5
	6	0+	0+	.001	.023	.103	.197	.209	.131	.044	.006	0+	0+	0+	0+	6
	7	0+	0+	0+	.006	.044	.131	.209	.197	.103	.023	.001	0+	0+	0+	7
	8	0+	0+	0+	.001	.014	.066	.157	.221	.180	.069	.006	0+	0+	0+	8
	9	0+	0+	0+	0+	.003	.024	.087	.184	.234	.154	.028	.003	0+	0+	9
	10	0+	0+	0+	0+	.001	.006	.035	.111	.218	.246	.100	.021	0+	0+	10
	11	0+	0+	0+	0+	0+	.001	.010	.045	.139	.268	.245	.111	.007	0+	11
	12	0+	0+	0+	0+	0+	0+	.002	.011	.054	.179	.367	.351	.115	0+	12
	13	0+	0+	0+	0+	0+	0+	0+	.001	.010	.055	.254	.513	.878	0+	13
14	0	.869	.488	.229	.044	.007	.001	0+	0+	0+	0+	0+	0+	0+	0+	0
	1	.123	.359	.356	.154	.041	.007	.001	0+	0+	0+	0+	0+	0+	0+	1
	2	.008	.123	.257	.250	.113	.032	.006	.001	0+	0+	0+	0+	0+	0+	2
	3	0+	.026	.114	.250	.194	.085	.022	.003	0+	0+	0+	0+	0+	0+	3
	4	0+	.004	.035	.172	.229	.155	.061	.014	.001	0+	0+	0+	0+	0+	4
	5	0+	0+	.008	.086	.196	.207	.122	.041	.007	0+	0+	0+	0+	0+	5
	6	0+	0+	.001	.032	.126	.207	.183	.092	.023	.002	0+	0+	0+	0+	6
	7	0+	0+	0+	.009	.062	.157	.209	.157	.062	.009	0+	0+	0+	0+	7
	8	0+	0+	0+	.002	.023	.092	.183	.207	.126	.032	.001	0+	0+	0+	8
	9	0+	0+	0+	0+	.007	.041	.122	.207	.196	.086	.008	0+	0+	0+	9
	10	0+	0+	0+	0+	.001	.014	.061	.155	.229	.172	.035	.004	0+	0+	10
	11	0+	0+	0+	0+	0+	.003	.022	.085	.194	.250	.114	.026	0+	0+	11
	12	0+	0+	0+	0+	0+	0+	.001	.006	.032	.113	.250	.257	.123	.008	12
	13	0+	0+	0+	0+	0+	0+	0+	.001	.007	.041	.154	.356	.359	.123	13
	14	0+	0+	0+	0+	0+	0+	0+	0+	.001	.007	.044	.229	.488	.869	14
15	0	.860	.463	.206	.035	.005	0+	0+	0+	0+	0+	0+	0+	0+	0+	0
	1	.130	.366	.343	.132	.031	.005	0+	0+	0+	0+	0+	0+	0+	0+	1
	2	.009	.135	.267	.231	.092	.022	.003	0+	0+	0+	0+	0+	0+	0+	2
	3	0+	.031	.129	.250	.170	.063	.014	.002	0+	0+	0+	0+	0+	0+	3
	4	0+	.005	.043	.188	.219	.127	.042	.007	.001	0+	0+	0+	0+	0+	4
	5	0+	.001	.010	.103	.206	.186	.092	.024	.003	0+	0+	0+	0+	0+	5
	6	0+	0+	.002	.043	.147	.207	.153	.061	.012	.001	0+	0+	0+	0+	6
	7	0+	0+	0+	.014	.081	.177	.196	.118	.035	.003	0+	0+	0+	0+	7
	8	0+	0+	0+	.003	.035	.118	.196	.177	.081	.014	0+	0+	0+	0+	8
	9	0+	0+	0+	0.001	.012	.061	.153	.207	.147	.043	.002	0+	0+	0+	9
	10	0+	0+	0+	0+	.003	.024	.092	.186	.206	.103	.010	.001	0+	0+	10
	11	0+	0+	0+	0+	0.001	.007	.042	.127	.219	.188	.043	.005	0+	0+	11
	12	0+	0+	0+	0+	0+	.002	.014	.063	.170	.250	.129	.031	0+	0+	12
	13	0+	0+	0+	0+	0+	0+	.003	.022	.092	.231	.267	.135	.009	0+	13
	14	0+	0+	0+	0+	0+	0+	0+	.005	.031	.132	.343	.366	.130	0+	14
	15	0+	0+	0+	0+	0+	0+	0+	0+	.005	.035	.206	.463	.860	0+	15

NOTA: 0+ representa una probabilidad positiva menor que 0.0005.

De Frederick C. Mosteller, Robert E. K. Rourke y George B. Thomas, Jr., *Probability with Statistical Applications*, 2a. ed., © 1970 Addison-Wesley Publishing Co., Reading, MA. Reproducido con permiso.

Puntuaciones z NEGATIVAS

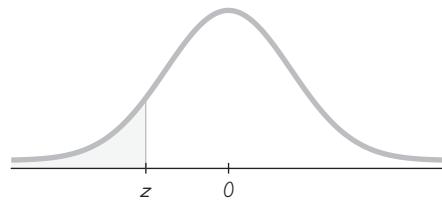


TABLA A-2 Distribución normal estándar (z): Área acumulativa desde la IZQUIERDA

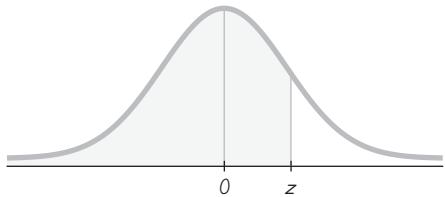
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.50 y menores	.0001									
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	* .0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	* .0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	* .0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	* .0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	* .0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	* .0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	* .0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	* .0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	* .0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	* .0495	.0485	.0475	* .0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	* .0606	.0594	.0582	* .0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	* .0735	.0721	.0708	* .0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	* .0885	.0869	.0853	* .0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	* .1056	.1038	.1020	* .1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	* .1251	.1230	.1210	* .1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	* .1469	.1446	.1423	* .1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	* .1711	.1685	.1660	* .1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	* .1977	.1949	.1922	* .1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	* .2266	.2236	.2206	* .2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	* .2578	.2546	.2514	* .2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	* .2912	.2877	.2843	* .2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	* .3264	.3228	.3192	* .3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	* .3632	.3594	.3557	* .3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	* .4013	.3974	.3936	* .3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	* .4404	.4364	.4325	* .4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	* .4801	.4761	.4721	* .4681	.4641

NOTA: Para valores de z por debajo de -3.49, utilice 0.0001 para el área.

*Utilice estos valores comunes que resultan por interpolación:

Puntuación

z	Área
-1.645	0.0500
-2.575	0.0050



Puntuaciones z POSITIVAS

TABLA A-2

(continuación) Área acumulativa desde la IZQUIERDA

<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	* .9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	↑ .9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	* .9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	↑ .9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	↑ .9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.50 y mayores	.9999									

NOTA: Para valores de *z* por encima de 3.49, utilice 0.9999 para el área.

*Utilice estos valores comunes que resultan por interpolación:

Puntuación

<i>z</i>	Área
1.645	0.9500
2.575	0.9950

Valores críticos comunes

Nivel de confianza	Valor crítico
0.90	1.645
0.95	1.96
0.99	2.575

TABLA A-3 Distribución t : Valores críticos t

	Área en una cola				
	0.005	0.01	0.025	0.05	0.10
Grados de libertad	0.01	Área en dos colas	0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316
26	2.779	2.479	2.056	1.706	1.315
27	2.771	2.473	2.052	1.703	1.314
28	2.763	2.467	2.048	1.701	1.313
29	2.756	2.462	2.045	1.699	1.311
30	2.750	2.457	2.042	1.697	1.310
31	2.744	2.453	2.040	1.696	1.309
32	2.738	2.449	2.037	1.694	1.309
34	2.728	2.441	2.032	1.691	1.307
36	2.719	2.434	2.028	1.688	1.306
38	2.712	2.429	2.024	1.686	1.304
40	2.704	2.423	2.021	1.684	1.303
45	2.690	2.412	2.014	1.679	1.301
50	2.678	2.403	2.009	1.676	1.299
55	2.668	2.396	2.004	1.673	1.297
60	2.660	2.390	2.000	1.671	1.296
65	2.654	2.385	1.997	1.669	1.295
70	2.648	2.381	1.994	1.667	1.294
75	2.643	2.377	1.992	1.665	1.293
80	2.639	2.374	1.990	1.664	1.292
90	2.632	2.368	1.987	1.662	1.291
100	2.626	2.364	1.984	1.660	1.290
200	2.601	2.345	1.972	1.653	1.286
300	2.592	2.339	1.968	1.650	1.284
400	2.588	2.336	1.966	1.649	1.284
500	2.586	2.334	1.965	1.648	1.283
750	2.582	2.331	1.963	1.647	1.283
1000	2.581	2.330	1.962	1.646	1.282
2000	2.578	2.328	1.961	1.646	1.282
Grande	2.576	2.326	1.960	1.645	1.282

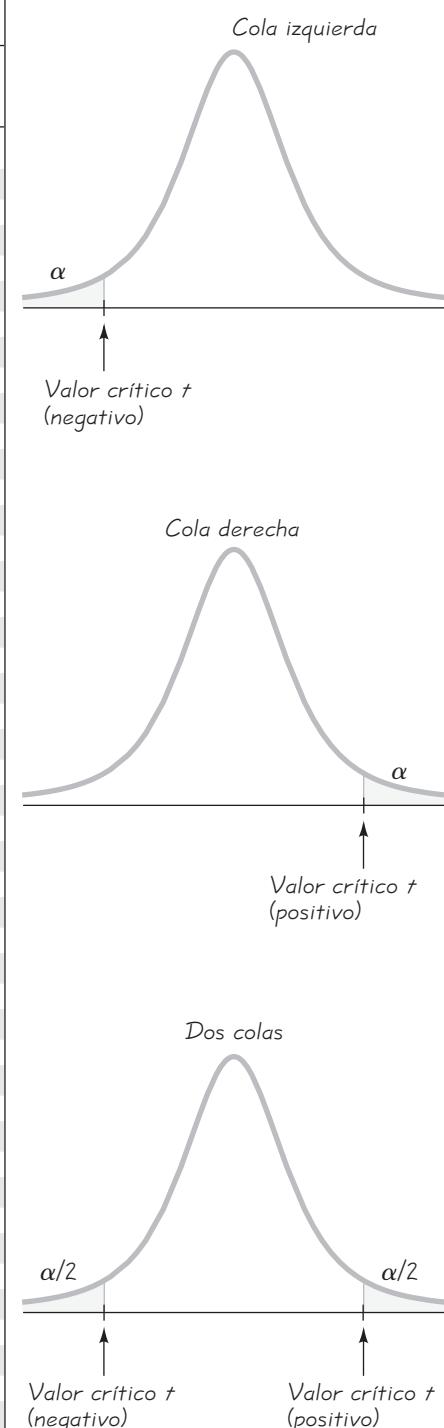


TABLA A-4 Distribución chi cuadrada (χ^2)

Grados de libertad	Área a la derecha del valor crítico									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

De Donald B. Owen, *Handbook of Statistical Tables*, © 1962 Addison-Wesley Publishing Co., Reading, MA. Reproducido con permiso del editor.

Grados de libertad

$n - 1$

para intervalos de confianza o pruebas de hipótesis con desviación estándar o varianza

$k - 1$

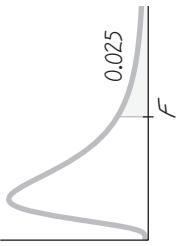
para experimentos multinomiales o bondad de ajuste con k categorías

$(r - 1)(c - 1)$

para tablas de contingencia con r renglones y c columnas

$k - 1$

para la prueba de Kruskal-Wallis con k muestras

TABLA A-5 Distribución F ($\alpha = 0.025$ en la cola derecha)

	Grados de libertad del numerador (g_1)								
	1	2	3	4	5	6	7	8	9
1	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
2	38.506	39.000	39.165	39.248	39.298	39.331	39.335	39.373	39.387
3	17.443	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473
4	12.218	10.649	9.9792	9.6045	9.3645	9.1973	9.0741	8.9796	8.9047
5	10.007	8.4336	7.7636	7.3879	7.1464	6.9777	6.8531	6.7572	6.6811
6	8.8131	7.2599	6.5988	6.2272	5.9876	5.8198	5.6955	5.5996	5.5234
7	8.0727	6.5415	5.8898	5.5226	5.2852	5.1186	4.9949	4.8993	4.8232
8	7.5709	6.0595	5.4160	5.0526	4.8173	4.6517	4.5286	4.4333	4.3572
9	7.2093	5.7147	5.0781	4.7181	4.4844	4.3197	4.1970	4.1020	4.0260
10	6.9367	5.4564	4.8256	4.4683	4.2361	4.0721	3.9498	3.8549	3.7790
11	6.7241	5.2559	4.6300	4.2751	4.0440	3.8807	3.7586	3.6638	3.5879
12	6.5538	5.0959	4.4742	4.1212	3.8911	3.7283	3.6065	3.5118	3.4358
13	6.4143	4.9653	4.3472	3.9959	3.7667	3.6043	3.4827	3.3880	3.3120
14	6.2979	4.8567	4.2417	3.8919	3.6634	3.5014	3.3799	3.2853	3.2093
15	6.1995	4.7650	4.1528	3.8043	3.5764	3.4147	3.2934	3.1987	3.1227
16	6.1151	4.6867	4.0768	3.7294	3.5021	3.3406	3.2194	3.1248	3.0488
17	6.0420	4.6189	4.0112	3.6648	3.4379	3.2767	3.1556	3.0610	2.9849
18	5.9781	4.5597	3.9539	3.6083	3.3820	3.2209	3.0999	3.0053	2.9291
19	5.9216	4.5075	3.9034	3.5587	3.3327	3.1718	3.0509	2.9563	2.8801
20	5.8715	4.4613	3.8587	3.5147	3.2891	3.1283	3.0074	2.9128	2.8365
21	5.8266	4.4199	3.8188	3.4754	3.2501	3.0895	2.9686	2.8740	2.7977
22	5.7863	4.3828	3.7829	3.4401	3.2151	3.0546	2.9338	2.8392	2.7628
23	5.7498	4.3492	3.7505	3.4083	3.1835	3.0232	2.9023	2.8077	2.7313
24	5.7166	4.3187	3.7211	3.3794	3.1548	2.9946	2.8738	2.7791	2.7027
25	5.6864	4.2909	3.6943	3.3530	3.1287	2.9685	2.8478	2.7531	2.6766
26	5.6586	4.2655	3.6697	3.3289	3.1048	2.9447	2.8240	2.7293	2.6528
27	5.6331	4.2421	3.6472	3.3067	3.0828	2.9228	2.8021	2.7074	2.6309
28	5.6096	4.2205	3.6264	3.2863	3.0626	2.9027	2.7820	2.6872	2.6106
29	5.5878	4.2006	3.6072	3.2674	3.0438	2.8840	2.7633	2.6686	2.5919
30	5.5675	4.1821	3.5894	3.2499	3.0265	2.8667	2.7460	2.6513	2.5746
40	5.4239	4.0510	3.4633	3.1261	2.9037	2.7444	2.6238	2.5289	2.4519
60	5.2856	3.9253	3.3425	3.0077	2.7863	2.6274	2.5068	2.4117	2.3344
120	5.1523	3.8046	3.2269	2.8943	2.6740	2.5154	2.3948	2.2994	2.2217
∞	5.0239	3.6889	3.1161	2.7858	2.5665	2.4082	2.2875	2.1918	2.1136

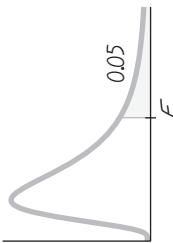
Grados de libertad del denominador (g_2)

TABLA A-5 Distribución F ($\alpha = 0.025$ en la cola derecha) (continuación)

	Grados de libertad del numerador (g_1)									
	10	12	15	20	24	30	40	60	120	∞
1	968.63	976.71	984.87	993.10	997.25	1001.4	1005.6	1009.8	1014.0	1018.3
2	39.398	39.415	39.431	39.448	39.456	39.465	39.473	39.481	39.490	39.498
3	14.419	14.337	14.253	14.167	14.124	14.081	14.037	13.992	13.947	13.902
4	8.8439	8.7512	8.6565	8.5599	8.5109	8.4613	8.4111	8.3604	8.3092	8.2573
5	6.6192	6.5245	6.4277	6.3286	6.2780	6.2269	6.1750	6.1225	6.0693	6.0153
6	5.4613	5.3662	5.2687	5.1684	5.1172	5.0652	5.0125	4.9589	4.9044	4.8491
7	4.7611	4.6658	4.5678	4.4667	4.4150	4.3624	4.3089	4.2544	4.1989	4.1423
8	4.2951	4.1997	4.1012	3.9995	3.9472	3.8940	3.8398	3.7844	3.7279	3.6702
9	3.9639	3.8682	3.7694	3.6669	3.6142	3.5604	3.5055	3.4493	3.3918	3.3329
10	3.7168	3.6209	3.5217	3.4185	3.3654	3.3110	3.2554	3.1984	3.1399	3.0798
11	3.5257	3.4296	3.3299	3.2261	3.1725	3.1176	3.0613	3.0035	2.9441	2.8828
12	3.3736	3.2773	3.1772	3.0728	3.0187	2.9633	2.9063	2.8478	2.7874	2.7249
13	3.2497	3.1532	3.0527	2.9477	2.8932	2.8372	2.7797	2.7204	2.6590	2.5955
14	3.1469	3.0502	2.9493	2.8437	2.7888	2.7324	2.6742	2.6142	2.5519	2.4872
15	3.0602	2.9633	2.8621	2.7559	2.7006	2.6437	2.5850	2.5242	2.4611	2.3953
16	2.9862	2.8890	2.7875	2.6808	2.6252	2.5678	2.5085	2.4471	2.3831	2.3163
17	2.9222	2.8249	2.7230	2.6158	2.5598	2.5020	2.4422	2.3801	2.3153	2.2474
18	2.8664	2.7689	2.6667	2.5590	2.5027	2.4445	2.3842	2.3214	2.2558	2.1869
19	2.8172	2.7196	2.6171	2.5089	2.4523	2.3937	2.3329	2.2696	2.2032	2.1333
20	2.7737	2.6758	2.5731	2.4645	2.4076	2.3486	2.2873	2.2234	2.1562	2.0853
21	2.7348	2.6368	2.5338	2.4247	2.3675	2.3082	2.2465	2.1819	2.1141	2.0422
22	2.6998	2.6017	2.4984	2.3890	2.3315	2.2718	2.2097	2.1446	2.0760	2.0032
23	2.6682	2.5699	2.4665	2.3567	2.2989	2.2389	2.1763	2.1107	2.0415	1.9677
24	2.6396	2.5411	2.4374	2.3273	2.2693	2.2090	2.1460	2.0799	2.0099	1.9353
25	2.6135	2.5149	2.4110	2.3005	2.2422	2.1816	2.1183	2.0516	1.9811	1.9055
26	2.5896	2.4908	2.3867	2.2759	2.2174	2.1565	2.0928	2.0257	1.9545	1.8781
27	2.5676	2.4688	2.3644	2.2533	2.1946	2.1334	2.0693	2.0018	1.9299	1.8527
28	2.5473	2.4484	2.3438	2.2324	2.1735	2.1121	2.0477	1.9797	1.9072	1.8291
29	2.5286	2.4295	2.3248	2.2131	2.1540	2.0923	2.0276	1.9591	1.8861	1.8072
30	2.5112	2.4120	2.3072	2.1952	2.1359	2.0739	2.0089	1.9400	1.8664	1.7867
40	2.3882	2.2882	2.1819	2.0677	2.0069	1.9429	1.8752	1.8028	1.7242	1.6371
60	2.2702	2.1692	2.0613	1.9445	1.8817	1.8152	1.7440	1.6668	1.5810	1.4821
120	2.1570	2.0548	1.9450	1.8249	1.7597	1.6899	1.6141	1.5299	1.4327	1.3104
∞	2.0483	1.9447	1.8326	1.7085	1.6402	1.5660	1.4835	1.3883	1.2684	1.0000

Grados de libertad del denominador (g_2)De Maxine Merrington y Catherine M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution", *Biometrika* 33 (1943): 80-84. Reproducido con permiso de Biometrika Trustees.

(continúa)

TABLA A-5 Distribución F ($\alpha = 0.05$ en la cola derecha)

	Grados de libertad del numerador (g_1)								
	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

Grados de libertad del denominador (g_2)

(continúa)

TABLA A-5 Distribución F ($\alpha = 0.05$ en la cola derecha) (continuación)

	Grados de libertad del numerador (g_1)									
	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487	19.496
3	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

Grados de libertad del denominador (g_2)

De Maxine Merrington y Catherine M. Thompson, "Tables of Percentage Points of the Inverted Beta (f) Distribution", *Biometrika* 33 (1943): 80-84. Reproducido con permiso de Biometrika Trustees.

TABLA A-6		Valores críticos del coeficiente de correlación de Pearson r
n	$\alpha = .05$	$\alpha = .01$
4	.950	.999
5	.878	.959
6	.811	.917
7	.754	.875
8	.707	.834
9	.666	.798
10	.632	.765
11	.602	.735
12	.576	.708
13	.553	.684
14	.532	.661
15	.514	.641
16	.497	.623
17	.482	.606
18	.468	.590
19	.456	.575
20	.444	.561
25	.396	.505
30	.361	.463
35	.335	.430
40	.312	.402
45	.294	.378
50	.279	.361
60	.254	.330
70	.236	.305
80	.220	.286
90	.207	.269
100	.196	.256

NOTA: Para probar $H_0: \rho = 0$ contra $H_1: \rho \neq 0$, rechace H_0 si el valor absoluto de r es mayor que el valor crítico en la tabla.

TABLA A-7 Valores críticos para la prueba del signo

n	α			
	.005 (una cola)	.01 (una cola) .02 (dos colas)	.025 (una cola) .05 (dos colas)	.05 (una cola) .10 (dos colas)
	.01 (dos colas)			
1	*	*	*	*
2	*	*	*	*
3	*	*	*	*
4	*	*	*	*
5	*	*	*	0
6	*	*	0	0
7	*	0	0	0
8	0	0	0	1
9	0	0	1	1
10	0	0	1	1
11	0	1	1	2
12	1	1	2	2
13	1	1	2	3
14	1	2	2	3
15	2	2	3	3
16	2	2	3	4
17	2	3	4	4
18	3	3	4	5
19	3	4	4	5
20	3	4	5	5
21	4	4	5	6
22	4	5	5	6
23	4	5	6	7
24	5	5	6	7
25	5	6	7	7

NOTAS:

1. * indica que no es posible obtener un valor en la región crítica.
2. Rechace la hipótesis nula si el número del signo menos frecuente (x) es menor que o igual al valor en la tabla.
3. Para valores de n mayores que 25, se utiliza una aproximación normal con

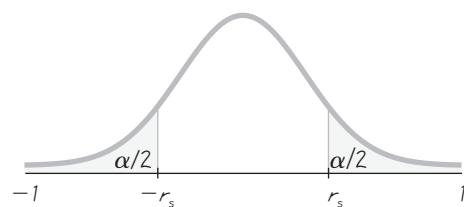
$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

		α			
		.005 (una cola)	.01 (una cola)	.025 (una cola)	.05 (una cola)
n		.01 (dos colas)	.02 (dos colas)	.05 (dos colas)	.10 (dos colas)
5		*	*	*	1
6		*	*	1	2
7		*	0	2	4
8		0	2	4	6
9		2	3	6	8
10		3	5	8	11
11		5	7	11	14
12		7	10	14	17
13		10	13	17	21
14		13	16	21	26
15		16	20	25	30
16		19	24	30	36
17		23	28	35	41
18		28	33	40	47
19		32	38	46	54
20		37	43	52	60
21		43	49	59	68
22		49	56	66	75
23		55	62	73	83
24		61	69	81	92
25		68	77	90	101
26		76	85	98	110
27		84	93	107	120
28		92	102	117	130
29		100	111	127	141
30		109	120	137	152

NOTAS:

1. * indica que no es posible obtener un valor en la región crítica.
2. Rechace la hipótesis nula si el estadístico de prueba T es menor que o igual al valor crítico encontrado en esta tabla. No rechace la hipótesis nula si el estadístico de prueba T es mayor que el valor crítico encontrado en la tabla.

De *Some Rapid Approximate Statistical Procedures*, Copyright © 1949, 1964 Lederle Laboratories Division of American Cyanamid Company. Reimpreso con permiso de la American Cyanamid Company.

**TABLA A-9**Valores críticos del coeficiente de correlación de rangos de Spearman r_s

n	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	.929
8	.643	.738	.833	.881
9	.600	.700	.783	.833
10	.564	.648	.745	.794
11	.536	.618	.709	.755
12	.503	.587	.678	.727
13	.484	.560	.648	.703
14	.464	.538	.626	.679
15	.446	.521	.604	.654
16	.429	.503	.582	.635
17	.414	.485	.566	.615
18	.401	.472	.550	.600
19	.391	.460	.535	.584
20	.380	.447	.520	.570
21	.370	.435	.508	.556
22	.361	.425	.496	.544
23	.353	.415	.486	.532
24	.344	.406	.476	.521
25	.337	.398	.466	.511
26	.331	.390	.457	.501
27	.324	.382	.448	.491
28	.317	.375	.440	.483
29	.312	.368	.433	.475
30	.306	.362	.425	.467

NOTAS:

1. Para $n > 30$, utilice $r_s = \pm z/\sqrt{n-1}$ donde z corresponde al nivel de significancia. Por ejemplo, si $\alpha = 0.05$, then $z = 1.96$.
2. Si el valor absoluto del estadístico de prueba r_s excede al valor crítico positivo, entonces rechace $H_0: \rho_s = 0$ y concluya que existe una correlación.

Basado en datos de "Biostatistical Analysis, 4th edition", © 1999, de Jerrold Zar, Prentice Hall, Inc., Upper Saddle River, Nueva Jersey, y "Distribution of Sums of Squares of Rank Differences to Small Numbers with Individuals", *The Annals of Mathematical Statistics*, vol. 9, núm. 2, con permiso del Institute of Mathematical Statistics.

TABLA A-10 Valores críticos para el número de rachas G

		Valor de n_2																		
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Valor de n_1	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
3	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
	6	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
4	1	1	1	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4
	6	8	9	9	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
5	1	1	2	2	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5
	6	8	9	10	10	11	11	12	12	12	12	12	12	12	12	12	12	12	12	12
6	1	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6
	6	8	9	10	11	12	12	13	13	13	13	14	14	14	14	14	14	14	14	14
7	1	2	2	3	3	3	4	4	4	5	5	5	5	5	6	6	6	6	6	6
	6	8	10	11	12	13	13	14	14	14	14	15	15	15	15	16	16	16	16	16
8	1	2	3	3	3	4	4	4	5	5	5	6	6	6	6	7	7	7	7	7
	6	8	10	11	12	13	14	14	14	15	15	16	16	16	16	17	17	17	17	17
9	1	2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8	8
	6	8	10	12	13	14	14	15	15	16	16	16	17	17	18	18	18	18	18	18
10	1	2	3	3	4	5	5	5	5	6	6	7	7	7	7	8	8	8	8	9
	6	8	10	12	13	14	15	15	16	16	17	17	18	18	18	19	19	19	19	20
11	1	2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9	9
	6	8	10	12	13	14	15	16	17	17	18	19	19	19	20	20	20	20	21	21
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	9	10	10
	6	8	10	12	13	14	16	16	17	18	19	19	20	20	21	21	21	21	22	22
13	2	2	3	4	5	5	6	6	7	7	8	8	8	9	9	9	10	10	10	10
	6	8	10	12	14	15	16	17	18	19	19	20	20	21	21	22	22	23	23	23
14	2	2	3	4	5	5	6	7	7	8	8	8	9	9	9	10	10	10	11	11
	6	8	10	12	14	15	16	17	18	19	20	20	21	22	22	23	23	23	24	24
15	2	3	3	4	5	6	6	7	7	8	8	8	9	9	10	10	11	11	11	12
	6	8	10	12	14	15	16	18	18	19	20	21	22	22	23	23	24	24	25	25
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	10	11	11	11	12	12
	6	8	10	12	14	16	17	18	19	20	21	21	22	23	23	24	25	25	25	25
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13	13
	6	8	10	12	14	16	17	18	19	20	21	22	23	23	24	25	25	26	26	26
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13	13
	6	8	10	12	14	16	17	18	19	20	21	22	23	24	25	25	26	26	27	27
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13	13
	6	8	10	12	14	16	17	18	20	21	22	23	23	24	25	26	26	27	27	27
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14	14
	6	8	10	12	14	16	17	18	20	21	22	23	24	25	25	26	27	27	28	28

NOTAS:

- Los valores en esta tabla son los valores críticos G , suponiendo una prueba de dos colas con un nivel de significancia de $\alpha = 0.05$.
- La hipótesis nula de aleatoriedad se rechaza si el número total de rachas G es menor que o igual al valor más bajo, o si es mayor que o igual al valor más alto.

De "Tables for Testing Randomness of Groupings in a Sequence of Alternatives", *The Annals of Mathematical Statistics*, vol. 14, núm. 1. Reproducido con permiso del Institute of Mathematical Statistics.

Apéndice B: Conjuntos de datos

- Conjunto de datos 1:** Resultados de examen de salud
Conjunto de datos 2: Temperaturas corporales de adultos saludables
Conjunto de datos 3: Alquitrán, nicotina y monóxido de carbono de cigarrillos
Conjunto de datos 4: Fumadores activos y pasivos
Conjunto de datos 5: Consumo de alcohol y tabaco en películas de dibujos animados para niños
Conjunto de datos 6: Osos (osos salvajes anestesiados)
Conjunto de datos 7: Pesos de álamos
Conjunto de datos 8: Temperaturas reales y pronosticadas
Conjunto de datos 9: Consumo de electricidad de una casa
Conjunto de datos 10: Precipitación pluvial en Boston para un año
Conjunto de datos 11: Géiser Old Faithful
Conjunto de datos 12: Pesos y volúmenes de bebidas de cola
Conjunto de datos 13: Pesos de una muestra de dulces M&M clásicos
Conjunto de datos 14: Pesos de monedas
Conjunto de datos 15: Cargas axiales de latas de aluminio
Conjunto de datos 16: Pesos de desechos de basura de una semana
Conjunto de datos 17: Distancias de *home runs*
Conjunto de datos 18: Casas vendidas en el Condado Dutchess, Nueva York

En el sitio Web www.pearsoneducacion.net/triola encontrará conjuntos de datos adicionales

Conjunto de datos 1: Resultados de examen de salud

La EDAD está dada en años, EST es estatura (pulgadas), PE es peso (libras), CINT es circunferencia de la cintura (cm), PULSO es frecuencia del pulso (en latidos por minuto), SIST es presión sanguínea sistólica (mmHg), DIAS es presión sanguínea diastólica (mmHg), COL es colesterol (mg), IMC es índice de masa corporal, MUS es longitud del muslo (cm), CODO es anchura del codo (cm), MUÑ es anchura de la muñeca (cm) y BRA es circunferencia del brazo (cm). Los datos son del Departamento de Salud y Servicios Humanos de Estados Unidos (National Center for Health Statistics, Third National Health and Nutrition Examination Survey).

STATDISK: El nombre del conjunto de datos para hombres es Mhealth.

Minitab: El nombre de la hoja de cálculo para hombres es MHEALTH.MTW.

Excel: El nombre del libro de trabajo para hombres es MHEALTH.XLS.

TI-83/84 Plus: El nombre de la App para datos de hombres es MHEALTH y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto para hombres: MAGE, MHT, MWT, MWAST, MPULS, MSYS, MDIAS, MCHOL, MBMI, MLEG, MELBW, MWRST, MARM.

Hombre	Edad	Est	Pe	Cint	Pulso	Sist	Dias	Col	Imc	Mus	Codo	Muñ	Bra
58	70.8	169.1	90.6	68	125	78	522	23.8	42.5	7.7	6.4	31.9	
22	66.2	144.2	78.1	64	107	54	127	23.2	40.2	7.6	6.2	31.0	
32	71.7	179.3	96.5	88	126	81	740	24.6	44.4	7.3	5.8	32.7	
31	68.7	175.8	87.7	72	110	68	49	26.2	42.8	7.5	5.9	33.4	
28	67.6	152.6	87.1	64	110	66	230	23.5	40.0	7.1	6.0	30.1	
46	69.2	166.8	92.4	72	107	83	316	24.5	47.3	7.1	5.8	30.5	
41	66.5	135.0	78.8	60	113	71	590	21.5	43.4	6.5	5.2	27.6	
56	67.2	201.5	103.3	88	126	72	466	31.4	40.1	7.5	5.6	38.0	
20	68.3	175.2	89.1	76	137	85	121	26.4	42.1	7.5	5.5	32.0	
54	65.6	139.0	82.5	60	110	71	578	22.7	36.0	6.9	5.5	29.3	
17	63.0	156.3	86.7	96	109	65	78	27.8	44.2	7.1	5.3	31.7	
73	68.3	186.6	103.3	72	153	87	265	28.1	36.7	8.1	6.7	30.7	
52	73.1	191.1	91.8	56	112	77	250	25.2	48.4	8.0	5.2	34.7	
25	67.6	151.3	75.6	64	119	81	265	23.3	41.0	7.0	5.7	30.6	
29	68.0	209.4	105.5	60	113	82	273	31.9	39.8	6.9	6.0	34.2	
17	71.0	237.1	108.7	64	125	76	272	33.1	45.2	8.3	6.6	41.1	
41	61.3	176.7	104.0	84	131	80	972	33.2	40.2	6.7	5.7	33.1	
52	76.2	220.6	103.0	76	121	75	75	26.7	46.2	7.9	6.0	32.2	
32	66.3	166.1	91.3	84	132	81	138	26.6	39.0	7.5	5.7	31.2	
20	69.7	137.4	75.2	88	112	44	139	19.9	44.8	6.9	5.6	25.9	
20	65.4	164.2	87.7	72	121	65	638	27.1	40.9	7.0	5.6	33.7	
29	70.0	162.4	77.0	56	116	64	613	23.4	43.1	7.5	5.2	30.3	
18	62.9	151.8	85.0	68	95	58	762	27.0	38.0	7.4	5.8	32.8	
26	68.5	144.1	79.6	64	110	70	303	21.6	41.0	6.8	5.7	31.0	
33	68.3	204.6	103.8	60	110	66	690	30.9	46.0	7.4	6.1	36.2	
55	69.4	193.8	103.0	68	125	82	31	28.3	41.4	7.2	6.0	33.6	
53	69.2	172.9	97.1	60	124	79	189	25.5	42.7	6.6	5.9	31.9	
28	68.0	161.9	86.9	60	131	69	957	24.6	40.5	7.3	5.7	32.9	
28	71.9	174.8	88.0	56	109	64	339	23.8	44.2	7.8	6.0	30.9	
37	66.1	169.8	91.5	84	112	79	416	27.4	41.8	7.0	6.1	34.0	
40	72.4	213.3	102.9	72	127	72	120	28.7	47.2	7.5	5.9	34.8	
33	73.0	198.0	93.1	84	132	74	702	26.2	48.2	7.8	6.0	33.6	
26	68.0	173.3	98.9	88	116	81	1252	26.4	42.9	6.7	5.8	31.3	
53	68.7	214.5	107.5	56	125	84	288	32.1	42.8	8.2	5.9	37.6	
36	70.3	137.1	81.6	64	112	77	176	19.6	40.8	7.1	5.3	27.9	
34	63.7	119.5	75.7	56	125	77	277	20.7	42.6	6.6	5.3	26.9	
42	71.1	189.1	95.0	56	120	83	649	26.3	44.9	7.4	6.0	36.9	
18	65.6	164.7	91.1	60	118	68	113	26.9	41.1	7.0	6.1	34.5	
44	68.3	170.1	94.9	64	115	75	656	25.6	44.5	7.3	5.8	32.1	
20	66.3	151.0	79.9	72	115	65	172	24.2	44.0	7.1	5.4	30.7	

(continúa)

Conjunto de datos 1: Resultados de examen de salud (*continuación*)



STATDISK: El nombre del conjunto de datos para mujeres es Fhealth.
Minitab: El nombre de la hoja de cálculo para mujeres es FHEALTH.MTW.
Excel: El nombre del libro de trabajo para mujeres es FHEALTH.XLS.
TI-83/84 Plus: El nombre de la App para datos de mujeres es FHEALTH y los nombres de los archivos son los mismos que para los archivos de texto.
Nombres de los archivos de texto para mujeres: FAGE, FHT, FWT, FWAST, FPULS, FSYS, FDIAS, FCHOL, FBMI, FLEG, FELBW, FWRST, FARM.

Mujer	Edad	Est	Pe	Cint	Pulso	Sist	Dias	Col	Imc	Mus	Codo	Muñ	Bra
	17	64.3	114.8	67.2	76	104	61	264	19.6	41.6	6.0	4.6	23.6
	32	66.4	149.3	82.5	72	99	64	181	23.8	42.8	6.7	5.5	26.3
	25	62.3	107.8	66.7	88	102	65	267	19.6	39.0	5.7	4.6	26.3
	55	62.3	160.1	93.0	60	114	76	384	29.1	40.2	6.2	5.0	32.6
	27	59.6	127.1	82.6	72	94	58	98	25.2	36.2	5.5	4.8	29.2
	29	63.6	123.1	75.4	68	101	66	62	21.4	43.2	6.0	4.9	26.4
	25	59.8	111.7	73.6	80	108	61	126	22.0	38.7	5.7	5.1	27.9
	12	63.3	156.3	81.4	64	104	41	89	27.5	41.0	6.8	5.5	33.0
	41	67.9	218.8	99.4	68	123	72	531	33.5	43.8	7.8	5.8	38.6
	32	61.4	110.2	67.7	68	93	61	130	20.6	37.3	6.3	5.0	26.5
	31	66.7	188.3	100.7	80	89	56	175	29.9	42.3	6.6	5.2	34.4
	19	64.8	105.4	72.9	76	112	62	44	17.7	39.1	5.7	4.8	23.7
	19	63.1	136.1	85.0	68	107	48	8	24.0	40.3	6.6	5.1	28.4
	23	66.7	182.4	85.7	72	116	62	112	28.9	48.6	7.2	5.6	34.0
	40	66.8	238.4	126.0	96	181	102	462	37.7	33.2	7.0	5.4	35.2
	23	64.7	108.8	74.5	72	98	61	62	18.3	43.4	6.2	5.2	24.7
	27	65.1	119.0	74.5	68	100	53	98	19.8	41.5	6.3	5.3	27.0
	45	61.9	161.9	94.0	72	127	74	447	29.8	40.0	6.8	5.0	35.0
	41	64.3	174.1	92.8	64	107	67	125	29.7	38.2	6.8	4.7	33.1
	56	63.4	181.2	105.5	80	116	71	318	31.7	38.2	6.9	5.4	39.6
	22	60.7	124.3	75.5	64	97	64	325	23.8	38.2	5.9	5.0	27.0
	57	63.4	255.9	126.5	80	155	85	600	44.9	41.0	8.0	5.6	43.8
	24	62.6	106.7	70.0	76	106	59	237	19.2	38.1	6.1	5.0	23.6
	37	60.6	149.9	98.0	76	110	70	173	28.7	38.0	7.0	5.1	34.3
	59	63.5	163.1	104.7	76	105	69	309	28.5	36.0	6.7	5.1	34.4
	40	58.6	94.3	67.8	80	118	82	94	19.3	32.1	5.4	4.2	23.3
	45	60.2	159.7	99.3	104	133	83	280	31.0	31.1	6.4	5.2	35.6
	52	67.6	162.8	91.1	88	113	75	254	25.1	39.4	7.1	5.3	31.8
	31	63.4	130.0	74.5	60	113	66	123	22.8	40.2	5.9	5.1	27.0
	32	64.1	179.9	95.5	76	107	67	596	30.9	39.2	6.2	5.0	32.8
	23	62.7	147.8	79.5	72	95	59	301	26.5	39.0	6.3	4.9	31.0
	23	61.3	112.9	69.1	72	108	72	223	21.2	36.6	5.9	4.7	27.0
	47	58.2	195.6	105.5	88	114	79	293	40.6	27.0	7.5	5.5	41.2
	36	63.2	124.2	78.8	80	104	73	146	21.9	38.5	5.6	4.7	25.5
	34	60.5	135.0	85.7	60	125	73	149	26.0	39.9	6.4	5.2	30.9
	37	65.0	141.4	92.8	72	124	85	149	23.5	37.5	6.1	4.8	27.9
	18	61.8	123.9	72.7	88	92	46	920	22.8	39.7	5.8	5.0	26.5
	29	68.0	135.5	75.9	88	119	81	271	20.7	39.0	6.3	4.9	27.8
	48	67.0	130.4	68.6	124	93	64	207	20.5	41.6	6.0	5.3	23.0
	16	57.0	100.7	68.7	64	106	64	2	21.9	33.8	5.6	4.6	26.4

Conjunto de datos 2: Temperaturas corporales (en grados Fahrenheit) de adultos saludables

Datos proporcionados por los doctores Steven Wasserman, Philip Mackowiak y Myron Levine de la University of Maryland.



STATDISK: El nombre del conjunto de datos para las temperaturas de las 12 A.M. del día 2 es Bodytemp.

Minitab: La hoja de cálculo para las temperaturas de las 12 A.M. del día 2 es BODYTEMP.MTW.

Excel: El nombre del libro de trabajo para las temperaturas de las 12 A.M. del día 2 es BODYTEMP.XLS.

TI-83/84 Plus: El nombre de la App para las temperaturas de las 12 A.M. del día 2 es BTEMP y el nombre del archivo es BTEMP.

Archivos de texto: El nombre del archivo de texto es BTEMP.

Sujeto	Edad	Género	Fuma	Temperatura día 1		Temperatura día 2	
				8 AM	12 AM	8 AM	12 AM
1	22	H	S	98.0	98.0	98.0	98.6
2	23	H	S	97.0	97.6	97.4	—
3	22	H	S	98.6	98.8	97.8	98.6
4	19	H	N	97.4	98.0	97.0	98.0
5	18	H	N	98.2	98.8	97.0	98.0
6	20	H	S	98.2	98.8	96.6	99.0
7	27	H	S	98.2	97.6	97.0	98.4
8	19	H	S	96.6	98.6	96.8	98.4
9	19	H	S	97.4	98.6	96.6	98.4
10	24	H	N	97.4	98.8	96.6	98.4
11	35	H	S	98.2	98.0	96.2	98.6
12	25	H	S	97.4	98.2	97.6	98.6
13	25	H	N	97.8	98.0	98.6	98.8
14	35	H	S	98.4	98.0	97.0	98.6
15	21	H	N	97.6	97.0	97.4	97.0
16	33	H	N	96.2	97.2	98.0	97.0
17	19	H	S	98.0	98.2	97.6	98.8
18	24	H	S	—	—	97.2	97.6
19	18	M	N	—	—	97.0	97.7
20	22	M	S	—	—	98.0	98.8
21	20	H	S	—	—	97.0	98.0
22	30	M	S	—	—	96.4	98.0
23	29	H	N	—	—	96.1	98.3
24	18	H	S	—	—	98.0	98.5
25	31	H	S	—	98.1	96.8	97.3
26	28	F	S	—	98.2	98.2	98.7
27	27	H	S	—	98.5	97.8	97.4
28	21	H	S	—	98.5	98.2	98.9
29	30	H	S	—	99.0	97.8	98.6
30	27	H	N	—	98.0	99.0	99.5
31	32	H	S	—	97.0	97.4	97.5
32	33	H	S	—	97.3	97.4	97.3
33	23	H	S	—	97.3	97.5	97.6

(continúa)

Conjunto de datos 2: Temperaturas corporales (*continuación*)

Sujeto	Edad	Género	Fuma	Temperatura día 1		Temperatura día 2	
				8 AM	12 AM	8 AM	12 AM
34	29	H	S	—	98.1	97.8	98.2
35	25	H	S	—	—	97.9	99.6
36	31	H	N	—	97.8	97.8	98.7
37	25	H	S	—	99.0	98.3	99.4
38	28	H	N	—	97.6	98.0	98.2
39	30	H	S	—	97.4	—	98.0
40	33	H	S	—	98.0	—	98.6
41	28	H	S	98.0	97.4	—	98.6
42	22	H	S	98.8	98.0	—	97.2
43	21	M	S	99.0	—	—	98.4
44	30	H	N	—	98.6	—	98.6
45	22	H	S	—	98.6	—	98.2
46	22	M	N	98.0	98.4	—	98.0
47	20	H	S	—	97.0	—	97.8
48	19	H	S	—	—	—	98.0
49	33	H	N	—	98.4	—	98.4
50	31	H	S	99.0	99.0	—	98.6
51	26	H	N	—	98.0	—	98.6
52	18	H	N	—	—	—	97.8
53	23	H	N	—	99.4	—	99.0
54	28	H	S	—	—	—	96.5
55	19	H	S	—	97.8	—	97.6
56	21	H	N	—	—	—	98.0
57	27	H	S	—	98.2	—	96.9
58	29	H	S	—	99.2	—	97.6
59	38	H	N	—	99.0	—	97.1
60	29	M	S	—	97.7	—	97.9
61	22	H	S	—	98.2	—	98.4
62	22	H	S	—	98.2	—	97.3
63	26	H	S	—	98.8	—	98.0
64	32	H	N	—	98.1	—	97.5
65	25	H	S	—	98.5	—	97.6
66	21	M	N	—	97.2	—	98.2
67	25	H	S	—	98.5	—	98.5
68	24	H	S	—	99.2	97.0	98.8
69	25	H	S	—	98.3	97.6	98.7
70	35	H	S	—	98.7	97.5	97.8
71	23	M	S	—	98.8	98.8	98.0
72	31	H	S	—	98.6	98.4	97.1
73	28	H	S	—	98.0	98.2	97.4
74	29	H	S	—	99.1	97.7	99.4
75	26	H	S	—	97.2	97.3	98.4
76	32	H	N	—	97.6	97.5	98.6
77	32	H	S	—	97.9	97.1	98.4
78	21	M	S	—	98.8	98.6	98.5
79	20	H	S	—	98.6	98.6	98.6

(continúa)

Conjunto de datos 4: Temperaturas corporales (*continuación*)

Sujeto	Edad	Género	Fuma	Temperatura día 1		Temperatura día 2	
				8 AM	12 AM	8 AM	12 AM
80	24	M	S	—	98.6	97.8	98.3
81	21	M	S	—	99.3	98.7	98.7
82	28	H	S	—	97.8	97.9	98.8
83	27	M	N	98.8	98.7	97.8	99.1
84	28	H	N	99.4	99.3	97.8	98.6
85	29	H	S	98.8	97.8	97.6	97.9
86	19	H	N	97.7	98.4	96.8	98.8
87	24	H	S	99.0	97.7	96.0	98.0
88	29	H	N	98.1	98.3	98.0	98.7
89	25	H	S	98.7	97.7	97.0	98.5
90	27	H	N	97.5	97.1	97.4	98.9
91	25	H	S	98.9	98.4	97.6	98.4
92	21	H	S	98.4	98.6	97.6	98.6
93	19	H	S	97.2	97.4	96.2	97.1
94	27	H	S	—	—	96.2	97.9
95	32	H	N	98.8	96.7	98.1	98.8
96	24	H	S	97.3	96.9	97.1	98.7
97	32	H	S	98.7	98.4	98.2	97.6
98	19	M	S	98.9	98.2	96.4	98.2
99	18	M	S	99.2	98.6	96.9	99.2
100	27	H	N	—	97.0	—	97.8
101	34	H	S	—	97.4	—	98.0
102	25	H	NN	—	98.4	—	98.4
103	18	H	N	—	97.4	—	97.8
104	32	H	S	—	96.8	—	98.4
105	31	H	S	—	98.2	—	97.4
106	26	H	N	—	97.4	—	98.0
107	23	H	N	—	98.0	—	97.0

Conjunto de datos 3: Alquitrán, nicotina y monóxido de carbono de cigarrillos

Todas las mediciones son en miligramos por cigarrillo, y todos los cigarrillos son de 100 mm de largo, con filtro, y no son del tipo mentolado o "light". Los datos son de la Federal Trade Commission.



- STATDISK:** El nombre del conjunto de datos es Cigaret.
Minitab: El nombre de la hoja de cálculo es CIGARET.MTW.
Excel: El nombre del libro de trabajo es CIGARET.XLS.
TI-83/84 Plus: El nombre de la App es CIGARET y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: TAR, NICOT, CO.

Marca	Alquitrán	Nicotina	CO
American Filter	16	1.2	15
Benson & Hedges	16	1.2	15
Camel	16	1.0	17
Capri	9	0.8	6
Carlton	1	0.1	1
Cartier Vendome	8	0.8	8
Chelsea	10	0.8	10
GPC Approved	16	1.0	17
Hi-Lite	14	1.0	13
Kent	13	1.0	13
Lucky Strike	13	1.1	13
Malibu	15	1.2	15
Marlboro	16	1.2	15
Merit	9	0.7	11
Newport Stripe	11	0.9	15
Now	2	0.2	3
Old Gold	18	1.4	18
Pall Mall	15	1.2	15
Players	13	1.1	12
Raleigh	15	1.0	16
Richland	17	1.3	16
Rite	9	0.8	10
Silva Thins	12	1.0	10
Tareyton	14	1.0	17
Triumph	5	0.5	7
True	6	0.6	7
Vantage	8	0.7	11
Viceroy	18	1.4	15
Winston	16	1.1	18

Conjunto de datos 4: Fumadores activos y pasivos

Todos los valores son mediciones de niveles de cotinina en suero (en ng/ml), un metabolito de la nicotina. (Cuando el cuerpo absorbe la nicotina, se produce la cotinina). Los datos son del Departamento de Salud y Servicios Humanos de Estados Unidos (National Center for Health Statistics, Third National Health and Nutrition Examination Survey).



STATDISK: El nombre del conjunto de datos es COTININE.
Minitab: El nombre de la hoja de cálculo es COTININE.MTW.
Excel: El nombre del libro de trabajo es COTININE.XLS.
TI-83/84 Plus: El nombre de la App es COTININE y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: NOETS, ETS, SMKR.

Fumadores (sujetos que reportan consumo de tabaco)

1	0	131	173	265	210	44	277	32	3
35	112	477	289	227	103	222	149	313	491
130	234	164	198	17	253	87	121	266	290
123	167	250	245	48	86	284	1	208	173

HTA (no fumadores expuestos al humo de tabaco ambiental)

384	0	69	19	1	0	178	2	13	1
4	0	543	17	1	0	51	0	197	3
0	3	1	45	13	3	1	1	1	0
0	551	2	1	1	1	0	74	1	241

SHTA (no fumadores sin exposición al humo de tabaco ambiental)

Conjunto de datos 5: Consumo de alcohol y tabaco en películas de dibujos animados para niños

La duración de las películas es en minutos, los tiempos de consumo de tabaco son en segundos, y los tiempos de consumo de alcohol son en segundos. Los datos se basan en "Tobacco and Alcohol Use in G-Rated Children's Animated Films", de Goldstein, Sobel y Newman, *Journal of the American Medical Association*, vol. 281, núm. 12.



STATDISK:

El nombre del conjunto de datos es Chmovie.

Minitab:

El nombre de la hoja de cálculo es CHMOVIE.MTW.

Excel:

El nombre del libro de trabajo es CHMOVIE.XLS.

TI-83/84 Plus:

El nombre de la App es CHMOVIE y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: CHLEN, CHTOB, CHALC.

Película	Compañía	Duración (min)	Consumo de tabaco (s)	Consumo de alcohol (s)
Blanca Nieves	Disney	83	0	0
Pinocho	Disney	88	223	80
Fantasía	Disney	120	0	0
Dumbo	Disney	64	176	88
Bambi	Disney	69	0	0
Los tres caballeros	Disney	71	548	8
Mickey y las habichuelas mágicas	Disney	76	0	4
La Cenicienta	Disney	74	37	0
Alicia en el país de las maravillas	Disney	75	158	0
Peter Pan	Disney	76	51	33
La dama y el vagabundo	Disney	75	0	0
La bella durmiente	Disney	75	0	113
101 dálmatas	Disney	79	299	51
La espada en la piedra	Disney	80	37	20
El libro de la selva	Disney	78	0	0
Los Aristógatos	Disney	78	11	142
Robin Hood	Disney	83	0	39
Los rescatadores	Disney	77	0	0
Winnie Pooh	Disney	71	0	0
El zorro y el sabueso	Disney	83	0	0
El corsario negro	Disney	80	0	34
Policias y ratones	Disney	73	165	414
Oliver y su pandilla	Disney	72	74	0
La sirenita	Disney	82	9	0
Los rescatadores en Cangurolandia	Disney	74	0	76
La bella y la bestia	Disney	84	0	123
Aladino	Disney	90	2	3
El rey león	Disney	89	0	0
Pocahontas	Disney	81	6	7
Toy Story	Disney	81	0	0
El jorobado de Notre Dame	Disney	90	23	46
James y el melocotón gigante	Disney	79	206	38
Hércules	Disney	92	9	13
NIMH, el ratoncito valiente	MGM	82	0	0
Todos los perros van al cielo	MGM	89	205	73
Todos los perros van al cielo 2	MGM	82	162	72
Jake y Jill en Villa de Juguete	MGM	74	0	0
Pulgarcita	Warner Bros	86	6	5
El jardín mágico de Stanley	Warner Bros	76	1	0
Space Jam	Warner Bros	81	117	0
Pippi Mediaslargas	Warner Bros	75	5	0
Los gatos no bailan	Warner Bros	75	91	0
Fievel y el Nuevo Mundo	Universal	77	155	74
En busca del valle encantado	Universal	70	0	0
Fievel va al Oeste	Universal	75	24	28
Rex, un dinosaurio en Nueva York	Universal	64	55	0
En busca del valle encantado 2	Universal	73	0	0
Balto	Universal	74	0	0
El bosque de colores	20th Century Fox	71	0	0
Anastasia	20th Century Fox	94	17	39

Conjunto de datos 6:

Oso (osos salvajes anestesiados)

La EDAD está en meses, MES es el mes de la medición (1 = enero), SEXO está codificado como 1 = macho y 2 = hembra, CABEZAL es la longitud de la cabeza (pulgadas), CABEZA es la anchura de la cabeza (pulgadas), CUELLO es la circunferencia del cuello (pulgadas), ESTAT es la estatura del cuerpo (pulgadas), PECHO es la circunferencia torácica (pulgadas) y el PESO está medido en libras. Los datos son de Gary Alt y Minitab, Inc.



STATDISK: El nombre del conjunto de datos es Bears.

Minitab: El nombre de la hoja de cálculo es BEARS.MTW.

Excel: El nombre del libro de trabajo es BEARS.XLS.

TI-83/84 Plus: El nombre de la App es BEARS y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: BAGE, BMNTH, BSEX, BHDLN, BHDWD, BNECK, BLEN, BCHST, BWGHT.

	Edad	Mes	Sexo	Cabezal	CabezaA	Cuello	Estat	Pecho	Peso
	19	7	1	11.0	5.5	16.0	53.0	26.0	80
	55	7	1	16.5	9.0	28.0	67.5	45.0	344
	81	9	1	15.5	8.0	31.0	72.0	54.0	416
	115	7	1	17.0	10.0	31.5	72.0	49.0	348
	104	8	2	15.5	6.5	22.0	62.0	35.0	166
	100	4	2	13.0	7.0	21.0	70.0	41.0	220
	56	7	1	15.0	7.5	26.5	73.5	41.0	262
	51	4	1	13.5	8.0	27.0	68.5	49.0	360
	57	9	2	13.5	7.0	20.0	64.0	38.0	204
	53	5	2	12.5	6.0	18.0	58.0	31.0	144
	68	8	1	16.0	9.0	29.0	73.0	44.0	332
	8	8	1	9.0	4.5	13.0	37.0	19.0	34
	44	8	2	12.5	4.5	10.5	63.0	32.0	140
	32	8	1	14.0	5.0	21.5	67.0	37.0	180
	20	8	2	11.5	5.0	17.5	52.0	29.0	105
	32	8	1	13.0	8.0	21.5	59.0	33.0	166
	45	9	1	13.5	7.0	24.0	64.0	39.0	204
	9	9	2	9.0	4.5	12.0	36.0	19.0	26
	21	9	1	13.0	6.0	19.0	59.0	30.0	120
	177	9	1	16.0	9.5	30.0	72.0	48.0	436
	57	9	2	12.5	5.0	19.0	57.5	32.0	125
	81	9	2	13.0	5.0	20.0	61.0	33.0	132
	21	9	1	13.0	5.0	17.0	54.0	28.0	90
	9	9	1	10.0	4.0	13.0	40.0	23.0	40
	45	9	1	16.0	6.0	24.0	63.0	42.0	220
	9	9	1	10.0	4.0	13.5	43.0	23.0	46
	33	9	1	13.5	6.0	22.0	66.5	34.0	154
	57	9	2	13.0	5.5	17.5	60.5	31.0	116
	45	9	2	13.0	6.5	21.0	60.0	34.5	182
	21	9	1	14.5	5.5	20.0	61.0	34.0	150
	10	10	1	9.5	4.5	16.0	40.0	26.0	65
	82	10	2	13.5	6.5	28.0	64.0	48.0	356
	70	10	2	14.5	6.5	26.0	65.0	48.0	316
	10	10	1	11.0	5.0	17.0	49.0	29.0	94
	10	10	1	11.5	5.0	17.0	47.0	29.5	86
	34	10	1	13.0	7.0	21.0	59.0	35.0	150
	34	10	1	16.5	6.5	27.0	72.0	44.5	270
	34	10	1	14.0	5.5	24.0	65.0	39.0	202
	58	10	2	13.5	6.5	21.5	63.0	40.0	202
	58	10	1	15.5	7.0	28.0	70.5	50.0	365
	11	11	1	11.5	6.0	16.5	48.0	31.0	79
	23	11	1	12.0	6.5	19.0	50.0	38.0	148
	70	10	1	15.5	7.0	28.0	76.5	55.0	446
	11	11	2	9.0	5.0	15.0	46.0	27.0	62
	83	11	2	14.5	7.0	23.0	61.5	44.0	236
	35	11	1	13.5	8.5	23.0	63.5	44.0	212
	16	4	1	10.0	4.0	15.5	48.0	26.0	60
	16	4	1	10.0	5.0	15.0	41.0	26.0	64
	17	5	1	11.5	5.0	17.0	53.0	30.5	114
	17	5	2	11.5	5.0	15.0	52.5	28.0	76
	17	5	2	11.0	4.5	13.0	46.0	23.0	48
	8	8	2	10.0	4.5	10.0	43.5	24.0	29
	83	11	1	15.5	8.0	30.5	75.0	54.0	514
	18	6	1	12.5	8.5	18.0	57.3	32.8	140

Conjunto de datos 7: Pesos de álamos

Los datos se tomaron de un estudio realizado por investigadores de Pennsylvania State University; los datos se obtuvieron de Minitab, Inc.



- STATDISK:** El nombre del conjunto de datos es Poplar.
Minitab: El nombre de la hoja de cálculo es POPLAR.MTW.
Excel: El nombre del libro de trabajo es POPLAR.XLS.
TI-83/84 Plus: El nombre de la App es POPLAR y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: Los nombres de los archivos son POPNO, POPFR, POPIR, POPFI.

Año 1

	Sin tratamiento	Fertilizante	Riego	Fertilizante y riego
Lugar 1 (fértil, húmedo)	0.15	1.34	0.23	2.03
	0.02	0.14	0.04	0.27
	0.16	0.02	0.34	0.92
	0.37	0.08	0.16	1.07
	0.22	0.08	0.05	2.38
Lugar 2 (arenoso, seco)	0.60	1.16	0.65	0.22
	1.11	0.93	0.08	2.13
	0.07	0.30	0.62	2.33
	0.07	0.59	0.01	1.74
	0.44	0.17	0.03	0.12

Año 2

	Sin tratamiento	Fertilizante	Riego	Fertilizante y riego
Lugar 1 (fértil, húmedo)	1.21	0.94	0.07	0.85
	0.57	0.87	0.66	1.78
	0.56	0.46	0.10	1.47
	0.13	0.58	0.82	2.25
	1.30	1.03	0.94	1.64
Lugar 2 (arenoso, seco)	0.24	0.92	0.96	1.07
	1.69	0.07	1.43	1.63
	1.23	0.56	1.26	1.39
	0.99	1.74	1.57	0.49
	1.80	1.13	0.72	0.95

Conjunto de datos 8: Temperaturas reales y pronosticadas

Las temperaturas están en grados Fahrenheit y las cantidades de precipitación en pulgadas. Todas las mediciones se registraron cerca de la casa del autor.



- STATDISK:** El nombre del conjunto de datos es Weather.
Minitab: El nombre de la hoja de cálculo es WEATHER.MTW.
Excel: El nombre del libro de trabajo es WEATHER.XLS.
TI-83/84 Plus: El nombre de la App es WEATHER y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: ACTHI, ACTLO, PHI1, PLO1, PHI3, PLO3, PHIS, PLO5 Y PREC.

Fecha	Real		Pronóstico		Pronóstico		Pronóstico		Pronóstico		Precip.
	máxima	mínima	1 día máxima	1 día mínima	3 días máxima	3 días mínima	5 días máxima	5 días mínima		(pulgadas)	
Sept. 1	80	54	78	52	79	52	80	56		0.00	
Sept. 2	77	54	75	53	86	63	80	57		0.00	
Sept. 3	81	55	81	55	79	59	79	59		0.00	
Sept. 4	85	60	85	62	83	59	80	56		0.00	
Sept. 5	73	64	76	53	80	63	79	64		0.00	
Sept. 6	73	51	75	58	76	61	82	57		0.00	
Sept. 7	80	59	79	66	80	63	76	61		0.00	
Sept. 8	72	61	74	66	79	67	73	63		0.47	
Sept. 9	83	68	75	66	76	66	77	59		1.59	
Sept. 10	81	62	80	53	79	58	83	61		0.07	
Sept. 11	75	53	75	51	78	58	77	58		0.01	
Sept. 12	78	52	79	55	75	54	79	56		0.00	
Sept. 13	80	56	80	53	74	48	74	50		0.01	
Sept. 14	71	56	70	53	73	55	75	52		0.00	
Sept. 15	73	54	72	60	73	59	76	54		0.00	
Sept. 16	78	64	79	63	76	60	78	62		0.06	
Sept. 17	75	62	75	60	76	56	76	58		0.01	
Sept. 18	63	55	67	43	73	52	75	60		2.85	
Sept. 19	63	48	64	43	75	53	77	55		0.00	
Sept. 20	70	40	69	46	68	53	71	50		0.00	
Sept. 21	77	47	77	50	77	51	74	54		0.00	
Sept. 22	82	49	81	55	83	54	73	56		0.00	
Sept. 23	81	53	81	51	78	57	75	53		0.01	
Sept. 24	76	51	80	53	75	54	79	56		0.00	
Sept. 25	77	54	78	54	77	51	74	53		0.00	
Sept. 26	76	58	76	50	72	46	71	44		0.00	
Sept. 27	74	48	76	60	74	56	70	45		0.01	
Sept. 28	66	61	70	49	74	49	73	52		1.99	
Sept. 29	66	57	69	41	68	41	72	48		0.67	
Sept. 30	62	53	68	45	72	50	69	47		0.21	
Oct. 1	71	51	75	49	72	49	70	47		0.02	
Oct. 2	68	46	71	47	73	52	68	46		0.01	
Oct. 3	66	44	68	42	66	43	67	46		0.05	
Oct. 4	71	39	69	44	68	44	61	40		0.00	
Oct. 5	58	43	56	29	62	38	64	42		0.00	

Conjunto de datos 9: Consumo de electricidad de una casa

Todas las mediciones se realizaron en la casa del autor.



- STATDISK:** El nombre del conjunto de datos es Electric.
Minitab: El nombre de la hoja de cálculo es ELECTRIC.MTW.
Excel: El nombre del libro de trabajo es ELECTRIC.XLS.
TI-83/84 Plus: El nombre de la App es ELECTRIC y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: Los nombres de los archivos de texto son KWH, COST, HDD, ADT.

Periodo	Consumo de electricidad (kWh)	Costo (dólares)	Energía necesaria para calefacción	Temperatura promedio diaria (°F)
Año 1: ene/feb	3375	321.94	2421	26
Año 1: mar/abr	2661	221.11	1841	34
Año 1: mayo/junio	2073	205.16	438	58
Año 1: julio/ago	2579	251.07	15	72
Año 1: sept/oct	2858	279.80	152	67
Año 1: nov/dic	2296	183.84	1028	48
Año 2: ene/feb	2812	244.93	1967	33
Año 2: mar/abr	2433	218.59	1627	39
Año 2: mayo/junio	2266	213.09	537	66
Año 2: julio/ago	3128	333.49	26	71
Año 2: sept/oct	3286	370.35	116	
Año 2: nov/dic	2749	222.79	1457	
Año 3: ene/feb	3427	316.18	253	
Año 3: mar/abr	578	77.39	1811	
Año 3: mayo/junio	3792	385.44	632	

Conjunto de datos 10: Cantidad de lluvia (en pulgadas) en Boston durante un año



	Lun	Mar	Miér	Jue	Vie	Sáb	Dom
STATDISK: El nombre del conjunto de datos es Bostrain.	0	0.44	0.14	0.04	0.04	0.64	0
	0.05	0	0	0	0.01	0.05	0
	0	0	0.64	0	0	0	0
Minitab: El nombre de la hoja de cálculo es BOSTRAIN.MTW.	0.01	0	0	0	0.3	0.05	0
	0	0	0.01	0.16	0	0	0.09
Excel: El nombre del libro de trabajo es BOSTRAIN.XLS.	0.12	0.06	0.18	0.39	0	0.1	0
	0	0	0	0	0.78	0.49	0
	0	0.02	0	0	0.01	0.17	0
TI-83/84 Plus:	1.41	0.65	0.31	0	0	0.54	0
El nombre de la App es BOSTRAIN y los nombres de los archivos son los mismos que para los archivos de texto.	0	0	0	0	0	0	0
	0	0	0	0	0	0.4	0.28
	0	0	0	0.3	0.87	0.49	0
	0.47	0	0	0	0	0	0
	0	0.09	0	0.24	0	0.05	0
	0	0.14	0	0	0.04	0.07	0
Nombres de los archivos de texto:	0.92	0.36	0.02	0.09	0.27	0	0
RNMON, RNTUE,	0.01	0	0.06	0	0	0	0.27
RNWED, RNTHU,	0.01	0	0	0	0	0	0.01
RNFRI, RNSAT,	0	0	0.27	0.08	0	0	0.33
RNSUN.	0	0	0	0	0	0	0
	0.03	0	0.08	0.14	0	0	0
	0	0.11	0.06	0.02	0	0	0
	0.01	0.05	0	0.01	0	0	0
	0	0	0	0	0.12	0	0
	0.11	0.03	0	0	0	0	0.44
	0.01	0.01	0	0	0.11	0.18	0
	0.49	0	0.64	0.01	0	0	0.01
	0	0	0.08	0.85	0.01	0	0
	0.01	0.02	0	0	0.03	0	0
	0	0	0.12	0	0	0	0
	0	0	0.01	0.04	0.26	0.04	0
	0	0	0	0	0	0.4	0
	0.12	0	0	0	0	0	0
	0	0	0	0	0.24	0	0.23
	0	0	0	0.02	0	0	0
	0	0	0	0.02	0	0	0
	0.59	0	0	0	0	0.68	0
	0	0.01	0	0	0	1.48	0.21
	0.01	0	0	0	0.05	0.69	1.28
	0	0	0	0	0.96	0	0.01
	0	0	0	0	0	0.79	0.02
	0.41	0	0.06	0.01	0	0	0.28
	0	0	0	0.08	0.04	0	0
	0	0	0	0	0	0	0
	0	0.74	0	0	0	0	0
	0.43	0.3	0	0.26	0	0.02	0.01
			0				

Conjunto de datos 11: Géiser Old Faithful

Las mediciones provienen de erupciones del géiser Old Faithful del Parque Nacional Yellowstone. Los errores de predicción se basan en los momentos predichos para la erupción, donde un valor negativo corresponde a una erupción que ocurrió antes del momento predicho.



STATDISK: El nombre del conjunto de datos es Oldfaith.

Minitab: El nombre de la hoja de cálculo es OLDFAITH.MTW.

Excel: El nombre del libro de trabajo es OLDFAITH.XLS.

TI-83/84

Plus: El nombre de la App es OLDFAITH y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres

de los
archivos

de texto: Los nombres de los archivos de texto son OFDUR, OFBEF, OFAFT, OFHT, OFERR.

	Duración (seg)	Intervalo previo (min)	Intervalo posterior (min)	Altura (pies))	Error de predicción (min)
	240	98	92	140	4
	237	92	95	140	-2
	250	95	92	148	1
	243	87	100	130	-7
	255	96	90	125	2
	120	90	65	110	-4
	260	65	92	136	0
	178	92	72	125	-2
	259	95	93	115	1
	245	93	98	120	-1
	234	98	94	120	4
	213	94	80	120	0
	255	93	93	150	-1
	235	93	83	140	-1
	250	96	89	136	2
	110	89	66	120	-5
	245	93	89	148	-1
	269	89	86	130	-5
	251	86	97	130	-8
	234	69	105	136	4
	252	105	92	130	13
	254	92	89	115	0
	273	89	93	136	-3
	266	93	112	130	1
	284	112	88	138	20
	252	95	105	120	3
	269	105	94	120	13
	250	94	90	120	2
	261	90	98	95	-2
	253	98	81	140	6
	255	81	101	125	-11
	280	69	94	130	4
	270	94	92	130	2
	241	92	106	110	0
	272	106	93	110	14
	294	93	96	125	1
	220	108	87	150	21
	253	87	97	130	-5
	245	97	86	120	5
	274	102	92	95	10

Conjunto de datos 12: Pesos y volúmenes de bebidas de cola

Los pesos están en libras y los volúmenes en onzas.



- STATDISK:** El nombre del conjunto de datos es Cola.
Minitab: El nombre de la hoja de cálculo es COLA.MTW.
Excel: El nombre del libro de trabajo es COLA.XLS.
TI-83/84 Plus: El nombre de la App es COLA, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: CRGWT, CRGVL, CDTWT, CDTV, PRGWT, PRGVL, PDTWT, PDTVL.

Peso Coca clásica	Volumen Coca clásica	Peso Coca dietética	Volumen Coca dietética	Peso Pepsi clásica	Volumen Pepsi clásica	Peso Pepsi dietética	Volumen Pepsi dietética
0.8192	12.3	0.7773	12.1	0.8258	12.4	0.7925	12.3
0.8150	12.1	0.7758	12.1	0.8156	12.2	0.7868	12.2
0.8163	12.2	0.7896	12.3	0.8211	12.2	0.7846	12.2
0.8211	12.3	0.7868	12.3	0.8170	12.2	0.7938	12.3
0.8181	12.2	0.7844	12.2	0.8216	12.2	0.7861	12.2
0.8247	12.3	0.7861	12.3	0.8302	12.4	0.7844	12.2
0.8062	12.0	0.7806	12.2	0.8192	12.2	0.7795	12.2
0.8128	12.1	0.7830	12.2	0.8192	12.2	0.7883	12.3
0.8172	12.2	0.7852	12.2	0.8271	12.3	0.7879	12.2
0.8110	12.1	0.7879	12.3	0.8251	12.3	0.7850	12.3
0.8251	12.3	0.7881	12.3	0.8227	12.2	0.7899	12.3
0.8264	12.3	0.7826	12.3	0.8256	12.3	0.7877	12.2
0.7901	11.8	0.7923	12.3	0.8139	12.2	0.7852	12.2
0.8244	12.3	0.7852	12.3	0.8260	12.3	0.7756	12.1
0.8073	12.1	0.7872	12.3	0.8227	12.2	0.7837	12.2
0.8079	12.1	0.7813	12.2	0.8388	12.5	0.7879	12.2
0.8044	12.0	0.7885	12.3	0.8260	12.3	0.7839	12.2
0.8170	12.2	0.7760	12.1	0.8317	12.4	0.7817	12.2
0.8161	12.2	0.7822	12.2	0.8247	12.3	0.7822	12.2
0.8194	12.2	0.7874	12.3	0.8200	12.2	0.7742	12.1
0.8189	12.2	0.7822	12.2	0.8172	12.2	0.7833	12.2
0.8194	12.2	0.7839	12.2	0.8227	12.3	0.7835	12.2
0.8176	12.2	0.7802	12.1	0.8244	12.3	0.7855	12.2
0.8284	12.4	0.7892	12.3	0.8244	12.2	0.7859	12.2
0.8165	12.2	0.7874	12.2	0.8319	12.4	0.7775	12.1
0.8143	12.2	0.7907	12.3	0.8247	12.3	0.7833	12.2
0.8229	12.3	0.7771	12.1	0.8214	12.2	0.7835	12.2
0.8150	12.2	0.7870	12.2	0.8291	12.4	0.7826	12.2
0.8152	12.2	0.7833	12.3	0.8227	12.3	0.7815	12.2
0.8244	12.3	0.7822	12.2	0.8211	12.3	0.7791	12.1
0.8207	12.2	0.7837	12.3	0.8401	12.5	0.7866	12.3
0.8152	12.2	0.7910	12.4	0.8233	12.3	0.7855	12.2
0.8126	12.1	0.7879	12.3	0.8291	12.4	0.7848	12.2
0.8295	12.4	0.7923	12.4	0.8172	12.2	0.7806	12.2
0.8161	12.2	0.7859	12.3	0.8233	12.4	0.7773	12.1
0.8192	12.2	0.7811	12.2	0.8211	12.3	0.7775	12.1

Conjunto de datos 13: Pesos de una muestra de dulces M&M clásicos



- STATDISK:** El nombre del conjunto de datos es M&M.
- Minitab:** El nombre de la hoja de cálculo es M&M.MTW.
- Excel:** El nombre del libro de trabajo es M&M.XLS.
- TI-83/84 Plus:** El nombre de la App es MM, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: Los nombres de los archivos de texto son RED, ORNG, YLLW, BROWN, BLUE, GREEN.

Rojo	Naranja	Amarillo	Café	Azul	Verde
0.751	0.735	0.883	0.696	0.881	0.925
0.841	0.895	0.769	0.876	0.863	0.914
0.856	0.865	0.859	0.855	0.775	0.881
0.799	0.864	0.784	0.806	0.854	0.865
0.966	0.852	0.824	0.840	0.810	0.865
0.859	0.866	0.858	0.868	0.858	1.015
0.857	0.859	0.848	0.859	0.818	0.876
0.942	0.838	0.851	0.982	0.868	0.809
0.873	0.863			0.803	0.865
0.809	0.888			0.932	0.848
0.890	0.925			0.842	0.940
0.878	0.793			0.832	0.833
0.905	0.977			0.807	0.845
	0.850			0.841	0.852
	0.830			0.932	0.778
	0.856			0.833	0.814
	0.842			0.881	0.791
	0.778			0.818	0.810
	0.786			0.864	0.881
	0.853			0.825	
	0.864			0.855	
	0.873			0.942	
	0.880			0.825	
	0.882			0.869	
	0.931			0.912	
				0.887	
				0.886	

Conjunto de datos 14: Pesos de monedas (en gramos)

Las monedas de un penique o centavo de dólar acuñadas antes de 1983 se fabricaron después de los centavos "indios" y "de trigo", y contenían un 97% de cobre y un 3% de zinc. Las monedas de un centavo acuñadas después de 1983 contienen un 3% de cobre y un 97% de zinc. Las monedas de 25 centavos de plata acuñadas antes de 1964 contienen un 90% de plata y un 10% de cobre. Las monedas de 25 centavos acuñadas después de 1964 contienen una aleación de cobre y níquel.



- STATDISK:** El nombre del conjunto de datos es Coins.
Minitab: El nombre de la hoja de cálculo es COINS.MTW.
Excel: El nombre del libro de trabajo es COINS.XLS.
TI-83/84 Plus: El nombre de la App es COINS, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: Los nombres de los archivos de texto son CPIND, CPWHT, CPPRE, CPPST, CPCAN, CQPRE, CQPST, CDOL.

Centavos indios	Centavos de trigo	Centavos antes de 1983	Centavos después de 1983	Centavos canadienses	Monedas de \$0.25 antes de 1964	Monedas de \$0.25 después de 1964	Monedas de dólar
3.0630	3.1366	3.1582	2.5113	3.2214	6.2771	5.7027	8.1008
3.0487	3.0755	3.0406	2.4907	3.2326	6.2371	5.7495	8.1072
2.9149	3.1692	3.0762	2.5024	2.4662	6.1501	5.7050	8.0271
3.1358	3.0476	3.0398	2.5298	2.8357	6.0002	5.5941	8.0813
2.9753	3.1029	3.1043	2.4950	3.3189	6.1275	5.7247	8.0241
	3.0377	3.1274	2.5127	3.2612	6.2151	5.6114	8.0510
	3.1083	3.0775	2.4998	3.2441	6.2866	5.6160	7.9817
	3.1141	3.1038	2.4848	2.4679	6.0760	5.5999	8.0954
	3.0976	3.1086	2.4823	2.7202	6.1426	5.7790	8.0658
	3.0862	3.0586	2.5163	2.5120	6.3415	5.6841	8.1238
	3.0570	3.0603	2.5222		6.1309	5.6234	8.1281
	3.0765	3.0502	2.5004		6.2412	5.5928	8.0307
	3.1114	3.1028	2.5248		6.1442	5.6486	8.0719
	3.0965	3.0522	2.5058		6.1073	5.6661	8.0345
	3.0816	3.0546	2.4900		6.1181	5.5361	8.0775
	3.0054	3.0185	2.5068		6.1352	5.5491	8.1384
	3.1934	3.0712	2.5016		6.2821	5.7239	8.1041
	3.1461	3.0717	2.4797		6.2647	5.6555	8.0894
	3.0185	3.0546	2.5067		6.2908	5.6063	8.0538
	3.1267	3.0817	2.5139		6.1661	5.5709	8.0342
	3.1524	3.0704	2.4762		6.2674	5.5591	
	3.0786	3.0797	2.5004		6.2718	5.5864	
	3.0131	3.0713	2.5170		6.1949	5.6872	
	3.1535	3.0631	2.4925		6.2465	5.6274	
	3.0480	3.0866	2.4876		6.3172	5.6157	
	3.0050	3.0763	2.4933		6.1487	5.6668	
	3.0290	3.1299	2.4806		6.0829	5.7198	
	3.1038	3.0846	2.4907		6.1423	5.6694	
	3.0357	3.0917	2.5017		6.1970	5.5454	
	3.0064	3.0877	2.4950		6.2441	5.6646	
	3.0936	2.9593	2.4973		6.3669	5.5636	
	3.1031	3.0966	2.5252		6.0775	5.6485	
	3.0408	2.9800	2.4978		6.1095	5.6703	
	3.0561	3.0934	2.5073		6.1787	5.6848	
	3.0994	3.1340	2.4658		6.2130	5.5609	
			2.4529		6.1947	5.7344	
			2.5085		6.1940	5.6449	
					6.0257	5.5804	
					6.1719	5.6010	
					6.3278	5.6022	

Conjunto de datos 15: Cargas axiales de latas de aluminio

Las cargas axiales están medidas en libras.



- STATDISK:** El nombre del conjunto de datos es Cans.
Minitab : El nombre de la hoja de cálculo es CANS.MTW.
Excel: El nombre del libro de trabajo es CANS.XLS.
TI-83/84 Plus: El nombre de la App es CANS, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: CN109, CN111.

Muestra	Latas de aluminio de 0.0109 pulgadas								Muestra	Latas de aluminio de 0.0111 pulgadas							
	Carga (libras)										Carga (libras)						
1	270	273	258	204	254	228	282		1	287	216	260	291	210	272	260	
2	278	201	264	265	223	274	230		2	294	253	292	280	262	295	230	
3	250	275	281	271	263	277	275		3	283	255	295	271	268	225	246	
4	278	260	262	273	274	286	236		4	297	302	282	310	305	306	262	
5	290	286	278	283	262	277	295		5	222	276	270	280	288	296	281	
6	274	272	265	275	263	251	289		6	300	290	284	304	291	277	317	
7	242	284	241	276	200	278	283		7	292	215	287	280	311	283	293	
8	269	282	267	282	272	277	261		8	285	276	301	285	277	270	275	
9	257	278	295	270	268	286	262		9	290	288	287	282	275	279	300	
10	272	268	283	256	206	277	252		10	293	290	313	299	300	265	285	
11	265	263	281	268	280	289	283		11	294	262	297	272	284	291	306	
12	263	273	209	259	287	269	277		12	263	304	288	256	290	284	307	
13	234	282	276	272	257	267	204		13	273	283	250	244	231	266	504	
14	270	285	273	269	284	276	286		14	284	227	269	282	292	286	281	
15	273	289	263	270	279	206	270		15	296	287	285	281	298	289	283	
16	270	268	218	251	252	284	278		16	247	279	276	288	284	301	309	
17	277	208	271	208	280	269	270		17	284	284	286	303	308	288	303	
18	294	292	289	290	215	284	283		18	306	285	289	292	295	283	315	
19	279	275	223	220	281	268	272		19	290	247	268	283	305	279	287	
20	268	279	217	259	291	291	281		20	285	298	279	274	205	302	296	
21	230	276	225	282	276	289	288		21	282	300	284	281	279	255	210	
22	268	242	283	277	285	293	248		22	279	286	293	285	288	289	281	
23	278	285	292	282	287	277	266		23	297	314	295	257	298	211	275	
24	268	273	270	256	297	280	256		24	247	279	303	286	287	287	275	
25	262	268	262	293	290	274	292		25	243	274	299	291	281	303	269	

Conjunto de datos 16: Pesos de desechos de basura de una semana

Los pesos están en libras. TAMAÑO es el tamaño del hogar. Datos proporcionados por Masakuza Tani, Proyecto Garbage, Universidad de Arizona.



- STATDISK:** El nombre del conjunto de datos es Garbage.
Minitab: El nombre de la hoja de cálculo es GARBAGE.MTW.
Excel: El nombre del libro de trabajo es GARBAGE.XLS.
TI-83/84 Plus: El nombre de la App es GARBAGE, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: HHSIZ, METAL, PAPER, PLAS, GLASS, FOOD, YARD, TEXT, OTHER, TOTAL.

Hogar	Tamaño	Metal	Papel	Plástico	Vidrio	Comida	Patio	Textos	Otros	Total
1	2	1.09	2.41	0.27	0.86	1.04	0.38	0.05	4.66	10.76
2	3	1.04	7.57	1.41	3.46	3.68	0.00	0.46	2.34	19.96
3	3	2.57	9.55	2.19	4.52	4.43	0.24	0.50	3.60	27.60
4	6	3.02	8.82	2.83	4.92	2.98	0.63	2.26	12.65	38.11
5	4	1.50	8.72	2.19	6.31	6.30	0.15	0.55	2.18	27.90
6	2	2.10	6.96	1.81	2.49	1.46	4.58	0.36	2.14	21.90
7	1	1.93	6.83	0.85	0.51	8.82	0.07	0.60	2.22	21.83
8	5	3.57	11.42	3.05	5.81	9.62	4.76	0.21	10.83	49.27
9	6	2.32	16.08	3.42	1.96	4.41	0.13	0.81	4.14	33.27
10	4	1.89	6.38	2.10	17.67	2.73	3.86	0.66	0.25	35.54
11	4	3.26	13.05	2.93	3.21	9.31	0.70	0.37	11.61	44.44
12	7	3.99	11.36	2.44	4.94	3.59	13.45	4.25	1.15	45.17
13	3	2.04	15.09	2.17	3.10	5.36	0.74	0.42	4.15	33.07
14	5	0.99	2.80	1.41	1.39	1.47	0.82	0.44	1.03	10.35
15	6	2.96	6.44	2.00	5.21	7.06	6.14	0.20	14.43	44.44
16	2	1.50	5.86	0.93	2.03	2.52	1.37	0.27	9.65	24.13
17	4	2.43	11.08	2.97	1.74	1.75	14.70	0.39	2.54	37.60
18	4	2.97	12.43	2.04	3.99	5.64	0.22	2.47	9.20	38.96
19	3	1.42	6.05	0.65	6.26	1.93	0.00	0.86	0.00	17.17
20	3	3.60	13.61	2.13	3.52	6.46	0.00	0.96	1.32	31.60
21	2	4.48	6.98	0.63	2.01	6.72	2.00	0.11	0.18	23.11
22	2	1.36	14.33	1.53	2.21	5.76	0.58	0.17	1.62	27.56
23	4	2.11	13.31	4.69	0.25	9.72	0.02	0.46	0.40	30.96
24	1	0.41	3.27	0.15	0.09	0.16	0.00	0.00	0.00	4.08
25	4	2.02	6.67	1.45	6.85	5.52	0.00	0.68	0.03	23.22
26	6	3.27	17.65	2.68	2.33	11.92	0.83	0.28	4.03	42.99
27	11	4.95	12.73	3.53	5.45	4.68	0.00	0.67	19.89	51.90
28	3	1.00	9.83	1.49	2.04	4.76	0.42	0.54	0.12	20.20
29	4	1.55	16.39	2.31	4.98	7.85	2.04	0.20	1.48	36.80
30	3	1.41	6.33	0.92	3.54	2.90	3.85	0.03	0.04	19.02
31	2	1.05	9.19	0.89	1.06	2.87	0.33	0.01	0.03	15.43
32	2	1.31	9.41	0.80	2.70	5.09	0.64	0.05	0.71	20.71
33	2	2.50	9.45	0.72	1.14	3.17	0.00	0.02	0.01	17.01
34	4	2.35	12.32	2.66	12.24	2.40	7.87	4.73	0.78	45.35
35	6	3.69	20.12	4.37	5.67	13.20	0.00	1.15	1.17	49.37
36	2	3.61	7.72	0.92	2.43	2.07	0.68	0.63	0.00	18.06

(continúa)

Conjunto de datos 16: Pesos de desechos de basura de una semana (*continuación*)

Hogar	Tamaño	Metal	Papel	Plástico	Vidrio	Comida	Patio	Textos	Otros	Total
37	2	1.49	6.16	1.40	4.02	4.00	0.30	0.04	0.00	17.41
38	2	1.36	7.98	1.45	6.45	4.27	0.02	0.12	2.02	23.67
39	2	1.73	9.64	1.68	1.89	1.87	0.01	1.73	0.58	19.13
40	2	0.94	8.08	1.53	1.78	8.13	0.36	0.12	0.05	20.99
41	3	1.33	10.99	1.44	2.93	3.51	0.00	0.39	0.59	21.18
42	3	2.62	13.11	1.44	1.82	4.21	4.73	0.64	0.49	29.06
43	2	1.25	3.26	1.36	2.89	3.34	2.69	0.00	0.16	14.95
44	2	0.26	1.65	0.38	0.99	0.77	0.34	0.04	0.00	4.43
45	3	4.41	10.00	1.74	1.93	1.14	0.92	0.08	4.60	24.82
46	6	3.22	8.96	2.35	3.61	1.45	0.00	0.09	1.12	20.80
47	4	1.86	9.46	2.30	2.53	6.54	0.00	0.65	2.45	25.79
48	4	1.76	5.88	1.14	3.76	0.92	1.12	0.00	0.04	14.62
49	3	2.83	8.26	2.88	1.32	5.14	5.60	0.35	2.03	28.41
50	3	2.74	12.45	2.13	2.64	4.59	1.07	0.41	1.14	27.17
51	10	4.63	10.58	5.28	12.33	2.94	0.12	2.94	15.65	54.47
52	3	1.70	5.87	1.48	1.79	1.42	0.00	0.27	0.59	13.12
53	6	3.29	8.78	3.36	3.99	10.44	0.90	1.71	13.30	45.77
54	5	1.22	11.03	2.83	4.44	3.00	4.30	1.95	6.02	34.79
55	4	3.20	12.29	2.87	9.25	5.91	1.32	1.87	0.55	37.26
56	7	3.09	20.58	2.96	4.02	16.81	0.47	1.52	2.13	51.58
57	5	2.58	12.56	1.61	1.38	5.01	0.00	0.21	1.46	24.81
58	4	1.67	9.92	1.58	1.59	9.96	0.13	0.20	1.13	26.18
59	2	0.85	3.45	1.15	0.85	3.89	0.00	0.02	1.04	11.25
60	4	1.52	9.09	1.28	8.87	4.83	0.00	0.95	1.61	28.15
61	2	1.37	3.69	0.58	3.64	1.78	0.08	0.00	0.00	11.14
62	2	1.32	2.61	0.74	3.03	3.37	0.17	0.00	0.46	11.70

Conjunto de datos 17: Distancias de *home runs*

Las distancias de *home runs* de Mark McGwire (1998), Sammy Sosa (1998) y Barry Bonds (2001) están dadas en pies.



- STATDISK:** El nombre del conjunto de datos es Homeruns.
Minitab: El nombre de la hoja de cálculo es HOMERUNS.MTW.
Excel: El nombre del libro de trabajo es HOMERUNS.XLS.
TI-83/84 Plus: El nombre de la App es HOMERUNS, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: MCGWR, SOSA, BONDS.

McGwire

360	370	370	430	420	340	460	410	440	410
380	360	350	527	380	550	478	420	390	420
425	370	480	390	430	388	423	410	360	410
450	350	450	430	461	430	470	440	400	390
510	430	450	452	420	380	470	398	409	385
369	460	390	510	500	450	470	430	458	380
430	341	385	410	420	380	400	440	377	370

Sosa

371	350	430	420	430	434	370	420	440	410
420	460	400	430	410	370	370	410	380	340
350	420	410	415	430	380	380	366	500	380
390	400	364	430	450	440	365	420	350	420
400	380	380	400	370	420	360	368	430	433
388	440	414	482	364	370	400	405	433	390
480	480	434	344	410	420				

Bonds

420	417	440	410	390	417	420	410	380	430
370	420	400	360	410	420	391	416	440	410
415	436	430	410	400	390	420	410	420	410
410	450	320	430	380	375	375	347	380	429
320	360	375	370	440	400	405	430	350	396
410	380	430	415	380	375	400	435	420	420
488	361	394	410	411	365	360	440	435	454
442	404	385							

Conjunto de datos 18: Casas vendidas en el Condado Dutchess, Nueva York



STATDISK: El nombre del conjunto de datos es **Homes**.
Minitab: El nombre de la hoja de cálculo es **HOMES.MTW**.
Excel: El nombre del libro de trabajo es **HOMES.XLS**.
TI-83/84 Plus: El nombre de la App es **HOMES**, y los nombres de los archivos son los mismos que para los archivos de texto.

Nombres de los archivos de texto: **HMSPL, HMLST, HMLA, HMACR, HMAGE, HMTAX, HMRMS, HMBRS, HMBTH.**

Precio de venta (dólares)	Precio de lista (dólares)	Área habitable (cientos de pies cuadrados)	Acres	Antigüedad (años)	Impuestos (dólares)	Habitaciones	Dormitorios	Baños (completos)
400000	414000	2704	2.27	27	4920	9	3	3
370000	379000	2096	0.75	21	4113	8	4	2
382500	389900	2737	1.00	36	6072	9	4	2
300000	299900	1800	0.43	34	4024	8	4	2
305000	319900	1066	3.60	69	3562	6	3	2
320000	319900	1820	1.70	34	4672	7	3	2
321000	328900	2700	0.81	35	3645	8	3	1
445000	450000	2316	2.00	19	6256	9	4	2
377500	385000	2448	1.50	40	5469	9	4	3
460000	479000	3040	1.09	20	6740	10	4	2
265000	275000	1500	1.60	39	4046	6	2	2
299000	299000	1448	0.42	44	3481	7	3	1
385000	379000	2400	0.89	33	4411	9	4	3
430000	435000	2200	4.79	6	5714	8	4	2
214900	219900	1635	0.25	49	2560	5	3	1
475000	485000	2224	11.58	21	7885	7	3	2
280000	289000	1738	0.46	49	3011	8	3	2
457000	499900	3432	1.84	14	9809	11	4	3
210000	224900	1175	0.94	64	1367	7	3	1
272500	274900	1393	1.39	44	2317	6	3	1
268000	275000	1196	0.83	44	3360	4	2	1
300000	319900	1860	0.57	32	4294	9	3	2
477000	479000	3867	1.10	19	9135	10	4	4
292000	294900	1800	0.52	47	3690	8	2	1
379000	383900	2722	1.00	29	6283	10	4	3
295000	299900	2240	0.90	144	3286	6	3	1
499000	499000	2174	5.98	62	3894	6	3	2
292000	299000	1650	2.93	52	3476	7	3	1
305000	299900	2000	0.33	36	4146	8	3	3
520000	529700	3350	1.53	6	8350	11	4	2
308000	320000	1776	0.63	42	4584	8	4	2
316000	310000	1850	2.00	25	4380	7	3	2
355500	362500	2600	0.44	46	4009	10	5	2
225000	229000	1300	0.62	49	3047	6	3	1
270000	290000	1352	0.68	24	2801	6	3	1
253000	259900	1312	0.68	44	4048	6	2	1
310000	314900	1664	1.69	53	2940	6	3	2
300000	309900	1700	0.83	33	4281	8	4	2
295000	295000	1650	2.90	34	4299	6	2	2
478000	479000	2400	2.14	6	6688	8	4	2

Apéndice C: Glosario

Alfa (α) Símbolo empleado para representar la probabilidad de un error tipo I. Véase también Nivel de significancia.

Análisis de varianza Método para analizar las varianzas poblacionales, que permite hacer pruebas de hipótesis acerca de medias de poblaciones.

Análisis de varianza de dos factores Análisis de varianza que implica datos clasificados según dos factores distintos.

Análisis de varianza de un factor Análisis de varianza que implica datos clasificados en grupos de acuerdo con un solo criterio.

Análisis exploratorio de datos (AED) Rama de la estadística que pone énfasis en la investigación de datos.

Anchura de clase La diferencia entre dos límites de clase inferiores consecutivos en una distribución de frecuencias.

ANOVA Véase Análisis de varianza.

Aproximación de la probabilidad por frecuencia relativa Valor de probabilidad estimado con base en observaciones reales.

Bajo control estadístico Véase Proceso estadísticamente estable.

Beta (β) Símbolo empleado para representar la probabilidad de un error tipo II.

Bimodal Se dice que un conjunto de datos es bimodal cuando tiene dos modas.

Bloque Grupo de individuos similares con respecto a las formas en que pueden afectar el resultado de un experimento.

Cambio marginal Para variables relacionadas por una ecuación de regresión, la magnitud del cambio en la variable dependiente, cuando una de las variables independientes cambia en una unidad y las demás variables independientes se mantienen constantes.

Celda Categoría empleada para separar datos cualitativos (o de atributo).

Censo Recolección de datos de cada elemento de una población.

Centroide El punto (\bar{x}, \bar{y}) , determinado a partir de un conjunto de datos bivariados.

CM (del error) Cuadrado medio del error; se usa en el análisis de varianza.

CM (del tratamiento) Cuadrado medio de tratamientos; se usa en el análisis de varianza.

CM (total) Cuadrado medio de la variación total; se usa en el análisis de varianza.

Coeficiente de confianza Probabilidad de que un parámetro de población esté contenido dentro de un intervalo de confianza particular; también se denomina nivel de confianza o grado de confianza.

Coeficiente de correlación Medida de la fuerza de la relación entre dos variables.

Coeficiente de correlación de rangos ordenados Medida de la fuerza de la relación entre dos variables; se basa en los rangos ordenados de los valores.

Coeficiente de correlación de rangos ordenados de Spearman Véase Coeficiente de correlación de rangos ordenados.

Coeficiente de correlación lineal Medida de la fuerza de la relación entre dos variables.

Coeficiente de correlación producto-momento de Pearson Véase Coeficiente de correlación lineal.

Coeficiente de determinación Cantidad de la variación de y que se explica con la línea de regresión.

Coeficiente de determinación ajustado Coeficiente de determinación múltiple R^2 , modificado para justificar el número de variables y del tamaño de la muestra.

Coeficiente de determinación múltiple Medida de qué tan bien una ecuación de regresión múltiple se ajusta a los datos muestrales.

Coeficiente de variación (CV) Razón entre la desviación estándar y la media, que se expresa como un porcentaje.

Complemento de un suceso Todos los resultados en los que el suceso original no ocurre.

Confusión Situación que se presenta cuando no es posible distinguir entre los efectos de dos o más variables.

Control estadístico de procesos (CEP) Uso de técnicas estadísticas, como gráficas de control, para analizar un proceso o sus salidas y así poder tomar medidas apropiadas con el fin de lograr y mantener un estado de control estadístico y mejorar la capacidad del proceso.

Corrección por continuidad Ajuste que se hace cuando una variable aleatoria discreta se aproxima con una variable aleatoria continua (sección 6-6).

Correlación Asociación estadística entre dos variables.

Cuartil medio La mitad de la suma de los cuartiles primero y tercero.

Cuartiles Los tres valores que dividen datos ordenados en cuatro grupos, con aproximadamente el 25% de los valores en cada grupo.

Curva de densidad Gráfica de una distribución de probabilidad continua.

Datos Información o números que describen alguna característica.

Datos apareados Se dice que en dos muestras hay datos apareados cuando cada valor de una de ellas está apareado con un valor correspondiente de la otra muestra.

Datos bivariados Datos ordenados en pares.

Datos categóricos Datos que pueden dividirse en diferentes categorías y que se distinguen por alguna característica no numérica.

Datos continuos Datos que se obtienen de un número infinito de valores posibles, que corresponden a puntos de una escala continua que abarca un rango de valores sin huecos, saltos ni interrupciones.

Datos cualitativos Datos que pueden dividirse en diferentes categorías que se distinguen por alguna característica no numérica.

Datos cuantitativos Datos que consisten en números que representan conteos o mediciones.

Datos de atributo Datos que pueden dividirse en diferentes categorías que se distinguen por alguna característica no numérica.

Datos de proceso Datos ordenados según alguna secuencia de tiempo, que miden una característica de bienes o servicios que resultan de alguna combinación de equipos, personas, materiales, métodos y condiciones.

Datos de series de tiempo Datos reunidos en diferentes puntos del tiempo.

Datos discretos Datos con la propiedad de que el número de valores posibles es un valor finito o que puede contarse, que resulta en 0 posibilidades, 1 posibilidad o 2 posibilidades, etcétera.

Datos numéricos Datos que consisten en números que representan conteos o mediciones.

Datos ordenados Datos acomodados en orden.

Desviación Magnitud de la diferencia entre un puntaje y la media; se expresa como $x - \bar{x}$.

Desviación absoluta La medida de variación que es igual a la suma de las desviaciones de cada valor respecto a la media, dividida entre el número de valores.

Desviación estándar Medida de variación igual a la raíz cuadrada de la varianza.

Desviación explicada Para un par de valores de un conjunto de datos bivariados, la diferencia entre el valor de y predicho y la media de todos los valores de y .

Desviación media absoluta Medida de variación que es igual a la suma de las desviaciones de cada valor respecto a la media, dividida entre el número de valores.

Desviación no explicada Para un par de valores en un conjunto de datos bivariados, la diferencia entre la coordenada y y el valor predicho.

Desviación total Suma de la desviación explicada y la desviación no explicada para un par dado de valores en un conjunto de datos bivariados.

Diagrama de árbol Representación gráfica de los diferentes resultados posibles en un suceso compuesto.

Diagrama de cuadro y bigotes Véase Gráfica de cuadro.

Diagrama de dispersión Representación gráfica de datos (x, y) apareados.

Diseño de bloques aleatorizado Diseño en el que se obtiene una medición para cada tratamiento aplicado a cada uno de varios individuos apareados según características similares.

Diseño rigurosamente controlado Diseño experimental en el que se obliga a que todos los factores sean constantes con el fin de eliminar los efectos de factores ajenos.

Diseño totalmente aleatorizado Procedimiento de un experimento en el que cada elemento tiene la misma

probabilidad de pertenecer a las diferentes categorías o tratamientos.

Distribución chi cuadrada Una distribución de probabilidad continua (se presenta en la sección 7-5).

Distribución de frecuencias Lista de valores de datos (individualmente o por grupos de intervalos) junto con sus frecuencias (o conteos) correspondientes.

Distribución de frecuencias acumulativas Distribución de frecuencias en la que cada clase y frecuencia representa los datos acumulativos hasta esa clase inclusive.

Distribución de frecuencias relativas Variación de la distribución básica de frecuencias en la que la frecuencia de cada clase se divide entre el total de todas las frecuencias.

Distribución de Poisson Distribución de probabilidad discreta que se aplica a ocurrencias de algún suceso durante un intervalo de tiempo, distancia, área, volumen u otra unidad similar que se especifique.

Distribución de probabilidad Conjunto de valores de una variable aleatoria junto con sus probabilidades correspondientes.

Distribución F Distribución de probabilidad continua; se explica en la sección 9-5.

Distribución muestral de medias muestrales Distribución de las medias muestrales que se obtiene al seleccionar repetidamente muestras de igual tamaño de la misma población.

Distribución muestral de proporciones Distribución de probabilidad de las proporciones muestrales, donde todas las muestras tienen el mismo tamaño muestral n .

Distribución normal Distribución de probabilidad con forma de campana, descrita algebraicamente con la fórmula 6-1 de la sección 6-1.

Distribución normal bivariada Distribución de datos apareados en la que, para cualquier valor fijo de una variable, los valores de la otra variable están distribuidos normalmente.

Distribución normal estándar Distribución normal con una media igual a cero y una desviación estándar igual a 1.

Distribución t Distribución normal que suele estar asociada con datos muestrales de una población con una desviación estándar desconocida.

Distribución t de Student Véase Distribución t .

Distribución uniforme Distribución de probabilidad en la que todos los valores de la variable aleatoria son igualmente probables.

Ecuación de regresión Ecuación algebraica que describe la relación entre variables.

Ecuación de regresión múltiple Ecuación que expresa una relación lineal entre una variable dependiente y dos o más variables independientes (x_1, x_2, \dots, x_k).

Efecto placebo Efecto que ocurre cuando un sujeto que no recibe tratamiento cree incorrectamente que sí lo está recibiendo y reporta una mejoría en sus síntomas.

Eficiencia Medida de la sensibilidad de una prueba no paramétrica en comparación con una prueba paramétrica correspondiente.

Enfoque clásico a la probabilidad Enfoque en el que la probabilidad de un evento se determina dividiendo el número de maneras en que éste puede suceder entre el número total de resultados posibles.

Error de muestreo Diferencia entre el resultado de una muestra y el resultado real de la población; se debe a fluctuaciones aleatorias en las muestras.

Error estándar de estimación Medida de la dispersión de puntos muestrales alrededor de la línea de regresión.

Error estándar de la media Desviación estándar de todas las posibles medias muestrales \bar{x} .

Error máximo de estimado Véase Margen de error.

Error tipo I Error que se comete al rechazar la hipótesis nula cuando ésta es verdadera.

Error tipo II Error que se comete al no rechazar la hipótesis nula cuando ésta es falsa.

Errores que no son de muestreo Errores debidos a factores externos no relacionados con el muestreo.

Espacio muestral Conjunto de todos los posibles resultados o sucesos de un experimento que no se pueden descomponer más.

Estadística Conjunto de métodos para planear experimentos, para obtener, organizar, resumir, presentar, analizar e interpretar datos, y para extraer conclusiones con base en esos datos.

Estadística descriptiva Métodos empleados para resumir las características clave de datos conocidos.

Estadística inferencial Métodos que implican el uso de datos muestrales para hacer generalizaciones o inferencias acerca de una población.

Estadístico Característica medida de una muestra.

Estadístico de prueba Estadístico muestral que se basa en los datos muestrales; sirve para tomar la decisión respecto a rechazar o no la hipótesis nula.

Estimado Valor específico o intervalo de valores que se usa para aproximar algún parámetro de población.

Estimado conjunto de p_1 y p_2 Probabilidad que se obtiene combinando los datos de dos proporciones muestrales y dividiendo el número total de éxitos entre el número total de observaciones.

Estimado conjunto de σ^2 Estimado de la varianza σ^2 que es común a dos poblaciones; se obtiene calculando un promedio ponderado de las dos varianzas muestrales.

Estimado de intervalo Rango de valores usado para estimar algún parámetro de población con un nivel de confianza específico; también se conoce como intervalo de confianza.

Estimado puntual Valor individual que sirve como estimado de un parámetro de población.

Estimador Estadístico de muestra (como la media de muestra \bar{x}), que sirve para aproximar un parámetro de población.

Estimador sin sesgo Estadístico muestral que tiende a acercarse al parámetro de población para cuya estimación se usa.

Estudio ciego Procedimiento utilizado en experimentos, en el que el sujeto no sabe si está recibiendo un tratamiento o un placebo.

Estudio de casos y controles Estudio en el que se reúnen datos del pasado (a través del examen de registros, entrevistas y otros recursos).

Estudio de cohorte Estudio de sujetos en grupos identificados que comparten factores comunes (denominados *cohorte*), en el que los datos se reunirán en el futuro.

Estudio doble ciego Procedimiento utilizado en un experimento, en el que el sujeto no sabe si está recibiendo un tratamiento o un placebo, y la persona que aplica el tratamiento tampoco lo sabe.

Estudio longitudinal Estudio de sujetos en grupos identificados que comparten factores comunes (llamados *cohorte*) y en el que los datos se reunirán en el futuro.

Estudio observacional Estudio en el que se observan y miden características específicas, pero no se intenta manipular o modificar a los sujetos en estudio.

Estudio prospectivo Estudio de sujetos en grupos identificados que comparten factores comunes (denominados *cohorte*), en el que los datos se reunirán en el futuro.

Estudio retrospectivo Estudio en el que se reúnen datos del pasado (a través del examen de registros, entrevistas y otros recursos).

Estudio transversal Estudio en el que los datos se observan, miden y reúnen en un punto del tiempo.

Experimento La aplicación de un tratamiento y la posterior observación de sus efectos sobre los sujetos.

Experimento binomial Experimento que tiene un número fijo de ensayos independientes y en el que cada resultado pertenece exactamente a una de dos categorías.

Experimento multinomial Experimento que tiene un número fijo de ensayos independientes, y en el que cada resultado pertenece exactamente a una de varias categorías.

Factor En análisis de varianza, propiedad o característica que nos permite distinguir unas poblaciones de otras.

Factor de corrección por población finita Factor para corregir el error estándar de la media cuando el tamaño de una muestra excede el 5% del tamaño de una población finita.

Fórmula de probabilidad binomial Expresión utilizada para calcular probabilidades en un experimento binomial (véase la fórmula 5-5 de la sección 5-3).

Fractiles Números que dividen los datos en partes de aproximadamente el mismo tamaño.

Frecuencia acumulativa Suma de las frecuencias de una clase y de todas las clases precedentes.

Frecuencia esperada Frecuencia teórica para una celda de una tabla de contingencias o tabla multinomial.

Frecuencia observada Conteo de frecuencia real registrado en una celda de una tabla de contingencias o tabla multinomial.

Frecuencia relativa Frecuencia de una clase, dividida entre el total de todas las frecuencias.

Fronteras de clase Valores que se obtienen de una distribución de frecuencias aumentando los límites de clase superior y reduciendo los límites de clase inferior en la misma cantidad, de forma que no haya huecos entre clases consecutivas.

Grado de confianza Probabilidad de que un parámetro de población esté contenido dentro de un intervalo de confianza particular; también se denomina nivel de confianza.

Grados de libertad Número de valores que pueden variar después de haberse impuesto ciertas restricciones a todos los valores.

Grados de libertad del denominador Grados de libertad que corresponden al denominador del estadístico de prueba F .

Grados de libertad del numerador Grados de libertad que corresponden al numerador del estadístico de prueba F .

Gráfica Gráfica de control que se usa para vigilar la media de un proceso.

Gráfica circular Representación gráfica de datos en forma de círculo que contiene divisiones radiales.

Gráfica de control Cualquiera de varios tipos de gráficas (capítulo 14) que describen alguna característica de un proceso con el fin de determinar si hay estabilidad estadística.

Gráfica de cuadro Representación gráfica de la dispersión de un conjunto de datos.

Gráfica de Pareto Gráfica de barras para datos cualitativos, con las barras dispuestas en orden según las frecuencias.

Gráfica de puntos Gráfica en la que cada valor de los datos se grafica como un punto sobre una escala de valores.

Gráfica de rachas Gráfica secuencial de valores de datos individuales a lo largo del tiempo, donde se usa un eje (casi siempre el vertical) para los valores de datos, y el otro eje (casi siempre el horizontal) para la secuencia de tiempo.

Gráfica de rango Gráfica de control basada en rangos muestrales; sirve para vigilar la variación en un proceso.

Gráfica de tallo y hojas Método para clasificar y acomodar datos para revelar su distribución.

Gráfica normal cuantilar Gráfica de puntos (x, y) , donde cada valor de x pertenece al conjunto original de datos muestrales, y cada valor y es una puntuación z correspondiente a un valor cuantilar de la distribución normal estándar.

Gráfica np Gráfica de control en la que se grafica el número de defectos con el fin de vigilar un proceso.

Gráfica p Gráfica de control que sirve para vigilar la proporción p de algún atributo en un proceso.

Gráfica R Gráfica de control basada en rangos muestrales; sirve para vigilar la variación en un proceso.

Gráfica s Gráfica de control basada en desviaciones estándar muestrales; sirve para vigilar la variación en un proceso.

Grupo de control Grupo de sujetos en un experimento a quienes no se les administra un tratamiento específico.

Grupo de tratamiento Grupo de sujetos que reciben algún tratamiento en un experimento.

Hipótesis Declaración o afirmación acerca de alguna propiedad de una población.

Hipótesis alternativa Afirmación que equivale a la negación de la hipótesis nula; se denota con H_1 .

Hipótesis nula Aseveración acerca de alguna característica de población, que por lo regular implica la ausencia de una diferencia; se denota con H_0 .

Histograma Gráfica de barras verticales que representa la distribución de frecuencia de un conjunto de datos.

Histograma de frecuencias relativas Variación del histograma básico en la que las frecuencias se sustituyen por frecuencias relativas.

Histograma de probabilidad Histograma en el que los resultados se listan a lo largo del eje horizontal y las probabilidades se listan a lo largo del eje vertical.

Interacción En el análisis de varianza de dos factores, el efecto que se observa cuando uno de ellos varía para diferentes categorías del otro factor.

Intercepto y Punto en el que una línea recta cruza el eje y .

Intervalo Nivel de medición de datos; caracteriza datos que pueden acomodarse en orden y para los que las diferencias entre los valores de los datos son significativas.

Intervalo de confianza Rango de valores empleado para estimar algún parámetro de población con un nivel de confianza específico; también se denomina estimado de intervalo.

Intervalo de predicción Estimado del intervalo de confianza de un valor predicho de y .

Límite de control Frontera que se usa en una gráfica de control para identificar puntos inusitados.

Límite de control inferior Frontera de una gráfica de control que separa los puntos inusitadamente bajos.

Límite de control superior Frontera que se usa en una gráfica de control para separar los puntos inusitadamente altos.

Límites de clase inferiores Los números más pequeños que pueden pertenecer a las diferentes clases de una distribución de frecuencias.

Límites de clase superiores Los números más grandes que pueden pertenecer a las diferentes clases de una distribución de frecuencias.

Límites del intervalo de confianza Dos números que se usan como fronteras superior e inferior de un intervalo de confianza.

Línea central Línea de una gráfica de control que representa un valor central de las mediciones características.

Línea de regresión Línea recta que mejor se ajusta a un conjunto de puntos que representan datos muestrales aparentados.

Margen de error Máxima diferencia probable (con probabilidad $1 - \alpha$) entre el estadístico muestral observado y el verdadero valor del parámetro poblacional.

Media La suma de un conjunto de valores, dividida entre el número de valores.

Media aritmética Suma de un conjunto de valores dividida entre el número de valores; normalmente se denomina media.

Media ponderada Media de un conjunto de valores a los que se han asignado diferentes grados de importancia.

Mediana Valor que está a la mitad de un conjunto de valores acomodados en orden por magnitud.

Medida de tendencia central Valor que pretende indicar el centro de los valores de una colección de datos.

Medida de variación Cualquiera de varias medidas diseñadas para reflejar la magnitud de la variación o dispersión de un conjunto de valores.

Método clásico de prueba de hipótesis Método para probar hipótesis que se basa en una comparación del estadístico de prueba con los valores críticos.

Método tradicional de prueba de hipótesis Método para probar hipótesis que se basa en una comparación del estadístico de prueba y los valores críticos.

Mitad del rango Mitad de la suma de los valores máximo y mínimo.

Moda Valor que se presenta con mayor frecuencia.

Modelo matemático Función matemática que se “ajusta” o que describe datos de la vida real.

Muestra Subconjunto de una población.

Muestra aleatoria Muestra seleccionada de tal manera que permite a cada miembro de la población tener la misma probabilidad de ser elegido.

Muestra aleatoria simple Muestra de cierto tamaño seleccionada de manera que toda posible muestra del mismo tamaño tenga la misma probabilidad de ser elegida.

Muestra autoselecciónada Muestra en la que los sujetos deciden por sí mismos ser incluidos; también se llama muestra de respuesta voluntaria.

Muestra de respuesta voluntaria Muestra en la que los sujetos deciden por sí mismos ser incluidos.

Muestra dependiente Muestra cuyos valores están relacionados con los valores de otra muestra.

Muestra independiente Muestra cuyos valores no están relacionados con los valores de otra muestra.

Muestras apareadas Dos muestras que son dependientes en el sentido de que los valores de los datos forman pares.

Muestreo de aceptación Elementos muestrales sin reemplazo y que permiten rechazar todo el lote con base en el número de defectos obtenidos.

Muestreo de conveniencia Muestreo en el que se seleccionan datos porque se encuentran disponibles.

Muestreo estratificado Muestreo en el que se obtienen muestras de cada estrato (clase).

Muestreo por conglomerados Tipo de muestreo en el que se divide el área de población en secciones (o conglomerados) y luego se seleccionan en forma aleatoria algunas de esas secciones; después se eligen *todos* los miembros de las secciones elegidas.

Muestreo sistemático Muestreo en el que se selecciona cada k -ésimo elemento.

Multimodal Se dice que un conjunto de datos es multimodal cuando tiene más de dos modas.

Nivel de confianza Probabilidad de que un parámetro de población esté contenido en un intervalo de confianza particular; también se llama grado de confianza.

Nivel de significancia Probabilidad de cometer un error tipo I al realizar una prueba de hipótesis.

Nominal Nivel de medición de datos; caracteriza datos que consisten únicamente en nombres, etiquetas o categorías.

Ojiva Representación gráfica de una distribución de frecuencias acumulativas.

Ordinal Nivel de medición de datos; caracteriza datos que podrían estar acomodados en orden, pero las diferencias entre los valores de los datos no pueden determinarse, o bien, carecen de sentido.

Parámetro Característica medida de una población.

Pares discordantes Pares de categorías en los que las dos categorías son diferentes; se usan en la prueba de McNemar.

Pendiente Medida de la inclinación de una línea recta.

Percentil Los 99 valores que dividen datos clasificados en 100 grupos, con aproximadamente el 1% de los valores en cada grupo.

Población Colección entera y completa de elementos por estudiar.

Polígono de frecuencias Representación gráfica de la distribución de los datos, que utiliza segmentos de línea recta conectados.

Posibilidades a favor Razón entre la probabilidad de que un suceso ocurra y la probabilidad de que no ocurra; suele expresarse como la proporción de dos enteros sin factores comunes.

Posibilidades de pago Razón de la ganancia neta (si se gana) en relación con lo apostado.

Posibilidades en contra Razón entre la probabilidad de que un suceso no ocurra y la probabilidad de que ocurra; suele expresarse en la forma $a:b$, donde a y b son enteros sin factores comunes.

Posibilidades reales a favor El recíproco de las posibilidades reales en contra de un suceso.

Posibilidades reales en contra El cociente $P(\bar{A})/P(A)$, generalmente expresado en la forma $a:b$ (o “ a es a b ”).

Potencia de una prueba La probabilidad $(1 - \beta)$ de rechazar una hipótesis nula falsa.

Probabilidad Medida de la posibilidad de que ocurra un suceso dado; se expresa como un número entre 0 y 1.

Probabilidad condicional La probabilidad de un suceso, dado que algún otro suceso ya ocurrió.

Probabilidad subjetiva Conjetura o estimado de una probabilidad con base en un conocimiento de las circunstancias relevantes.

Procedimientos de comparación múltiple Procedimientos para identificar cuáles medias específicas son diferentes, después de concluir que tres o más medias no son iguales.

Proceso estadísticamente estable Proceso que sólo tiene variación natural, sin patrones, ciclos ni puntos inusitados.

Promedio Cualquiera de varias medidas diseñadas para revelar la tendencia central de un conjunto de datos.

Propiedad de mínimos cuadrados Propiedad que afirma que, para una línea de regresión, la suma de los cuadrados de las desviaciones verticales de los puntos muestrales, respecto a la línea de regresión, es la más pequeña posible.

Prueba de bondad de ajuste Prueba para determinar qué tan bien alguna distribución de frecuencias observada se ajusta a una distribución teórica.

Prueba de cola derecha Prueba de hipótesis en la que la región crítica se ubica en el área extrema derecha de la distribución de probabilidad.

Prueba de cola izquierda Prueba de hipótesis en la que la región crítica está situada en el área extrema izquierda de la distribución de probabilidad.

Prueba de dos colas Prueba de hipótesis en la que la región crítica se divide entre las áreas extremas izquierda y derecha de la distribución de probabilidad.

Prueba de hipótesis Método para probar afirmaciones acerca de poblaciones; también se llama prueba de significancia.

Prueba de homogeneidad Prueba de la aseveración de que diferentes poblaciones tienen la misma proporción de alguna característica.

Prueba de independencia Prueba de la hipótesis nula que afirma que, en una tabla de contingencias, la variable de renglón y la variable de columna no están relacionadas.

Prueba de Kruskal-Wallis Prueba de hipótesis no paramétrica que sirve para comparar tres o más muestras independientes; también se llama prueba H.

Prueba de McNemar Usa conteos de frecuencia de datos nominales apareados de dos categorías para poner a prueba la hipótesis nula de que las frecuencias de pares discordantes ocurren en la misma proporción.

Prueba de rachas Método no paramétrico que sirve para detectar aleatoriedad.

Prueba de rangos con signo (o de rangos signados) de Wilcoxon Prueba de hipótesis no paramétrica que se utiliza para comparar dos muestras dependientes.

Prueba de significancia Véase Prueba de hipótesis.

Prueba de suma de rangos de Wilcoxon Prueba de hipótesis no paramétrica que se utiliza para comparar dos muestras independientes.

Prueba del signo Prueba de hipótesis no paramétrica que sirve para comparar muestras de dos poblaciones.

Prueba H Prueba no paramétrica de Kruskal-Wallis.

Prueba U de Mann-Whitney Prueba de hipótesis que equivale a la prueba de suma de rangos de Wilcoxon para dos muestras independientes.

Pruebas de distribución libre Pruebas que no requieren una distribución específica, como la distribución normal. Véase también Pruebas no paramétricas.

Pruebas no paramétricas Procedimientos estadísticos para hacer pruebas de hipótesis o estimar parámetros, en los que no es preciso hacer suposiciones acerca de la naturaleza o forma de las distribuciones de las poblaciones; también se denominan pruebas de distribución libre.

Pruebas paramétricas Procedimientos estadísticos basados en parámetros de población para probar hipótesis o estimar parámetros.

Punto influyente Punto que afecta fuertemente la gráfica de una línea de regresión.

Punto medio de clase En una clase de una distribución de frecuencias, el valor que está a la mitad entre el límite de clase superior y el límite de clase inferior.

Puntuación estándar Número de desviaciones estándar que un valor dado está por arriba o por debajo de la media; también se llama puntuación z.

Puntuación z Número de desviaciones estándar que un valor dado está por arriba o por debajo de la media.

Racha Secuencia de datos que presentan la misma característica; se usa en la prueba de rachas para detectar aleatoriedad.

Rango Medida de variación que es la diferencia entre los valores máximo y mínimo.

Rango de percentiles 10-90 Diferencia entre los percentiles décimo y nonagésimo.

Rango intercuartilar La diferencia entre los cuartiles primero y tercero.

Rango ordenado Posición numérica de un elemento en un conjunto muestral acomodado en orden.

Rango semi-intercuartilar La mitad de la diferencia entre los cuartiles primero y tercero.

Razón Nivel de medición de los datos; caracteriza datos que pueden ser acomodados en orden, para los que las diferencias entre los valores son significativas y existe un punto de partida cero inherente.

Región crítica El conjunto de todos los valores del estadístico de prueba que harían que se rechazara la hipótesis nula.

Regla de combinaciones Regla para determinar el número de combinaciones diferentes de elementos seleccionados.

Regla de conteo fundamental Regla que dice que, para una secuencia de dos sucesos en la que el primer suceso puede ocurrir de m maneras y el segundo de n maneras, los sucesos juntos pueden ocurrir en un total de $m \cdot n$ maneras.

Regla de la multiplicación Regla para determinar la probabilidad de que ocurra el suceso A en un ensayo y de que ocurra el suceso B en un segundo ensayo.

Regla de la suma Regla para determinar la probabilidad de que, en un solo ensayo, ocurra el suceso A o el suceso B , o bien, de que ocurran ambos.

Regla de permutaciones Regla para determinar el número de arreglos diferentes de elementos seleccionados.

Regla del suceso infrecuente Si, bajo un supuesto dado, la probabilidad de un resultado específico observado es muy pequeña, se concluye que posiblemente el supuesto no sea correcto.

Regla empírica Regla que usa la desviación estándar para dar información sobre datos que tienen una distribución normal (sección 3-3).

Regla factorial Regla que afirma que n objetos distintos se pueden acomodar de $n!$ maneras distintas.

Regla práctica del rango Regla que se basa en el principio de que, para un conjunto típico de datos, la diferencia entre el valor mínimo típico y el valor máximo típico es de aproximadamente cuatro desviaciones estándar ($4s$).

Regresión logística Método usado en la regresión múltiple cuando la variable ficticia es la variable de respuesta (y).

Regresión múltiple Estudio de relaciones lineales entre tres o más variables.

Regresión por pasos Proceso que implica el uso de diferentes combinaciones de variables hasta obtener el mejor modelo; se usa en regresión múltiple.

Réplica Repetición de un experimento.

Residual Diferencia entre un valor muestral y observado y el valor de y que se predice con una ecuación de regresión.

Resumen de cinco números Valor mínimo, valor máximo, mediana y el primer y tercer cuartiles de un conjunto de datos.

SC (del error) Suma de cuadrados que representa la variabilidad que se supone es común a todas las poblaciones consideradas; se usa en el análisis de varianza.

SC (del tratamiento) Medida de la variación entre las medias muestrales; se usa en el análisis de varianza.

SC (total) Medida de la variación total (alrededor de \bar{x}) en todos los datos muestrales combinados; se usa en el análisis de varianza.

Selección aleatoria Selección de elementos muestrales de forma que todos los elementos disponibles para ser seleccionados tengan la misma probabilidad de ser elegidos.

Sesgado No simétrico y que se extiende más hacia un lado que hacia el otro.

Sesgo negativo Sesgado hacia la izquierda.

Sesgo positivo Sesgado hacia la derecha.

Simetría Propiedad de datos cuya distribución puede dividirse en dos mitades que son aproximadamente imágenes especulares al trazar una línea vertical por la mitad.

Simulación Proceso que se comporta de forma similar a algún experimento, por lo que se obtienen resultados similares.

Suceso Resultado de un experimento.

Suceso compuesto Combinación de sucesos simples.

Suceso simple Resultado experimental que no puede descomponerse más.

Sucesos dependientes Sucesos para los cuales la ocurrencia de cualquier suceso individual afecta las probabilidades de ocurrencia de los demás sucesos.

Sucesos disjuntos o mutuamente excluyentes Sucesos que no pueden ocurrir de manera simultánea.

Sucesos independientes Sucesos para los cuales la ocurrencia de cualquiera de ellos no afecta las probabilidades de ocurrencia de los demás.

Sucesos mutuamente excluyentes Sucesos que no pueden ocurrir de manera simultánea.

Tabla de contingencias Tabla de frecuencias observadas en la que los renglones corresponden a una variable de clasificación y las columnas corresponden a otra variable de clasificación; también se denomina tabla de dos factores.

Tabla de dos factores Véase Tabla de contingencia.

Tabla de frecuencias Lista de categorías de valores junto con sus frecuencias correspondientes.

Tamaño de muestra Número de elementos de una muestra.

Teorema de Chebyshev Teorema que usa la desviación estándar para dar información acerca de la distribución de los datos.

Teorema del límite central Teorema que afirma que las medias muestrales tienden a distribuirse normalmente, con una media m y una desviación estándar σ / \sqrt{n} .

Tratamiento Propiedad o característica que permite distinguir unas poblaciones de otras; se usa en el análisis de varianza.

Unidades experimentales Sujetos de un experimento.

Valor crítico Valor que separa la región crítica de los valores del estadístico de prueba que no conducirían al rechazo de la hipótesis nula.

Valor de probabilidad Véase Valor P .

Valor esperado Para una variable aleatoria discreta, la media de los resultados.

Valor P Probabilidad de que un estadístico de prueba, en una prueba de hipótesis, sea al menos tan extremo como el que en realidad se obtuvo.

Valores extremos Valores poco comunes, en el sentido de que están muy lejos de la mayoría de los datos.

Valores predichos Valores de una variable dependiente que se obtienen usando valores de variables independientes en una ecuación de regresión.

Variabilidad de muestreo Variación de un estadístico en distintas muestras.

Variable aleatoria Variable (casi siempre representada con x) que tiene un solo valor numérico (determinado por el azar) para cada resultado de un experimento.

Variable aleatoria continua Variable aleatoria con valores infinitos, que se puede asociar con los puntos en un intervalo lineal continuo.

Variable aleatoria discreta Variable aleatoria que tiene un número finito de valores, o bien, un número de valores que pueden contarse.

Variable de respuesta Variable y en una ecuación de regresión o en una ecuación de regresión múltiple.

Variable dependiente Variable y de una ecuación de regresión o de regresión múltiple.

Variable dicotómica Variable que tiene dos valores discretos posibles.

Variable ficticia o indicadora Variable dicotómica con valores posibles de 0 y 1. Se emplea en la regresión múltiple.

Variable independiente La variable x de una ecuación de regresión, o una de las variables x de una ecuación de regresión múltiple.

Variable interventora Variable que afecta las variables que se están estudiando, pero que no está incluida en el estudio.

Variables de predicción Variables independientes en una ecuación de regresión.

Variación aleatoria Tipo de variación en un proceso que se debe al azar; el tipo de variación inherente a cualquier proceso que no puede generar todos los bienes o servicios exactamente de la misma forma todo el tiempo.

Variación assignable Tipo de variación en un proceso, que es el resultado de causas identificables.

Variación debida al error Véase Variación dentro de muestras.

Variación debida al tratamiento Véase Varianza entre muestras.

Variación dentro de muestras En el análisis de varianza, la variación que se debe al azar.

Variación explicada Suma de los cuadrados de las desviaciones explicadas para todos los pares de datos bivariados de una muestra.

Variación no explicada Suma de los cuadrados de las desviaciones no explicadas para todos los pares de datos bivariados en una muestra.

Variación total Suma de los cuadrados de la desviación total para todos los pares de datos bivariados de una muestra.

Varianza Medida de variación que es igual al cuadrado de la desviación estándar.

Varianza entre muestras En el análisis de varianza, la variación entre las diferentes muestras.

Apéndice D: Bibliografía

*El asterisco significa que se recomienda especialmente la lectura del libro. Los demás libros se recomiendan como textos de referencia.

- Bennett, D. 1998. *Randomness*. Cambridge: Harvard University Press.
- *Best, J. 2001. *Damned Lies and Statistics*. Berkeley: University of California Press.
- *Campbell, S. 2004. *Flaws and Fallacies in Statistical Thinking*. Mineola. N.Y: Dover Publications.
- *Crossen, C. 1994. *Tainted Truth: The Manipulation of Fact in America*. Nueva York: Simon & Schuster.
- *Freedman, D., R. Pisani, R. Purves y A. Adhikari. 1997. *Statistics*. 3a. edición, Nueva York: Norton.
- *Gonick, L. y W. Smith. 1993. *The Cartoon Guide to Statistics*. Nueva York: HarperCollins.
- Halsey, J. y E. Reda. 2006. *Excel Student Laboratory Manual and Workbook*. Boston: Addison-Wesley.
- *Heyde, C. y E. Seneta (eds.). 2001. *Statisticians of the Centuries*. Nueva York: Springer-Verlag.
- *Hollander, M. y F. Proschan. 1984. *The Statistical Exorcist: Dispelling Statistics Anxiety*. Nueva York: Marcel Dekker.
- *Holmes, C. 1990. *The Honest Truth About Lying with Statistics*. Springfield, Ill.: Charles C. Thomas.
- *Hooke, R. 1983. *How to Tell the Liars from the Statisticians*. Nueva York: Marcel Dekker.
- *Huff, D. 1993. *How to Lie with Statistics*. Nueva York: Norton.
- Humphrey, P. 2006. *Graphing Calculator Manual for the TI-83 Plus, TI-84 Plus, and the TI-89*, Boston: Addison-Wesley.
- *Jaffe, A. y H. Spirer. 1987. *Misused Statistics*. Nueva York: Marcel Dekker.
- *Kimble, G. 1978. *How to Use (and Misuse) Statistics*. Englewood Cliffs, N.J.: Prentice-Hall.
- Kotz, S. y D. Stroup. 1983. *Educated Guessing-How to Cope in an Uncertain World*. Nueva York: Marcel Dekker.
- *Loyer, M. 2006. *Student Solutions Manual to Accompany Elementary Statistics*. 10a. edición, Boston: Addison-Wesley.
- *Moore. D. 2001. *Statistics: Concepts and Controversies*. 5a. edición, San Francisco: Freeman.
- Morgan, J. 2006. *SAS Student Laboratory Manual and Workbook*. 3a. edición, Boston: Addison-Wesley.
- *Paulos, J. 2001. *Innumeracy: Mathematical Illiteracy and Its Consequences*. Nueva York: Hill and Wang.
- Peck, R. 2006. *SPSS Student Laboratory Manual and Workbook*. Boston: Addison-Wesley.
- *Reichard, R. 1974. *Tire Figure Finaglers*. Nueva York: McGraw-Hill.
- *Reichmann, W. 1962. *Use and Abuse of Statistics*. Nueva York: Oxford University Press.
- *Rossmann, A. 1996. *Workshop Statistics: Discovery with Data*. Nueva York: Springer.
- *Salsburg, D. 2000. *The Lady, Tasting Tea: How Statistics Revolutionized the Twentieth Century*. Nueva York: W. H. Freeman.
- Sheskin, D. 1997. *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton. Fla.: CRC Press.
- Simon, J. 1992. *Resampling: The New Statistics*. Belmont, Calif.: Duxbury Press.
- *Stigler, S. 1986. *The History of Statistics*, Cambridge, Mass.: Harvard University Press.
- *Tanur, J., (ed.) 1989. *Statistics: A Guide to the Unknown*. 3a edición, Belmont, Calif.: Wadsworth.
- Triola, M. 2006. *Minitab Student Laboratory Manual and Workbook*. 10a. edición, Boston: Addison-Wesley.
- Triola, M. 2006. *STATDISK 10.0 Student Laboratory Manual and Workbook*. 10a. edición, Boston: Addison-Wesley.
- Triola, M. y L. Franklin. 1994. *Business Statistics*. Boston: Addison-Wesley.
- Triola, M. y M. Triola. 2006. *Biostatistics for the Biological and Health Sciences*. Boston: Addison-Wesley.
- *Tufte, E. 2001. *The Visual Display of Quantitative Information*. 2a. edición, Cheshire, Conn.: Graphics Press.
- Tukey, J. 1977. *Exploratory Data Analysis*. Boston: Addison-Wesley.
- Zwillinger, D. y S. Kokoska. 2000. *CRC Standard Probability and Statistics Tables and Formulae*. Boca Raton, Fla.: CRC Press.

Apéndice E: Soluciones de los ejercicios impares (y de TODOS los siguientes ejercicios: Los ejercicios de conocimientos estadísticos y pensamiento crítico que aparecen al final de los capítulos, los ejercicios de repaso y los ejercicios de repaso acumulativo)

Capítulo 1 Respuestas

Sección 1-2

1. Un parámetro es una medición numérica que describe alguna característica de una población; mientras que un estadístico es una medición numérica que describe alguna característica de una muestra.
3. Los datos discretos resultan cuando el número de valores posibles es finito o puede contarse (donde el número de valores posibles es 0, 1 o 2, etc.), mientras que los datos continuos resultan de un número infinito de valores posibles, que corresponden alguna escala continua que cubre un rango de valores sin huecos, interrupciones o saltos.
5. Estadístico
7. Parámetro
9. Discretos
11. Continuos
13. Nominal
15. Nominal
17. De intervalo
19. Ordinal
21. Muestra: los 25 senadores seleccionados. Población: los 100 senadores actualmente en servicio. Es probable que la muestra sea representativa.
23. Muestra: los 1059 adultos seleccionados. Población: todos los adultos. Es probable que la muestra sea representativa.
25. Al carecer de un punto inicial natural, las temperaturas tienen un nivel de medición de intervalo, de manera que operaciones como "el doble de" carecen de significado.
27. Ordinal o de intervalo son respuestas razonables, aunque ordinal es más lógico debido a que es probable que las diferencias entre valores no tengan un significado. Por ejemplo, es probable que la diferencia entre un alimento con calificación de 1 y un alimento con calificación de 2 no sea igual a la diferencia entre un alimento con calificación de 9 y uno con calificación de 10.

Sección 1-3

1. Una muestra de respuesta voluntaria (o muestra autoseleccionada) es aquella en la que los sujetos deciden participar. Es inadecuada porque las personas con intereses especiales son más propensas a participar y la muestra tiende a estar sesgada.
3. No. Puesto que la tasa de respuesta es muy baja y que se trata de una muestra de respuesta voluntaria, es muy probable que esté sesgada.
5. Las personas más altas tienen la ventaja de estar más cerca de la canasta, por lo que la gente alta suele tener un mejor desempeño y mayores probabilidades de jugar.
7. Quizás los oficiales de policía sean más propensos a detener a individuos de grupos minoritarios que a individuos blancos, de manera que los primeros reciben más multas que los últimos.

9. Puesto que el estudio fue financiado por una empresa de dulces y por la Chocolate Manufacturers Association, existe la posibilidad real de que los investigadores estuvieran motivados a obtener resultados favorables para el consumo de chocolate.
11. No, ella utilizó una muestra de respuesta voluntaria.
13. No se incluyen las personas con números privados ni a las que no cuentan con teléfono, de manera que la muestra podría estar sesgada.
15. Los motociclistas que murieron.
17. No necesariamente, porque se tomarían en cuenta los tamaños poblacionales. Sería mucho mejor obtener una muestra aleatoria nacional de individuos asalariados y calcular el promedio de sus ingresos.
19. Faltan los hogares sin hijos. Los resultados no serían representativos.
21. a. 15%
b. 0.567
c. 170
d. 78.9%
23. a. 540
b. 5%
25. El 62% del 8% de 1875 es tan sólo 93.
27. Todos los porcentajes de éxito deben ser múltiplos de 5. Los porcentajes dados no pueden ser correctos.

Sección 1-4

1. En una muestra aleatoria, cada individuo tiene las mismas probabilidades de ser elegido. En una muestra aleatoria simple, todas las muestras del mismo tamaño tienen las mismas probabilidades de ser elegidas.
3. Un estudio ciego es un método en el que un sujeto (o investigador) en un experimento no sabe si tal sujeto está recibiendo un tratamiento o un placebo. Es importante el uso de los estudios ciegos para que los resultados no se vean distorsionados debido a un efecto placebo, en el que los sujetos creen que experimentan mejorías sólo por el hecho de ser tratados.
5. Estudio observacional
7. Estudio observacional
9. Prospectivo
11. Transversal
13. Sistemático
15. Estratificado
17. Por conglomerados
19. De conveniencia
21. Aleatorio
23. Por conglomerados
25. Sí, no. Los estudiantes individuales (incluso aquellos que pertenecen a clases grandes y a clases pequeñas) tienen las mismas probabilidades de ser elegidos. Algunas muestras no son posibles, como las que incluyen estudiantes de al menos seis clases diferentes.
27. No, no. El segundo M&M no tiene probabilidades de ser elegido. Las muestras que incluyen al segundo M&M no tienen probabilidades de ser incluidas.
29. Sí, sí. Cada estudiante tiene la misma probabilidad, y cada muestra de tamaño seis tiene las mismas probabilidades de ser elegida.

31. Las respuestas varían.
33. Prospectivo: el experimento se inició y los resultados se siguieron a lo largo del tiempo. Aleatorizado: los sujetos fueron asignados a los diferentes grupos a través de un proceso de selección aleatoria, por el cual tenían las mismas probabilidades de pertenecer a cada grupo. Doble ciego: los sujetos no sabían a cuál de los tres grupos pertenecían, y tampoco lo sabían las personas que evaluaron los resultados. Placebo controlado: había un grupo de sujetos que recibió un placebo; al comparar el grupo placebo con los dos grupos de tratamiento es posible comprender mejor los efectos de los tratamientos.

Capítulo 1 Conocimientos estadísticos y pensamiento crítico

1. No. Si una muestra se obtiene de forma inapropiada, como el uso de una muestra de respuesta voluntaria, tiene muchas probabilidades de estar sesgada y de ser una muestra inadecuada, incluso si es muy grande.
2. a. Cuantitativo
b. Continuo
c. Los corredores que terminaron ese maratón en particular.
3. No, porque se trata de una muestra de respuesta voluntaria. Es probable que esté sesgada.
4. No. A las personas se les pregunta el valor de sus automóviles y es probable que exageren. Asimismo, los resultados deben ser ponderados para que representen los distintos números de propietarios de automóviles en los diferentes estados.

Capítulo 1 Ejercicios de repaso

1. No, porque se trata de una muestra de respuesta voluntaria y es probable que no sea representativa de la población.
2. La respuesta varía.
3. a. De razón
b. Ordinal
c. Nominal
d. De intervalo
4. a. Discretos
b. De razón
c. Estratificado
d. Estadístico
e. Los valores más grandes, porque representan a los accionistas que podrían adquirir el control de la compañía.
5. a. Sistemático; representativo
b. De conveniencia; no representativo
c. Por conglomerados; no representativo
d. Aleatorio; representativo
e. Estratificado; no representativo
6. a. Diseñar el experimento de tal forma que los sujetos no sepan si están utilizando Sleepeze o un placebo, y también diseñarlo para que aquellos que observan y evalúan a los sujetos tampoco sepan cuáles están utilizando Sleepeze y cuáles un placebo.

- b. El estudio ciego ayudará a distinguir entre la eficacia del Sleepeze y el efecto placebo, en el que los sujetos y los evaluadores tienden a creer que las mejorías se deben únicamente al tratamiento que se está aplicando.
 - c. Los sujetos se asignan a diferentes grupos por medio de un proceso de *selección aleatoria*.
 - d. Los sujetos se *asignan con mucho cuidado* a los diferentes grupos, de manera que éstos sean similares en los aspectos importantes.
 - e. La réplica se utiliza cuando el experimento se repite. Es importante tener una muestra de sujetos que sea lo suficientemente grande para poder ver la verdadera naturaleza de cualquier efecto. También es importante no confundirse por la conducta errática de muestras que son demasiado pequeñas.
7. a. Parámetro
b. Discretos
c. 743,005
 8. a. Si no contienen grasa en lo absoluto, entonces tienen 100% menos que cualquier otra cantidad con grasa, de manera que el dato del 125% no puede ser correcto.
b. 45
c. 16.1%

Capítulo 1 Ejercicios de repaso acumulativo

1. 3.02754 g
2. -0.64516129
3. -6.6423420
4. 266.77778
5. 0.5555556
6. 7
7. 2.6457513
8. 0.89735239
9. 0.0009765625
10. 1,099,511,627,776 (La mayoría de las calculadoras no muestran los últimos dígitos, de manera que un resultado como 1,099,511,600,000 es aceptable).
11. 13,841,287,201 (La mayoría de las calculadoras no muestran los últimos dígitos, de manera que un resultado como 13,841,287,000 es aceptable).
12. 0.000014272477

Capítulo 2 Respuestas

Sección 2-2

1. Una distribución de frecuencias lista valores de datos (ya sea de manera individual o por grupos de valores), junto con sus conteos de frecuencia correspondientes. Sirve para organizar y resumir datos.
3. Para valores tales como 10, puede pertenecer a cualquiera de dos clases, pero cada valor debe pertenecer sólo a una clase. Se debe evitar el traslape de los límites de clase.
5. Anchura de clase: 5. Marcas de clase: 37, 42, 47, 52, 57, 62, 67. Fronteras de clase: 34.5, 39.5, 44.5, 49.5, 54.5, 59.5, 64.5, 69.5.
7. Anchura de clase: 5.0. Marcas de clase: 62.45, 67.45, 72.45, 77.45, 82.45, 87.45, 92.45, 97.45, 102.45, 107.45. Fronteras de clase: 59.95, 64.95, 69.95, 74.95, 79.95, 84.95, 89.95, 94.95, 99.95, 104.95, 109.95.

9. Sí

11. El hombre más alto mide entre 105.0 pulgadas y 109.9 pulgadas, que es mayor a una estatura de 8 pies. Probablemente ese valor es incorrecto. Despues de borrar ese error, la distribución parece ser aproximadamente normal.

13. Temperatura mínima diaria (°F)	Frecuencia relativa
35–39	3%
40–44	9%
45–49	14%
50–54	31%
55–59	20%
60–64	20%
65–69	3%

15. Temperatura mínima diaria (°F)	Frecuencia acumulativa
Menor que 40	1
Menor que 45	4
Menor que 50	9
Menor que 55	20
Menor que 60	27
Menor que 65	34
Menor que 70	35

17. Debido a que hay un número desproporcionadamente mayor de ceros y cinco, parece que las estaturas fueron reportadas en lugar de medidas. En consecuencia, es probable que los resultados no sean muy precisos.

x	Frecuencia
0	9
1	2
2	1
3	3
4	1
5	15
6	2
7	0
8	3
9	1

19. La distribución no parece ser normal. La mayoría de los días no hay precipitación pluvial. La distribución no es simétrica y hay muy pocos días con grandes cantidades de precipitación.

Precipitación	Frecuencia
0.00–0.19	44
0.20–0.39	6
0.40–0.59	1
0.60–0.79	0
0.80–0.99	0
1.00–1.19	0
1.20–1.39	1

21. Al parecer la distribución es aproximadamente normal.

IMC	Frecuencia
15.0–20.9	10
21.0–26.9	15
27.0–32.9	11
33.0–38.9	2
39.0–44.9	2

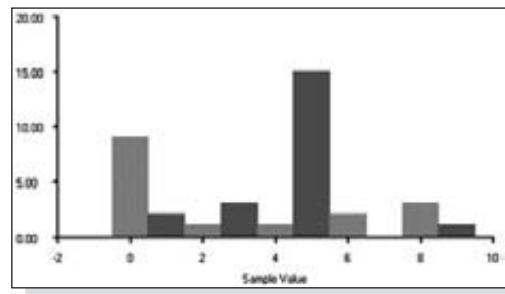
23. La distribución parece ser aproximadamente normal.

Peso	Frecuencia
2.9500–2.9999	2
3.0000–3.0499	3
3.0500–3.0999	22
3.1000–3.1499	7
3.1500–3.1999	1

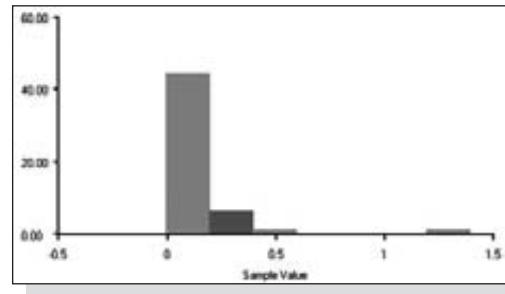
25. Las respuestas varían dependiendo de la anchura de clase y punto de inicio que se elijan. Las distribuciones de frecuencias relativas no son demasiado diferentes, con excepción del valor extremo de 504 libras, incluido en la lista de cargas axiales de las latas con 0.0111 pulgadas de ancho.
27. 46-90, 91-181, 182-362, 363-724, 725-1448, 1449-2896.

Sección 2-3

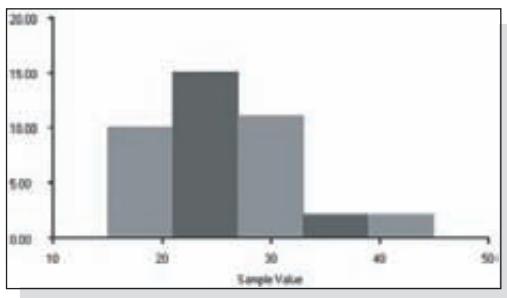
- Distribución de los datos
- No se puede observar la verdadera naturaleza de la distribución.
- 18
- En comparación con lo remeros, los timoneles tienden a ser muy ligeros. Lo más probable es que los dos miembros de la tripulación con los pesos más bajos sean timoneles.
- Los dígitos cero y cinco ocurren desproporcionadamente con mayor frecuencia que los demás, de manera que parece que las estaturas fueron reportadas y no medidas.



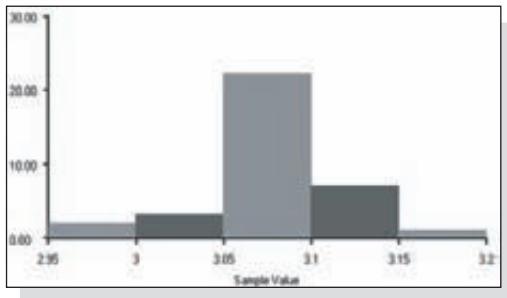
11. Los datos no parecen estar distribuidos normalmente.



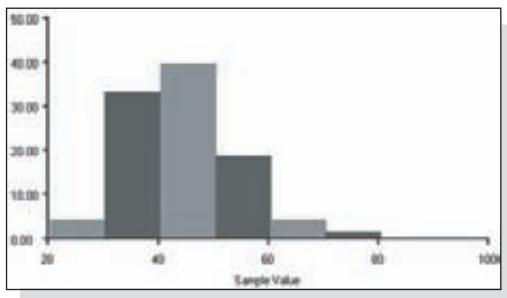
13. Parece que los datos tienen una distribución aproximadamente normal.



15. Parece que los datos tienen una distribución aproximadamente normal.



17. Las formas de las distribuciones no son radicalmente diferentes, pero parece que las edades de las actrices se centran alrededor de los 36 años, mientras que las edades de los actores parecen concentrarse alrededor de los 43 años. Al parecer las actrices ganan Óscars cuando son más jóvenes que los actores.

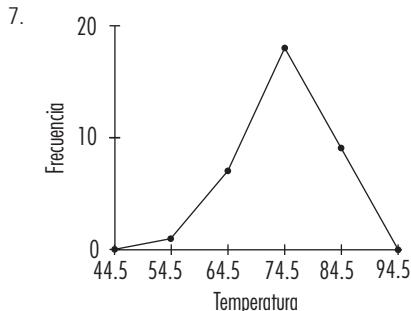
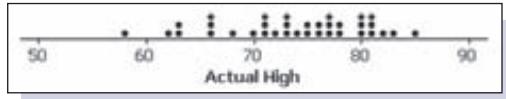


19. Los histogramas de frecuencias relativas continuos pueden variar hasta cierto punto, dependiendo de las fronteras de clase elegidas. Las gráficas deben mostrar que las latas con una anchura de 0.0109 pulgadas tienen cargas axiales que se concentran alrededor de 270 libras, mientras que las latas con una anchura de 0.0111 pulgadas tienen cargas que se concentran alrededor de 280 libras. No existe una diferencia notable.

Sección 2-4

- Los datos deben graficarse de manera que se puedan ver las características que no pueden observarse al hacer un examen sencillo de la lista de los datos. Las gráficas revelan características como la naturaleza de la distribución y la presencia de valores extremos.
- Los datos deben ordenarse en alguna secuencia temporal, y la gráfica de series del tiempo sirve para revelar tendencias o patrones a lo largo del tiempo.

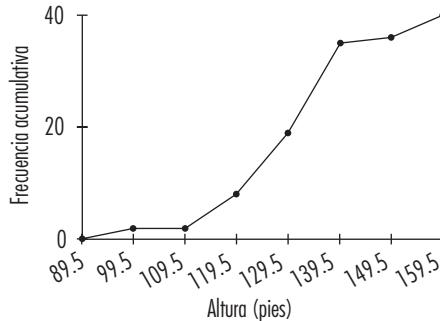
5. Las temperaturas máximas reales van de los 58 grados a los 85 grados, y la mayoría de las lecturas se concentra en el rango de 70-80 grados.



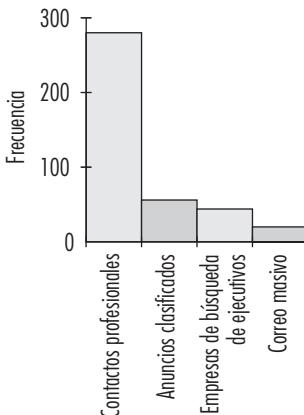
9. Parece que la distribución es aproximadamente normal.

9	55
10	00
11	000055
12	000000005555
13	0000000066668
14	000088
15	00

11. Erupciones por debajo de 120 pies: 7.

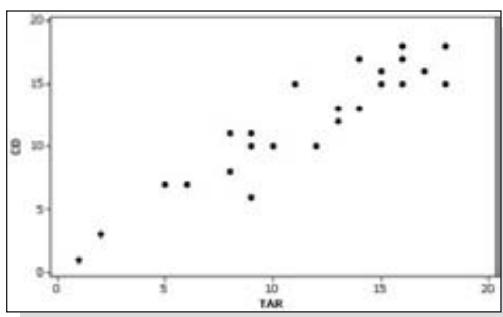


13. Al parecer los contactos profesionales son el método más eficaz para obtener un empleo.

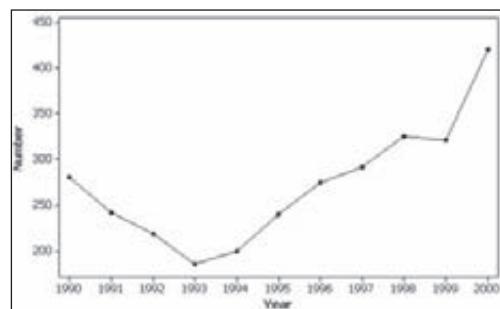




17. Al parecer existe una relación en la que mayores cantidades de alquitrán están asociadas con mayores cantidades de monóxido de carbono.



19. Al parecer existe una tendencia reciente al incremento de fallas de aterrizaje.



21. 10,000; 2.4%
 23. 100,000; 500 millas
 25. Las edades de las actrices son evidentemente menores que las de los actores.

Actresses' Ages (units)	Stem (tens)	Actors' Ages (units)
99999988887777666655554421	2	9
9988887765555554443333221100	3	0012224455666777888999
965322111110	4	0001111122222333344455567778999
40	5	112222346677
310	6	00222
4	7	6
0	8	

Capítulo 2 Conocimientos estadísticos y pensamiento crítico

- En general el histograma es más eficaz porque ofrece una imagen, y una imagen visual es mucho más fácil de entender que una tabla con números.
- Si los dos conjuntos de datos tienen números de valores sumamente diferentes, se dificulta más la comparación de las distribuciones de frecuencias, ya que debemos comparar números con magnitudes muy distintas. Puesto que la distribución de frecuencias relativas utiliza porcentajes, los números pueden compararse. Por lo tanto, las distribuciones de frecuencias relativas son mejores.
- Sería mejor una gráfica de series de tiempo, ya que los precios de venta han cambiado drásticamente con el paso del tiempo. El histograma ocultaría el factor del tiempo.

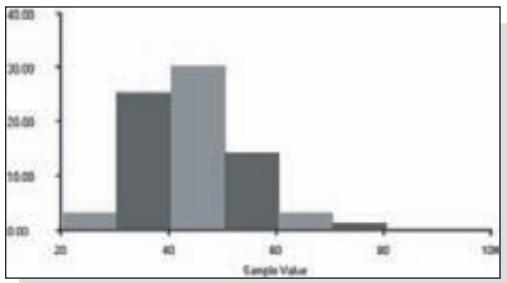
- Las barras del histograma tienen un inicio relativamente bajo, luego se incrementan hasta una altura máxima y luego disminuye nuevamente. El histograma es simétrico, ya que la mitad izquierda es casi una imagen en espejo de la mitad derecha.

Capítulo 2 Ejercicios de repaso

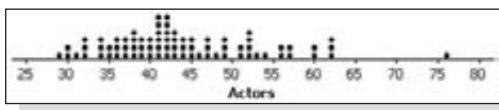
- Las edades de los actores son mayores que las edades de las actrices.

Edad de los actores	Frecuencia
21–30	3
31–40	25
41–50	30
51–60	14
61–70	3
71–80	1

2. El histograma muestra que los actores son mayores que las actrices.



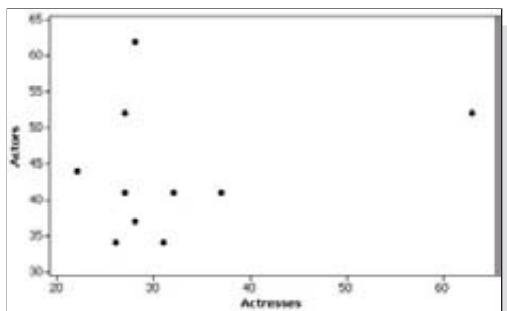
3. La gráfica de puntos muestra que los actores son mayores que las actrices cuando ganan Óscares.



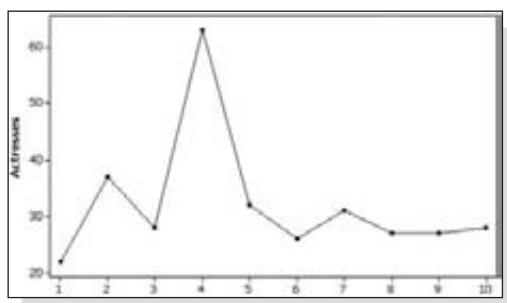
4. La gráfica de tallo y hojas indica que los actores son mayores que las actrices cuando ganan Óscars.

2	9
3	001222444556667778888999
4	000111112222233334455567778999
5	112222346677
6	00222
7	6

5. Al parecer no existe una asociación entre las edades de las actrices y de los actores cuando ganan Óscars.



6. La gráfica no revela una tendencia. Las edades no cambian sistemáticamente con el paso del tiempo.



Capítulo 2 Ejercicios de repaso acumulativo

- No, los números se utilizan para identificar la ranura en la que cae la pelota. Se podrían haber utilizado letras u otros símbolos.
- Los resultados se encuentran a nivel de medición nominal. (Los conteos de frecuencias representan un nivel de medición de razón).
- Las primeras siete clases representan cinco ranuras cada una, mientras que la última clase representa tres ranuras. Por lo tanto, la frecuencia esperada de la última clase debe ser alrededor de $3/5$ de la frecuencia típica para cada una de las primeras siete clases. La frecuencia de 25 no es drásticamente diferente de la frecuencia esperada por el azar.
- Sí. El promedio debe ser aproximadamente 18, de manera que un promedio de cinco sugiere que la ruleta no se está comportando de forma aleatoria. Al apostar a números pequeños, el jugador podría incrementar las probabilidades de ganar.
- No. La encuesta de respuesta voluntaria tiende a estar sesgada, de manera que los resultados probablemente no reflejen la población de todos los propietarios de automóviles.

Capítulo 3 Respuestas

Sección 3-2

- Utilizan diferentes métodos para proporcionar valores centrales o intermedios de un conjunto de datos.
- No. Debido a que los números no miden o cuentan algo, la media sería un estadístico sin significado.
- $\bar{x} = 58.3$ segundos; mediana = 55.5 segundos; moda = 49 segundos; mitad del rango = 62.0 segundos. La muestra de estudiantes no es una muestra representativa de la población.
- $\bar{x} = 1.9$; mediana = 2.0; moda = 1; mitad del rango = 2.5. La moda de 1 indica correctamente que los chicharos de color amarillo claro ocurren con mayor frecuencia que cualquier otro fenotipo, pero las otras medidas de tendencia central no tienen significado con esos datos a nivel de medición nominal.
- $\bar{x} = 133.9$; mediana = 132.5; moda = 130; mitad del rango = 135. Dado que se midió a la misma persona, los valores parecen variar de manera considerable.
- $\bar{x} = 0.807$ mm; mediana = 0.840 mm; moda = 0.84 mm; mitad del rango = 0.780 mm. La muestra de Pocatello no sería representativa de la población de Estados Unidos.
- Cat on a Hot Tin Roof.* $\bar{x} = 3.9$; mediana = 3.0. *The Cat in the Hat*: $\bar{x} = 3.1$; mediana = 3.0. Con base en las medias, las palabras en el primer libro parecen ser más largas que las del segundo libro.
- Un día: $\bar{x} = 0.5$ grados; mediana = 0.0 grados. Cinco días: $\bar{x} = -0.4$ grados; mediana = -1.0 grados. Al parecer no existe una diferencia sustancial en la precisión.
- Jefferson Valley: $\bar{x} = 7.15$ min; mediana = 7.20 min. Providence: Los mismos resultados que Jefferson Valley. Aunque las medidas de tendencia central son las mismas, los tiempos de espera del Providence son mucho más variados que los tiempos de espera del Jefferson Valley.
- Antes de 1983: $\bar{x} = 3.08950$ g; mediana = 3.09065 g. Despues de 1983: $\bar{x} = 2.50251$ g; mediana = 2.50110 g. Parece que los pesos de las monedas acuñadas después de 1983 son sustancialmente más bajos que los pesos de las monedas acuñadas antes de 1983.

21. Un día: $\bar{x} = -0.6$ grados; mediana = 0 grados. Cinco días: $\bar{x} = -0.7$ grados; mediana = -1.0 grados. El pronóstico de un día parece ser un poco más preciso. Hay mayor confianza en los resultados de un conjunto de datos grande que en los resultados de un conjunto de datos pequeño.
23. Antes de 1983: $\bar{x} = 3.07478$ g; mediana = 3.07630 g. Después de 1983: $\bar{x} = 2.49910$ g; mediana = 2.50040 g. La misma conclusión.
25. 53.4 grados, que se acerca a la media de 53.8 grados calculada a partir de la lista original de datos.
27. 46.8 mi/h, que es muy cercano al valor calculado a partir de la lista original de datos.
29. a. 182.9 libras
b. 171.0 libras
c. 159.2 libras
- Los resultados difieren en cantidades importantes, lo que sugiere que la media del conjunto original de pesos se ve muy afectada por los valores extremos.
31. La media es de al menos 3.42 años.
33. 48.0 mi/h
35. 109.35 volts
23. Antes de 1983: rango = 0.19890 g; $s^2 = 0.00153$ g²; $s = 0.03910$ g.
Después de 1983: rango = 0.07690 g; $s^2 = 0.00027$ g²; $s = 0.01648$ g.
Los centavos acuñados antes de 1983 tienen pesos con mayor variación que los centavos acuñados después de 1983.
25. 7.0 se acerca mucho a 6.9, calculado a partir de la lista original de datos.
27. 4.1 se acerca mucho a 4.0, calculado a partir de la lista original de datos.
29. La respuesta varía, pero si se utiliza un mínimo de 23 años y un máximo de 70 años, se estima que la desviación estándar sería de alrededor de 12 años.
31. 1830 kWh, 3846 kWh. Sí, 578 kWh es menor que el valor mínimo común.
33. a. 68%
b. 99.7%
35. Hombres: 3.1%. Huevos: 5.1%. Las cantidades relativas de variación no difieren de manera importante.
37. El efecto del valor extremo sería muy grande.

Sección 3-3

- Es una medida de la cantidad que los valores se desvían (o varían) de la media.
- 85 es un valor inusualmente alto porque está a más de dos desviaciones estándar por arriba de la media.
- Rango = 26.0 s; $s^2 = 89.6$ s²; $s = 9.5$ s. La muestra es muy pequeña.
- Rango = 3.0; $s^2 = 0.9$; $s = 0.9$. Debido a que los datos se encuentran a un nivel nominal de medición, esos resultados no tienen sentido.
- Rango = 30.0; $s^2 = 81.8$; $s = 9.0$. De manera ideal, la desviación estándar sería cero.
- Rango = 0.280 mm; $s^2 = 0.009$ mm²; $s = 0.094$ mm. El estimado no es razonable porque la muestra no es representativa.
- Cat on a Hot Tin Roof*: rango = 10.0; $s^2 = 6.8$; $s = 2.6$. *The Cat in the Hat*: rango = 3.0; $s^2 = 0.8$; $s = 0.9$. Existe una variación mucho menor en la longitud de las palabras en *The Cat in the Hat*.
- Un día: rango = 11.0 grados; $s^2 = 6.9$ grados²; $s = 2.6$ grados. Cinco días: rango = 15.0 grados; $s^2 = 20.6$ grados²; $s = 4.5$ grados. Al parecer existe una mayor variación en los errores del pronóstico a cinco días.
- Jefferson Valley: rango = 1.20 min; $s^2 = 0.23$ min²; $s = 0.48$ min. Providence: rango = 5.80 min; $s^2 = 3.32$ min²; $s = 1.82$ min. La variación en una sola fila es mucho menor que en las filas múltiples.
- Antes de 1983: rango = 0.11840 g; $s^2 = 0.00145$ g²; $s = 0.03812$ g. Despues de 1983: rango = 0.04750 g; $s^2 = 0.00023$ g²; $s = 0.01506$ g. Los centavos acuñados antes de 1983 tienen pesos que varían más que los centavos acuñados después de 1983.
- Un día: rango = 14.0 grados; $s^2 = 6.5$ grados²; $s = 2.5$ grados. Cinco días: rango = 24.0 grados; $s^2 = 31.2$ grados²; $s = 5.6$ grados. Al parecer hay una mayor variación en los errores del pronóstico a cinco días. La conclusión es la misma, pero con mayor confianza porque los tamaños muestrales son más grandes.

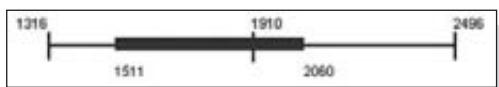
Sección 3-4

- El valor está por debajo de la media en una cantidad igual a dos desviaciones estándar.
- Alrededor del 25% de los valores están por debajo de 15 y alrededor del 75% de los valores están por arriba de 15.
- a. 6 cm.
b. 6/7 o 0.86
c. 0.86
d. Común
- a. 4.5 pulgadas
b. 2.14
c. -2.14
d. Infrecuente
- a. -1.13
b. 0.65
c. 0
- 2.67; infrecuente
- En una prueba psicológica, ya que $z = -0.50$ es mayor que $z = -2.00$.
- 5
- 75
- 25
- 28
- 29
- 21
- a. 11.5
b. 33.75
c. 24
d. Sí; sí
e. No; no
- a. P_{10} , P_{50} , P_{80}
b. 25, 27, 29, 31, 33.5, 35, 38, 41, 49
c. 27, 31, 35, 41

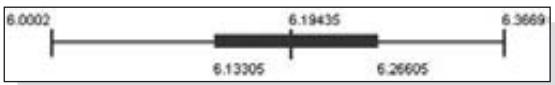
Sección 3-5

- 2 es el valor mínimo, 5 es el primer cuartil, 10 es la mediana, 12 es el tercer cuartil y 20 es el valor máximo.
- Sigma tiene menor variación. Debido a que sigma tiene menor variación, los estimados de los costos de reparación tenderán a ser más precisos, de tal modo que los costos serán más predecibles.

5.



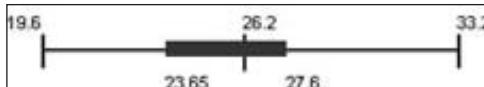
7.



9. Aproximadamente simétrica



11.



- Seleccione los acumuladores representados por la gráfica de cuadro superior. Éstos tienen la mejor combinación de la media más alta y la variación más baja.
- Ligero: 0.03; extremo: 0.05, 0.11, 0.12, 0.41, 0.43, 0.47, 0.49, 0.59, 0.92, 1.41.

Capítulo 3 Conocimientos estadísticos y pensamiento crítico

- No. La reducción de la variación implica que los tiempos de reparación son más cercanos entre sí, pero la media no cambia necesariamente.
- No debe ignorarse la desviación estándar. Es posible tener una media mayor que 10 años con una desviación estándar tan alta que algunos acumuladores fallarán poco tiempo después de haber sido instalados. Esa situación sería inaceptable.
- La media y la desviación estándar cambiarán de manera drástica, pero probablemente la mediana cambiará muy poco.
- Debido a que los datos provienen de una muestra de respuesta voluntaria, es muy probable que los valores no sean representativos de la población de todos los automóviles propiedad de estadounidenses. Asimismo, es muy probable que los participantes exageren los verdaderos valores de sus automóviles.

Capítulo 3 Ejercicios de repaso

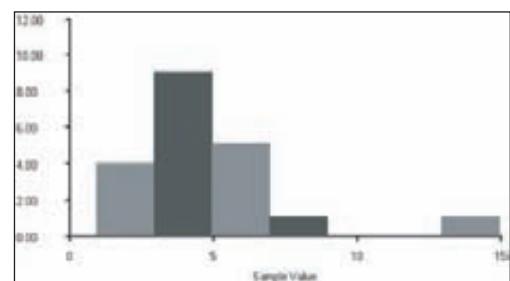
- 4.54 pies
- 3.95 pies
- 1.8 pies, 3.7 pies, 5.1 pies
- 7.75 pies
- 11.90 pies
- 2.65 pies

- 7.03 pies^2
- 3.25 pies
- 5.15 pies
- 1.85 pies
- a. 3.46
- Sí, porque son más de dos desviaciones estándar por arriba de la media.
- Ningún otro valor es infrecuente.

- 3.
- | x | Frecuencia |
|-----------|------------|
| 1.0–2.9 | 4 |
| 3.0–4.9 | 9 |
| 5.0–6.9 | 5 |
| 7.0–8.9 | 1 |
| 9.0–10.9 | 0 |
| 11.0–12.9 | 0 |
| 13.0–14.9 | 1 |

x	Frecuencia
1.0–2.9	4
3.0–4.9	9
5.0–6.9	5
7.0–8.9	1
9.0–10.9	0
11.0–12.9	0
13.0–14.9	1

4. Sesgada



- 1.8
- 3.95
- 5.15
- 13.7
- The puntuación del 19 es mejor porque $z = -0.20$ es mayor que $z = -0.67$.
- a. La respuesta varía, pero un estimado de alrededor de 4 años es razonable.
- b. La respuesta varía, pero si se utiliza un mínimo de 0 años y un máximo de 20 años, el resultado es $s = 5.0$ años.
- Mínimo: 64 pulgadas; máximo: 74 pulgadas. Una estatura de 72 pulgadas no es infrecuente porque cae dentro del rango de los valores comunes.

Capítulo 3 Ejercicios de repaso acumulativo

- Continuos
- De razón
- a. La moda es más apropiada porque identifica la elección más común. Las otras medidas de tendencia central no se aplican a datos que tienen un nivel de medición nominal.
- De conveniencia
- Por conglomerados
- La desviación estándar debe disminuirse

Capítulo 4 Respuestas

Sección 4-2

1. Existe una posibilidad en 20,358,520. Un triunfo como este sería infrecuente porque su probabilidad es demasiado baja.
3. La afirmación es correcta. Puesto que la probabilidad es demasiado baja, el suceso ocurrirá en promedio sólo una vez en cada 1000 ensayos, de manera que el suceso es infrecuente.
5. a. 1
b. $1/10$ o 0.1
c. 0
7. $-1, 2, 5/3, \sqrt{2}$
9. a. $3/8$
b. $3/8$
c. $1/8$
11. 0.738; sí
13. 0.908; sí
15. a. 0.27
b. El resultado coincide con la aseveración de que el 24% son azules.
17. a. $1/365$
b. Sí
c. Él ya lo sabía.
d. 0
19. 0.34; no; sí
21. 0.159; no; sí
23. a. 0.135
b. No
25. a. hh, hm, mh, mm
b. $1/4$
c. $1/2$
27. a. Café/café, café/azul, café/café, café/azul
b. 0
c. 1
29. 423:77 o aproximadamente 5.5:1 o 11:2
31. a. $18/38$ o 0.474
b. 10:9
c. \$18
d. \$20
33. $1/10$ o 0.1
35. 1

Sección 4-3

1. Los dos sucesos no pueden ocurrir al mismo tiempo.
3. Es probable que la muestra de conveniencia no sea representativa de la población; podría estar sesgada y, por lo tanto, proporcionar resultados incorrectos.
5. a. No
b. No
c. Yes
7. a. 0.95
b. 0.9975
9. 0.410

11. 0.920
13. 0.6
15. 0.49
17. 0.14
19. 0.87
21. Renglón superior: 5, 3. Renglón inferior: 4, 2.
23. 0.290
25. No. He aquí un ejemplo: A = suceso de seleccionar un hombre menor de 30 años de edad, B = seleccionar una mujer, C = seleccionar un hombre mayor de 18 años de edad.
27. $P(A \text{ o } B \text{ o } C) = P(A) + P(B) + P(C) - P(A \text{ y } B) - P(A \text{ y } C) - P(B \text{ y } C) + P(A \text{ y } B \text{ y } C)$

Sección 4-4

1. La ocurrencia de uno de los sucesos no afecta la probabilidad del otro suceso.
3. Muestreo sin reemplazo. El segundo resultado no es independiente del primero.
5. a. Independiente
b. Independiente
c. Dependiente
7. $1/8$; no
9. a. 0.00238
b. 0.00200
c. Si se seleccionaran casos para un estudio de seguimiento, no tendría mucho sentido seleccionar el mismo elemento dos veces, de manera que conviene hacer una selección sin reemplazo.
11. 0.922
13. $1/4096$; sí. La probabilidad de obtener 12 niñas en 12 nacimientos por el azar es tan baja, que el resultado de 12 niños sugiere que el método es eficaz.
15. a. 0.025
b. 0.000625
c. 0.999375
d. Sí
17. 0.0195
19. a. 0.590
b. 0.348
c. 0.348
d. Los resultados son iguales cuando se redondea a tres decimales.
21. a. 0.992
b. 0.973
c. 0.431
23. 0.0192

Sección 4-5

1. El número de defectos es 1 o 2 o 3 o 4 o 5 o 6 o 7 o 8 o 9 o 10.
3. No. Se ignoró la proporción de usuarios masculinos de tarjetas de crédito.
5. Los seis aspirantes resultan negativos.
7. Al menos uno tiene el gene.

9. 0.999
11. 15/16; sí
13. 0.5; no
15. 0.865
17. 0.999875; no logrará mucha confiabilidad, ya que la probabilidad de llegar a clases a tiempo aumenta de 0.95 (con un reloj despertador) a 0.999875 (con tres relojes despertadores).
19. 0.271
21. 0.535
23. 0
25. a. 0.431
b. 0.569
27. 1/12; 35

Sección 4-6

1. Una simulación es un proceso que se comporta de la misma forma que cierto procedimiento. Si se utiliza una simulación, la respuesta no suele ser la respuesta correcta exacta.
3. No, porque la gente generalmente favorece algunos números sobre otros, de manera que no selecciona números con un proceso que sea verdaderamente aleatorio.
5. Repita este procedimiento 20 veces: genere aleatoriamente un entero entre 1 y 100, y considere que un resultado entre 1 y 95 representa un hombre, mientras que un resultado entre 96 y 100 es una mujer.
7. Genere aleatoriamente 500 enteros entre 1 y 100, y considere que un resultado de 1 o 2 es un teléfono celular defectuoso, mientras que cualquier resultado entre 3 y 100 es un teléfono celular sin defectos.
9. a. La respuesta varía.
b. Los resultados podrían mostrar que sería infrecuente obtener un resultado consistente de 10 mujeres.
11. a. La respuesta varía.
b. Sería muy infrecuente no encontrar defectos en 500 teléfonos celulares.
13. La respuesta varía, pero probablemente los resultados muestran que es fácil obtener 12 éxitos cuando la tasa de éxitos es del 20%, de manera que no existen evidencias firmes de que el fármaco sea efectivo.
15. Si se cambia de decisión, $P(\text{ganar}) = 2/3$. Si no se cambia de decisión, $P(\text{ganar}) = 1/3$.
17. No; no

Sección 4-7

1. Con las permutaciones, distintos ordenamientos de los mismos elementos se cuentan de forma separada, pero con las combinaciones no se cuentan de manera separada.
3. No. Los métodos de esta sección no son adecuados para calcular la frecuencia relativa de una palabra en un texto típico en inglés.
5. 120
7. 10,626
9. 2652
11. 4060
13. 1/324,632

15. 1/2,760,681
17. 1/69,090,840
19. 1/3003; sí
21. 1,048,576
23. a. 120
b. 1/120
25. 1/7776 o 0.000129; sí
27. a. 11,880
b. 495
29. 14,348,907
31. 144
33. a. 63
b. $0.5^{63} \approx 1.08 \times 10^{-19}$
c. 5,738,831,575 o alrededor de 5.7 mil millones
35. 1/41,416,353
37. 2,095,681,645,538 (alrededor de 2 billones)
39. a. Calculadora: 3.0414093×10^{64} ; aproximación: 3.0363452×10^{64}
b. 615
41. 293

Capítulo 4 Conocimientos estadísticos y pensamiento crítico

1. Un valor de probabilidad es un número entre 0 y 1. Cuando se lanza una moneda legal, la probabilidad de obtener cara es de 0.5.
2. No, 0.27 es lo suficientemente alto, por lo que el azar es una explicación razonable.
3. No. El razonamiento supone que los dos resultados (vida, no vida) son igualmente probables, pero no lo son.
4. Dos sucesos son disjuntos si no pueden ocurrir simultáneamente. Dos sucesos son independientes si el resultado de uno no afecta la probabilidad del otro.

Capítulo 4 Ejercicios de repaso

1. 0.8
2. 0.32
3. 0.97
4. 0.85
5. 0.460
6. 0.638
7. 0.469
8. 0.188
9. a. 1/365
b. 31/365
c. La respuesta varía, pero probablemente sea pequeña, como 0.02.
d. Sí
10. 0.130
11. 1/4096 o 0.000244; puesto que la probabilidad de obtener 12 niñas por azar es muy pequeña, parece que el azar no es una explicación razonable, de manera que los resultados sustentan la aseveración de que el método es efectivo.
12. 0.979

13. $1/10$ o 0.1
14. 0.0777
15. a. 0.0027832
b. 0.00000775
c. 0.9944413
16. 10,000,000,000,000

Capítulo 4 Ejercicios de repaso acumulativo

1. a. 4.0
b. 4.0
c. 2.2
d. 4.7
e. Sí
f. $6/7$
g. 0.729
h. $1/262,144$; sí
2. a. 76 grados
b. 0.25
c. 0.75
d. $1/16$
e. No. Las temperaturas de ambos días podrían estar determinadas por un patrón climático que afecte varios días consecutivos, de manera que un día con una temperatura alta aumenta la probabilidad de que el siguiente día haya una temperatura elevada.

Capítulo 5 Respuestas

Sección 5-2

1.	x	$P(x)$
	1	$1/6$
	2	$1/6$
	3	$1/6$
	4	$1/6$
	5	$1/6$
	6	$1/6$

3. No. Puesto que la suma de las probabilidades es 2.1, es imposible que los resultados ocurran con las probabilidades dadas. Una distribución de probabilidad no se describe por medio de la lista de resultados con sus probabilidades correspondientes.
5. a. Continua
b. Discreta
c. Continua
d. Discreta
e. Discreta
7. Distribución de probabilidad con $\mu = 0.7$ y $\sigma = 0.7$.
9. No se trata de una distribución de probabilidad porque $\sum P(x) = 1.2 \neq 1$.
11. Distribución de probabilidad con $\mu = 5.8$ y $\sigma = 1.1$. No es infrecuente que un equipo gane cuatro juegos, ya que la probabilidad es alta (0.1818).

13. a. 0.003
b. 0.004
c. El resultado del inciso b) es relevante.
d. Sí. Debido a que la probabilidad de seleccionar aleatoriamente a cinco o menos mexicano-estadounidenses es tan baja (0.004), es poco probable que eso ocurra por el azar.
15. a. 0.206 (que es la probabilidad de 8 o menos mexicano-estadounidenses)
b. No. Puesto que la probabilidad de 8 o menos mexicano-estadounidenses es alta (0.206), ese suceso podría ocurrir fácilmente por el azar.
17. a. -26 centavos
b. -26 centavos
c. No apostar, porque el valor esperado sin apuesta es 0, que es mejor que -26 centavos.
19. a. Vive: -250 dólares (una pérdida); muere: \$99,750 (una ganancia)
b. -100 dólares
c. \$150
d. El valor negativo esperado es un precio relativamente bajo a pagar por la seguridad económica de sus herederos.
21. $\mu = 1.5$; $\sigma = 0.9$ No es infrecuente la obtención de tres niñas, ya que la probabilidad de obtener tres niñas es alta ($1/8$), lo que indica que este resultado podría ocurrir fácilmente por el azar.
23. $\mu = 4.5$; $\sigma = 2.9$. El histograma de probabilidad es plano.
25. Una distribución de frecuencias resume los resultados reales observados, mientras que una distribución de probabilidad describe cuáles podrían ser los resultados a largo plazo.
27. $\mu = 0.6$; $\sigma = 0.6$.

Sección 5-3

1. Si p representa la probabilidad de una respuesta correcta, entonces x es un conteo del número de respuestas correctas. p y x deben referirse al mismo suceso.
3. La tabla A-1 sólo incluye valores seleccionados de p , y se detiene en $n = 15$.
5. No binomial; los resultados pertenecen a más de dos categorías.
7. No binomial; los resultados pertenecen a más de dos categorías.
9. Binomial
11. Binomial
13. a. 0.128
b. IIC, ICI, CII; 0.128 para cada una
c. 0.384
15. 0.857
17. 0+
19. 0.113
21. 0.264
23. 0.234
25. 0.0006; sí
27. 0.2639; no
29. a. 0.107
b. 0.893

- c. 0.375 (o 0.376)
 - d. No, porque con una tasa del 20% la probabilidad de a lo sumo uno es alta (es mayor que 0.05).
31. 0.751
 33. 0.0874; no
 35. 0.000201; sí
 37. 0.0524
 39. 0.000535

Sección 5-4

1. Sí, porque el máximo valor común es 60.
3. 1.44 mujeres²
5. $\mu = 80.0$, $\sigma = 6.9$, mínimo = 66.1, máximo = 93.9
7. $\mu = 373.0$, $\sigma = 16.7$, mínimo = 339.6, máximo = 406.4
9. a. $\mu = 8.0$, $\sigma = 2.0$
 - b. No. Al hacer conjjeturas, el número común de respuestas correctas está entre 4 y 12, de manera que no sería infrecuente obtener 10 respuestas correctas.
11. a. $\mu = 20.0$, $\sigma = 4.0$
 - b. No, porque 25 M&M anaranjados está dentro del rango de valores comunes (12 a 28). La tasa aseverada del 20% no necesariamente parece incorrecta, ya que por lo general la tasa resultará entre 12 y 28 M&M anaranjados (en un total de 100), y el número observado de M&M anaranjados se encuentra dentro de ese rango.
13. a. $\mu = 162.5$, $\sigma = 9.0$
 - b. Sí, 295 niñas está muy fuera del rango de valores comunes (144.5 a 180.5), y al parecer el método es efectivo.
15. a. $\mu = 200.2$, $\sigma = 13.6$
 - b. No, está dentro del rango de valores comunes (173.0 a 227.4).
17. a. $\mu = 611.2$, $\sigma = 15.4$
 - b. No, parece que muchas más personas afirman haber votado, que la proporción de individuos que en realidad lo hicieron.
 - c. No, parece que muchas más personas afirman haber votado, que la proporción de individuos que en realidad lo hicieron.
19. a. $\mu = 16.4$, $\sigma = 4.0$
 - b. No, porque 19 se encuentra dentro del rango de los valores comunes (8.4 a 24.4).
 - c. No
21. a. Mínimo = 40.0, máximo = 60.0
 - b. Sí. El histograma de probabilidad tiene forma de campana.
 - c. La probabilidad es de 0.95 (porque 40 y 60 están a dos desviaciones estándar a partir de la media).

Sección 5-5

1. La variable aleatoria x es el número de ocurrencias de un suceso durante cierto intervalo, las ocurrencias son aleatorias y son independientes una de otra, y se distribuyen de manera uniforme a lo largo del intervalo.
3. Únicamente la media.
5. 0.175

- 7. 0.0126
- 9. a. 0.000912
 - b. 0.999
 - c. 0.0296
- 11. a. 62.2
 - b. 0.0155 (0.0156 si se utiliza la media redondeada)
- 13. a. 0.728
 - b. 0.231
 - c. 0.0368
 - d. 0.00389
 - e. 0.000309
- Si se utilizan las probabilidades calculadas, las frecuencias esperadas son 266, 84, 13, 1.4 y 0.1, y coinciden bastante bien con las frecuencias reales.
- 15. 4.82×10^{-64} tan pequeño que, para propósitos prácticos, podemos considerarlo cero.

Capítulo 5 Conocimientos estadísticos y pensamiento crítico

1. Una distribución de probabilidad es una gráfica, tabla o fórmula que proporciona la probabilidad para cada valor de la variable aleatoria.
2. $\sum P(x) = 1$ y $0 \leq P(x) \leq 1$ para cada valor individual de x .
3. La distribución de probabilidad incluye una variable aleatoria que tiene un número finito de valores o un número contable de valores, donde "contable" se refiere al hecho de que podría haber una cantidad infinita de valores, pero que se puede asociar con un proceso de conteo. Otro tipo de distribución de probabilidad es continua.
4. No. Hay muchas distribuciones de probabilidad discreta que no satisfacen los requisitos de una distribución binomial o de una distribución de Poisson.

Capítulo 5 Ejercicios de repaso

1. a. $\sum P(x) = 0.999 = 1$ (con un pequeño error por redondeo), y cada valor de x está entre 0 y 1.
 - b. 2.0
 - c. 1.3
 - d. 0.033
 - e. 2.0
 - f. 0.892 (o 0.893)
 - g. No, porque es fácil (con probabilidad 0.892) obtener al menos una correcta al tratar de adivinar la respuesta a cada pregunta.
2. a. 1.8
 - b. 1.8
 - c. 1.2
 - d. 0.172
 - e. Debido a que la probabilidad de que no haya televisores (de un total de 12) sintonizando *Cold Case* es de 0.142, no es infrecuente.
3. a. 0.00361
 - b. Al parecer esta empresa es muy diferente, debido a que el suceso de al menos cuatro despidos es muy poco probable, con una probabilidad de 0.00361.

4. a. 7/365
 b. 0.981
 c. 0.0188
 d. 0.0002
 e. No, porque este suceso es muy infrecuente.

Capítulo 5 Ejercicios de repaso acumulativo

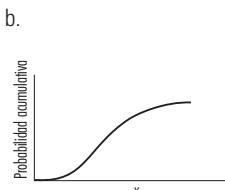
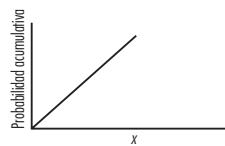
1. a. $\bar{x} = 4.4, s = 3.0$
 b. Reemplace los conteos de frecuencias con las siguientes frecuencias relativas: 9%, 18%, 6%, 14%, 10%, 5%, 6%, 8%, 15%, 10%.
 c. $\mu = 4.5, \sigma = 2.9$
 d. La frecuencia esperada para cada celda es 8. Al comparar las frecuencias observadas en la tabla con las frecuencias esperadas (todas de 8), observamos que hay algunas diferencias evidentes, pero el desacuerdo general no es extremo.
2. a. $\mu = 15.0, \sigma = 3.7$
 b. 12 casos positivos caen dentro del rango de valores comunes (7.6 a 22.4), de manera que 12 no es infrecuentemente bajo. Debido a que 12 casos positivos podrían ocurrir con facilidad con un programa ineficaz, no tenemos suficiente justificación para afirmar que programa sea efectivo.
3. a. Sí, porque satisface los requisitos de que la suma de las probabilidades debe ser 1, y que cada probabilidad individual es un valor entre 0 y 1.
 b. La tabla corresponde a la población que consiste únicamente en los 25 sujetos encuestados. En definitiva no se trata de la población de todos los tarjetahabientes de Estados Unidos.
 c. Sí, él utiliza una muestra de conveniencia que sólo describe a los 25 amigos que encuestó. Si hubiera utilizado una muestra aleatoria, los resultados serían mucho más representativos de una población más grande.
 d. 1.7
 e. 1.0

Capítulo 6 Respuestas

Sección 6-2

1. El término “normal” tiene un significado especial en la estadística. Se refiere a una distribución específica en forma de campana que puede describirse por medio de la fórmula 5-1.
3. La media y la desviación estándar tienen los valores de $\mu = 0$ y $\sigma = 1$.
5. 0.75
 7. 0.5
 9. 0.1587
 11. 0.8413
 13. 0.1056
 15. 0.9599
 17. 0.1359
 19. 0.0157 (con tecnología: 0.0156)
 21. 0.9220

23. 0.8412 (con tecnología: 0.8413)
 25. 0.0001 (con tecnología: 0.0002)
 27. 0.5000
 29. 0.6826 (con tecnología: 0.6827)
 31. 0.9974 (con tecnología: 0.9973)
 33. 0.9500
 35. 0.0100
 37. 1.28
 39. -1.96, 1.96
 41. a. 68.26% (con tecnología: 68.27%)
 b. 95.00%
 c. 91.74% (con tecnología: 99.73%)
 d. 81.85% (con tecnología: 81.86%)
 e. 4.56% (con tecnología: 4.55%)
 43. a. 1.23
 b. 1.50
 c. 1.52
 d. -2.42
 e. -0.13
 45. a.



Sección 6-3

1. La distribución normal estándar tiene una media de 0 y una desviación estándar de 1, pero una distribución no estándar tiene un valor diferente para uno o para los dos parámetros.
3. No, porque los dígitos generados no tienen una distribución normal. La probabilidad de un dígito menor que 5 es 0.5.
5. 0.9772
 7. 0.4972 (con tecnología: 0.4950)
 9. 80.8
 11. 105.9 (con tecnología: 105.8)
 13. 0.52%
 15. a. Hombres: 0.01% (con tecnología: 0.004%); mujeres: 0.01% (con tecnología: 0.00%)
 b. 73.6 pulgadas
 17. 2405 g (con tecnología: 2403 g)
 19. 0.2005 (con tecnología: 0.2015)

21. a. 0.0038; o ha ocurrido un suceso muy infrecuente o el esposo no es el padre.
b. 242 días
23. 4.0 pulgadas; 8.0 pulgadas
25. a. $\bar{x} = 118.9$, $s = 10.5$, el histograma es aproximadamente normal.
b. 101.6, 136.2
27. a. Las puntuaciones z son números reales que no tienen unidades de medida.
b. $\mu = 0$; $\sigma = 1$; la distribución es normal.
29. a. 75; 5
b. No, la conversión también debe explicar la variación.
c. 31.4, 27.6, 22.4, 18.6
d. El inciso c, porque la variación está incluida en la conversión.

Sección 6-4

1. Una distribución muestral de un estadístico es la distribución de todos los valores de ese estadístico, cuando se toman de la misma población todas las muestras posibles del mismo tamaño.
3. Media, proporción, varianza.
5. No. Debido a la variabilidad muestral, las proporciones muestrales variarán de manera natural de la proporción poblacional verdadera, incluso si el muestreo se realiza con un procedimiento perfectamente válido.
7. a. 10-10; 10-6; 10-5; 6-10; 6-6; 6-5; 5-10; 5-6; 5-5; las medias se listan en el inciso b).
b. Las nueve muestras tienen medias de 10.0, 8.0, 7.5, 8.0, 6.0, 5.5, 7.5, 5.5, 5.0, y cada una tiene una probabilidad de 1/9. (La distribución muestral se puede describir de forma condensada listando las diferentes medias de 10.0, 8.0, 7.5, 6.0, 5.5, 5.0 con sus correspondientes probabilidades de 1/9, 2/9, 2/9, 1/9, 2/9, 1/9.).
c. 7.0
d. Sí; sí
9. a. Las medias de 20.0, 20.5, 21.0, 31.5, 32.0, 33.5, 34.0, 43.0, 45.0, 47.0 tienen las probabilidades correspondientes de 4/25, 4/25, 1/25, 4/25, 2/25, 4/25, 2/25, 1/25, 2/25, 1/25.
b. 30.2 (billones de dólares).
c. Sí; sí
11. a. Las proporciones de 0, 0.5, 1 tienen las probabilidades correspondientes de 1/16, 6/16, 9/16.
b. 0.75
c. Sí; sí
13. a. Las medias de 3, 6, 9, 13, 16, 23 tienen las probabilidades correspondientes de 1/9, 2/9, 1/9, 2/9, 2/9, 1/9.
b. No
15. La fórmula produce $P(0) = 0.25$, $P(0.5) = 0.5$, y $P(1) = 0.25$, que describe la distribución muestral de las proporciones. La fórmula es sólo una forma diferente de presentar la misma información de la tabla, que describe la distribución muestral.

Sección 6-5

1. Es la desviación estándar de las medias muestrales, que se denota por $\sigma_{\bar{x}}$ o σ / \sqrt{n} .

3. $\mu_{\bar{x}}$, $\sigma_{\bar{x}}$
5. a. 0.5636
b. 0.8315
7. a. 0.1566 (con tecnología: 0.1565)
b. 0.4077
c. Si la población original tiene una distribución normal, el teorema del límite central proporciona buenos resultados para cualquier tamaño muestral.
9. a. 0.5675 (con tecnología: 0.5684)
b. 0.7257 (con tecnología: 0.7248)
c. Probablemente el teleférico esté diseñado para transportar con seguridad una carga un poco mayor que 2004 libras, pero los operadores actuarían bien al evitar una carga de 12 hombres, especialmente si parecen ser pesados.
11. a. 0.0001 (con tecnología: 0.0000 con redondeo)
b. No, pero los consumidores no están siendo engañados porque las latas tienen más líquido y no menos.
13. a. 0.0062
b. 0.0001 (con tecnología: 0.0000 con redondeo)
c. El inciso a). El comportamiento de las luces estroboscópicas individuales es más importante que el comportamiento de los lotes de 60 luces estroboscópicas.
15. a. 0.0274 (con tecnología: 0.072)
b. 0.0001
c. Debido a que la población original tiene una distribución normal, la distribución muestral de la media de muestras se distribuirá normalmente para cualquier tamaño muestral.
d. No, la media puede ser menor que 140 aunque los valores individuales estén por arriba de 140.
17. a. 0.5302 (con tecnología: 0.5317)
b. 0.7323 (con tecnología: 0.7326)
c. El inciso a), porque los asientos estarán ocupados por mujeres individuales y no por grupos de mujeres.
19. a. 15
b. 0.9998 (con tecnología: 1)
c. El inciso a), porque las monedas individuales rechazadas podrían provocar pérdidas en las ventas y menores ganancias.
21. a. 267.5 pulgadas.
b. Debido a que los jugadores de fútbol colegial tienden a ser más grandes que hombres elegidos al azar, la media y la desviación estándar dadas no aplican. La banca debe ser más larga que la longitud de 267.5 pulgadas que se calculó en el inciso a).
23. a. $\mu = 8.0$, $\sigma = 5.4$
b. 2, 3, 2, 6, 2, 8, 2, 11, 2, 18, 3, 6, 3, 8, 3, 11, 3, 18, 6, 8, 6, 11, 6, 18, 8, 11, 8, 18, 11, 18
c. 2.5, 4.0, 5.0, 6.5, 10.0, 4.5, 5.5, 7.0, 10.5, 7.0, 8.5, 12.0, 9.5, 13.0, 14.5
d. $\mu_{\bar{x}} = 8.0$, $\sigma_{\bar{x}} = 3.4$

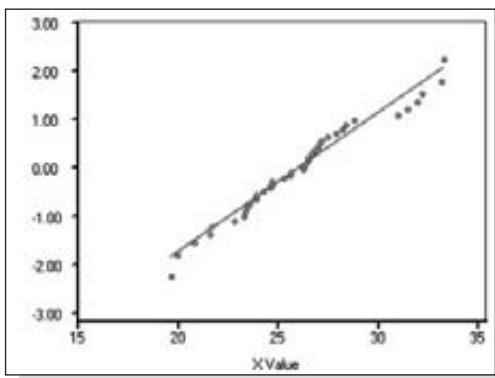
Sección 6-6

1. Aproximadamente con forma de campana (normal).
3. No. Con $n = 6$ y $p = 0.001$, los requisitos de $np \geq 5$ y $nq \geq 5$ no se satisfacen.

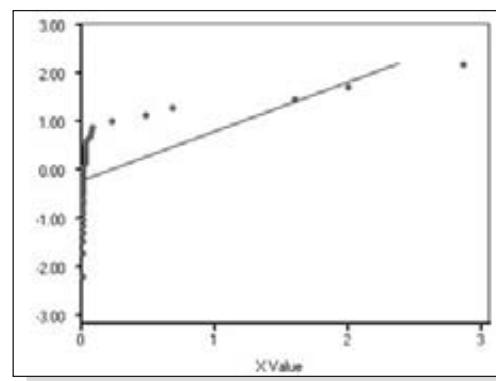
5. El área a la derecha de 15.5
7. El área de la izquierda de 11.5
9. El área a la izquierda de 4.5
11. El área entre 7.5 y 10.5
13. 0.227; la aproximación normal no es adecuada.
15. 0.887; aproximación normal: 0.8869 (con tecnología: 0.8861)
17. 0.1292 (con tecnología: 0.1303). No es infrecuente.
19. 0.0001 (con tecnología: 0.0000 con redondeo). Los resultados sugieren que la gente encuestada no respondió con exactitud.
21. 0.2676 (con tecnología: 0.2665); no
23. 0.2709 (con tecnología: 0.2697); parece que los reportes de los medios son incorrectos.
25. 0.0080 (con tecnología: 0.0097); sí
27. 0.6368 (con tecnología: 0.6375); es probable que el grupo sea suficiente, pero la probabilidad debe ser mucho más alta. Sería mejor incrementar el grupo de voluntarios.
29. 0.0001 (con tecnología: 0.0000); sí
31. La probabilidad de obtener al menos 18 totales mayores de \$100 es de 0.0051 (con tecnología: 0.0050), que es muy bajo. Esto sugiere que el gasto es poco común y que debe verificarse.
33. 6; 0.4602
35. a. 0.821
b. 0.9993
c. 0.0000165
d. 0.552
37. La respuesta varía según la tecnología, pero generalmente incluye un número sumamente grande de ensayos n .

Sección 6-7

1. Una gráfica cuantílica normal se puede usar para determinar si los datos muestrales provienen de una población con una distribución normal.
3. Debe haber 100 puntos configurados para que se acerquen razonablemente al patrón de una línea recta, y no hay un patrón sistemático que no sea el patrón de una línea recta.
5. No es normal. Existe un patrón sistemático que no es el de una línea recta.
7. Normal
9. Normal
11. No es normal
13. Normal

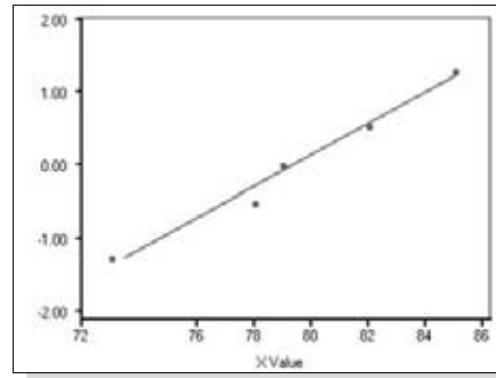


15. No es normal



17. Las estaturas parecen ser normales, pero no así los niveles de colesterol. Los niveles de colesterol se ven muy afectados por la dieta, y las dietas pueden variar de manera tan drástica que no producen resultados distribuidos normalmente.

19. -1.28, -0.52, 0, 0.52, 1.28; normal



21. No, la transformación a puntuaciones z implica la resta de una constante y la división entre una constante, de manera que la gráfica de los puntos (x, z) siempre producirá una línea recta, sin importar la naturaleza de la distribución.

Capítulo 6 Conocimientos estadísticos y pensamiento crítico

1. Una distribución normal es la distribución de probabilidad de una variable aleatoria continua, con las propiedades de que una gráfica de la distribución es simétrica y con forma de campana, y puede describirse por medio de la fórmula 6-1.
2. Las medias muestrales tenderán a distribuirse normalmente.
3. No, la muestra podría estar sesgada. Tal vez tendría la forma de distribución correcta, pero la media y la desviación estándar podrían ser muy diferentes de los parámetros poblacionales.
4. Cuando se obtienen muestras aleatorias simples de tamaño n de una población con media μ y desviación estándar σ , conforme el tamaño de la muestra aumenta, la distribución de las medias muestrales se aproxima a una distribución normal, la media de todas las medias muestrales es μ , y la desviación estándar de todas las medias es σ/\sqrt{n} .

Capítulo 6 Ejercicios de repaso

1. a. 0.3830 (con tecnología: 0.3829)
b. 0.9544 (con tecnología: 0.9545)
c. 1.28
2. a. 3.67% (con tecnología: 3.71%)
b. 0.0001 (con tecnología: 0.000178)
c. 72.6 pulgadas
3. a. 0.0222 (con tecnología: 0.0221)
b. 0.2847 (con tecnología: 0.2836)
c. 0.6720 (con tecnología: 0.6715)
d. 254.6
4. 0.0004. La probabilidad tan baja sugiere que debe descartarse el azar como una explicación para el hecho de que 18 de las respuestas correctas fueron de mujeres. Al parecer existe una fuerte evidencia para el argumento de que la pregunta está sesgada a favor de los hombres.
5. 0.1112 (con tecnología: 0.1113) o 0.1071 si se utiliza la aproximación binomial. La probabilidad no es muy baja (como de 0.05 o menos), por lo que no parece haber evidencia suficiente para aseverar que la empresa está discriminando con base en el género.
6. Un histograma muestra una gráfica que difiere mucho de la normalidad, de manera que las duraciones no parecen provenir de una población con una distribución normal.

Capítulo 6 Ejercicios de repaso acumulativo

1. a. 10.5 g
b. 8.5 g
c. 10.2 g
d. 103.7 g^2
e. -0.74
f. 62.5%
g. 0.7704 (con tecnología: 0.7689)
h. De razón
i. Continuos
j. No. La muestra no es la dieta típica estadounidense, y los estadounidenses no consumen los artículos muestrales en las mismas cantidades.
2. a. 0.001
b. 0.271
c. El requisito de que $np \geq 5$ no se satisface, lo que indica que la aproximación normal produciría errores demasiado grandes.
d. 5.0
e. 2.1
f. No, el 8 está a dos desviaciones estándar de la media y está dentro del rango de valores que podría ocurrir fácilmente por el azar.

Capítulo 7 Respuestas

Sección 7-2

1. Es una puntuación z estándar que se puede utilizar para distinguir entre estadísticos muestrales que tienen probabilidades de ocurrir y aquellos que son poco probables. El número $z_{\alpha/2}$ es una puntuación z que separa un área de $\alpha/2$ en la cola derecha de la distribución normal estándar.

3. El estimado puntual no revela ninguna información acerca de la precisión del estimado. Al proporcionar un rango de valores asociados con una probabilidad, un intervalo de confianza revela información sobre la precisión.
5. 2.575 (con tecnología: 2.5758293)
7. 2.33 (con tecnología: 2.3263479)
9. 0.333 ± 0.111
11. 0.246 ± 0.040
13. 0.879; 0.011
15. 0.660; 0.053
17. 0.0429
19. 0.0309 (con tecnología: 0.0308)
21. $0.357 < p < 0.443$
23. $0.219 < p < 0.281$
25. 2401
27. 842
29. $0.866 < p < 0.949$; el método parece ser efectivo porque la proporción de niñas es mucho mayor que 0.5.
31. $0.496 < p < 0.514$; no, porque la proporción fácilmente podría ser igual a 0.5. La proporción no es mucho menor que 0.5 la semana anterior al Día de acción de gracias.
33. a. $0.226 < p < 0.298$
b. No, el intervalo de confianza incluye a 0.25, de manera que el porcentaje verdadero fácilmente podría ser igual al 25%.
35. a. $0.0267\% < p < 0.0376\%$
b. No, porque 0.0340% está incluido en el intervalo de confianza.
37. $0.347 < p < 0.432$ (utilizando $x = 339$); sí, porque la proporción seleccionada parece ser mucho menor que 0.791 (o 79.1%).
39. $0.931 < p < 0.949$ (utilizando $x = 4019$); sí, con base en el intervalo de confianza, podemos tener confianza en que la proporción es mayor que 0.35 o 35%.
41. 4145 (con tecnología: 4147)
43. 1509 (con tecnología: 1504)
45. $0.183 < p < 0.357$; sí, porque el intervalo de confianza incluye a 0.24.
47. Miércoles: $0.178 < p < 0.425$; domingo: $0.165 < p < 0.412$. Los intervalos de confianza no son muy diferentes. Al parecer no hay mayor precipitación en alguno de los días.
49. 406
51. $0.0395 < p < 0.710$; no
53. a. El requisito de al menos 5 éxitos y de al menos 5 fracasos no se satisface, por lo que no se puede utilizar la distribución normal.
b. 0.15

Sección 7-3

1. Existe una confianza del 95% de que los límites de 2.5 y 6.0 contienen la media poblacional verdadera μ .
3. No, porque la muestra no es adecuada. Es muy posible que la muestra de conveniencia no sea representativa de la población de todos los visitantes potenciales.

5. 1.96
 7. 1.75
 9. 2.94
 11. 12.875 (con tecnología: 12.879)
 13. $\$61,605 < \mu < \$72,795$ (con tecnología:
 $\$61,606 < \mu < \$72,794$)
 15. $667 < \mu < 709$
 17. 97
 19. 390
 21. 67.3849
 23. 67.3849 ± 0.7295
 25. 4.3 años < μ < 5.3 años; no, porque el intervalo de confianza es un estimado de la media, y esto no implica que 95% de las veces se encuentren dentro de los límites del intervalo de confianza.
 27. 55.4 s < μ < 61.2 s; sí, porque los límites del intervalo de confianza contienen a 60 s.
 29. 128.7 < μ < 139.2; de manera ideal, todas las mediciones serían iguales, por lo que no habría un estimado del intervalo.
 31. 5.61161 g < μ < 5.66698 g; el intervalo de confianza sugiere que las monedas acuñadas cubren las especificaciones y que el proceso de producción es adecuado.
 33. 217
 35. 6907
 37. 80,770 (con tecnología: 80,767); el tamaño de la muestra es demasiado grande para ser práctico, por lo que debe reducirse al incrementar un margen de error aceptable.
 39. $105 < \mu < 115$
- valor extremo, lo que sugiere que el supuesto de una población distribuida normalmente no es correcto.
23. $589.7 < \mu < 731.0$; no, habrá una gran cantidad de solicitantes que no califican.
25. $0.075 < \mu < 0.168$; no, tal vez el requisito se cumpla, pero también es muy posible que la media no sea menor que 0.165 g/mi.
27. a. $65.8 < \mu < 73.0$
 b. $72.3 < \mu < 80.3$
 c. Debido a que los dos intervalos de confianza se traslanan, no podemos concluir que las dos medias poblacionales sean diferentes.
29. $26.2 < \mu < 96.3$; el intervalo de confianza cambia en una cantidad importante. Los intervalos de confianza son muy sensibles a los valores extremos. Para determinar si se trata de errores, los valores extremos se deben examinar y descartar de manera cuidadosa. Si un valor extremo es un valor correcto, sería muy útil construir el intervalo de confianza incluyendo y sin incluir el valor extremo, para poder ver los efectos.
31. $0.8462 \text{ g} < \mu < 0.8668 \text{ g}$; $0.8474 < \mu < 0.8656$; el segundo intervalo de confianza es más estrecho, lo que indica que tenemos un estimado más preciso cuando la muestra relativamente grande proviene de una población finita relativamente pequeña.
33. a. No, un solo valor no tiene variación.
 b. El valor crítico t no se puede obtener porque $gl = 0$ no están disponibles en la tabla de valores críticos de t .
 c. $-104.16 \text{ pies} < \mu < 128.16 \text{ pies}$; no

Sección 7-4

1. La cantidad se refiere a un promedio, que probablemente sea la media o la mediana, pero el margen de error es apropiado para una proporción y no para una media o una mediana. El margen de error debería ser una cantidad en dólares y no en puntos porcentuales.
 3. Tenemos una confianza del 99% de que los límites de 114.4 y 123.4 contienen la media verdadera de la presión sanguínea sistólica.
 5. $t_{\alpha/2} = 2.201$
 7. No se aplica ni la distribución normal ni la distribución t .
 9. $t_{\alpha/2} = 1.653$
 11. $z_{\alpha/2} = 2.33$
 13. 1.6 kg; 1.4 kg < μ < 4.6 kg
 15. $113.583 < \mu < 122.417$; tenemos una confianza del 95% de que los límites de 113.583 y 122.417 contienen la media poblacional verdadera de la puntuación del CI de los estudiantes de estadística.
 17. $3002 \text{ g} < \mu < 3204 \text{ g}$; el peso medio de los bebés nacidos de madres que consumieron cocaína es mucho menor que el peso medio de los bebés nacidos de madres que no utilizaron cocaína. Parece que la cocaína afecta el peso al nacer.
 19. a. $-3.5^\circ < \mu < 0.9^\circ$
 b. Sí, incluye el 0. No, porque el intervalo de confianza indica que la diferencia media podría ser de 0.
 21. $-0.471 < \mu < 3.547$; es probable que el intervalo de confianza no sea un buen estimado, ya que la puntuación de 5.40 parece ser un

Sección 7-5

1. Tenemos una confianza del 95% de que los límites de 2.25 pulgadas y 3.52 pulgadas contienen el valor verdadero de la desviación estándar de las estaturas de todas las mujeres.
 3. No, el intervalo de confianza sería un estimado de la variación entre las medias muestrales y no de la variación entre individuos. Es probable que los individuos tengan una variación mucho mayor.
 5. 13.844, 41.923
 7. 20.707, 66.766
 9. $\$15,006 < \sigma < \$23,385$
 11. $55 < \sigma < 86$ (con tecnología: $56 < \sigma < 87$)
 13. 767
 15. 1401; el tamaño muestral es común para encuestas y sondeos, pero en muchas situaciones el tamaño muestral sería impráctico porque es demasiado grande.
 17. $586 \text{ g} < \sigma < 717 \text{ g}$; no, porque los límites del intervalo de confianza contienen a 696 g.
 19. $0.54^\circ \text{F} < \sigma < 0.77^\circ \text{F}$ (con tecnología: $0.53^\circ \text{F} < \sigma < 0.75^\circ \text{F}$); sí
 21. $64.9 < \sigma < 173.2$
 23. $1.195 < \sigma < 4.695$; sí, es probable que el intervalo de confianza no sea un buen estimado porque el valor de 5.40 parece ser un valor extremo, lo que sugiere que el supuesto de una población distribuida normalmente no es correcto.

25. a. $2.62 < \sigma < 4.71$ (con tecnología: $2.65 < \sigma < 4.79$)
 b. $4.71 < \sigma < 8.46$ (con tecnología: $4.76 < \sigma < 8.61$)
 c. Los dos intervalos de confianza no se traslapan, de manera que parece que las dos poblaciones tienen diferentes cantidades de variación. (Con tecnología: los intervalos de confianza se traslan ligeramente, de manera que, con un nivel de confianza del 99%, parece que las dos poblaciones tienen cantidades de variación que no difieren mucho).
 27. 152.3644 y 228.4771 se acercan a los valores del STATDISK.

Capítulo 7 Conocimientos estadísticos y pensamiento crítico

1. Los dos valores críticos de $z = -1.96$ y $z = 1.96$ separan cada uno áreas de 0.025 de la cola, de manera que el área entre $z = -1.96$ y $z = 1.96$ es 0.95 o 95%.
2. El intervalo de confianza es un indicador mucho mejor de la precisión del estimado de 27.44 libras. Nos indica qué tan bueno es el estimado.
3. $0.62 < p < 0.68$
4. Tenemos una confianza del 95% de que los límites de 0.62 y 0.68 contienen el valor verdadero de la proporción poblacional.

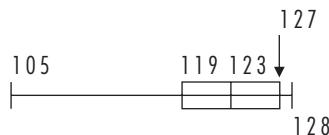
Capítulo 7 Ejercicios de repaso

1. $91.5\% < p < 94.6\%$ (utilizando $x = 934$); tenemos una confianza del 95% de que los límites de 91.5% y 94.6% contienen el valor verdadero del porcentaje poblacional.
2. 157 (utilizando $\hat{p} = 0.93$ y $\hat{q} = 0.07$)
3. $10.06 \text{ onzas} < \mu < 11.72 \text{ onzas}$; el intervalo de confianza no contiene a 12, y la media muestral es 10.89, lo que indica que la media es menor que 12. Los consumidores están recibiendo suficiente café. Asimismo, la variación es demasiado alta, y algunas cantidades son demasiado bajas.
4. $1.15 \text{ onzas} < \sigma < 2.41 \text{ onzas}$; no, 0.25 onzas no es un valor posible de σ . Definitivamente la máquina requiere modificaciones para reducir la variación.
5. a. 2401
 b. 1164
 c. 2401
6. a. $5.47 \text{ años} < \mu < 8.55 \text{ años}$
 b. $2.92 \text{ años} < \sigma < 5.20 \text{ años}$
 c. 1484
 d. No, la muestra no sería representativa de la población de todos los propietarios de automóviles.
7. a. $15.1\% < p < 21.6\%$ (utilizando $x = 144$)
 b. Sí, parece que la tasa de tabaquismo de los graduados universitarios es significativamente menor que la tasa de la población general.
8. a. $\$5403 < \mu < \$12,605$
 b. \$12,605

Capítulo 7 Ejercicios de repaso acumulativo

1. a. 121.0 libras
 b. 123.0 libras
 c. 119 libras, 128 libras

- d. 116.5 libras
 e. 23.0 libras
 f. 58.8 libras²
 g. 7.5 libras
 h. 119.0 libras
 i. 123.0 libras
 j. 127.0 libras
 k. De razón
 l.



- m. $112.6 \text{ libras} < \mu < 129.4 \text{ libras}$
 n. $4.5 \text{ libras} < \sigma < 18.4 \text{ libras}$
 o. 95
 p. Los pesos individuales de las supermodelos no parecen diferir mucho de los pesos de mujeres seleccionadas al azar, ya que todos están dentro de 1.31 desviaciones estándar a partir de la media de 143 libras. Sin embargo, cuando se les considera un grupo, su media es significativamente menor que la media de 143 libras [vea el inciso (m)].
2. a. 0.0089
 b. $0.260 < p < 0.390$
 c. Debido a que los límites del intervalo de confianza no contienen a 0.25, es poco probable que el experto esté en lo correcto.
3. a. 9.00%
 b. $7.40\% < p < 10.6\%$
 c. 2653 (con tecnología: 2654).

Capítulo 8 Respuestas

Sección 8-2

1. La prueba de hipótesis no puede utilizarse con datos de encuesta para "demostrar" que un porcentaje poblacional es igual a algún valor específico. La aseveración correcta sería que no existe evidencia suficiente para concluir que la tasa del 50% es incorrecta o que, con base en la encuesta, el porcentaje parece ser de alrededor del 50%.
3. Un estadístico de prueba se basa en los datos muestrales, pero el valor crítico se basa en la distribución que se está utilizando y en el nivel de significancia que se eligió.
5. No existe suficiente evidencia para sustentar la aseveración (porque es muy fácil obtener por el azar 11 caras en 20 lanzamientos).
7. Al parecer existe evidencia suficiente para sustentar la aseveración.
9. $H_0: p = 0.25$. $H_1: p > 0.25$.
11. $H_0: \mu = 121 \text{ libras}$. $H_1: \mu \neq 121 \text{ libras}$.
13. $H_0: \sigma = 15$. $H_1: \sigma < 15$.
15. $H_0: \mu = 0.8535 \text{ g}$. $H_1: \mu < 0.8535 \text{ g}$.
17. $z = \pm 1.96$

19. $z = 2.33$
 21. $z = \pm 1.645$
 23. $z = -2.05$
 25. -2.21
 27. 7.77
 29. 0.1587
 31. 0.0500
 33. 0.0668
 35. 0.0802 (con tecnología: 0.0801)
 37. Existe suficiente evidencia para sustentar la aseveración de que la proporción de hombres golfistas es menor que 0.5.
 39. No existe suficiente evidencia para sustentar la aseveración de que la proporción de M&M rojos es diferente de 0.13.
 41. Error tipo I: rechazar la aseveración de que la proporción de demandas por negligencia médica resueltas es de 0.25, cuando la proporción en realidad es de 0.25.
 Error tipo II: no rechazar la aseveración de que la proporción de demandas por negligencia médica resueltas es de 0.25, cuando la proporción en realidad es diferente de 0.25.
 43. Error tipo I: rechazar la aseveración de que la proporción de asesinatos aclarados por un arresto es de 0.62, cuando la proporción real es de 0.62.
 Error tipo II: no rechazar la aseveración de que la proporción de asesinatos aclarados por un arresto es de 0.62, cuando la proporción es en realidad diferente de 0.62.
 45. Valor $P = 0.9999$. Con una hipótesis alternativa de $p > 0.5$, es imposible que un estadístico muestral de 0.27 se localice en la región crítica. Ninguna proporción muestral menor que 0.5 puede sustentar nunca la aseveración de que $p > 0.5$.
 47. a. 0.7852 (con tecnología: 0.7857)
 b. 0.2148 (con tecnología: 0.2143)

Sección 8-3

1. Normal, porque la distribución normal es una aproximación adecuada para una distribución binomial, siempre y cuando los requisitos se cumplan.
 3. No. Debido a que se trata de una muestra de respuesta voluntaria, no es adecuada para hacer inferencias acerca de la población general.
 5. a. $z = -0.12$
 b. $z = \pm 1.96$
 c. 0.9044 (con tecnología: 0.9025)
 d. No existe evidencia suficiente para sustentar al rechazo de la aseveración de que los chícharos con flores verdes ocurren con una tasa del 25%.
 e. No, una prueba de hipótesis no puede demostrar que una proporción es igual a cierto valor aseverado.
 7. $H_0: p = 0.29$. $H_1: p > 0.29$. Estadístico de prueba: $z = 0.73$. Valor crítico: $z = 1.645$ (suponiendo un nivel de significancia de 0.05). Valor $P = 0.2313$. No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que más del 29% de los crímenes federales se debieron a crímenes por drogas.
 9. $H_0: p = 0.5$. $H_1: p > 0.5$. Estadístico de prueba: $z = 14.70$. Valor crítico: $z = 2.33$. Valor $P = 0.0001$ (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la probabilidad de que un bebé sea niña es mayor que 0.5. Parece que el método es efectivo.

11. $H_0: p = 0.5$. $H_1: p > 0.5$. Estadístico de prueba: $z = 4.20$. Valor crítico: $z = 2.33$. Valor $P = 0.0001$ (con tecnología: 0.0000). Se rechaza H_0 . Existe suficiente evidencia para sustentar la aseveración de que más del 50% de los choques de automóviles ocurren a 5 millas del hogar. Los resultados podrían ser cuestionables porque tal vez el patrocinador tenga intereses personales en los resultados.
 13. $H_0: p = 0.15$. $H_1: p > 0.15$. Estadístico de prueba: $z = 1.60$. Valor crítico: $z = 1.645$. Valor $P = 0.0548$ (con tecnología: 0.0543). No se rechaza H_0 . No existe suficiente evidencia para sustentar la aseveración de que más del 15% de los hogares estadounidenses utilizan el correo electrónico. La conclusión no es válida en la actualidad porque las características de la población (uso del correo electrónico) cambian con rapidez al paso del tiempo.
 15. $H_0: p = 0.000340$. $H_1: p \neq 0.000340$. Estadístico de prueba: $z = -0.66$. Valores críticos: $z = \pm 2.81$. Valor $P = 0.5092$ (con tecnología: 0.5122). No se rechaza H_0 . No existe suficiente evidencia para sustentar la aseveración de que la tasa difiere del 0.0340%. Los usuarios de teléfonos celulares no deben preocuparse por el cáncer cerebral o del sistema nervioso.
 17. $H_0: p = 0.01$. $H_1: p \neq 0.01$. Estadístico de prueba: $z = 2.19$. Valores críticos: $z = \pm 1.96$. Valor $P = 0.0286$ (con tecnología: 0.0284). Se rechaza $H_0: p = 0.01$. Existe suficiente evidencia para sustentar el rechazo de la aseveración de que el 1% de las ventas tienen sobreprecios. Debido a que el 1.62% de los artículos muestreados tienen sobreprecio, parece que la tasa de error es peor, y no mejor, con los escáneres.
 19. $H_0: p = 1/3$. $H_1: p < 1/3$. Estadístico de prueba: $z = -0.91$. Valor crítico: $z = -1.645$. Valor $P = 0.1814$ (con tecnología: 0.1826). No se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que menos de 1/3 de todos los adultos nunca bebe. Parece que la redacción de la pregunta produce respuestas honestas, aunque la naturaleza del tema podría provocar que las personas no sean muy honestas.
 21. $H_0: p = 0.5$. $H_1: p \neq 0.5$. Estadístico de prueba: $z = -2.00$. Valores críticos: $z = \pm 1.96$. Valor $P = 0.0456$ (con tecnología: 0.0458). Se rechaza H_0 . Existe evidencia suficiente para sustentar el rechazo de la aseveración de que el porcentaje de adultos que nunca o casi nunca vuela es igual al 50%. Sí, hasta cierto punto hubiera sido mejor obtener datos muestrales sobre hábitos de vuelo observados, en lugar de confiar en la honestidad de los sujetos. Es posible que se hayan sentido avergonzados al decir que nunca o casi nunca vuelan.
 23. $H_0: p = 0.012$. $H_1: p > 0.012$. Estadístico de prueba: $z = 3.42$ (utilizando $x = 35$). Valor crítico: $z = 2.33$. Valor $P = 0.0003$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que más del 1.2% de los usuarios de Clarinex experimentan fatiga. Parece que la fatiga es una reacción adversa del Clarinex.
 25. $H_0: p = 0.24$. $H_1: p \neq 0.24$. Estadístico de prueba: $z = 0.70$. Valores críticos: $z = \pm 1.96$ (suponiendo un nivel de significancia de 0.05). Valor $P = 0.4840$ (con tecnología: 0.4824). No se rechaza H_0 . No existe evidencia suficiente para sustentar el rechazo de la aseveración de que el 24% de los dulces M&M son azules.
 27. $H_0: p = 0.5$. $H_1: p < 0.5$. Estadístico de prueba: $z = -2.87$. Valor crítico: $z = -1.645$. Valor $P = 0.0021$ (con tecnología: 0.0020). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que $p < 0.5$. El resultado sugiere que la temperatura máxima pronosticada difiere en 2° menos del 50% de las veces.

29. a. $H_0: p = 0.10$. $H_1: p \neq 0.10$. Estadístico de prueba: $z = 2.00$. Valores críticos: $z = \pm 1.96$. Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que la proporción de ceros es de 0.1.
- b. $H_0: p = 0.10$. $H_1: p \neq 0.10$. Estadístico de prueba: $z = 2.00$. Valor P : 0.0456 (con tecnología: 0.0452). Existe evidencia suficiente para sustentar el rechazo de la aseveración de que la proporción de ceros es de 0.1.
- c. $0.0989 < p < 0.139$; puesto que 0.1 está incluido dentro del intervalo de confianza, no se rechaza H_0 : $p = 0.10$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que la proporción de ceros es de 0.1.
- d. Tanto el método tradicional como el método del valor P conducen al rechazo de la aseveración, pero el método del intervalo de confianza no lleva al rechazo.
31. a. Con $n = 80$ y $p = 0.0025$, las condiciones $np \geq 5$ y $nq \geq 5$ no se satisfacen.
- b. 0.0000000165 se calcula utilizando la distribución de probabilidad binomial.
- c. Si la probabilidad de que un hombre tenga ceguera al color es de 0.0025, entonces la probabilidad de obtener al menos 7 hombres con ceguera al color, de un total de 80, es sumamente baja, de manera que parece que la tasa de ceguera al color es en realidad mayor que 0.0025 o 0.25%. Existe evidencia suficiente para justificar la aseveración de que la tasa de la ceguera al color en los hombres es mayor que el 0.25%.
33. a. 0.7224 (con tecnología: 0.7219).
- b. 0.2776 (con tecnología: 0.2781).
- c. La potencia de 0.7224 indica que existe una probabilidad razonable de tomar la decisión correcta al rechazar la hipótesis nula falsa. Sería mejor si la potencia fuese más alta, como mayor que 0.8 o 0.9.
15. $H_0: \mu = 0$. $H_1: \mu < 0$. Estadístico de prueba: $z = -2.77$. Valor crítico: $z = -1.645$. Valor P : 0.0028. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que el cambio medio del peso es menor que 0. Parece que la dieta es efectiva, pero la pérdida media de peso de sólo 2.1 libras sugiere que la dieta no es práctica.
17. $H_0: \mu = 140$ mm Hg. $H_1: \mu < 140.0$ mmHg. Estadístico de prueba: $z = -2.27$. Valor crítico: $z = -1.645$. Valor P : 0.0116. Se rechaza H_0 . Existe suficiente evidencia para sustentar la aseveración de que la media es menor que 140.0 mmHg. Sí.
19. $H_0: \mu = 5.670$ g. $H_1: \mu \neq 5.670$ g. Estadístico de prueba: $z = -2.86$. Valores críticos: $z = \pm 2.575$. Valor P : 0.0042 (con tecnología: 0.0043). Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 5.670 g. Parece que estas monedas tienen pesos que no cumplen con las especificaciones, pero esto puede deberse al desgaste por el uso.
21. Potencia = 0.9582 (con tecnología: 0.9580); $\beta = 0.0418$ (con tecnología: 0.0420). El alto valor de la potencia indica que si $\mu = 108$, entonces la prueba es muy efectiva al reconocer que $\mu > 100$.

Sección 8-4

1. No, los requisitos se pueden satisfacer cuando $n \leq 30$, siempre y cuando la población tenga una distribución normal. (Asimismo, se debe conocer σ y la muestra debe ser aleatoria simple).
3. 0.98 o 98%
5. Los requisitos se satisfacen.
7. No, se debe conocer la desviación estándar.
9. a. $z = 0.23$
b. $z = \pm 1.96$
c. 0.8180 (con tecnología: 0.8182)
d. No se rechaza H_0 .
e. No existe evidencia suficiente para justificar el rechazo de la aseveración de que el peso medio de dulces M&M es igual a 0.8535 g.
11. $H_0: \mu = 30.0^\circ\text{C}$. $H_1: \mu > 30.0^\circ\text{C}$. Estadístico de prueba: $z = 1.732$. Valor crítico: $z = 1.645$. Valor P : 0.0416. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 30.0°C .
13. $H_0: \mu = 60$ s. $H_1: \mu \neq 60$ s. Estadístico de prueba: $z = -1.13$. Valores críticos: $z = \pm 1.96$. Valor P : 0.2584 (con tecnología: 0.2577). No se rechaza H_0 . No existe suficiente evidencia para justificar el rechazo de la aseveración de que la media es igual a 60 s. La percepción general media es razonablemente precisa, pero los valores individuales podrían ser erróneos por cantidades bastante grandes.

Sección 8-5

1. 4; 4
3. Si $\bar{x} > 12$ onzas, la prueba de hipótesis nunca podrá sustentar la aseveración de que $\mu < 12$ onzas. Si $\bar{x} > 12$ onzas, entonces no puede ser significativamente menor que 12 onzas.
5. t de Student
7. Ni distribución normal ni t de Student
9. Tabla: $0.005 < \text{valor } P < 0.01$; con tecnología: 0.00641.
11. Tabla: $\text{valor } P < 0.01$; tecnología: 0.0000
13. $H_0: \mu = 120$. $H_1: \mu > 120$. Estadístico de prueba: $t = 3.464$. Valor crítico: $t = 1.796$. Valor $P < 0.005$ (con tecnología: 0.0026). Existe evidencia suficiente para sustentar la aseveración de que la media del CI es mayor que 120.
15. $H_0: \mu = 110$. $H_1: \mu > 110$. Estadístico de prueba: $t = 3.74$. Valor crítico: $t = 1.711$. Valor P : 0.001. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 110.
17. $H_0: \mu = 98.6$. $H_1: \mu < 98.6$. Estadístico de prueba: $t = -6.642$. Valor crítico: $t = -1.660$ (aproximadamente). Valor P : menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la media es menor que 98.6.
19. $H_0: \mu = 3103$ g. $H_1: \mu < 3103$ g. Estadístico de prueba: $t = -8.612$. Valor crítico: $t = -2.345$. El valor P es menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que el peso medio es menor que 3103 g. Parece que los pesos al nacer se ven afectados por el uso de cocaína por parte de la madre.
21. $H_0: \mu = 0$. $H_1: \mu > 0$. Estadístico de prueba: $t = 8.447$. Valor crítico: $t = 2.528$. El valor P es menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la media es positiva. Parece que el tratamiento es efectivo.
23. $H_0: \mu = 63.6$ pulgadas. $H_1: \mu > 63.6$ pulgadas. Estadístico de prueba: $t = 13.200$. Valor crítico: $t = 2.896$. El valor P es menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que las supermodelos tienen una estatura media mayor que 63.6 pulgadas. Sí, la evidencia es fuerte.

25. $H_0: \mu = 1.5$. $H_1: \mu > 1.5$. Estadístico de prueba: $t = 0.049$. Valor crítico: $t = 2.015$. El valor P es mayor que 0.10 (con tecnología: 0.4814). No se rechaza H_0 . No existe evidencia suficiente para justificar la aseveración de que la media es mayor que $1.5 \mu\text{g}/\text{m}^3$. El supuesto de una distribución normal es cuestionable porque parece que 5.40 es un valor extremo.
27. $H_0: \mu = 1.8 \text{ g}$. $H_1: \mu \neq 1.8 \text{ g}$. Estadístico de prueba: $t = -1.297$. Valores críticos: $t = \pm 2.145$ (suponiendo $\alpha = 0.05$). El valor P es mayor que 0.20 (con tecnología: 0.2155). No se rechaza H_0 . No existe evidencia suficiente para justificar la aseveración de que los murciélagos tienen un peso medio igual a 1.8 g.
29. $H_0: \mu = 5.670 \text{ g}$. $H_1: \mu \neq 5.670 \text{ g}$. Estadístico de prueba: $t = -3.135$. Valores críticos: $t = \pm 2.024$ (aproximadamente, suponiendo $\alpha = 0.05$). El valor P es menor que 0.01 (con tecnología: 0.0033). Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 5.670 g.
31. $H_0: \mu = 60$. $H_1: \mu > 60$. Estadístico de prueba: $t = 5.262$. Valor crítico: $t = 1.686$ (aproximadamente, suponiendo $\alpha = 0.05$). El valor P es menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la población tiene un pulso medio mayor que 60.
33. $H_0: \mu = 100 \text{ s}$. $H_1: \mu \neq 100 \text{ s}$. Estadístico de prueba: $t = 1.999$. Valores críticos: $t = \pm 2.040$. $0.05 < \text{valor } P < 0.10$ (con tecnología: 0.0545). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 100. Utilizando el método alternativo con 15.0 para σ , obtenemos el estadístico de prueba $z = 2.00$, los valores críticos $z = \pm 1.96$, el valor $P = 0.0456$, de manera que rechazamos H_0 y concluimos que existe evidencia suficiente para sustentar el rechazo de la aseveración de que la media es igual a 100. Las conclusiones son diferentes.
35. El estadístico de prueba cambia a $t = 0.992$ y el valor P cambia a 0.182. Un valor extremo puede cambiar el estadístico de prueba y el valor P de manera sustancial. Aunque en este caso la conclusión no se modifica, podría ocurrir en otros casos.
37. Potencia = 0.4127, $P(\text{error tipo II}) = \beta = .5873$. No, la potencia de 0.4127 indica que la probabilidad de rechazar $\mu = 1.8 \text{ g}$ (cuando $\mu = 1.7 \text{ g}$) no es muy alta. La potencia debe ser igual o mayor que 0.8.
9. $H_0: \sigma = 696 \text{ g}$. $H_1: \sigma \neq 696 \text{ g}$. Estadístico de prueba: $\chi^2 = 162.317$. Los valores críticos ya están dados. Valor P mayor que 0.05 (con tecnología: 0.1591). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la desviación estándar de los pesos al nacer de hijos nacidos de madres que consumen cocaína difiere de 696 g.
11. $H_0: \sigma = 0.056 \text{ g}$. $H_1: \sigma > 0.056 \text{ g}$. Estadístico de prueba: $\chi^2 = 1225.765$. Valor crítico: $\chi^2 = 63.691$. Valor P menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que los dulces M&M con cacahuate tienen pesos que varían más que los pesos de los dulces M&M sencillos. Los pesos de los dulces M&M con cacahuate varían más que los pesos de los cacahuetes varían mucho.
13. $H_0: \sigma = 83$. $H_1: \sigma \neq \text{espacio } 83$. Estadístico de prueba: $\chi^2 = 20.024$. Valores críticos: $\chi^2 = 6.262, 27.488$. Valor P mayor que 0.20 (con tecnología: 0.3420). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la desviación estándar difiera de 83. Al parecer los solicitantes de la sucursal no tienen calificaciones de crédito que varían más que las del banco principal.
15. $H_0: \sigma = 29$. $H_1: \sigma < 29$. Estadístico de prueba: $\chi^2 = 0.540$. Valor crítico: $\chi^2 = 1.646$. Valor P menor que 0.005 (con tecnología: 0.0002). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que los pesos de las supermodelos varían menos que los pesos de las mujeres en general.
17. $H_0: \sigma = 0.4$. $H_1: \sigma > 0.4$. Estadístico de prueba: $\chi^2 = 114.506$. Valor crítico: $\chi^2 = 11.071$. Valor P menor que 0.005 (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para justificar la aseveración de que la desviación estándar es mayor que 0.4. Parece que el valor muestral de 5.40 es un valor extremo, y el requisito de una distribución normal es cuestionable.
19. $H_0: \sigma = 0.068 \text{ g}$. $H_1: \sigma \neq 0.068 \text{ g}$. Estadístico de prueba: $\chi^2 = 32.355$. Valores críticos: $\chi^2 = 24.433, 59.342$ (aproximadamente). Valor P mayor que 0.20 (con tecnología: 0.4695). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que la desviación estándar es igual a 0.068 g. Parece que la variación de los pesos es la deseada.
21. El valor aproximado de 152.3645 se acerca al valor dado por STATDISK de 152.8222, y el valor aproximado de 228.4771 se acerca al valor dado por STATDISK de 228.9638. La aproximación es bastante buena.
23. Sí, porque un valor extremo afecta de manera drástica el valor de la desviación estándar muestral.

Sección 8-6

- La prueba chi cuadrada no es resistente a los alejamientos de la normalidad, de manera que es muy sensible a estos alejamientos y tiene un mal desempeño con distribuciones muy alejadas de la normalidad. Cuando se pone a prueba el requisito de una distribución normal, utilizamos criterios mucho más estrictos para determinar que una distribución tiene forma de campana.
- No, es probable que la distribución de los resultados sea uniforme y no normal, como se requiere.
- Estadístico de prueba: $\chi^2 = 20.250$. Valores críticos: $\chi^2 = 2.700, 19.023$. Valor P entre 0.02 y 0.02 (con tecnología: 0.0329). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que $\sigma \neq 2.00$.
- Estadístico de prueba: $\chi^2 = 8.889$. Valor crítico: $\chi^2 = 8.260$. Valor P entre 0.01 y 0.025 (con tecnología: 0.158). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que $\sigma < 15$.

Capítulo 8 Conocimientos estadísticos y pensamiento crítico

- La pérdida media de peso es estadísticamente significativa, pero no parece tener una significancia práctica. No debe recomendarse porque produce una pérdida de peso muy pequeña.
- La población está limitada a los participantes. Debido a que se trata de una muestra de respuesta voluntaria, los resultados de una prueba de hipótesis no se aplican a la población de todos los estadounidenses adultos.
- 0.001, porque corresponde a la evidencia más firme de un tratamiento efectivo. Debido a que el valor P de 0.001 es tan bajo, indica que es muy poco probable que los resultados se puedan obtener por el azar, suponiendo que el tratamiento no tiene efecto alguno.

4. No existe evidencia suficiente para sustentar la aseveración de que la cantidad media de bebida de cola sea mayor que 12 onzas. Suponiendo que la cantidad media de bebidas de cola en las latas es igual a 12 onzas, existe una probabilidad de 0.25 de obtener una media muestral con un valor igual o mayor que la media muestral que se obtuvo.

Capítulo 8 Ejercicios de repaso

1. a. $H_1: \mu < \$10,000$; distribución t
b. $H_1: p > 0.5$; distribución normal
c. $H_1: \mu \neq 100$; distribución normal
d. $H_1: \sigma > 1.8$ s; los métodos de este capítulo no deben utilizarse.
e. $H_1: \sigma < 15$; distribución chi cuadrada
2. $H_0: p = 0.5$. $H_1: p < 0.5$. Estadístico de prueba: $z = -1.47$. Valor crítico: $z = -1.645$. Valor P : 0.0708. No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que menos de la mitad de todo los ejecutivos identifican “poco o ningún conocimiento sobre la empresa” como el error de entrevista más común.
3. $H_0: p = 0.62$. $H_1: p < 0.62$. Estadístico de prueba: $z = -2.06$. Valor crítico: $z = -2.33$. Valor P : 0.0197. No se rechaza H_0 . No existe evidencia suficiente para justificar la aseveración de que menos del 62% de las novias gastan menos de \$750 en su vestido de novia. Los resultados no serían válidos si se obtuvieran de una muestra de respuesta voluntaria.
4. $H_0: \mu = 3.5$ g. $H_1: \mu \neq 3.5$ g. Estadístico de prueba: $t = 9.723$. Valores críticos: $t = \pm 1.994$ (aproximadamente, suponiendo que $\alpha = 0.05$). Valor P menor que 0.01 (con tecnología: 0.0000). Se rechaza H_0 . Existe suficiente evidencia para justificar el rechazo de la aseveración de que la media es igual al 3.5 g. Al parecer la diferencia no es lo suficientemente grande para crear problemas de salud a los consumidores.
5. $H_0: \mu = 12$ onzas. $H_1: \mu < 12$ onzas. Estadístico de prueba: $t = -4.741$. Valor crítico: $t = -1.714$ (suponiendo que $\alpha = 0.05$). El valor P es menor que 0.005 (con tecnología: 0.00004). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la media es menor que 12 onzas. El argumento de Windsor no es válido.
6. $H_0: p = 0.019$. $H_1: p > 0.019$. Estadístico de prueba: $z = 0.65$. Valor crítico: $z = 2.33$. Valor P : 0.2578 (con tecnología: 0.2582). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que el porcentaje de pacientes con síntomas de gripe tratados sea mayor al 1.9%. Al parecer los síntomas de gripe no son una reacción adversa al tratamiento.
7. $H_0: \sigma = 15$. $H_1: \sigma \neq 15$. Estadístico de prueba: $\chi^2 = 2.765$. Valores críticos: $\chi^2 = 4.404, 23.337$. Valor P menor que 0.01 (con tecnología: 0.0060). Se rechaza H_0 . Me existe evidencia suficiente para justificar el rechazo de la aseveración de que $\sigma = 15$. Parece que la desviación estándar de los profesores de estadística es menor que 15.
8. $H_0: \mu = 100$. $H_1: \mu \neq 100$. Estadístico de prueba: $z = -0.75$. Valor crítico: $z = \pm 1.645$. Valor P : 0.4532 (con tecnología: 0.4532). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 100. Con base en esos resultados, la calculadora funciona bien.
9. $H_0: \mu = 98.6^\circ\text{F}$. $H_1: \mu < 98.6^\circ\text{F}$. Estadístico de prueba: $t = -1.349$. Valor crítico: $t = -1.796$. Valor P mayor que 0.10 (con tecnología:

0.1023). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la temperatura media corporal es menor que 98.6°F .

10. $H_0: \sigma = 2.52$ pulgadas. $H_1: \sigma \neq 2.52$ pulgadas. Estadístico de prueba: $\chi^2 = 46.107$. Valores críticos: $\chi^2 = 24.433, 59.342$ (aproximadamente). Valor P mayor que 0.20 (con tecnología: 0.4038). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que $\sigma = 2.52$ pulgadas. Los diseños que se basan en una desviación estándar tan pequeña podrían provocar que muchas mujeres no sean capaces de utilizar los productos.
11. $H_0: p = 0.610$. $H_1: p > 0.610$. Estadístico de prueba: $z = 1.60$. Valor crítico: $z = 1.645$. Valor P : 0.0548 (con tecnología: 0.0552). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la tasa de “strikes” sea mayor que 61.0%.
12. $H_0: \mu = 40$ mg. $H_1: \mu > 40$ mg. Estadístico de prueba: $t = 2.746$. Valor crítico: $t = 2.821$. Valor P mayor que 0.01 (con tecnología: 0.0113). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que el contenido medio corporal de nicotina sea mayor de 40 mg.

Capítulo 8 Ejercicios de repaso acumulativo

1. a. 71.5 pulgadas
b. 71.5 pulgadas
c. 2.1 pulgadas
d. 4.6 pulgadas²
e. 8.0 pulgadas
f. 70.4 pulgadas $< \mu < 72.6$ pulgadas.
g. $H_0: \mu = 69.0$ pulgadas. $H_1: \mu > 69.0$ pulgadas. Estadístico de prueba: $t = 4.908$ (o 4.875 si se utiliza una mayor precisión con los datos originales). Valor crítico: $t = 1.746$. Valor P menor que 0.005 (con tecnología: 0.0001). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que el peso medio es mayor que 69.0 pulgadas. Al parecer los presidentes son más altos que el hombre común.
2. Sí, los puntos se acercan mucho al patrón de una línea recta, y no existen otros patrones diferentes de éste.
3. a. 0.4840 (con tecnología: 0.4852)
b. 0.0266 (con tecnología: 0.0269)
c. 0.4681 (con tecnología: 0.4670)
d. 634

Capítulo 9 Respuestas

Sección 9-2

1. Está representada por \bar{p} y es la proporción que consiste en el número total de éxitos en ambas muestras, dividido entre el número total de ensayos para ambas muestras. Se utiliza como un estimado de la proporción poblacional común en la prueba de hipótesis, pero no se utiliza para construir intervalos de confianza para la diferencia entre las dos proporciones.
3. Tenemos una confianza del 95% de que los límites 0.200 y 0.300 contienen el valor verdadero de la diferencia $p_1 - p_2$. Esto significa que, si repitiéramos el proceso de muestreo muchas veces, los límites del intervalo de confianza incluirían la diferencia verdadera entre las proporciones poblacionales en el 95% de los casos.

5. 748
7. 68
9. a. $150/900 = 0.167$
b. 3.00
c. ± 1.96
d. 0.0026 (con tecnología: 0.0027)
11. $H_0: p_1 = p_2$. $H_1: p_1 \neq p_2$. Estadístico de prueba: $z = 1.10$. Valor P : 0.2719. Valores críticos para $\alpha = 0.05$: $z = \pm 1.96$. No se rechaza H_0 . No existe evidencia suficiente para justificar la aseveración de que existe una diferencia significativa de las proporciones de los triunfos locales. Con base en esos resultados, parece que la ventaja de ser el equipo local es casi igual en el básquetbol y en el fútbol.
13. a. $0.0224 < p_1 - p_2 < 0.264$
b. Sí. Los límites del intervalo de confianza no incluyen a 0, lo que sugiere que existe una diferencia entre las dos proporciones. Parece que tanto el método XSORT como el método YSORT son efectivos porque ambos producen resultados que difícilmente ocurrían por el azar.
15. $H_0: p_1 = p_2$. $H_1: p_1 > p_2$. Estadístico de prueba: $z = 9.59$. Valor P : 0.0000. Valor crítico: $z = 2.33$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la tasa de rechazos es más baja con la encuesta rigurosa, la cual tiene más probabilidades de producir resultados precisos.
17. $H_0: p_1 = p_2$. $H_1: p_1 > p_2$. Estadístico de prueba: $z = 2.17$. Valor P : 0.0150 (con tecnología: 0.0151). Valor crítico: $z = 1.645$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que, entre los que afirman que la vigilancia del correo electrónico es muy poco ética, la proporción de empleados es mayor que la proporción de jefes.
19. $H_0: p_1 = p_2$. $H_1: p_1 \neq p_2$. Estadístico de prueba: $z = 2.43$. Valor P : 0.0150 (con tecnología: 0.0149). Valores críticos para $\alpha = 0.05$: $z = \pm 1.96$. Valores críticos para $\alpha = 0.01$: $z = \pm 2.575$. La diferencia es significativa a nivel 0.05, pero no a nivel 0.01.
21. Utilice las proporciones muestrales de $35/1655$ y $20/1652$. $H_0: p_1 = p_2$. $H_1: p_1 > p_2$. Estadístico de prueba: $z = 2.03$. Valor P : 0.0212 (con tecnología: 0.0210). Valor crítico: $z = 1.645$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la incidencia de la fatiga es mayor entre los individuos que usan Clarinex. Aunque la incidencia de la fatiga parece ser significativamente mayor entre los usuarios de Clarinex, la tasa del 2.1% no es mucho mayor que la tasa del 1.2%, de manera que la fatiga no parece ser una preocupación importante para los usuarios del medicamento.
23. Utilice las proporciones muestrales de $880/1068$ y $725/1064$ para obtener $0.106 < p_1 - p_2 < 0.179$. Parece que no existe una diferencia significativa. Factor principal: la respuesta varía, pero una razón podría ser que los propietarios de su casa tienden a vivir en los suburbios (donde tienen mayores posibilidades de conducir), mientras que aquellos que rentan suelen vivir en áreas urbanas (donde tienen mayores posibilidades de utilizar el transporte público). Los niveles de ingreso podrían ser otro factor.
25. $H_0: p_1 = p_2$. $H_1: p_1 \neq p_2$. Estadístico de prueba: $z = -9.83$. Valor P : 0.0002 (con tecnología: 0.0000). Valores críticos: $z = \pm 1.96$. Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que ambos géneros tienen la misma tasa de uso del cinturón de seguridad. Al parecer no existe una brecha entre géneros.
27. $-0.135 < p_1 - p_2 < 0.0742$ (utilizando $x_1 = 49$ y $x_2 = 70$); al parecer no existe una brecha entre géneros.
29. Entre semana: 92/261, a fines de semana: 35/104. $H_0: p_1 = p_2$. $H_1: p_1 < p_2$. Estadístico de prueba: $z = 0.29$. Valor P : 0.6141 (con tecnología: 0.6136). Valor crítico utilizando nivel de significancia de 0.05: $z = -1.645$. No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la proporción de días del fin de semana con lluvia sea mayor que la proporción de días entre semana con lluvia.
31. $H_0: p_1 = p_2$. $H_1: p_1 < p_2$. Estadístico de prueba: $z = -0.60$. Valor P : 0.2743 (con tecnología: 0.2739). Valor crítico utilizando nivel de significancia de 0.05: $z = -1.645$. No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la proporción de películas infantiles que muestran el consumo de alcohol sea menor que la proporción que muestra el consumo de tabaco. Los resultados no se aplican al conjunto de datos 5 porque las muestras no son independientes.
33. a. $0.0227 < p_1 - p_2 < 0.217$; puesto que los límites del intervalo de confianza no contienen a 0, parece que no se puede rechazar $p_1 = p_2$.
b. $0.491 < p_1 < 0.629$; $0.371 < -p_2 < 0.509$; debido a que los intervalos de confianza se traslanan, parece que no se puede rechazar $p_1 = p_2$.
c. $H_0: p_1 = p_2$. $H_1: p_1 \neq p_2$. Estadístico de prueba: $z = 2.40$. Valor P : 0.0164. Valores críticos: $z = \pm 1.96$. Se rechaza H_0 . Existe evidencia suficiente para rechazar $p_1 = p_2$.
d. Se rechaza $p_1 = p_2$. Método menos efectivo: utilizar el traslape entre los dos intervalos de confianza individuales.
35. $H_0: p_1 - p_2 = 0.50$. $H_1: p_1 - p_2 \neq 0.50$. Estadístico de prueba: $z = 1.65$ (utilizando $x_1 = 151$ y $x_2 = 17$). Valor P : 0.0990 (con tecnología: 0.0985). Valores críticos: $z = \pm 1.96$. No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que la tasa de filtración de virus de los guantes de vinilo es 50 puntos porcentuales mayor que la tasa de filtración de virus de los guantes de látex.
37. a. Para las proporciones 0, 0.5, 0.5, 1, $\mu = 0.5$ y $\sigma^2 = 1/8$.
b. Para las diferencias 0, -0.5, -0.5, -1, 0.5, 0, 0, -0.5, 0.5, 0, 0, -0.5, 1, 0.5, 0.5, 0, $\mu = 0$ y $\sigma^2 = 1/4$.
c. $\sigma_{\hat{p}_D}^2 + \sigma_{\hat{p}_0}^2 = 1/8 + 1/8 = 1/4$, y $\sigma_{(\hat{p}_D - \hat{p}_0)}^2 = 1/4$.

Sección 9-3

1. No, 80 sería un valor extremo, lo que sugiere que el requisito de distribuciones normales no se satisface.
3. El valor crítico de 1.717 es más conservador que 1.682 pulgadas en el sentido de que el rechazo de la hipótesis nula de medias iguales requiere de una diferencia mayor entre las medias muestrales. La evidencia muestral debe ser más fuerte con el valor crítico de 1.717.
5. Independiente
7. Datos apareados
9. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 \neq \mu_2$. Estadístico de prueba: $t = 1.674$. Valores críticos: $t = \pm 1.968$ (aproximadamente). El valor P está entre 0.05 y 0.10 (con tecnología: 0.0946). No se rechaza H_0 . Al parecer la equinacea no tuvo efecto alguno.
11. $6.7 < \mu_1 - \mu_2 < 12.1$. Parece que la puntuación del CI no se ve afectada por el peso al nacer.
13. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 > \mu_2$. Estadístico de prueba: $t = 2.790$. Valor crítico: $t = 2.390$ (aproximadamente). El valor P es menor que 0.005 (con tecnología: 0.0031). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que los consumidores frecuentes tienen una media más baja que los consumidores ocasionales. Puesto que parece que la marihuana tiene un efecto adverso en las habilidades mentales, debería ser un motivo de preocupación importante.

15. $-0.65 < \mu_1 - \mu_2 < 3.03$ (con tecnología: $-0.61 < \mu_1 - \mu_2 < 2.99$). Puesto que el intervalo de confianza incluye a 0, no debemos concluir que las dos medias poblacionales son diferentes. Parece que el tratamiento no es efectivo, de manera que no se debe prescribir la paroxetina.
17. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 > \mu_2$. Estadístico de prueba: $t = 0.132$. Valor crítico: $t = 1.729$. Valor $P > 0.10$ (con tecnología: 0.4480). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que los magnetos son efectivos para reducir el dolor. Es válido argumentar que los magnetos podrían ser efectivos con tamaños muestrales más grandes.
19. $-0.01 < \mu_1 - \mu_2 < 0.23$; puesto que este CI incluye a 0, al parecer no existe una diferencia significativa entre las dos medias poblacionales, de manera que el trastorno obsesivo-compulsivo no tiene una base biológica. (Con tecnología, $gl = 18$ y $0.01 < \mu_1 - \mu_2 < 0.21$, que no contienen a cero, lo que sugiere que hay una diferencia significativa y parece que el trastorno obsesivo-compulsivo tiene una base biológica. Este es un caso infrecuente en el que el estimado sencillo y conservador de gl conduce a una conclusión diferente a la de la fórmula 9-1 más exacta).
21. $1.46 < \mu_1 - \mu_2 < 3.52$ (tecnología: $1.47 < \mu_1 - \mu_2 < 3.51$). Puesto que el intervalo de confianza no incluye a cero, parece que hay una diferencia significativa entre las dos medias poblacionales. Al parecer los que fueron tratados con alcohol cometieron significativamente más errores.
23. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 > \mu_2$. Estadístico de prueba: $t = 3.992$. Valor crítico: $t = 2.132$. El valor P se encuentra entre 0.005 y 0.01 (con tecnología: 0.0035). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que las casas con vista al mar valen más que las casas que están junto al mar. La gran diferencia entre las medias compensa el hecho de que las muestras sean pequeñas.
25. Con filtro: $n_1 = 21$, $\bar{x}_1 = 13.3$, $s_1 = 3.7$. Sin filtro: $n_2 = 8$, $\bar{x}_2 = 24.0$, $s_2 = 1.7$. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 < \mu_2$. Estadístico de prueba: $t = -10.585$. Valor crítico: $t = -1.895$. Valor $P < 0.005$ (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la cantidad media de alquitrán en cigarrillos largos con filtro es menor que la cantidad media de alquitrán en cigarrillos largos sin filtro.
27. $n_1 = 40$, $\bar{x}_1 = 6.192672$, $s_1 = 0.086995$, $n_2 = 40$, $\bar{x}_2 = 5.639298$, $s_2 = 0.0619368$. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 \neq \mu_2$. Estadístico de prueba: $t = 32.773$. Valores críticos: $t = \pm 2.024$ (suponiendo un nivel de significancia de 0.05). Valor $P < 0.01$ (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que las dos poblaciones tienen medias iguales. La diferencia es muy significativa, aunque la muestra es relativamente pequeña.
29. $-1.3 < \mu_1 - \mu_2 < 3.1$. Los resultados son similares a los del ejercicio 12, de manera que en este caso no serán afectados por el supuesto de desviaciones estándar iguales.
31. El estadístico de prueba es el mismo: $t = 0.132$. El valor critico cambia de $t = 1.729$ to $t = 1.686$. Valor $P > 0.10$ (con tecnología: 0.4479). En este caso, las conclusiones son iguales.
33. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 \neq \mu_2$. Estadístico de prueba: $t = \pm 2.093$. Valor $P > 0.20$ (con tecnología: 0.4904). El estadístico de prueba cambia en una cantidad importante. Aunque en este caso la conclusión es la misma, los resultados se ven drásticamente afectados.
35. a. $50/3$
 b. $2/3$
 c. $52/3$
 d. $50/3 + 2/3 = 52/3$
 e. El rango de los valores x y y es igual al rango de los valores x más el rango de los valores y .
37. $gl = 18$ (en vez de 9), los valores críticos se convierten en $t = \pm 2.878$ (en vez de ± 3.250), los límites del intervalo de confianza se convierten en 0.007 y 0.213, y el valor P es menor que 0.01 (no está entre 0.01 y 0.02). Con el uso de la fórmula 9-1, el intervalo de confianza es un poco más estrecho, el valor crítico es un poco más pequeño y el valor P también es un poco más pequeño. Con $gl = 9$ no parece que el trastorno obsesivo-compulsivo tenga una base biológica; con $gl = 18$ de la fórmula 9-1, parece que el trastorno obsesivo-compulsivo si tiene una base biológica. Utilizar el más pequeño de $n_1 - 1$ y $n_2 - 1$ para gl es más conservador (el uso de la fórmula 9-1) en el sentido de que los datos muestrales necesitan ser más extremos para ser considerados significativos, como se puede deducir de las conclusiones diferentes.

Sección 9-4

- No. Aunque los datos son apareados, las dos variables miden cantidades diferentes por lo que no tiene sentido poner a prueba una aseveración acerca de la media de las diferencias.
- \bar{d} Representa la media de las diferencias entre los valores muestrales apareados, y s_d denota la desviación estándar de esas diferencias.
- a. -0.4
 b. 2.1
 c. -0.431
 d. ± 2.776
- $-3.0 < \mu_d < 2.2$
- a. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = -0.41$. Valor $P = 0.691$. No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que el astemizole tenga un efecto. No tome astemizole para el malestar provocado por el movimiento.
 b. 0.3455; le existe evidencia suficiente para sustentar la aseveración de que el astemizole evite el malestar por movimiento.
- a. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = 0.218$. Valores críticos: $t = \pm 2.776$. Valor $P > 0.20$ (con tecnología: 0.8379). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que las diferencias tienen una media de cero. Parece que los valores pronosticados no son razonablemente precisos.
 b. $-2.3 < \mu_d < 2.7$. El intervalo de confianza incluye a cero, por lo que no podemos rechazar la aseveración de que las diferencias tienen una media de cero.
- $1.7 < \mu_d < 13.9$. Puesto que el intervalo de confianza no contiene a cero, parece que las diferencias no tienen una media igual a cero. Los resultados sugieren que no hay una presión sanguínea estable que sea medida de manera consistente y con exactitud.
- $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = -0.132$. Valores críticos: $t = \pm 2.306$ (utilizando un nivel de significancia de 0.05). Valor $P > 0.20$ (con tecnología: 0.8981). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que las diferencias tienen una media de 0. Los resultados sugieren

- que las puntuaciones de los individuos del grupo control no cambiaron mucho.
17. a. $0.69 < \mu_d < 5.56$.
 - b. $H_0: \mu_d = 0$. $H_1: \mu_d > 0$. Estadístico de prueba: $t = 3.036$. Valor crítico: $t = 1.895$. El valor P está entre 0.005 y 0.01 (con tecnología: 0.0095). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que las medidas sensoriales son más bajas después de la hipnosis.
 - c. Sí.
 19. a. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = -1.690$. Valores críticos: $t = \pm 2.228$. El valor P está entre 0.10 y 0.20 (con tecnología: 0.1218). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencia entre las cosechas de los dos tipos de semillas.
 - b. $-78.2 < \mu_d < 10.7$
 - c. No
 21. a. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = 0.155$. Valores críticos: $t = \pm 2.032$. Valor $P > 0.20$ (con tecnología: 0.8775). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay una diferencia entre las temperaturas mínimas reales y las pronosticadas.
 - b. $-1.7 < \mu_d < 2.0$
 - c. La misma conclusión que en el ejemplo. Parece que las temperaturas pronosticadas son exactas.
 23. a. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = -5.354$. Valores críticos: $t = \pm 2.024$. Valor $P < 0.01$ (con tecnología: 0.0000). Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencias entre los precios de venta y los precios de lista. Parece que la diferencia es significativa.
 - b. $-\$10,011.50 < \mu_d < -\4518.50 (con tecnología: $-\$10,009.80 < \mu_d < -\$4,520.20$)
 25. Puesto que los valores deben estimarse a partir de la gráfica, las respuestas variarán. $H_0: \mu_d = 0$. $H_1: \mu_d < 0$. Estadístico de prueba: $t = -4.334$. Valor crítico: $t = -1.860$. Valor $P < 0.005$ (con tecnología: 0.0012). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que el curso de preparación sirve para aumentar las puntuaciones.
 27. a. Estadístico de prueba: $t = 1.861$. Valor crítico: $t = 1.833$. El valor P está entre 0.05 y 0.10 (con tecnología: 0.0620). No se rechaza H_0 . Existe evidencia suficiente para sustentar que $\mu_d > 0$.
 - b. Estadístico de prueba: $t = 1.627$. Valor crítico: $t = 1.833$. El valor P está entre 0.05 y 0.10 (con tecnología: 0.0620). No se rechaza H_0 . No existe evidencia suficiente para sustentar que $\mu_1 > \mu_2$.
 - c. Sí, la conclusión se ve afectada por la prueba que se utilice.
- evidencia suficiente para sustentar la aseveración de que las poblaciones de tratamiento y placebo tienen varianzas diferentes.
7. $H_0: \sigma_1 = \sigma_2$. $H_1: \sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 2.9233$. Valor $P = 0.0021$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la Coca-Cola regular y la Coca-Cola de dieta tienen pesos con desviaciones estándar diferentes. Probablemente la diferencia se deba al azúcar de la Coca-Cola regular.
 9. $H_0: \sigma_1 = \sigma_2$. $H_1: \sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 1.0133$. El valor crítico se encuentra entre 1.8752 y 2.0739. (Con tecnología: el valor crítico superior es $F = 1.9611$ y el valor $P = 0.9690$). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la Coca-Cola de dieta y la Pepsi de dieta tengan pesos con desviaciones estándar diferentes.
 11. $H_0: \sigma_1^2 = \sigma_2^2$. $H_1: \sigma_1^2 > \sigma_2^2$. Estadístico de prueba: $F = 2.1267$. El valor crítico F se encuentra entre 2.1555 y 2.2341. (Con tecnología: Valor $P = 0.0543$.) No se rechaza H_0 . No existe evidencia suficiente para justificar la aseveración de que las reducciones del dolor en el grupo de tratamiento simulado varíen más que las reducciones del dolor en el grupo de tratamiento magnético.
 13. $H_0: \sigma_1^2 = \sigma_2^2$. $H_1: \sigma_1^2 > \sigma_2^2$. Estadístico de prueba: $F = 9.3364$. El valor crítico F se encuentra entre 2.0540 y 2.0960. (Con tecnología: Valor $P = 0.0000$.) Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que el grupo de tratamiento tiene puntuaciones que varían más que las puntuaciones del grupo placebo.
 15. $H_0: \sigma_1 = \sigma_2$. $H_1: \sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 1.9729$. El valor crítico superior de F se encuentra entre 1.8752 y 2.0739, pero debe ser más cercano a 1.8752. (Con tecnología: Valor $P = 0.0368$.) Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que las dos poblaciones de monedas tienen la misma desviación estándar.
 17. $H_0: \sigma_1^2 = \sigma_2^2$. $H_1: \sigma_1^2 \neq \sigma_2^2$. Estadístico de prueba: $F = 2.5906$. El valor crítico superior de F es 2.7006 (suponiendo un nivel de significancia de 0.05). (Con tecnología: Valor $P = 0.0599$.) No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que las dos muestras provienen de poblaciones con la misma varianza.
 19. a. Estadístico de prueba: $F = 2.2080$. El valor crítico superior de F se encuentra entre 1.6668 y 1.8752. (Con tecnología: Valor $P = 0.0052$.) Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que las cantidades de precipitación del miércoles y del domingo tienen la misma desviación estándar.
 - b. Debido a que los valores más bajos incluyen una gran cantidad de ceros, ni las cantidades de precipitación del miércoles ni las del domingo se distribuyen normalmente.
 - c. Puesto que parece que las poblaciones no se distribuyen normalmente, la conclusión dada en el inciso a) no es necesariamente válida. No se aplican los métodos de la sección 9-5.
 21. $c_1 = 1$, $c_2 = 0$, el valor crítico es 5. No se rechaza $\sigma_1^2 = \sigma_2^2$.
 23. a. $F_L = 0.2484$, $F_R = 4.0260$
 - b. $F_L = 0.2315$, $F_R = 5.5234$
 - c. $F_L = 0.1810$, $F_R = 4.3197$

Sección 9-5

1. La prueba F es demasiado sensible a las distribuciones que se alejan de la normalidad. Alternativas: Prueba del conteo de cinco y prueba de Levene-Brown-Forsythe.
3. Es la distribución muestral de los valores de s_1^2/s_2^2 .
5. $H_0: \sigma_1^2 = \sigma_2^2$. $H_1: \sigma_1^2 \neq \sigma_2^2$. Estadístico de prueba: $F = 4.0000$. Valor crítico superior: $F = 2.1819$. Valor $P = 0.0005$. Se rechaza H_0 . Existe

Capítulo 9 Conocimientos estadísticos y pensamiento crítico

1. Utilice los métodos de la sección 9-2 para hacer inferencias acerca de la diferencia entre la proporción de hombres votantes que están a favor del candidato y la proporción de mujeres votantes que también están.

2. Sí. Cada muestra de tamaño 200 tiene las mismas probabilidades de ser elegida.
3. No es adecuado. Los estados tienen diferentes poblaciones, por lo que se deben utilizar medias ponderadas. Un procedimiento más adecuado consiste en el uso de muestras aleatorias simples seleccionadas de la población de Estados Unidos.
4. Si las muestras no son independientes, son dependientes. A las muestras dependientes a menudo se les conoce como datos apareados. En los datos apareados, cada valor en una muestra se aparea con un valor específico de la otra muestra.

Capítulo 9 Ejercicios de repaso

1. a. $H_0: p_1 = p_2$. $H_1: p_1 > p_2$. Estadístico de prueba: $z = 0.64$. Valor crítico: $z = 1.645$. Valor $P = 0.2611$ (con tecnología: 0.2603). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que la proporción de individuos negros detenidos por la policía es mayor que la proporción de individuos blancos.
b. 90%: $-0.0251 < p_1 - p_2 < 0.0551$; debido a que el intervalo de confianza contiene a cero, no parece haber una diferencia significativa entre las dos proporciones.
2. $H_0: \mu_d = 0$. $H_1: \mu_d > 0$. Estadístico de prueba: $t = 2.701$. Valor crítico: $t = 1.812$ (suponiendo un nivel de significancia de 0.05). El valor P se encuentra entre 0.01 y 0.025 (con tecnología: 0.0111). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que los estudiantes varones de estadística exageran sus estaturas.
3. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 > \mu_2$. Estadístico de prueba: $t = 5.529$. Valor crítico: $t = 1.796$. Valor $P < 0.005$ (con tecnología: 0.0000). Se rechaza H_0 . Existe suficiente evidencia para sustentar la aseveración de que es más fácil leer *Harry Potter* que *La Guerra y la Paz*.
4. $H_0: \sigma_1 = \sigma_2$. $H_1: \sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 2.8176$. El valor crítico superior de F se encuentra entre 3.4296 y 3.5257. (Con tecnología: valor $P = 0.1000$). No se rechaza H_0 . No existe evidencia suficiente para justificar al rechazo de la aseveración de que las páginas de los dos libros tienen puntuaciones de facilidad de lectura de Flesch con la misma desviación estándar.
5. a. $H_0: p_1 = p_2$. $H_1: p_1 < p_2$. Estadístico de prueba: $z = -2.82$. Valor crítico: $z = -1.645$. Valor $P = 0.0024$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración. Parece que se debe calentar a los pacientes quirúrgicos por rutina.
b. 90%
c. $-0.205 < p_1 - p_2 < -0.0543$
d. No, es probable que las conclusiones sean diferentes.
6. a. $-27.80 < \mu_1 - \mu_2 < 271.04$
(con tecnología: $-17.32 < \mu_1 - \mu_2 < 260.56$)
b. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 \neq \mu_2$. Estadístico de prueba: $t = 1.841$. Valores críticos: $t = \pm 2.262$. Valor P entre 0.05 y 0.10 (con tecnología: 0.0824). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencia entre las dos medias poblacionales.
c. No
7. $H_0: \sigma_1 = \sigma_2$. $H_1: \sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 1.2922$. Valor crítico superior: $F = 4.0260$. (Con tecnología: valor $P = 0.7087$). No se rechaza H_0 . No existe evidencia suficiente para sustentar la aseveración de que las dos poblaciones tienen diferentes cantidades de variación.

8. a. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = -1.532$. Valores críticos: $t = \pm 2.228$. El valor P está entre 0.10 y 0.20 (con tecnología: 0.1565). No se rechaza H_0 . No existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencia. Parece que no hay diferencia.
b. $-2.7 < \mu_d < 0.5$
c. No, no hay una diferencia significativa.

Capítulo 9 Ejercicios de repaso acumulativo

1. a. Media: 69.5 mi/h; mediana: 69.5 mi/h; $s = 3.4$ mi/h; $s^2 = 11.6$ mi²/h²; rango: 9.0 mi/h.
b. $H_0: \mu = 65$. $H_1: \mu > 65$. Estadístico de prueba: $t = 3.765$. Valor crítico: $t = 1.686$ (aproximadamente). Valor $P < 0.005$ (con tecnología: valor $P = 0.0003$). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 65 mi/h.
c. Con el uso de un histograma o de una gráfica cuantil normal, la distribución no es muy normal, pero su forma tampoco difiere mucho de la de una campana.
d. $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 \neq \mu_2$. Estadístico de prueba: $t = 1.265$. Valores críticos: $t = \pm 2.093$ (suponiendo un nivel de significancia de 0.05). Valor $P > 0.20$ (con tecnología: 0.2167). No se rechaza H_0 . No existe evidencia suficiente para justificar al rechazo de la aseveración de que la velocidad media de los carriles que van hacia el norte es igual a la velocidad media de los carriles que van hacia el sur. La distribución es normal, pero no hay valores extremos y la distribución no está muy sesgada, de manera que la robustez de la prueba t probablemente lo haga aceptable en esta situación.
2. a. 1/512 o 0.00195. La probabilidad indica que es muy poco probable que obtenga nueve caras por el azar, de manera que parece que tiene la capacidad de lanzar la moneda y que caiga en cara.
b. Independiente. Los resultados no están apareados de ninguna forma.
c. $H_0: p = 0.5$. $H_1: p > 0.5$. Estadístico de prueba: $z = 2.36$. Valor crítico: $z = 2.33$. Valor $P = 0.0091$ (con tecnología: 0.0092). Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que se puede lanzar una moneda y que resultan caras con mayor frecuencia de lo esperado por el azar.
3. Debe haber un error, porque tasas del 13.7% y 10.6% no son posibles con tamaños muestrales de 100.

Capítulo 10 Respuestas

Sección 10-2

1. No, porque los números del seguro social no satisfacen los requisitos necesarios. Debido a que funcionan como etiquetas de identificación que no miden ni cuentan algo, no son datos cuantitativos.
3. Una correlación es una relación entre dos variables. Una variable intervencional es aquella que afecta a las variables que se estudian, pero que no está incluida en el estudio.
5. a. Sí, porque el valor absoluto del estadístico de prueba excede a los valores críticos $r = \pm 0.707$.
b. 0.986

7. a. No, porque el valor absoluto del estadístico de prueba no excede a los valores críticos $r = \pm 0.444$ (aproximadamente).
b. 0.0177
9. El diagrama de dispersión sugiere que existe una correlación lineal. Con $r = 0.994$ y los valores críticos de $r = \pm 0.878$ (para un nivel de significancia de 0.05), existe una correlación lineal significativa.
11. a. Al parecer hay una correlación lineal.
b. $r = 0.906$. Valores críticos: $r = \pm 0.632$ (para un nivel de significancia de 0.05). Existe una correlación lineal.
c. $r = 0$. Valores críticos: $r = \pm 0.666$ (para un nivel de significancia de 0.05). Parece que no hay una correlación lineal.
d. El efecto de un solo par de valores puede ser muy grande y cambiar la conclusión.
13. $r = 0.269$. Valores críticos: $r = \pm 0.707$. No es una correlación lineal.
15. $r = 0.926$. Valores críticos: $r = \pm 0.754$. Sí hay una correlación lineal. Otros factores incluyen la calidad de la película, el tipo de película y las estrellas que actúan.
17. $r = 0.983$. Valores críticos: $r = \pm 0.632$. Existe una correlación lineal.
19. $r = 0.658$. Valores críticos: $r = \pm 0.532$. Existe una correlación lineal. Otro problema es la exactitud de las mediciones, ya que éstas parecen variar mucho. Se podría realizar un estudio para determinar si la presión sanguínea del sujeto realmente varía considerablemente, o si las mediciones son erróneas debido a otros factores.
21. $r = -0.118$. Valores críticos: $r = \pm 0.707$. No existe una correlación lineal. Menor: Susan Lucci. Mayor: Kelsey Grammer.
23. $r = 0.183$. Valores críticos: $r = \pm 0.707$. No existe una correlación lineal. Parece que el tiempo del triunfo no se ve afectado por la temperatura.
25. $r = 0.874$. Valores críticos: $r = \pm 0.707$. Existe una correlación lineal. Sí.
27. $r = -0.038$. Valores críticos: $r = \pm 0.576$. No existe una correlación lineal. Parece que la hipótesis no es correcta.
29. $r = 0.995$. Valores críticos: $r = \pm 0.312$. Existe una correlación lineal.
31. a. $r = 0.961$. Valores críticos: $r = \pm 0.361$ aproximadamente. Existe una correlación lineal. Sí.
b. $r = 0.863$. Valores críticos: $r = \pm 0.361$ aproximadamente. Existe una correlación lineal. Sí.
c. El alquitrán, porque tiene una correlación más alta con la nicotina.
33. Con un coeficiente de correlación lineal muy cercano a 0, debemos concluir que no existen a correlación lineal, de manera que no parece haber una asociación entre la edad y la puntuación, como sugiere la conclusión de manera incorrecta. 35. Aunque no existe una correlación lineal, las variables podrían estar relacionadas de una manera no lineal.
35. Aunque no hay correlación lineal, las variables deben estar relacionadas de alguna forma no lineal.
37. a. 0.972
b. 0.905
c. 0.999 (el más grande)
d. 0.992
e. -0.984
39. $0.386 < \rho < 0.753$

Sección 10-3

1. En una ecuación de regression $\hat{y} = b_0 + b_1x$, la variable x es la variable predictor y la variable y es la variable de respuesta. Reciben estos nombres porque el valor de x generalmente se utiliza para predecir el valor de y , de manera que el valor de y responde al valor de x .
3. podrían aplicar diferentes condiciones, de manera que el valor predicho se alejaría por una importante cantidad.
5. 109
7. 3.86 calorías por gramo de cereal.
9. $\hat{y} = 0.857 + 0.643x$
11. a. $\hat{y} = 0.264 + 0.906x$
b. $\hat{y} = 2 + 0x$ (o $\hat{y} = 2$)
c. Los resultados son muy diferentes, lo que indica que un punto puede afectar de manera drástica la ecuación de regresión.
13. $\hat{y} = 54.3 + 0.246x$; 86 min
15. $\hat{y} = -164 + 3.47x$; -\$25,200,000 (que es una pérdida).
17. $\hat{y} = -252 + 12.4x$; 368 libras.
19. $\hat{y} = -14.4 + 0.769x$; 93
21. $\hat{y} = 6.76 - 0.0111x$; 6.5 millones.
23. $\hat{y} = 145 + 0.0316x$; 147.077
25. $\hat{y} = 27.6 + 0.0523x$; 79.9°F
27. $\hat{y} = 79.9 - 0.113x$; 73 latidos por minuto.
29. $\hat{y} = 99.2 + 0.979x$; \$391,699
31. a. $\hat{y} = 0.154 + 0.0651x$; 1.1
b. $\hat{y} = 0.192 + 0.0606x$; 1.1
33. Con $\beta_1 = 0$, la línea de regresión es horizontal, de manera que distintos valores de x producen el mismo valor de y , y no existe una correlación entre x y y .
35. La ecuación $\hat{y} = -49.9 + 27.2x$ es mejor porque tiene $r = 0.997$, que es mayor que $r = 0.963$ para $\hat{y} = -103.2 + 134.9x \ln x$.

Sección 10-4

1. Un intervalo de predicción es un intervalo de confianza (o estimado de un intervalo) de un valor predicho de y .
3. Dados datos apareados, \bar{y} es la media de los valores y observados y, para un valor observado de x , podemos calcular el valor predicho de y (denotado por \hat{y}). La desviación explicada es $\hat{y} - \bar{y}$, y la variación explicada es la suma de los cuadrados de los valores de la desviación explicada.
5. 0.09; 9%
7. 0.812; 81.2%
9. 0.961; sí.
11. 1.3
13. a. 147.3965
b. 18.03204
c. 165.4286
d. 0.891
e. 1.899055
15. a. 628.9603
b. 824.254

- c. 1453.214
- d. 0.433
- e. 8.287812
- 17. a. 25 mi/gal
- b. $19.5 \text{ mi/gal} < y < 30.9 \text{ mi/gal}$
- 19. a. 78
- b. $57.7 < y < 98.2$
- 21. $56.6 \text{ min} < y < 97.3 \text{ min}$
- 23. $68.6 \text{ min} < y < 94.6 \text{ min}$
- 25. $13.4 < \beta_0 < 56.1; 0.138 < \beta_1 < 0.330$

Sección 10-5

1. Una ecuación de regresión múltiple expresa una relación lineal entre una variable de respuesta (y) y dos o más variables predictoras (x). Difiere de las ecuaciones de regresión analizadas en la sección 10-3, ya que esa sección incluyó casos con una sola variable predictoría (x), mientras que una ecuación de regresión múltiple tiene dos o más variables predictoras (x).
3. No, porque los datos son categóricos (o cualitativos). Los métodos de esta sección requieren datos cuantitativos. (Puede haber excepciones con la regresión logística).
5. PESO = $-272 - 0.87 \text{ CABEZAL} + 0.55 \text{ ESTAT} + 12.2 \text{ PECHO}$ o $\hat{y} = -272 - 0.87x_1 + 0.55x_2 + 12.2x_3$, donde x_1 representa la longitud de la cabeza, x_2 representa la estatura y x_3 representa el tamaño del pecho.
7. Sí, porque la significancia general de 0.000 es pequeña y la R^2 de 0.924 es alta.
9. El tamaño de la cintura, porque tiene el valor P más bajo de 0.000 y el valor más alto de R^2 ajustada de 0.785.
11. $\hat{y} = -206 + 2.66 \text{ EST} + 2.15 \text{ CINT}$ o $\hat{y} = -206 + 2.66x_1 + 2.15x_2$ (donde x_1 representa la estatura y x_2 representa el tamaño de la cintura). El uso de las variables predictoras de la estatura y el tamaño de la cintura produce el valor P más bajo de 0.000 y la R^2 ajustada más alta de 0.870. El uso de las variables predictoras de la estatura, el tamaño de la cintura y el nivel del colesterol también produce un valor P de 0.000 y una R^2 de 0.870, pero es mejor utilizar dos variables predictoras en lugar de tres.
13. a. $\hat{y} = 0.154 + 0.0651x$
 b. $\hat{y} = 0.192 + 0.0606x$
 c. $\hat{y} = 0.182 + 0.0818x_1 - 0.0186x_2$, donde x_1 representa la cantidad de alquitran y x_2 representa la cantidad de monóxido de carbono
 d. Inciso (a). Los tres producen un valor P de 0.000, pero los valores de la R^2 ajustada son 0.924, 0.736 y 0.928, respectivamente. Elija el inciso a) porque sólo tiene una variable predictoria, mientras que el inciso c) tiene una R^2 un poco más elevada, aunque el inciso c) tiene dos variables predictorias. Si los valores de la R^2 ajustada no son muy diferentes, es mejor utilizar la ecuación de regresión con un menor número de variables predictorias.
 e. Sí, porque el valor P es bajo (0.000) y la R^2 ajustada es alta (0.924).
15. $\hat{y} = 99.2 + 0.979x$, donde x representa el precio de lista. Con un valor P de 0.000 y una R^2 de 0.990, es probable que esta ecuación de regresión sea muy buena para predecir los precios de venta con base en los precios de lista.

17. Para $H_0: \beta_1 = 0$, el estadístico de prueba es $t = 5.453$, el valor P es 0.00282, y los valores críticos son $t = \pm 2.571$, de manera que se rechaza H_0 y se concluye que se debe mantener el coeficiente de regresión de $\beta_1 = 0.245$. Para $H_0: \beta_2 = 0$, el estadístico de prueba es $t = -0.606$, el valor P es 0.571, y los valores críticos son $t = \pm 2.571$, de manera que no se rechaza H_0 y se concluye que se debe omitir el coeficiente de regresión de $\beta_2 = -0.098$. Parece que la ecuación de regresión debe mantener la duración como variable predictoria, pero se debe omitir la estatura.
19. $\hat{y} = 3.06 + 82.4x_1 + 2.91x_2$, donde x_1 representa el sexo y x_2 representa la edad. Hembra: 61 libras; macho: 144 libras. Parece que el sexo del oso no tiene efecto alguno sobre su peso. La ecuación de regresión indica que el peso predicho de una oso macho es aproximadamente 82 libras más que el peso predicho de una hembra, si las demás características permanecen constantes.

Sección 10-6

1. Un modelo matemático es una función matemática que ajusta o describe datos del mundo real.
3. El año 3000 sobrepasa por mucho el alcance de los datos disponibles. Las condiciones podrían cambiar, por lo que el modelo no debería utilizarse para el año 3000.
5. Lineal: $y = 2x + 3$
7. Cuadrática: $y = 2x^2 - 1$
9. Potencia: $y = 18.1x^{0.455}$ (donde 1983 se codifica con 1). Con $R^2 = 0.789$, parece que el modelo es adecuado.
11. Cuadrática: $y = 4.90x^2 - 0.0286x + 0.00476$. El modelo produce una distancia de 705 m, pero el edificio sólo tiene 50 m de alto, de manera que la distancia no puede ser mayor que 50 m.
13. a. Exponencial: $y = 2^{2(x-1)}$ [o $y = (0.629961)(1.587401)^x$ para un valor inicial de 1 que se duplica cada 1.5 años]
 b. Exponencial: $y = (1.38)(1.42)^x$
 c. Parece que la ley de Moore funciona bastante bien. Con $R^2 = 0.984$, parece que el modelo es muy bueno.
15. a. 6641.8
 b. 73.2
 c. La suma cuadrática de los cuadrados de los residuales (73.2) es menor que la suma de los cuadrados de los residuales del modelo lineal (6641.8).

Capítulo 10 Conocimientos estadísticos y pensamiento crítico

1. Con la correlación investigamos y existe una relación entre variables, pero con la regresión tratamos de identificar la relación con una ecuación.
2. No. Podría haber una relación que no es lineal.
3. No. El coeficiente de correlación lineal se puede utilizar para establecer que existe asociación entre la ingesta del fármaco y el nivel del colesterol, pero no se puede emplear para establecer que el fármaco *causa* niveles más bajos de colesterol.
4. No necesariamente. Es muy posible que exista una correlación lineal significativa, pero los valores predichos no son muy exactos.

Capítulo 10 Ejercicios de repaso

1. a. $r = 0.090$. Valores críticos: $r = \pm 0.602$ (suponiendo un nivel de significancia de 0.05). No existe una correlación lineal.
1. b. $\hat{y} = 35.7 + 0.166x$. El mejor número predicho de muertes por causas naturales es 47 (la media).
2. a. $r = 0.389$. Valores críticos: $r = \pm 0.707$ (suponiendo un nivel de significancia de 0.05). No existe una correlación lineal.
2. b. $\hat{y} = 105 + 0.108x$.
2. c. 128.75 pies (la media)
3. a. $r = 0.915$. Valores críticos: $r = \pm 0.632$. Existe una correlación lineal.
3. b. 84%
3. c. $\hat{y} = -51.7 + 0.113x$
3. d. \$287.30
4. a. $r = 0.085$. Valores críticos: $r = \pm 0.632$. No existe una correlación lineal.
4. b. 0.7%
4. c. $\hat{y} = 235 + 0.243x$
4. d. \$247.30 (la media)
5. $\hat{y} = -128 + 0.123x_1 + 0.955x_2$. $R^2 = 0.942$; R^2 ajustada = 0.925; valor $P = 0.000$. La ecuación se puede utilizar para predecir el costo, y es mejor que las ecuaciones de regresión de los ejercicios 3 y 4 porque tiene la R^2 ajustada más alta y el valor P más bajo.

Capítulo 10 Ejercicios de repaso acumulativo

1. $r = -0.431$. Valores críticos: $r = \pm 0.707$ (suponiendo un nivel de significancia de 0.05). No hay una correlación lineal.
2. $\hat{y} = 13,403 + 67.2x$; 9950 (la media)
3. Sí, pero la prueba no tendría sentido porque las dos variables miden cantidades diferentes.
4. $41.9 < \mu < 60.9$
5. Los valores del DJIA cambian con el tiempo. El intervalo de confianza sería un estimado de la media basado en los años incluidos en la muestra, pero no necesariamente sería un buen estimado del valor del DJIA del siguiente año.
6. Sí. Una gráfica cuantil normal muestra que los puntos se asemejan razonablemente al patrón de una línea recta, y no existe un patrón diferente de éste.
7. $\bar{x} = 51.4$; $s = 11.4$
8. 0.1587

Capítulo 11 Respuestas

Sección 11-2

1. Hacer una prueba de la “bondad de ajuste” implica poner a prueba la hipótesis de que los datos muestrales coinciden con o se ajustan a una distribución particular que se ha identificado.
3. Una frecuencia observada es un conteo del número de valores muestrales que corresponden a una de las categorías consideradas. Una frecuencia esperada es el número de valores muestrales esperados para una categoría, asumiendo que las frecuencias siguen la distribución aseverada.

5. a. $H_0: p_1 = p_2 = p_3$
5. b. 10, 10, 10
5. c. $\chi^2 = 15.000$
5. d. $\chi^2 = 5.991$
5. e. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las tres categorías son igualmente probables.
7. a. $gl = 37$, de manera que $\chi^2 = 51.805$ (aproximadamente).
7. b. $0.10 < \text{valor } P < 0.90$
7. c. No existe evidencia suficiente para justificar el rechazo de la aseveración de que las ranuras de la ruleta son igualmente probables.
9. Estadístico de prueba: $\chi^2 = 5.860$. Valor crítico: $\chi^2 = 11.071$. (Con tecnología: valor $P = 0.3201$). No existe evidencia suficiente para sustentar la aseveración de que los resultados no son igualmente probables. Parece que los resultados son igualmente probables, de manera que el dado cargado no parece comportarse de manera diferente a un lado legal.
11. Estadístico de prueba: $\chi^2 = 30.017$. Valor crítico: $\chi^2 = 12.592$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las muertes por accidentes automovilísticos ocurren con la misma frecuencia los distintos días de la semana. Quizás la conducta de beber alcohol los viernes sea la causa de un número excepcionalmente grande de muertes las primeras horas del sábado.
13. Estadístico de prueba: $\chi^2 = 47.200$. Valor crítico: $\chi^2 = 19.675$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las muertes en motocicleta ocurren con la misma frecuencia los distintos meses. Quizás las muertes disminuyan en los meses de invierno, cuando el clima frío se relaciona con un uso mucho menor de las motocicletas.
15. Estadístico de prueba: $\chi^2 = 4.211$. Valor crítico: $\chi^2 = 19.675$. (Con tecnología: valor $P = 0.9633$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que las actrices ganadoras del Óscar nacen en los diferentes meses con la misma frecuencia.
17. Estadístico de prueba: $\chi^2 = 7.382$. Valor crítico: $\chi^2 = 19.675$. (Con tecnología: valor $P = 0.7674$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que las bodas se llevan a cabo en los diferentes meses con la misma frecuencia. Los resultados refutan la creencia de que la mayoría de las bodas se llevan a cabo en junio.
19. Estadístico de prueba: $\chi^2 = 7.346$. Valor crítico: $\chi^2 = 7.815$. (con tecnología: valor $P = 0.0616$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que las frecuencias observadas coinciden con las proporciones teóricas. Con base en esos resultados, no existe evidencia suficiente para sustentar la aseveración de que las series de siete juegos ocurran con mayor frecuencia de lo esperado.
21. Estadístico de prueba: $\chi^2 = 6.682$. Valor crítico: $\chi^2 = 11.071$. (Con tecnología: valor $P = 0.2454$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que la distribución del color es como afirma Mars, Inc.
23. Estadístico de prueba: $\chi^2 = 51.270$. Valor crítico: $\chi^2 = 9.488$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que los participantes tienen la misma distribución que la población de Estados Unidos. Si los participantes del estudio no son representativos de la población, los resultados podrían ser confusos, porque algunos grupos podrían tener tasas de cáncer diferentes de los otros, y esto podría ser dar los resultados.

25. Estadístico de prueba: $\chi^2 = 14.421$. Valor crítico: $\chi^2 = 15.507$. (Con tecnología: valor $P = 0.0714$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que los dígitos provienen de una población de dígitos líder que coincide con la ley de Benford. Al parecer sí coinciden.
27. En lugar de usar una prueba de cola derecha, utilice una prueba de cola izquierda. Si los datos muestrales se ajustan a la distribución aseverada de manera casi perfecta, el estadístico de prueba se ubicará cerca del lado izquierdo de cero.
29. a. 0.296, 0.444, 0.222, 0.037
 b. 88.9, 133.3, 66.7, 11.1
 c. Estadístico de prueba: $\chi^2 = 23.241$. valor crítico: $\chi^2 = 7.815$. Se rechaza la aseveración de que las frecuencias observadas se ajustan a una distribución binomial con $n = 3$ y $p = 1/3$.

Sección 11-3

- El estadístico de la prueba chi cuadrada mide la cantidad de desacuerdo entre las frecuencias observadas y las frecuencias que se esperarían si la hipótesis nula de independencia (o proporciones iguales) fuese verdadera.
- Contingencia significa *dependencia*, y se refiere a la prueba de dependencia (con independencia) entre las variables de renglón y de columna.
- Estadístico de prueba: $\chi^2 = 0.413$. Valor $P = 0.521$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que el hecho de ser detenido por la policía es independiente de la raza y del origen étnico. No hay evidencia suficiente para sustentar una afirmación de discriminación racial.
- Estadístico de prueba: $\chi^2 = 153.462$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la independencia entre tratamiento y el desarrollo de la gripe. Parece que la vacuna no es efectiva.
- Estadístico de prueba: $\chi^2 = 5.516$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.0188$). Existe evidencia suficiente para justificar el rechazo de la aseveración de independencia entre el género y la dominancia de la mano izquierda.
- Estadístico de prueba: $\chi^2 = 32.273$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que si un sujeto miente, esto es independiente de la indicación del polígrafo. Los resultados sugieren que el polígrafo es efectivo, pero no siempre es correcto.
- Estadístico de prueba: $\chi^2 = 42.557$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que la sentencia es independiente de la declaración de inocencia. Los resultados sugieren que debe fomentar una declaración de culpabilidad.
- Estadístico de prueba: $\chi^2 = 153.739$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que la proporción de caras es igual cuando se lanza que cuando se gira la moneda.
- Estadístico de prueba: $\chi^2 = 65.524$. Valor crítico: $\chi^2 = 7.815$ (suponiendo un nivel de significancia de 0.05). (Con tecnología: valor $P = 0.0000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que la ocupación es independiente de que la causa de muerte sea homicidio. Parece que los cajeros son los más vulnerables al homicidio.

- Estadístico de prueba: $\chi^2 = 1.358$. Valor crítico: $\chi^2 = 7.815$ (suponiendo un nivel de significancia de 0.05). (Con tecnología: valor $P = 0.7154$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que la cantidad de cigarrillos fumados es independiente del uso del cinturón de seguridad. Los datos dados no sustentan la teoría.
- Estadístico de prueba: $\chi^2 = 9.971$. Valor crítico: $\chi^2 = 9.488$ (suponiendo a nivel de significancia de 0.05). (Con tecnología: valor $P = 0.0409$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las lesiones son independientes del color del casco. Las conclusiones son las mismas del ejemplo.
- Sin corrección de Yates: $\chi^2 = 0.413$. Con corrección de Yates: $\chi^2 = 0.270$. La corrección de Yates disminuye el estadístico de prueba, de manera que los datos muestrales deben ser más extremos para considerarse significativos.

Sección 11-4

- La prueba de McNemar sólo se utiliza con valores que provienen de datos apareados, mientras que los métodos de la sección 11-3 no se aplican a ese tipo de datos.
- Los pares discordantes de resultados provienen de pares de categorías en las que ambas son diferentes.
- 144
- 130
- b, c
- 6.635
- Estadístico de prueba: $\chi^2 = 2.382$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.123$). No existe evidencia suficiente para justificar el rechazo de la hipótesis nula que plantea que las siguientes dos proporciones son iguales: 1) la proporción de sujetos sin cura en el pie tratado con fungicida, y con cura en el pie tratado con placebo; 2) la proporción de sujetos con cura en el pie tratado con fungicida y sin cura en el pie tratado con un placebo. Parece que el tratamiento con fungicidas no es efectivo.
- Estadístico de prueba: $\chi^2 = 6.750$. Valor crítico: $\chi^2 = 3.841$ (suponiendo un nivel de significancia de 0.05). (Con tecnología: valor $P = 0.009$). Se rechaza la hipótesis nula de que las siguientes dos proporciones son iguales: 1) la proporción de tumores identificados incorrectamente con IRM e identificados correctamente con TEP y TC; 2) la proporción de tumores identificados correctamente con IRM e identificados incorrectamente con TEP y TC. Al parecer las tecnologías TEP y TC son más precisas.
- El estadístico de prueba sin corregir es 21.333. El valor sin corregir es un poco más grande que el valor corregido de 20.021. La conclusión es la misma en este caso. Podría haber casos en los que el estadístico de prueba no corregido conduzca al rechazo de la hipótesis nula, mientras que el estadístico de prueba corregido no.
- Cuando se redondea a tres decimales, se obtiene el mismo valor P de 0.289. Con un valor P de 0.289, no se rechaza la hipótesis nula de que las siguientes dos proporciones son iguales: 1) la proporción de sujetos con un pie curado con el tratamiento de Pedacream y el otro pie sin cura con el tratamiento de Fungacream; 2) la proporción de sujetos sin cura en el pie que recibió el tratamiento de Pedacream y el otro pie curado con el tratamiento de Fungacream. Al parecer no existe una diferencia significativa entre los dos tratamientos.

Capítulo 11 Conocimientos estadísticos y pensamiento crítico

- Los datos categóricos (o datos cualitativos) son aquellas que pueden separarse en distintas categorías, que se distinguen por alguna característica no numérica.
- Él está utilizando una muestra de conveniencia que puede estar sesgada y no ser representativa de una población, que no sea la de los compañeros de clase seleccionados
- a, c
- Existe el requisito de que todas las frecuencias *esperadas* deben ser de al menos 5, pero no existe el requisito de que las frecuencias observadas en todas las celdas deban ser de al menos 5. Los métodos podrían emplearse si todos los requisitos se satisfacen, incluso si hay celdas con frecuencias observadas menores que 5.

Capítulo 11 Ejercicios de repaso

- Estadístico de prueba: $\chi^2 = 7.417$. Valor crítico: $\chi^2 = 12.592$. (Con tecnología: valor $P = 0.284$). No existe evidencia suficiente para justificar el rechazo de la afirmación de que los choques fatales ocurren con la misma frecuencia los días de la semana. Al parecer los choques son causados por individuos que beben diariamente.
- Estadístico de prueba: $\chi^2 = 4.698$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.030$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que la respuesta es independiente de ser sujeto es trabajador o jefe. La conclusión cambia si se utiliza un nivel de significancia de 0.01. Parece que los trabajadores y los jefes no coinciden en este tema.
- Estadístico de prueba: $\chi^2 = 119.330$. Valor crítico: $\chi^2 = 5.991$. (Con tecnología: valor $P = 0.000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que el tipo de crimen es independiente del hecho de que el criminal sea un extraño. La policía debería dar mayor importancia a los conocidos y a los parientes cuando investiga homicidios.
- Estadístico de prueba: $\chi^2 = 13.225$. Valor crítico: $\chi^2 = 3.841$. (Con tecnología: valor $P = 0.000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las dos proporciones dadas son iguales. Parece que el tratamiento es efectivo.

Capítulo 11 Ejercicios de repaso acumulativo

- $\bar{x} = 79.6$; mediana: 80.0; rango: 20.0; $s^2 = 44.8$; $s = 6.7$: en resumen de los cinco números: 70, 74, 80, 84.5, 90.
- a. 0.240
b. 0.518
c. 0.633
d. 0.232
e. 0.274
- Tabla de contingencia: vea la sección 11-3. Estadístico de prueba: $\chi^2 = 0.021$. Valor crítico: $\chi^2 = 7.815$ (suponiendo a nivel de significancia de 0.05). (Con tecnología: valor $P = 0.999$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que los hombres y las mujeres eligen las diferentes respuestas en las mismas proporciones.
- Utilice una correlación; vea la sección 10-2. Estadístico de prueba: $r = 0.989$. Valores críticos: $r = \pm 0.950$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para sustentar

la aseveración de que existe una relación entre la memoria y las puntuaciones en la prueba de razonamiento.

- Utilice la prueba para datos apareados; vea la sección 9-4.
 $\bar{d} = 5.75$; $s_d = 0.957$. Estadístico de prueba: $t = 12.011$. Valores críticos: $t = \pm 3.182$ (suponiendo un nivel de significancia de 0.05). (Con tecnología: valor $P = 0.001$). Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que el entrenamiento no tiene efecto alguno. Al parecer el entrenamiento sí tiene un efecto.
- Haga una prueba para la diferencia entre dos muestras independientes; vea la sección 9-3. Estadístico de prueba: $t = 1.265$. Valores críticos: $t = \pm 3.182$ (suponiendo un nivel de significancia de 0.05). (Con tecnología: los valores críticos son $t = \pm 2.447$ y el valor $P = 0.253$). No se rechaza H_0 : $\mu_1 = \mu_2$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que los hombres y las mujeres tienen la misma puntuación.

Capítulo 12 Respuestas

Sección 12-2

- El análisis de varianza de un factor es un método que sirve para probar la igualdad de tres o más medias poblacionales al analizar las varianzas muestrales. El método se llama análisis de varianza de un factor porque se utiliza una sola característica para categorizar las poblaciones.
- La varianza entre muestras es una medida de variación entre medias muestrales. La varianza dentro de las muestras es una medida de variación basada en las varianzas muestrales.
- a. $\mu_1 = \mu_2 = \mu_3 = \mu_4$
b. Al menos una de las cuatro medias es diferente de las demás.
c. $F = 8.448$
d. $F = 3.2874$
e. 0.0016
f. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las cuatro poblaciones tienen medias iguales.
g. La eliminación del valor extremo cambia el estadístico de prueba de $F = 5.7314$ a $F = 8.448$, y el valor P cambia de 0.0073 a 0.0016. La conclusión no cambia.
- Estadístico de prueba: $F = 0.3521$. Valor crítico: $F = 2.6626$. Valor $P = 0.7877$. No se rechaza H_0 : $\mu_1 = \mu_2 = \mu_3 = \mu_4$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que la pérdida media de peso es igual en todas las dietas. Parece que las dietas no son muy efectivas.
- Estadístico de prueba: $F = 0.443$. Valor crítico: $F = 2.3113$. Valor $P = 0.8173$. No se rechaza H_0 : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que los tres colores de los dulces M&M tienen la misma media. No se requiere de una acción correctiva.
- Estadístico de prueba: $F = 0.9922$. Valor crítico: $F = 3.2389$. Valor $P = 0.4216$. No se rechaza H_0 : $\mu_1 = \mu_2 = \mu_3 = \mu_4$. No existe evidencia suficiente para sustentar la aseveración de que los automóviles más grandes son más seguros.
- Estadístico de prueba: $F = 2.4749$. Valor crítico: $F = 3.0984$. Valor $P = 0.0911$. No se rechaza H_0 : $\mu_1 = \mu_2 = \mu_3 = \mu_4$. No existe evidencia suficiente para justificar el rechazo de la aseveración de medias iguales. Se puede considerar que los grupos son muestras de la misma población.

15. Estadístico de prueba: $F = 1713.725$. Valor crítico: $F = 2.6802$ (aproximadamente). Valor $P = 0.0000$. Se rechaza $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las cuatro categorías de centavos tienen el mismo peso medio. Las máquinas que cuentan monedas no pueden tratar los pesos de la misma forma.
17. a. Estadístico de prueba: $t = -0.57139$. Valores críticos: $t = \pm 2.306$. (Con tecnología: valor $P = 0.583$). No se rechaza $H_0: \mu_1 = \mu_2$.
- b. Estadístico de prueba: $F = 0.3265$. Valor crítico: $F = 5.3177$. Valor $P = 0.583$. No se rechaza $H_0: \mu_1 = \mu_2$.
- c. Estadísticos de prueba: $t^2 = F = 0.3265$. Valores críticos: $t^2 = F = 5.3177$.
19. Los resultados de la prueba de Tukey indican que la media de la muestra 4 es significativamente diferente de las otras tres medias muestrales. La conclusión es la misma que la obtenida con los resultados de la prueba de Bonferroni.

Sección 12-3

1. "Análisis de varianza" se refiere al método utilizado, el cual se basa en dos estimados diferentes de la varianza poblacional y común asumida. "Dos factores" se refiere a la inclusión de dos factores diferentes, los cuales son propiedades o características utilizadas para distinguir unas poblaciones de otras. El análisis de varianza de dos factores se utiliza para probar un efecto de cada uno de los factores diferentes, así como también para probar un efecto de interacción entre los dos factores.
3. Si existen interacción entre factores, entonces no debemos considerar los efectos de alguno de los factores sin tomar en cuenta los efectos del otro factor.
5. Estadístico de prueba: $F = 2.79$. Valor $P = 0.056$. Suponiendo un nivel de significancia de 0.05, no se rechaza la hipótesis nula de no interacción. Al parecer no existe un efecto significativo de la interacción entre el terreno y el tratamiento.
7. Estadístico de prueba: $F = 2.57$. Valor $P = 0.072$. No se rechaza la hipótesis nula de que el tratamiento no tiene efecto sobre el peso. No existe evidencia suficiente para sustentar la aseveración de que el tratamiento tiene un efecto sobre el peso.
9. Estadístico de prueba: $F = 5.03$. Valor $P = 0.031$. Suponiendo un nivel de significancia de 0.05, se rechaza la hipótesis nula de que el género no tiene un efecto sobre las puntuaciones de la prueba SAT. Existe evidencia suficiente para sustentar la aseveración de que el género tiene un efecto sobre las puntuaciones de la prueba SAT.
11. Estadístico de prueba: $F = 3.87$. Valor $P = 0.000$. Se rechaza la hipótesis nula de que la selección del sujeto tiene un efecto sobre la puntuación de la prueba de audición.
13. Para la interacción, el estadístico de prueba es $F = 0.36$ y el valor P es 0.701, de manera que no existe un efecto de interacción significativo. Para el género, el estadístico de prueba es $F = 0.09$ y el valor P es 0.762, de manera que no hay un efecto significativo del género. Para la edad, el estadístico de prueba es $F = 0.36$ y el valor P es 0.701, de manera que no hay un efecto significativo de la edad.
15. a. Los estadísticos de prueba, los valores críticos, los valores P y las conclusiones no cambian.
- b. Los estadísticos de prueba, los valores críticos, los valores P y las conclusiones no cambian.

- c. Los estadísticos de prueba, los valores críticos, los valores P y las conclusiones no cambian.
- d. Un valor extremo puede afectar drásticamente y cambiar todos los resultados y las conclusiones.

Capítulo 12 Conocimientos estadísticos y pensamiento crítico

- El método para probar la igualdad de medias se basa en dos estimados diferentes de una varianza poblacional común.
- El análisis de varianza de un factor.
- Prueba t para dos medias independientes (como en la sección 9-3).
- No, porque los datos son cualitativos (o categóricos) en lugar de cuantitativos.

Capítulo 12 Ejercicios de repaso

- Estadístico de prueba: $F = 46.90$. Valor $P = 0.000$. Se rechaza $H_0: \mu_1 = \mu_2 = \mu_3$. Existe evidencia suficiente para justificar el rechazo de la aseveración de medias poblacionales iguales.
- Estadístico de prueba: $F = 9.469$. Valor crítico: $F = 3.3158$ (aproximadamente). (Con tecnología: valor $P = 0.000562$). Se rechaza $H_0: \mu_1 = \mu_2 = \mu_3$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que los tres libros tienen la misma puntuación media. Parece que los tres libros no tienen el mismo nivel de lectura.
- Estadístico de prueba: $F = 0.19$. Valor $P = 0.832$. No se rechaza la hipótesis nula de no interacción. Al parecer no hay un efecto significativo de la interacción entre el género y la carrera.
- Estadístico de prueba: $F = 0.78$. Valor $P = 0.395$. No se rechaza la hipótesis nula de que el género no tiene efecto alguno sobre la longitud estimada. No hay evidencia suficiente para sustentar la aseveración de que la longitud estimada esté afectada por el género.
- Estadístico de prueba: $F = 0.13$. Valor $P = 0.876$. No se rechaza la hipótesis nula de que la carrera no tiene un efecto sobre la longitud estimada. No existe evidencia suficiente para sustentar la aseveración de que la longitud estimada está afectada por la carrera.
- Para la interacción, el estadístico de prueba es $F = 0.8733$ y el valor P es 0.3685, de manera que no hay un efecto significativo de interacción. Para el género, el estadístico de prueba es $F = 0.0178$ y el valor P es 0.8960, de manera que no hay un efecto significativo del género. Para el tabaquismo, el estadístico de prueba es $F = 3.0119$ y el valor P es 0.1082, por lo tanto no hay un efecto significativo del tabaquismo.
- a. Estadístico de prueba: $F = 1.0000$. Valor $P = 0.4226$. No existe evidencia suficiente para sustentar la aseveración de que las cantidades de gases invernadero emitidos estén afectados por el tipo de transmisión.
- b. Estadístico de prueba: $F = 7.0000$. Valor $P = 0.1250$. No existe evidencia suficiente para sustentar la aseveración de que las cantidades de gases invernadero estén afectadas por el número de cilindros.
- c. Quizás los gases invernadero están afectados por el tipo de transmisión y/o el número de cilindros; sin embargo, los datos muestrales dados no proporcionan evidencia suficiente para sustentar esta afirmación.

8. Estadístico de prueba: $F = 3.242$. El valor crítico F está entre 3.0718 y 3.1504. (Con tecnología: valor $P = 0.0449$). Con un nivel de significancia de 0.05, se rechaza $H_0: \mu_1 = \mu_2 = \mu_3$. Existe evidencia suficiente para sustentar el rechazo de la aseveración de que los tres grupos tienen la misma longevidad media. Parece que los tiempos de supervivencia tienen medias que no son iguales.

Capítulo 12 Ejercicios de repaso acumulativo

1. a. 15.1, 13.1, 22.7
b. 9.5, 9.0, 18.6
c. Estadístico de prueba: $t = -1.471$. Valores críticos: $t = \pm 2.160$ (suponiendo nivel de significancia de 0.05). (Con tecnología: valor $P = 0.1613$). No se rechaza $H_0: \mu_1 = \mu_2$. No existe evidencia suficiente para sustentar la aseveración de que existe una diferencia entre la media de los dos grupos.
d. Normal, porque el histograma tiene aproximadamente forma de campana.
e. $11.9 \text{ años} < \mu < 18.2 \text{ años}$
2. a. 960.5, 980.0, 1045.0; no
b. 914.5, 1010.5, 1008.5; no
c. 174.6, 239.6, 224.1; no
d. Estadístico de prueba: $t = -0.294$. Valores críticos: $t = \pm 2.093$ (suponiendo nivel de significancia de 0.05). (Con tecnología: valor $P = 0.7704$). No se rechaza $H_0: \mu_1 = \mu_2$.
e. $878.8 < \mu < 1042.2$
f. Estadístico de prueba: $F = 0.8495$. Valor $P = 0.4330$. No se rechaza $H_0: \mu_1 = \mu_2 = \mu_3$. No existe evidencia suficiente para justificar el rechazo de la afirmación de que las tres poblaciones tienen la misma puntuación media en el SAT.
3. Utilizando la distribución normal como aproximación de la distribución binomial: 0.1020. (Resultado exacto utilizando la tecnología: 0.0995). Puesto que la probabilidad de obtener 19 o menos descendientes con ojos azules es tan alta, no hay evidencia suficiente para concluir que la proporción de una cuarta parte sea incorrecta.
4. a. 0.3372 (con tecnología: 0.3365)
b. 0.0455 (con tecnología: 0.0457)
c. $1/8$ o 0.125

Capítulo 13 Respuestas

Sección 13-2

1. La prueba del signo no cumple con el requisito de que los datos muestrales provengan de la población con una distribución específica.
3. Puesto que la proporción de niñas es menor que 0.5, no existe evidencia suficiente para sustentar la aseveración de que el método incrementa la probabilidad de que un bebé sea niña. Con una proporción de niñas menor que 0.5, no hay forma en que podamos sustentar la aseveración de que la proporción poblacional sea significativamente mayor que 0.5.
5. El estadístico de prueba de $x = 4$ es menor o igual al valor crítico de 4. Existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencia.
7. El estadístico de prueba de $z = -0.50$ no es menor o igual que el valor crítico de -1.96 . No existe evidencia suficiente para rechazar la aseveración de que no hay diferencia.

9. El estadístico de prueba de $x = 1$ no es menor o igual al valor crítico de 0. No hay evidencia suficiente para rechazar la aseveración de que no hay efecto alguno. Con base en los datos muestrales, parece que los viernes 13 el número de admisiones hospitalarias no se vea afectado.
11. El estadístico de prueba de $x = 5$ no es menor o igual al valor crítico de 2. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas reportadas y las estaturas medidas.
13. El estadístico de prueba de $x = 1$ es menor o igual al valor crítico de 2. Existe evidencia suficiente para sustentar la aseveración de que la población tiene una mediana menor que 98.6°F .
15. El estadístico de prueba de $z = -3.64$ es menor o igual que el valor crítico de $z = -2.33$. (Con tecnología: valor $P = 0.0001$). Existe evidencia suficiente para sustentar la aseveración de que, con el método YSORT, la probabilidad de un niño es mayor que 0.5. Parece que el método funciona.
17. El estadístico de prueba de $z = -0.48$ no es menor o igual que el valor crítico de $z = -2.33$. (Con tecnología: valor $P = 0.3157$). No existe evidencia suficiente para sustentar la aseveración de que, entre los usuarios de Internet, menos del 50% lo utilice para hacer planes de viaje. Los resultados son importantes para los agentes de viajes, porque su negocio se podría ver afectado.
19. Primer método: $z = -1.90$; se rechaza H_0 .
Segundo método: $z = -1.73$; se rechaza H_0 .
Tercer método: $z = 0$; no se rechaza H_0 .
21. Se convierte $x = 18$ en el estadístico de prueba $z = -2.31$. Valor crítico: $z = -2.33$. (Con tecnología: valor $P = 0.0104$). No existe evidencia suficiente para sustentar la aseveración de un sesgo por género. Si se utiliza la distribución binomial en lugar de la aproximación normal, el valor P es 0.0099, que es menor que 0.01; por lo tanto, existe evidencia suficiente para sustentar la aseveración de un sesgo por género. Si se utiliza la aproximación normal, el estadístico de prueba cae apenas a fuera de la región crítica; si se utiliza la distribución binomial, el estadístico de prueba cae apenas dentro de la región crítica.

Sección 13-3

1. La prueba de rangos con signo de Wilcoxon se debe utilizar cuando la población de diferencias tienen la distribución aproximadamente simétrica, y el número de datos muestrales apareados es pequeño ($n \leq 30$), y los pares de valores tienen diferencias que no provienen de una población con una distribución aproximadamente normal.
3. Al utilizar los rangos en lugar de los signos de las diferencias, la prueba de rangos con signo de Wilcoxon incluye información sobre la magnitud de las diferencias, y no sólo el signo de las diferencias. La prueba de rangos con signo de Wilcoxon utiliza más información y tiene más probabilidades de reflejar la verdadera naturaleza de los datos.
5. Estadístico de prueba: $T = 0$. Valor crítico: $T = 4$. Se rechaza la hipótesis nula de que la población de las diferencias tienen una mediana de 0.
7. Estadístico de prueba: $T = 1.5$. Valor crítico: $T = 1$. No se rechaza la hipótesis nula de que la población de las diferencias tiene una mediana de 0. Con base en los datos muestrales, parece que los viernes 13 el número de admisiones hospitalarias no se vea afectado.

9. Estadístico de prueba: $T = 34$. Valor crítico: $T = 14$. No se rechaza la hipótesis nula de que la población de las diferencias tiene una mediana de 0. No existe evidencia suficiente para sustentar la aseveración de que existe una diferencia entre las estaturas reportadas y las estaturas medidas de varones entre 12 y 16 años.
11. Estadístico de prueba: $T = 158.5$. Valor crítico: $T = 127$ (suponiendo un nivel de significancia de 0.05). No se rechaza la hipótesis nula de que la población de las diferencias tiene una mediana de 0. Parece que la diferencia entre las temperaturas máximas reales y pronosticadas no es significativa.
13. Se convierte $T = 661$ en el estadístico de prueba $z = -5.67$. (Con tecnología: valor $P = 0.0000$). Valores críticos: $z = \pm 1.96$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que los adultos saludables tienen una temperatura corporal media igual a 98.6°F .

Sección 13-4

1. La prueba de rangos con signo de Wilcoxon se aplica a datos muestrales apareados, mientras que la prueba de suma de rangos de Wilcoxon se aplica a dos muestras independientes que no están apareadas de ninguna forma.
3. A diferencia de la prueba paramétrica de la sección 9-3, la prueba no paramétrica de suma de rangos de Wilcoxon no requiere de poblaciones distribuidas normalmente, por lo que se puede utilizar en más situaciones.
5. $R_1 = 104$, $R_2 = 172$, $\mu_R = 132$, $\sigma_R = 16.248$, estadístico de prueba: $z = -1.72$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.0848$). No se rechaza la hipótesis nula de que las poblaciones tienen la misma mediana.
7. $\mu_R = 150$, $\sigma_R = 17.321$, $R = 96.5$, $z = -3.09$. Estadístico de prueba: $z = -3.09$. Valores críticos: $z = \pm 2.575$. (Con tecnología: valor $P = 0.0020$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las dos muestras provienen de poblaciones con la misma mediana. Con base en esos resultados, parece que el trastorno obsesivo-compulsivo tiene una base biológica.
9. $\mu_R = 648$, $\sigma_R = 46.4758$, $R = 616.5$, $z = -0.68$. Estadístico de prueba: $z = -0.68$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.4979$). No se rechaza la hipótesis nula de que las dos muestras provienen de poblaciones con la misma mediana. Los resultados no sugieren que los taxis sean más nuevos.
11. $\mu_R = 1620$, $\sigma_R = 103.923$, $R = 1339.5$, $z = -2.70$. Estadístico de prueba: $z = -2.70$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.0070$). Se rechaza la hipótesis nula de que las dos muestras provienen de poblaciones con la misma mediana.
13. $z = -0.98$; el estadístico de prueba es el mismo valor pero con el signo opuesto.

Sección 13-5

1. La prueba de Kruskal-Wallis no requiere que las poblaciones tengan distribuciones normales o cualquier otra distribución específica.
3. La puntuación de eficiencia de 0.95 indica que, si todos los demás factores permanecen constantes, la prueba de Kruskal-Wallis requiere de 100 observaciones muestrales para lograr los mismos resultados

que 95 observaciones con la prueba paramétrica del análisis de varianza de un factor, suponiendo que se cubren los requisitos más estrictos para utilizar la prueba paramétrica.

5. Estadístico de prueba: $H = 1.1914$. Valor crítico: $\chi^2 = 7.815$. (Con tecnología: valor $P = 0.754$). No existe evidencia suficiente para sustentar la aseveración de que las heridas en la cabeza para las cuatro categorías de peso tengan medidas con medianas diferentes. (Parece que las medianas son iguales). Los datos dados no proporcionan evidencia suficiente para concluir que los automóviles más pesados sean más seguros en un choque.
7. Estadístico de prueba: $H = 14.7485$. Valor crítico: $\chi^2 = 5.991$. (Con tecnología: valor $P = 0.0006$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las medianas de las lecturas de voltaje son iguales en los tres tipos de días diferentes. Parece que los días soleados dan como resultado mayores cantidades de energía.
9. Estadístico de prueba: $H = 6.0317$. Valor crítico: $\chi^2 = 7.815$. (Con tecnología: valor $P = 0.1101$). No existe evidencia suficiente para justificar el rechazo de la aseveración de que los diferentes grupos tienen la misma presión sanguínea mediana. Se puede considerar que los grupos son muestras de la misma población.
11. Estadístico de prueba: $H = 74.8519$. Valor crítico: $\chi^2 = 7.815$. (Con tecnología: valor $P = 0.000$). Existe evidencia suficiente para justificar el rechazo de la aseveración de que las cuatro categorías de centavos tienen la misma mediana de peso. Las máquinas contadora de monedas no pueden tratar los pesos de la misma manera.
13. 14.840 (utilizando $T = 6, 6, 24$); no

Sección 13-6

1. La correlación de rangos no requiere de distribuciones normales o de cualquier otro tipo de distribución específica.
3. El subíndice s se utiliza para poder distinguir el coeficiente de correlación de rangos del coeficiente de correlación lineal r . El subíndice no representa la desviación estándar s . Se utiliza para honrar a Charles Spearman, que creó el método de correlación de rangos.
5. $r_s = 1$ y al parecer existe una correlación entre x y y .
7. a. ± 0.587
b. ± 0.570
c. ± 0.255
d. ± 0.290
9. $r_s = 0.345$. Valores críticos: $r_s = \pm 0.648$. No hay correlación. Al parecer no existe una correlación entre el DJIA y el número de automóviles vendidos.
11. $r_s = 0.855$. Valores críticos: $r_s = \pm 0.648$. Sí hay una correlación. Al parecer existe una correlación entre el salario y el estrés.
13. $r_s = 0.857$. Valores críticos: $r_s = \pm 0.738$. Sí hay una correlación. Al parecer existe una correlación entre el número de chirridos por minuto y la temperatura.
15. $r_s = 0.561$. Valores críticos: $r_s = \pm 0.700$. No hay correlación. Al parecer no existe una correlación entre el número de impresiones de audiencia y el número de álbumes vendidos. Parece que las ventas no están muy afectadas por el número de impresiones de audiencia.

17. a. $r_s = 0.918$. Valores críticos: $r_s = \pm 0.368$. Hay una correlación entre el alquitrán y la nicotina.
 b. $r_s = 0.739$. Valores críticos: $r_s = \pm 0.368$. Hay una correlación entre el monóxido de carbono y la nicotina.
 c. El alquitrán es la mejor opción, ya que tiene una correlación más alta con la nicotina.
19. $r_s = 1$. Valores críticos: $r_s = \pm 0.618$ (suponiendo un nivel de significancia de 0.05). Existen a correlación lineal. Con el uso de la correlación lineal (sección 10-2), $r_s = 0.572$, los valores críticos son $r_s = \pm 0.602$ (suponiendo un nivel de significancia de 0.05), y no existe una correlación *lineal*. El diagrama de dispersión muestra un patrón, pero no es el patrón de una línea recta. Los resultados de los dos métodos son diferentes.

Sección 13-7

1. No, porque la mezcla cambió el orden en que los sujetos fueron entrevistados.
3. No. Podría haber problemas con el proceso de selección de datos. Por ejemplo, es probable que una muestra de respuesta voluntaria parezca aleatoria, pero no sería adecuada para la mayor parte de los análisis estadísticos.
5. $n_1 = 4$, $n_2 = 8$, $G = 8$, valores críticos: 3, 10; no se rechaza la aleatoriedad.
7. $n_1 = 6$, $n_2 = 6$, $G = 7$, valores críticos: 3, 11; no se rechaza la aleatoriedad.
9. $n_1 = 13$, $n_2 = 7$, $G = 11$, valores críticos: 5, 15; no se rechaza la aleatoriedad.
11. $n_1 = 14$, $n_2 = 19$, $G = 20$, valores críticos: 11, 23; no se rechaza la aleatoriedad. Parece que las ligas ganan en una secuencia aleatoria.
13. $n_1 = 8$, $n_2 = 8$, $G = 2$, valores críticos: 4, 14; se rechaza la aleatoriedad. Los resultados sugieren que existe una tendencia al alza.
15. $n_1 = 35$, $n_2 = 19$, $G = 23$, $\mu_G = 25.6296$, $\sigma_G = 3.3137$. Estadístico de prueba: $z = -0.79$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.4274$). No se rechaza la aleatoriedad. Parece que la secuencia de géneros es aleatoria.
17. $n_1 = 49$, $n_2 = 51$, $G = 43$, $\mu_G = 50.98$, $\sigma_G = 4.9727$. Estadístico de prueba: $z = -1.60$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.1085$). No se rechaza la aleatoriedad. Parece que los dígitos tienen una secuencia aleatoria.
19. El mínimo es 2, el máximo es 4. No se pueden obtener los valores críticos de 1 y 6, de manera que la hipótesis nula de aleatoriedad nunca podrá ser rechazada.

Capítulo 13 Conocimientos estadísticos y pensamiento crítico

1. Se trata de una prueba de hipótesis que no requiere que los datos muestrales tengan una distribución normal o cualquier otro tipo de distribución específica.
2. No hay diferencia. Son distintos nombres para la misma categoría de prueba de hipótesis que no requiere que las poblaciones tengan distribuciones normales o cualquier otro tipo de distribución específica.
3. Un rango de su número asignado a un elemento muestral individual, de acuerdo con su orden en la lista. Al primer elemento se le asigna un rango de 1, al segundo elemento se le asigna un rango de 2, y así sucesivamente.

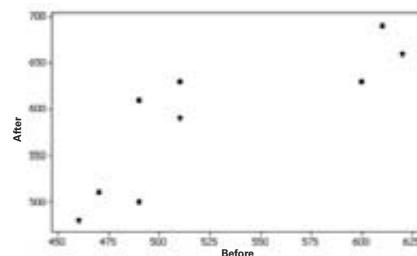
4. La eficiencia es una medida de lo fuerte que debe ser la evidencia muestral para que la prueba no paramétrica produzca los mismos resultados que una prueba paramétrica correspondiente. Por ejemplo, la prueba del signo tiene una eficiencia de 0.63, lo que significa que, bajo las mismas condiciones, la prueba del signo requiere de 100 observaciones muestrales para lograr los mismos resultados que 63 observaciones muestrales analizadas con una prueba paramétrica correspondiente.

Capítulo 13 Ejercicios de repaso

1. El estadístico de prueba $x = 2$ es menor o igual que el valor crítico de 2. Existe evidencia suficiente para justificar el rechazo de la aseveración de que no existe una diferencia entre los tiempos del primero y del segundo ensayo.
2. Estadístico de prueba: $T = 5.5$. Valor crítico: $T = 21$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que no existe una diferencia de los tiempos del primero y el segundo ensayo.
3. Correlación de rangos: $r_s = 0.709$. Valores críticos: $r_s = \pm 0.648$. Al parecer existe una correlación entre la clasificación de las escuelas de negocios y la clasificación de las escuelas de leyes.
4. Correlación de rangos: $r_s = -0.810$. Valores críticos: $r_s = \pm 0.738$. Parece que existe una correlación entre la clasificación y la calificación promedio en la prueba MCAT.
5. Prueba del signo: se convierte $x = 15$ en el estadístico de prueba $z = -1.42$. Valor crítico: $z = -1.645$. (Con tecnología: valor $P = 0.0774$). No existe evidencia suficiente para sustentar la aseveración de un sesgo a favor de los hombres.
6. Prueba de rangos con signo de Wilcoxon: $\mu_R = 162$, $\sigma_R = 19.442$, $R = 89.5$, $z = -3.73$. Estadístico de prueba: $z = -3.73$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.0002$). Se rechaza la hipótesis nula de que las dos muestras provienen de poblaciones con la misma mediana. Existe evidencia suficiente para justificar el rechazo de la aseveración de que los consumidores de cerveza y los consumidores de licor tienen los mismos niveles de CAS.
7. Prueba de rachas: $n_1 = 22$, $n_2 = 18$, $G = 18$, $\mu_G = 20.8$, $\sigma_G = 3.0894$. Estadístico de prueba: $z = -0.91$. Valores críticos: $z = \pm 1.96$. (Con tecnología: valor $P = 0.3648$). No se rechaza la aleatoriedad. Parece que los dígitos pares y los dígitos nones ocurren de manera aleatoria.
8. Prueba de Kruskal-Wallis: estadístico de prueba: $H = 4.234$. Valor crítico: $\chi^2 = 7.815$. (Con tecnología: valor $P = 0.237$). No existe evidencia suficiente para sustentar la aseveración de que las mediciones de heridas no son iguales en las cuatro categorías. No existe evidencia suficiente para sustentar la aseveración de que los automóviles más pesados son más seguros.

Capítulo 13 Ejercicios de repaso acumulativo

1. $\bar{x} = 528.9$; mediana = 510.0; rango = 160.0; $s = 63.1$; $s^2 = 3986.1$.
2. Parece que sí hay una correlación.



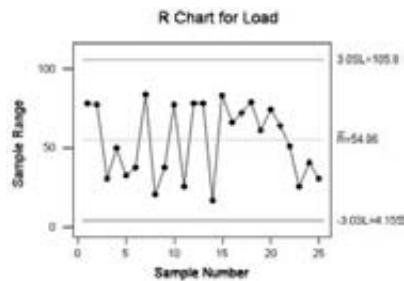
3. $r = 0.833$. Valores críticos: $r = \pm 0.666$ (suponiendo un nivel de significancia de 0.05). Existe una correlación lineal, pero eso no necesariamente significa que el curso sea efectivo.
4. $r_s = 0.907$. Valores críticos: $r_s = \pm 0.700$ (suponiendo un nivel de significancia de 0.05). Existe una correlación.
5. $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$. Estadístico de prueba: $t = -4.334$. Valores críticos: $t = \pm 2.306$ (suponiendo un nivel de significancia de 0.05). (Con tecnología: valor $P = 0.0025$). Se rechaza H_0 . Existe evidencia suficiente para justificar el rechazo de la aseveración de que la diferencia media entre las calificaciones antes y después del curso es igual a cero.
6. El estadístico de prueba $x = 0$ es menor o igual que el valor crítico de 1 (suponiendo un nivel de significancia de 0.05). Hay suficiente evidencia para rechazar la aseveración de que no existe diferencia entre las calificaciones antes y después.
7. $T = 0$. Valor crítico: $T = 6$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para rechazar la aseveración de ausencia de diferencia entre las calificaciones antes y después del curso.
8. Como las diferencias parecen provenir de una población con distribución normal, la prueba t en el ejercicio 5 es muy útil. El curso de preparación parece ser efectivo, ya que las calificaciones después de tomarlo son significativamente más altas que las calificaciones antes de tomarlo.

Capítulo 14 Respuestas

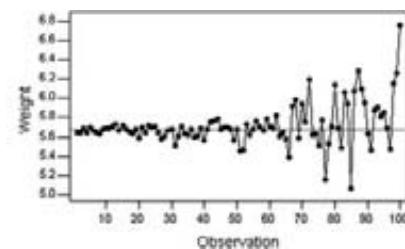
Sección 14-2

1. Los *datos de proceso* son aquellas que se ordenan de acuerdo con una secuencia de tiempo. Son mediciones de una característica de bienes o servicios, que resultan de cierta combinación de equipo, personas, materiales, métodos y condiciones.
3. Una gráfica de control de una característica de proceso (como la media o la variación) consiste en valores que se grafican de manera secuencial con el paso del tiempo, e incluye una línea central, así como un límite inferior de control y un límite superior de control.
5. a. Bajo control estadístico
b. No aplicable
c. La variación es demasiado grande, de manera que algunas latas están demasiado llenas mientras otros están demasiado vacías.
7. a. Fuera de control estadístico.
b. Hay un valor excepcionalmente bajo.
c. La cantidad baja de 10.5 onzas es inaceptable, por lo que el proceso no se está comportando como debe.

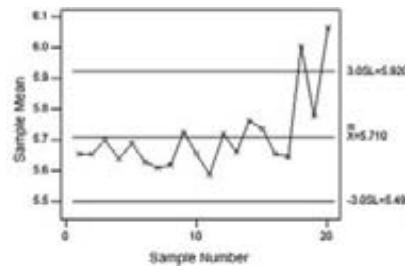
9. Parece que la variación del proceso está bajo control estadístico.



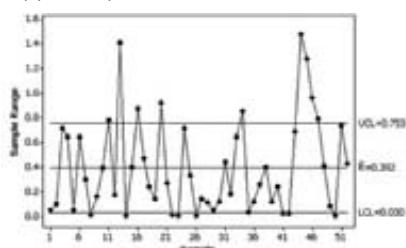
11. Existe un patrón de variación creciente, por lo que el proceso está fuera de control estadístico. La variación creciente producida cada vez más defectos.



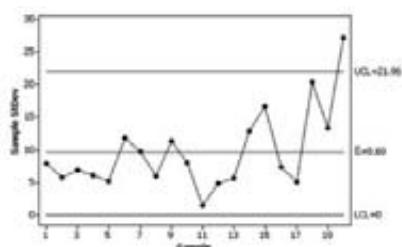
13. Hay un patrón de variación creciente, hay puntos que se ubican fuera del límite superior de control, y existen ocho puntos consecutivos que están por debajo de la línea central, de manera que la media del proceso está fuera de control estadístico. Este proceso necesita acción correctiva.



15. Parece que la variación del proceso está fuera de control estadístico. Hay puntos que se salen de los límites de control.

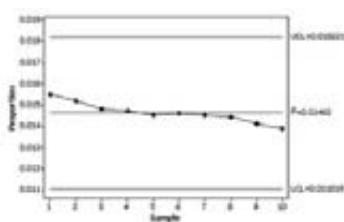


17. Parece que la variación del proceso está fuera de control estadístico. Hay un punto que se sale del límite superior de control, y parece que existe una tendencia ascendente.

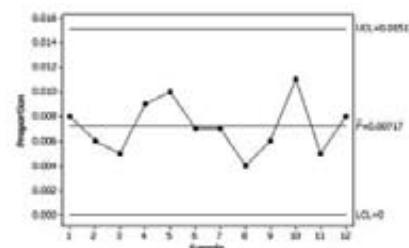


Sección 14-3

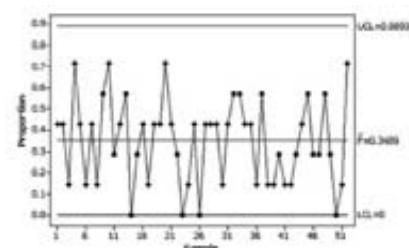
1. Es una gráfica de control de la proporción de algún atributo, y se utiliza para supervisar la proporción y determinar si el proceso está bajo control estadístico.
3. 0.250 y 0
5. Bajo control estadístico.
7. Al parecer el proceso está fuera de control estadístico porque existe el patrón de una tendencia ascendente y hay un punto que está fuera del límite superior de control.
9. El proceso está fuera de control estadístico debido a un patrón descendente. Posibles explicaciones: ha habido una disminución sustancial en la tasa de nacimientos entre las adolescentes; el incremento del periodo de vida produce una proporción menor de mujeres en edad reproductiva.



11. El proceso está bajo control estadístico. Aunque parece que no es indispensable una acción correctiva, la empresa debe trabajar para alcanzar la meta de cero defectos.



13. El proceso es estadísticamente estable.

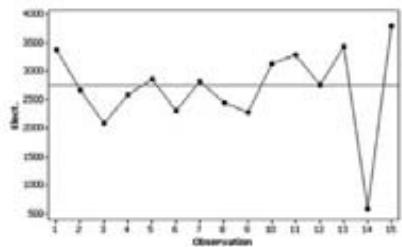


Capítulo 14 Conocimientos estadísticos y pensamiento crítico

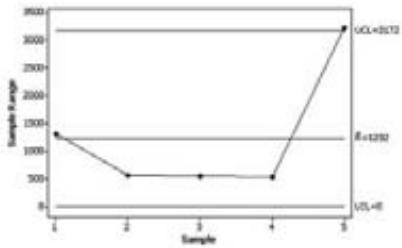
1. Si los datos cambian con el tiempo, características importantes como la media, la desviación estándar o la proporción de defectos no son estables y sus valores cambian. Si es importante controlar una o más de las características, entonces es necesario vigilar los datos.
2. El control estadístico de procesos consiste en métodos que se utilizan para vigilar los datos al paso del tiempo y para determinar si estos están bajo control estadístico, es decir, que los datos no cubren cualquiera de los criterios específicos que definen que están fuera de control.
3. Si el proceso de embotellamiento se sale de control estadístico, sus costos podrían aumentar debido a un creciente número de defectos, y los mayores costos podrían obligar el cierre de la planta por la pérdida de utilidades.
4. El proceso podría salirse de control estadístico debido a una media cambiante, a una mayor variación o a ambos.

Capítulo 14 Ejercicios de repaso

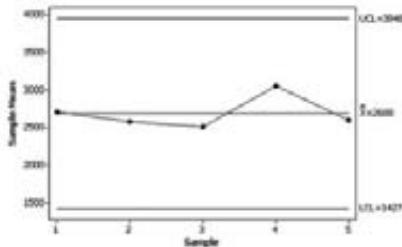
1. El proceso no está bajo control estadístico, ya que existe un punto excepcionalmente bajo.



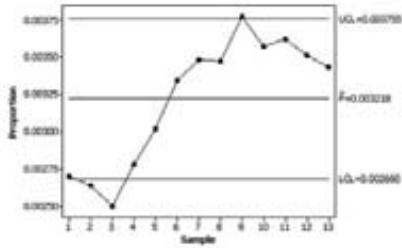
2. Debido a que existe un punto que rebasa el límite superior de control, la variación del proceso está fuera de control estadístico.



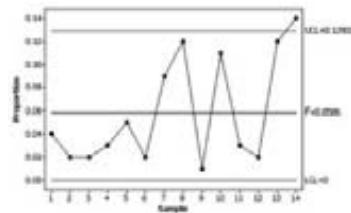
3. La media del proceso está bajo control estadístico.



4. El proceso está fuera de control porque hay un cambio hacia arriba y hay puntos que están fuera de los límites de control.

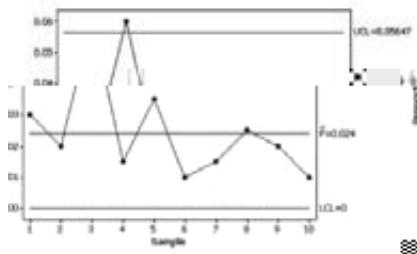


5. El proceso está fuera de control estadístico porque hay un punto que cae fuera de los límites de control.



Capítulo 14 Ejercicios de repaso acumulativo

1. El proceso está fuera de control estadístico porque hay un punto que rebasa el límite superior de control.



2. $0.0173 < p < 0.0307$

3. $H_0: p = 0.01$. $H_1: p > 0.01$. Estadístico de prueba: $z = 6.29$. Valor crítico: $z = 1.645$. Valor $P: 0.0000$. Se rechaza H_0 . Existe evidencia suficiente para sustentar la aseveración de que la tasa de defectos es mayor que el 1%.

4. a. 1/256
b. 1/256
c. 1/128

Créditos

Fotografías

Al inicio de capítulos

- Capítulo 1: Seis grados de Kevin Bacon © AFP/Getty Images
Capítulo 2: Premios de la Academia: Edad de las mejores actrices © AP Photo
Capítulo 3: Premios de la Academia: Mejores actores © Getty Images
Capítulo 4: Prueba de drogas al solicitar empleo © Corbis
Capítulo 5: Selección de miembros del jurado: Benjamin Spock © Digital Vision
Capítulo 6: Seguridad de los pasajeros © Thinkstock
Capítulo 7: Terapia de contacto © Beth Anderson
Capítulo 8: Contactos personales para conseguir empleo © PhotoDisc Red
Capítulo 9: Tratamiento para el síndrome del túnel carpiano © Stockbyte Platinum
Capítulo 10: Géiser Old Faithful © Digital Vision
Capítulo 11: Detección de fraudes en cheques © Corbis
Capítulo 12: Álamos © Corbis
Capítulo 13: Clasificación de universidades © Comstock
Capítulo 14: Altímetros de aviación © Digital Vision
Capítulo 15: © Medio Images

Las siguientes fotografías al margen tienen derechos reservados por PhotoDisc, Inc.

- Capítulo 1
p. 5, El estado de la estadística
p. 13, Entrevista de *Literary Digest*
p. 17, Sesgo de una publicación
- Capítulo 3
p. 79, Paradoja del tamaño de la clase
p. 94, Confiabilidad y validez
p. 99, Más acciones, menos riesgo
- Capítulo 4
p. 140, Apuestas
p. 141, ¿Qué tan probable?
p. 151, Los niños y las niñas no tienen la misma probabilidad
p. 171, Muestreo compuesto
p. 174, Probabilidad de un evento que jamás ha ocurrido
p. 180, Saturación de números
p. 181, Ganar la lotería
p. 183, ¿Cuántas veces hay que barajar?
p. 184, La secretaria aleatoria
p. 185, Muy pocos códigos de barras
- Capítulo 5
p. 202, Elección de números de lotería
p. 215, ¿Hay alguien en casa?
- Capítulo 6
p. 271, Predicción de números de lotería ganadores
p. 293, ¿Es seguro el paracaísmo?
p. 303, Los estados y las loterías locales
- Capítulo 7
p. 339, Estimación del tamaño de las poblaciones silvestres
- Capítulo 8
p. 388, Detectores de mentiras
p. 392, Un tamaño grande de una muestra no es suficiente
p. 419, Comerciales

- Capítulo 9
p. 458, El margen de error de quien lleva la delantera
p. 460, ¿La aspirina ayuda a prevenir los ataques cardíacos?
p. 473, Uso de la estadística para identificar ladrones
p. 487, ¿Las bolsas de aire salvan vidas?
p. 497, Menor variación; alta calidad
p. 499, Investigación sobre gemelos
- Capítulo 10
p. 522, Los alumnos califican a los maestros
p. 543, Posposición de la muerte
p. 567, Predicción de éxito
p. 568, Componer música con regresión múltiple
- Capítulo 11
p. 609, Un falso positivo durante ocho años
p. 612, Ventaja de los equipos locales
- Capítulo 12
p. 656, Encuestas y psicólogos
- Capítulo 13
p. 690, Diferencias de géneros en pruebas de consumo de drogas
p. 711, Vínculo directo entre tabaquismo y cáncer
- Fotografías diversas
- Capítulo 1
p. 26, Estudio *Prospective National Children*; derechos reservados por PhotoDisc Red
p. 29, ¿Debe creerse en un estudio estadístico?; derechos reservados por Jeff Greenberg/PhotoEdit
- Capítulo 2
p. 47, Gráficas de crecimiento actualizadas; derechos reservados por Kelly/Mooney Photography/Corbis
p. 52, Datos faltantes; derechos reservados por Brand X Pictures
p. 59, El poder de una gráfica; derechos reservados por Stockbyte Gold
- Capítulo 3
p. 96, ¿Dónde están los bateadores de 0.400?; derechos reservados por AP/Wide World Photos
p. 121, Una propina de valor extremo; derechos reservados por Reuters NewMedia, Inc./Corbis
p. 126, Detectores de mentiras; derechos reservados por Getty/Image Bank
p. 145, El cumpleaños más común: 5 de octubre; derechos reservados por Dora Kingsley
- Capítulo 4
p. 164, Calificaciones perfectas del SAT; derechos reservados por Taxi/Getty Images
- Capítulo 6
p. 261, Pruebas clínicas que arrojan conclusiones antes de lo previsto; derechos reservados por Seits/Photo Researchers, Inc.
p. 270, ¿Predominan los niños o las niñas en una familia?; derechos reservados por Anthony Redpath/Corbis
- Capítulo 7
p. 326, Encuestas falseadas; derechos reservados por PhotoDisc Red
p. 356, Estimaciones para mejorar los censos; derechos reservados por Corbis
p. 357, Estimaciones del tamaño de una multitud; derechos reservados por Beth Anderson

Capítulo 8
 p. 393, Gane \$1,000,000 por sus poderes extrasensoriales; derechos reservados por PhotoDisc Blue
 p. 408, Proceso de aprobación de un fármaco; derechos reservados por Digital Vision
 p. 428, ¿Los zurdos mueren más pronto?; derechos reservados por PhotoDisc Blue
 p. 430, Meta-análisis; derechos reservados por Beth Anderson

Capítulo 9
 p. 459, El autor como testigo; derechos reservados por Bill Wisser/Getty Images
 p. 474, Píldoras dietéticas caras; derechos reservados por lance Manning/Corbis
 p. 476, Súper Bowl; derechos reservados por Comstock

Capítulo 10
 p. 545, Predicciones de precios de condominios; derechos reservados por Comstock
 p. 558, El Súper Bowl como factor de predicción de la Bolsa de Valores; derechos reservados por PhotoDisc Red
 p. 570, Salarios de la NBA y desempeño; derechos reservados por Marc Aspin/Corbis Saba
 p. 572, Congelar al jugador; derechos reservados por PhotoDisc Red

Capítulo 12
 p. 643, Ética en los reportajes; derechos reservados por AP/Wide World

Capítulo 13
 p. 719, Rachas de éxito en los deportes; derechos reservados por 2002 NBAE, Fotografía de Noren Trotman/NBAE/Getty Images

Capítulo 14
 p. 739, Variación assignable costosa; derechos reservados por NASA
 p. 740, Los muestreos continuos pueden ser nocivos; derechos reservados por Davis Lees/Corbis
 p. 743, Gráficas de control para detectar sobornos; derechos reservados por Owen Franken/Corbis

Ilustraciones

Las siguientes ilustraciones al margen fueron proporcionadas por Bob Giuliani.

Capítulo 1
 p. 16, Detección de datos falsos
 p. 24, Los efectos del experimentador

Capítulo 2
 p. 44, Autores identificados
 p. 63, Florence Nightingale

Capítulo 3
 p. 110, El Índice del Costo de la Risa

Capítulo 4
 p. 139, Probabilidades que cambian la intuición
 p. 152, Vocabulario de Shakespeare
 p. 177, El mono mecanógrafo

Capítulo 6
 p. 262, Encuestas sobre temas delicados

Capítulo 7
 p. 340, Los números de serie de los tanques capturados revelan el tamaño de la población
 p. 351, Extractos de una circular del Departamento del Transporte
 p. 355, Estimación del azúcar de las naranjas

Capítulo 8
 p. 427, Ética en los experimentos

Capítulo 9
 p. 462, La pena de muerte como método de disuasión
 p. 486, Crest y las muestras dependientes

Capítulo 12
 p. 642, Resistencia a las encuestas

Capítulo 14
 p. 736, El efecto Flynn
 p. 749, Six Sigma en la industria

Las siguientes ilustraciones al margen fueron proporcionadas por James Bryant.

Capítulo 1
 p. 7, Medición de la desobediencia

Capítulo 3
 p. 119, Minería de datos

Capítulo 4
 p. 160, Motores independientes en las aeronaves
 p. 163, Convicto por probabilidad
 p. 170, ¿Coincidencias?
 p. 182, Selección de códigos personales de seguridad

Capítulo 6
 p. 286, El teorema del límite central poco claro

Capítulo 8
 p. 439, Preguntas tendenciosas en las encuestas

Capítulo 9
 p. 461, Experimento sobre poliomielitis

Capítulo 10
 p. 519, Lectura de las palmas de las manos

Capítulo 11
 p. 613, El encuestador puede afectar los resultados

Capítulo 13
 p. 680, Asistencia a clase y calificaciones

Capítulo 14
 p. 748, Control de calidad en Perstop

Las siguientes ilustraciones fueron proporcionadas por Darwin Hennings.

Capítulo 5
 p. 231, Filas de espera

Capítulo 6
 p. 285, El efecto placebo

Capítulo 9
 p. 477, Mejores resultados con clases de menor tamaño

Capítulo 10
 p. 577, Estadística: Puestos de trabajo y empleadores

Figura 2.9
 p. 63, Derechos reservados por 2003 Consumers Union of U.S., Inc. Yonkers, NY 10703-1057, organización no lucrativa. Reimpresa con permiso sólo para fines educativos, Consumer Reports®, abril de 2003. No se permite su uso o reproducción con fines comerciales. <<http://www.consumerreports.org/>>

Ilustración:

p. 64, Bajas de soldados en el ejército de Napoleón durante la campaña en Rusia (1812-1813); Edward R. Tufte, *The Visual Display of Quantitative Information* (Cheshire, CT: Graphics Press, 1983). Reimpresa con permiso

Índice

A

Accidentes en vehículos automotores, 476
Aeronaves
equipo de navegación, 351
problema
de asientos, 315
de peso, 245, 293-294
de producción de altímetros, 733, 735-736, 740-741,
742-744, 749-750
Aleatoriedad, prueba de rachas para, 717-722
Análisis
de varianza (ANOVA), 634-672. *Véase también* Varianza
CM (tratamiento), 645
de dos factores, 655-660
de dos muestras, 636
de un factor, 637-649
definición, 636
diseño del experimento, 645-646
distribución F , 636-637, 640-643
grados de libertad, 642
interacción, 655-657
SC (total), 644-645
tamaño de la muestra, 641-646
tratamiento (factor), 638
exploratorio de datos (AED), 119-126
Anchura de clase, 43
Apuestas en el hipódromo, 142
Asistencia a clases, 680
Aspirina, 460
Atributos, gráficas de control, 748-750

B

Bloques experimentales, 24
Burger, Joanna, 383

C

Calculadora. *Véase* Resultados de programas
de cómputo/calculadoras
Calidad de la caja de velocidades, 497
Calificaciones
de la prueba SAT, 567
de un curso, 7-8
Cambio marginal, 547-548
Cáncer pulmonar, 711
Carvalho, Barbara, 243
Castaneda vs. Partida, 199, 201, 206, 208, 215-217, 219-220,
225-227
Causalidad vs. correlación, 17, 525
Censo, 45, 326, 356
Centro, 42, 76-86
definición, 77
media, 77-78, 83-84, 85
mediana, 78-79, 85
mejor medida, 84

moda, 80, 85
rango medio, 80-81, 85
Centroide, 530-531
Chiang-Stein, Christy, 197
Children's Defense Fund, 322
Clase
anchura de, 43
límites de, 43, 44
punto medio, 43
CM (tratamiento), 645
Cociente de posibilidades, 150-151
Códigos
de barras, 185
personales de seguridad, 182
Coeficiente
de correlación
de rangos de Spearman, 708
lineal, 518, 520-524
producto-momento de Pearson, 518, 520-524
de determinación, 559
ajustado de Wald, 331
múltiple, 568-569
de variación, 103-104
Coincidencia, 170
Comerciales, 419
Complemento, 144
Comportamiento de votación, 174, 459
Confiabilidad, 94
Confusión, 23
del inverso, 171
Consumer's Report, 63
Conteo, 179-186
regla
de las combinaciones, 185
de las permutaciones, 182-184, 185
factorial, 181-182
fundamental de conteo, 180
Contraseñas, 182
Control de calidad. *Véase* Gráficas de control
Corrección
por población finita, 286-287
por continuidad, 294-297
Correlación, 517-532
centroide, 530-531
coeficiente de correlación lineal, 518, 520-524
definición, 517
diagramas de dispersión, 518, 519, 530
errores, 525
orden de rangos, 708-713
prueba formal de hipótesis, 525-532
variación explicada, 524-525
vs. causalidad, 17, 525
vs. linealidad, 525
Crema dental, 486

- Cuadrado
 medio (CM), 645
 de raíz, 91
- Cuartiles, 111-115
- Curva de densidad (función de densidad de probabilidad), 248-249
- CVDVT, 42, 65, 763
- D
- Datos
 apareados, 469, 517, 615
 diseño experimental, 488-489
 inferencias, 484-489
 intervalo de confianza, 458
 prueba(s) y, 485-488, 621-625, 680-682, 689-693
- bimodales, 80
- bivariados, 517
- características, 42
- centrales, 42, 76-86. *Véase también* Centro
- confiabilidad, 94
- continuos (numéricos), 6-7
- cualitativos (categóricos, de atributos), 6
- cuantitativos, 6
- de intervalo, 8, 9
- de proceso, 734-735
- de razón, 8-9
- de tiempo, 42
- definición, 45
- discretos, 6-7
- distribución, 42
- falsos, 16
- faltantes, 16, 52
- intervalo de, 8, 9
- multimodales, 80
- nominales, 7, 9, 682-683
 prueba del signo de, 682-683
- normales, 100-101
- numéricos, 6-7
- ordinales, 7-8, 9
- prueba de rachas, 717-722
- rango, 93
- sesgados, 85-86
- simétricos, 85
- tipos de, 5-9
- validez de los, 94
- valor extremo, 42, 119-120, 121
- variación, 42
- visualización, 56-65
- Decimales, conversión desde porcentajes, 15
- Desviación. *Véase* Desviación estándar; variación
- estándar, 204-205
 cálculo de, 94-96
 comparación de dos muestras, 495-501
 definición, 94
 desviación media absoluta, 102-103
 distribución binomial, 225-227
 estimación de, 98-99, 102-103
 fundamentos, 101-103
 interpretación de, 98-101
 intervalo de confianza, 440
- población, 96
 prueba de una aseveración, 436-440, 445
- regla empírica, 101-102
 regla práctica del intervalo, 98-99
 tamaño de la muestra, 344-345
 teorema de Chebyshev, 101
- explicada, 557-559
 media absoluta, 102-103
 no explicada, 557-559
 total, 558
- Detectores de mentiras, 388
- Diagrama
 de árbol, 160
 de dispersión, 60-61, 518, 519, 530
 cuadrantes, 530
 valor extremo/punto de influencia, 458
 de Venn, 154-155
- Diseño experimental rigurosamente controlado, 646
- Dispersión. *Véase* Variación
- Distorsiones, 17
- Distribución
 binomial, 213-220
 aproximación de la distribución normal, 291-298
 aproximación de Poisson, 232-233
 definición, 214
 fórmula de probabilidad binomial, 216-217, 219-220
 media, varianza y desviación estándar, 225-227
 notación, 214-215
 tecnología, 218-219, 220
 chi cuadrada, 364-367, 590
 prueba de hipótesis, 437-439, 445, 592-600
 valor crítico, 365-367
- de frecuencias
 acumulativas, 46
 clases y, 43-44
 construcción de, 44-45
 definición, 43
 interpretación de la, 46-48
 media, 83-84
 relativa, 45-46
- de Poisson, 230-233
- de probabilidad, 198-242. *Véase también* Distribución normal;
 Distribución normal estándar
 binomial, 213-220
 definición, 201
 desviación estándar, 204-2 105
 distribución de Poisson, 230-233
 histograma de probabilidad, 203-204
 media, 204-205
 requisitos, 203-204
 resultados infrecuentes, 205-208, 236
 valor esperado, 208-209
 variables aleatorias, 201-209
 varianza, 204-205
 F , 495-498, 636-637, 640-643
 hipergeométrica, 224
 lognormal, 306
 muestral, 269-278
 estimadores, 276-277
 media, 274-275

- proporción, 270-274
- teoría del límite central, 280-287
- multinomial, 224
- normal, 46-48, 100-101, 245-316. *Véase también* Distribución normal
 - normal estándar, 247-257
 - aplicaciones, 259-265
 - aproximación de la distribución binomial, 291-298
 - bivariada, 520
 - cálculo de la, 250-254-255, 262-266
 - curva de densidad, 249-250
 - definición, 246, 249
 - distribución de muestras, 269-278
 - evaluación, 302-306
 - gráfica cuantil normal, 302-306
 - histograma, 53
 - no estándar, 260
 - teorema del límite central, 280-287
 - transformación de datos, 306-307
 - t* de Student, 350-351, 353-354, 427
 - uniforme, 247-248
 - valor *P*, 429-430
 - vs. distribución *z*, 354-357
- Documentos federalistas*, 44
- Duragesic, 586

- E
- Efecto
 - Flynn, 736
 - Hawthorne, 24
 - placebo, 24, 285
 - Rosenthal, 24
- Ejército de Napoleón, 63, 64
- Encuesta(s)
 - autoselecciónadas, 12-13
 - de correo electrónico, 329-330
 - de la “cámara-policía”, 381
 - de *Literary Digest*, 13
 - devueltas, 215
 - efectos por género en las, 612-613
 - ética y, 389
 - Gallup, 21
 - medios por los que se realizan las, 613
 - muestras de respuesta voluntaria, 12-13
 - preelección, 458
 - proyecto, 759-760
 - redacción de preguntas, 656
 - resistencia a las, 642
 - sensibilidad del encuestado, 262
 - sobre clonación, 141-143
 - sobre el correo electrónico, 329-330
- Error(es)
 - control de, 399-400
 - de muestreo, 30-31
 - estándar
 - de estimación, 560-561
 - de la media, 282
 - no de muestreo, 30
 - notación, 399
 - tipo I, 398-400
 - tipo II, 398-400
- Estadística, 4, 5
 - confusa, 11-17, 29, 322, 392
 - descriptiva, 76
 - empleos en, 577
 - en educación, 766
 - indiferenciado, 76
 - periodismo, 113
- Estadístico de prueba, 391-392, 703
 - ANOVA de un factor, 640
 - ANOVA, 645
 - correlación de rangos, 710
 - correlación lineal, 527-529
 - datos apareados, 485
 - desviación estándar/varianza, 392, 436
 - dos medias, 470, 474-475
 - dos proporciones, 457
 - dos varianzas, 496
 - media, 391-392, 419, 426, 470, 474-475
 - nivel de significancia, 392-393
 - proporción, 391-392, 408, 413
 - prueba de homogeneidad, 611-613
 - prueba de independencia, 607-608
 - prueba de Kruskal-Wallis, 703
 - prueba de la suma de rangos de Wilcoxon, 696-697
 - prueba de la suma de rangos ordenados de Wilcoxon, 690
 - prueba de McNemar, 622
 - prueba de rachas para detectar aleatoriedad, 719
 - prueba del signo, 680
 - pruebas de bondad de ajuste, 593-594, 599-600
 - región crítica, 392, 393
 - valor crítico, 393-394
 - valor *P*, 394, 396
- Estimación, 318-382
 - desviación estándar, 98-99
 - distribución
 - chi cuadrada, 364-367
 - muestral, 276-277
 - t* de Student, 350, 353-357
 - estimado puntual, 321, 330-331, 338-339, 357, 367
 - grados de libertad, 351, 364
 - intervalos de confianza, 322-323. *Véase también* Intervalos de confianza (estimación del intervalo)
 - margen de error, 325-326. *Véase también* Margen de error
 - media, 338-359
 - de muestra, 338-339, 350
 - probabilidad, 140-141
 - proporción, 320-332
 - proporción de muestra, 321, 325
 - tamaño de la muestra, 328-331, 338, 342-345, 370-371
 - valor crítico, 324-325. *Véase también* Valor crítico
 - varianza, 363-371
 - de la muestra, 367-370
- Estimado
 - agrupado, 476-477
 - del intervalo. *Véase* Intervalo de confianza (estimado del intervalo)
 - puntual, 321, 330-331, 357, 367
- Estimador, 97, 339, 367
 - sin sesgo, 97, 339, 367

- Estudio
 ciego, 23-24
 de cáncer y tabaquismo, 711
 de la anemia falciforme, 261
 de la bolsa de aire, 487
 de la hidroxiurea, 261
 de la polio, 23-24, 26, 461
 de la vacuna de Salk, 23-24, 26, 461
 de Lipitor, 59
 de los osos, 719-720
 de posposición de la muerte, 543
 doble ciego, 24
 observacional, 21, 22
 vs. ensayo clínico, 23
 para el propio beneficio, 17
 prospectivo (longitudinal o cohorte), 22
 retrospectivo (caso-control), 22
 sobre desobediencia, 7
 sobre mentiras, 126
 transversal, 22
- Ética, 427
- Evaluaciones de maestros, 522
- Exactitud, 17
- Excel. Véase Resultados de programas de cómputo/calculadora
- Éxitos, 458
- Expansión. Véase Variación
- Experimento(s), 21-31
 bloque, 24
 aleatorizado, 24, 25
 ciego, 23-24
 completamente aleatorizado, 25, 646
 confusos, 23
 definición, 21
 diseño de, 21-31, 488-489
 errores de muestreo, 30-31
 estrategias de muestreo, 26-30
 ética, 427
 multinomiales
 bondad de ajuste, 591-600
 definición, 591
 frecuencias esperadas, 592-598
 valor P , 594, 598-600
 potencia de, 402-403
 réplica de, 25-26
 rigurosamente controlado, 25
 tamaño de la muestra, 25-26
 tipos de, 21-22
 totalmente aleatorizado, 25, 646
 variables, 23-25
- F
 Fabricación de discos compactos, 568
 Falsificación de datos, 326
 Falso negativo, 137
 Falso positivo, 137, 609
 Falta de respuesta, 16
 Fluoruro, 486
 Fórmula de probabilidad binomial, 216-217, 219-220
 Foy, Jeffrey, 673
 Fracción, conversión desde porcentajes, 15
- Frecuencia, 43
 esperada, 592-598, 608-610
 relativa, 45
- G
 Gillespie, Angela, 731
 Grados de libertad, 351-352, 470, 642
 definición, 351
 distribución chi cuadrada, 364, 437
 Gráfica(s), 56-65
 circular, 59-60
 confusas, 13-14
 cuantil normal, 302-306
 de control, 732-756
 atributos, 748-750
 datos de proceso, 734-735
 de crecimiento, 47
 de Pareto, 59, 60
 de puntos(diagrama de cuadro y bigotes) 58, 120-122
 modificadas, 123-126
 de rachas, 735-738
 de rango, 738-741
 de series de tiempo, 61-62
 de tallo y hojas, 58-59
 diagrama de dispersión, 60-61, 518, 519, 530
 diseño de, 62-65
 gráficas de cuadro, 120-126
 gráficas residuales, 550-552
 histograma, 51-54, 203-204
 de frecuencias relativas, 52
 de probabilidad, 203-204
 interpretación de, 741-742
 límites superiores/inferiores de control, 739, 741-742
 línea central, 739
 media, 743-744
 ojiva, 57-58
 p , 748-750
 polígono de frecuencias, 57
 propósitos, 65
 R, 738-741
 residuales, 550-552
 series de tiempo, 61-62
 variación, 738-741
 \bar{x} , 743-744
- Graunt, John, 5
- Guerra de Crimea, 63, 65
- H
 Haskell, Mark D., 587
 Hipótesis. Véase también Prueba de hipótesis
 alternativa, 390-391
 de investigación, 390
 definición, 386
 nula, 389-391, 394-400
 aceptar/no rechazar, 396-397
 números, precisos, 17
 Histograma, 51-54, 203-204
 de frecuencia relativa, 52
 de probabilidad, 203-204
 Holzman, Robert S., 135

- I**
- Índice
 del Costo de la Risa, 110
 Inferencias, 763-765
 a partir de dos muestras, 454-512
 datos apareados, 484-489
 desviación estándar/varianza, 495-501
 distribución F , 495-498
 intervalo de confianza, 458, 460-463, 470-471,
 474-475, 488
 medias, 469-478
 proporciones, 456-463
 prueba de hipótesis e, 458-460, 472-474, 475-478,
 485-487
 Informe oral, 758
 Interacción, 655-657
 Intercepto y , 542-544
 Intervalo(s)
 de confianza de la puntuación de Wilson, 331
 de predicción, 560-562
 estimado del
 comparación de datos, 323, 359
 datos apareados, 485
 definición, 322
 desviación estándar, 440
 dos proporciones, 458
 estimado puntual, 330-331, 357
 fundamentos, 342, 370
 interpretación, 322-323
 intervalos de predicción, 560-562
 límites, 340
 margen de error, 326-328, 340, 351-352, 357
 media (desviación estándar desconocida), 352, 430-431
 media, 339-342, 351-353, 421-422, 470-471, 474-476,
 477
 método “bootstrap”, 380-381
 proporción, 322-323, 326-328, 410-413, 460-463
 regla del redondeo, 326, 341
 traslape de, 359, 368-369
 varianza, 367-370
 Inverso, confusión de, 171
 Investigación de gemelos, 499
- J**
- Jugador de fútbol americano, congelar a un, 572
- L**
- Lebbos, Nabil, 633
 Ley de los números grandes, 141-144
 Límites
 inferiores
 de clase, 43
 de control, 793, 741-742
 Línea central, 739
 Linealidad versus correlación, 525
 Lycett, Mark T., 513
- M**
- Madison, James, 44
 Mano dominante, 428
- Margen de error, 458
 definición, 325-326
 encuestas, 458
 media, 339-342, 352, 357
 proporción, 325-326, 328
 Mars Climate Orbiter, 739
 Media, 85, 204-205, 338-359. Véase también Estimación;
 Desviación estándar
 ajustada, 90
 aleatoriedad por arriba o por debajo, 721-722
 aritmética. Véase Media
 armónica, 91
 cuadrática, 91
 definición, 77
 desviación estándar desconocida, 349-359, 426-431
 de población. Véase Media
 de tendencia central, 77
 distribución
 binomial, 225-227
 de frecuencias, 83-84
 muestral, 274-275
 error estándar, 282
 estimado puntual, 350, 357
 geométrica, 91
 gráfica de control, 743-744
 inferencias sobre dos medias, 469-478
 intervalo de confianza, 339-342, 351-353, 470-471, 474-476,
 477
 margen de error, 339-342, 352, 357
 media muestral, 282, 338-339
 ponderada, 84
 prueba
 de comparación múltiple, 647-649
 de hipótesis, 475-476, 477
 de rangos, 646-647
 de una aseveración, 418-422, 426-431, 445
 tamaño de la muestra, 338-339, 342-345
 valor esperado, 208-209
 Mediana, 78-79, 85
 aleatoriedad por arriba y por debajo, 721-722
 prueba del signo, 684-685
- Medición
 a nivel
 de intervalo, 8, 9
 de razón, 8-9
 nominal, 7, 9
 ordinal, 7-8, 9
 Mesnick, Sarah, 39
 Meta-análisis, 430
- Método(s)
 de contar cinco, 501
 del valor P , 401
 distribución t de Student, 429-430
 experimento multinomial, 594, 598-600
 prueba de hipótesis, 394-397, 409-412, 419-421, 429-430,
 440, 458-460, 488, 501, 610-611
 regresión múltiple, 569
 tablas de contingencia, 610-611
 no paramétricos, 674-730
 correlación de rangos y, 708-713

- datos apareados y, 680-683
 definición, 676
 desventajas del, 676-677
 empates en rangos del, 678
 mediana de una sola población del, 684-685
 pruebas y, 678-686, 689-693, 695-699, 702-705,
 717-722
 rangos del, 677-678
 resumen del, 725
 tasa de eficiencia del, 677
 ventajas del, 676
- Miedo a volar, 512
- Milgram, Stanley, 7
- Minitab. *Véase* Resultados de programas de cómputo/calculadoras
- Miringoff, Lee, 243
- Mitad del intervalo, 80-81, 85
- Moda, 80, 85
- Modelo(s), 576-579
 de población, 578-579
 matemáticos, 576-579
- Morrison, Kathleen, 513
- Muestra(s)
 aleatoria, 26-27
 simple, 26-27
 de probabilidad, 26-27
 de respuesta voluntaria (autoselecciónada), 12-13
 definición, 4, 45
 dependiente, 469
 desviación estándar, 367, 426. *Véase también* Desviación estándar
 diseño de múltiples etapas, 29-30
 espacio de, 139
 inadecuada, 12-13
 independiente, 469
 media, 282, 338-339
 pequeña, 13
 probabilidad, 26-27
 varianza dentro de, 640
 varianza entre, 640
 varianza, 97-98, 367
- Muestreo
 aleatorio, 26-27, 28
 simple, 28
 “bootstrap”, 380-381
 compuesto, 171
 con reemplazo, 163, 277
 continuo, 740
 de conveniencia, 27, 28
 de varias etapas, 29-30
 distribución, 269-278
 errores, 30-31
 estratificado, 27, 28
 vs. muestreo por racimos, 29
 por racimos, 27, 28
 vs. muestreo estratificado, 29
 repetido “Bootstrap”, 380-381
 sin reemplazo, 163, 277
 sistemático, 27, 28
 variabilidad, 275
- N
 Naipes, barajar, 183
 Naranjas, 355
 Nightingale, Florence, 63-65
 Nivel de confianza (coeficiente de confianza, grado de confianza), 322
 Nivel de significancia, 392-393, 403
 Normalidad, 338, 426
 Números
 aleatorios, 175-176, 175-177, 282-283
 precisos, 17
 seriales de tanques, 340
 telefónicos, 180
- O
 O’Toole, Dan, 757
 Obermayer, Joel B., 317
 Ojiva, 57-58
- P
 Parámetro, 5
 Pares discordantes, 624-625
 Pena de muerte, 462
 Pendiente
 b_1 , 542-544
 ecuación de regresión, 542-534
 Pensamiento crítico, 11-17, 29, 76
 aprobación de fármaco, 671
 asientos de aeronaves, 315
 bondad de ajuste, 72
 cargas axiales, 750
 configuración de teclado, 133
 discriminación de género, 38
 distribución de frecuencias, 46-48
 Duragesic, 586
 encuesta de la cámara vigilante, 381
 fraude, 632
 identificación de cáncer de vejiga, 451
 lotería de reclutamiento, 729
 medidas de tendencia central, 81-82
 métodos de selección de género, 241
 miedo a volar, 512
 prueba de embarazo, 196
 Percentiles, 111-115
 Percepción extrasensorial, 393
 Periodismo, 15
 Pictogramas, 14, 15
 Píldoras dietéticas, 474
 Población
 definición, 4
 desviación estándar, 96
 finita, 286-287
 tamaño, 330
 Polígono de frecuencias, 57
 Porcentajes
 confusos, 14-15
 conversión
 a decimales, 15
 a proporciones, 321
 principios, 15

- Posibilidades, 145-146
 Posición relativa, 110-116
 cuartiles, 111-115
 percentiles, 111-115
 puntuación z , 110-111. Véase también Puntuación z
 Potencia, 400, 402-403
 Precios de condominios, 545
 Predicción, ecuación de regresión, 544-547
 Preguntas
 orden de, 16
 sin respuesta, 16
 tendenciosas, 16
 Probabilidad, 138-146
 “al menos uno”, 168-169
 aproximación de la frecuencia relativa, 139, 140
 complementos, 168-171
 condicional, 161, 169-171
 conteo, 179-186
 ley de los números grandes, 141-144
 método clásico, 140
 notación, 139, 159
 posibilidades, 145-146
 puntuación z , 249-256
 redondeo, 144-145
 regla, 139-140
 de la multiplicación, 159-164
 de la suma, 151-155
 de las combinaciones, 185
 de las permutaciones, 182-184
 del suceso infrecuente, 138
 factorial, 181-182
 simulaciones, 174-177
 subjetiva, 140-141
 sucesos complementarios, 144-145, 155
 Probabilidades de ganar, 145-146
 Problema(s)
 de apuestas, 140, 142, 146, 175
 de bateo de béisbol, 96
 de búsqueda de empleo, 385, 408
 de calificación de universidades, 675, 710-712
 de choque de automóvil, 72
 de control de calidad, 427-428, 438-439
 de los dulces M&M, 427-428
 de cotinina, 180-181
 de detección de fraudes, 589, 597-598
 de discriminación por edad en los premios Óscar 41, 45, 48, 75, 81, 99, 112-115, 119-122, 471-474
 de diseño de tablero de automóviles, 264-265
 de duración de la clase, 247-249
 de ensayos clínicos, 183, 185, 261
 de estatura, 46-47, 103-104, 110-111, 303-306
 de fútbol, 271-272
 de género
 agrupación, 612-613
 discriminación, 38
 métodos de selección, 164, 174, 184, 241, 386-387, 388-389, 682-683
 probabilidades de nacimiento, 143, 144, 168, 170
 pruebas médicas, 690
 Senado de Estados Unidos, 273
 de inversión de acciones, 99
 de la Coca Cola vs. Pepsi Cola, 498-501
 de la fila de los clientes bancarios, 92-92, 95-96
 de la máquina *pinball*, 712-713
 de la precipitación pluvial, 720-721
 de la prueba de embarazo, 196
 de la prueba de marihuana, 137, 152-154, 160-161, 170-171
 de la prueba del termómetro, 251-254, 255-256
 de la terapia de contacto, 319, 321, 327-328
 de las bombas de la Segunda Guerra Mundial, 231-232
 de los polizontes, 357
 de medición del plomo, 78, 79, 81
 de pesos de monedas, 47, 48, 369-370
 de prueba de drogas, 137, 152-154, 160-161, 170-171
 de selección
 de plantas, 163
 de un jurado, 199, 201, 206, 208, 215-217, 219-220, 225-227
 de semillas de maíz, 681-682, 691-693
 de tabaquismo, 180-181, 711
 de temperatura corporal, 285-286, 684-685
 de tiros libres, 141
 de básquetbol, 141
 de uso de Internet, 296-297
 de la lotería, 161, 181, 202, 209, 233, 271, 292, 303
 de la puntuación de CI, 101-102, 261-262, 344-345, 736
 del Día de Acción de Gracias, 144
 del asiento expulsor del avión de propulsión a chorro, 160
 del bombardeo, 231, 232
 del casco de motocicleta, 609-610
 del cumpleaños, 145, 175-176
 del fertilizante, 24, 25, 488-489
 del géiser Old Faithful, 515-516, 523-524, 527-529, 544, 546-547, 559, 561, 562, 567-568
 del índice de masa corporal, 358, 697-698
 del informe del peso,
 del lanzamiento de monedas, 613-614
 del meteorito, 142
 del peso, 103-104
 al nacer, 356-357
 de álamos, 635, 639-640, 703-705
 de la lata de bebida de cola, 498-501
 del pie de atleta, 623-624
 del pronóstico de temperatura, 485-488
 del pulso, 47, 48, 99, 124-126, 339, 341-342
 del síndrome del túnel carpiano, 455, 458-460, 461-462
 del taxi acuático, 260-261, 263-264, 283-285
 del vendedor viajero, 182
 genéticos, 142, 411-413
 Procedimientos, 763-765
 Proceso
 de aprobación de fármaco, 408
 estadísticamente estable, 736, 742-743
 Programas seis sigma, 749
 Promedio, 81
 Pronóstico de temperatura, 542
 Propiedad de los mínimos cuadrados, 548-550
 Proporción, 320-332
 agrupada muestral, 457
 de población. Véase Proporción

- definición, 321
 distribución muestral de proporciones, 270-274
 estadístico de prueba, 408, 457
 estimado puntual, 321, 330-331
 inferencias acerca de dos proporciones, 456-463
 intervalo de confianza, 322-323, 326-328, 460-463
 margen de error, 328
 muestral, 321, 325
 notación, 321
prueba
 de hipótesis, 407-413, 458-460
 de una aseveración, 407-413, 445
 tamaño de la muestra, 328-331
 valores críticos, 324-325
Proyecto(s), 758-763
 de Internet, 38
 análisis de varianza, 672
 distribución de probabilidad, 242
 gráficas, 72
 de control, 750
 intervalos de confianza, 382
 probabilidades, 196
 prueba de hipótesis, 451, 512
 pruebas no paramétricas, 730
 regresión lineal, 586
 resumen de datos, 134
 tablas de contingencia, 632
 teorema del límite central, 316
Prueba
 de bondad de ajuste, 72, 591-600
 de cola
 derecha, 394, 395, 396, 529
 izquierda, 394, 395, 396, 529
 de comparación múltiple, 647-649
 de Bonferroni, 647-649
 de correlación de rangos de Spearman, 708-713
 de correlación de rangos ordenados, 708-713
 de dos colas, 394, 395, 528, 529
 de homogeneidad, 611-614
 de hipótesis, *Véase también* Estadístico de prueba
 componentes, 389-394
 conclusiones, 394-400
 correlación lineal, 525-532
 datos apareados, 485-488
 de cola
 derecha, 394, 395, 396, 529
 izquierda, 394, 395, 396, 529
 de dos colas, 394, 395, 528, 529
 desviación estándar/varianza, 436-440, 445, 495-501
 distribución chi cuadrada, 437-439, 445, 592-600
 errores, 398-400
 estadístico de prueba, 391-392. *Véase también*
 Estadístico de prueba
 intervalo de confianza, 395, 400, 410-413, 421-422, 430-431, 440
 media, 418-422, 455, 475-476, 477 (desviación estándar desconocida), 426-431
 método del valor *P*, 394-397, 401, 409-412, 419-421, 429-430, 440, 458-460, 488, 501, 610-611
 método tradicional, 401, 410, 412, 421, 427-428, 460, 485-487, 495-500, 609-610
 nivel de significancia, 392-393, 403
 potencia, 400, 402-403
 proporción, 407-413, 445, 458-460
 región crítica, 392, 393
 resumen, 445
 simulaciones, 450
 valor crítico, 393-394
 de independencia, 607-608
 de Kruskal-Wallis, 702-705
 de la sífilis, 171
 de la suma de rangos de Wilcoxon, 695-699
 de Levene-Brown-Forsythe, 501
 de McNemar, 621-625
 de rachas para detectar aleatoriedad, 717-722
 de rangos con signo de Wilcoxon, 689-693
 de razón, 9
 de significancia. *Véase* Prueba de hipótesis
 de una cola, 529
 del signo, 678-686
 exacta de Fisher, 614-615
 H , 702-705
 U de Mann-Whitney, 695
 de rango, 646-648
 de rendimiento académico, 567
Punto(s)
 medio, de clase, 43
 de influencia, 548
Puntuación estándar. *Véase* Puntuación *z*
Puntuación *z*, 110-111
 cálculo, 254-255
 definición, 110
 distribución normal estándar, 250-254
 notación, 324
 valor crítico, 324-325
 valores infrecuentes, 111
 vs. área, 262
 vs. distribución *t* de Student, 354-357
Q
 Quiromancia, 519
R
Rango, 93
 intercuartilar, 123
 ordenado, 8, 677-678
Rechazos, 16
Región crítica (región de rechazo), 392
Regla
 de combinaciones, 185
 de la multiplicación, 159-164
 de la suma, 151-155, 164
 de las permutaciones, 182-184, 185
 del redondeo, 81, 98
 coeficiente de correlación lineal, 521-522
 desviación estándar, 205
 intervalo de confianza, 326, 341
 media, 205
 probabilidades, 144-145

- proporción, 326
- regresión, 543
- tamaño de la muestra, 329, 343
- varianza, 205
- del suceso infrecuente, 138, 207-208, 285, 386
- empírica, 100-101
- factorial, 181-182
 - formal de la multiplicación, 162
 - formal de la suma, 153
- fundamental de conteo, 180
- intuitiva de la multiplicación, 162-163
- intuitiva de la suma, 153
- práctica de intervalo, 98-99, 205-206
- Regresión, 541-552
 - cambio marginal, 547-548
 - coeficiente de determinación, 559
 - ecuación/línea de regresión, 542-544
 - error estándar de estimación, 560-561
 - interpretación, 547-548
 - lineamientos, 547
 - logística, 572
 - múltiple, 566-573
 - por pasos, 571
 - predicciones, 544-547
 - propiedades de los mínimos cuadrados, 548-550
 - puntos de influencia, 548
 - residuales, 548-550
 - valores extremos, 548
- Réplica, 25-26
- Residual, 548-550
- Resultados de programas de cómputo/calculadoras, 37
 - ANOVA
 - de dos factores, 661
 - de un factor, 649
 - correlación, 532
 - de rangos, 713
 - datos apreados, 489
 - desviación estándar/varianza, 501
 - distribución
 - de Poisson, 233
 - de probabilidad, 218-219, 220
 - normal estándar, 250-251, 265
 - dos medias, 478
 - dos proporciones de población, 463
 - dos varianzas, 501
 - estadística descriptiva, 86, 116
 - experimentos multinomiales, 600
 - generación de números aleatorios, 175-176
 - gráfica(s), 66
 - cuantilar normal, 307
 - de puntos, 127
 - de rachas, 735-736
 - p , 750
 - de control, 744-745, 750
 - histograma, 53-54
 - intervalo de confianza, 332, 357, 358
 - modelos, 579
 - probabilidades binomiales, 218-219
 - prueba
 - de hipótesis, 422, 431, 440, 450
- de Kruskal-Wallis, 705
- de McNemar, 621-625
- de rachas, 721-722
- del signo, 686
- de Wilcoxon, 693, 699
- puntuación z , 257
- regresión, 552
 - múltiple, 573
- tablas de contingencia, 615
- tamaño de la muestra, 332, 345, 370-371
- teorema del límite central, 314
- variación, 563
- Resumen de cinco cifras, 121
- Riesgo relativo, 150
- Robo, 473
 - de identidad, 180
- S
 - Saccucci, Michael, 452-453
 - Salarios
 - de la NBA, 570
 - de maestros, 82-83
 - SC (total), 644-645
 - Seguridad en paracaídas, 293
 - Sehlinger Bob, 73
 - Seis grados de separación, 3, 30
 - Selección aleatoria, 4, 45
 - Sesgo, 85-86
 - de publicación, 17
 - Shakespeare, William, 152
 - Símbolo de factorial (!), 181
 - Simulaciones, 143, 174-177, 450
 - de dados, 176-177
 - Soborno, 743
 - Sondeo. Véase Encuestas
 - de empuje, 439
 - Spock, Benjamin, 199
 - STATDISK. Véase Resultados de programas de cómputo/calculadoras
 - Sucesos
 - complementarios, 144-145, 155
 - compuestos, 151
 - definición, 139
 - dependientes, 162, 163-164
 - disjuntos, 153-155
 - independientes, 162, 163-164
 - mutuamente excluyentes, 153-155
 - simples, 139
 - Suma de cuadrados, SC (total), 644-645
 - Súper Bowl, 476, 558
 - T
 - Tablas de contingencia, 606-615
 - definición, 606
 - frecuencia esperada, 608-610
 - prueba de homogeneidad, 611-614
 - prueba de independencia, 607-608
 - valores P , 610-611
 - Tamaño
 - de la clase, 79, 477

- de la muestra (n), 25-26, 77
 - desigual, 643-646
 - desviación estándar, 344-345
 - igual, 641-643
 - media, 338-339, 342-345
 - proporción, 328-331
 - regla del redondeo, 329
 - varianza, 370-371
- de la multitud, 357
- de la población de la fauna, 339
- Tasa de ataques cardiacos, 460
- Temas, proyectos, 760-763
- Teorema
 - de Bayes, 190
 - de Chebyshev, 101
 - del límite central, 259-271, 280-287
 - aplicaciones, 282-286
 - borroso, 286
 - corrección de población finita, 286-287
 - correcciones por continuidad, 294-297
- Teoría de las colas, 231
- Terapia con parches de nicotina, 330-331
- TI-83/84 Plus. *Véase* Resultados de programas de cómputo/calculadoras
- Tiempo, cambio de datos con el, 42, 78
- Tratamiento (factor), 638

- U**
- “Uno al menos”, 168-169
- Unidades experimentales (sujetos), 21

- V**
- Validez, 94
- Valor
 - esperado, 208-209
- Valor crítico, 324-325, 351-352, 393-394
 - correlación lineal y, 324, 325-326, 365-367, 419, 426, 485, 487, 496-497, 527-529
- notación de, 324
 - prueba de independencia y, 607
 - prueba de Kruskal-Wallis y, 703
 - prueba de McNemar y, 622
 - prueba del signo y, 685
 - pruebas de bondad del ajuste y, 593
 - pruebas de Wilcoxon y, 690, 696
 - varianza y, 365-367
- Valores
 - extremos, 42, 119-120, 129, 121, 548
 - leves, 129
 - infrecuentes, 111
- Variabilidad de muestreo, 275
- Variable(s), 21-22

- X**
- \bar{x} , distribución muestral, 281-287

- Z**
- z , puntuación, *Véase* Puntuación z

Tabla de símbolos

\bar{A}	Complemento del suceso A	H	Estadístico de prueba Kruskal-Wallis
H_0	Hipótesis nula	R	Suma de rangos para una muestra; se utiliza en la prueba de suma de rangos de Wilcoxon
H_1	Hipótesis alternativa	μ_R	Rango medio esperado; se utiliza en la prueba de suma de rangos de Wilcoxon
α	Alfa; probabilidad de un error tipo I o el área de la región crítica	σ_R	Desviación estándar de rangos esperada; se utiliza en la prueba de suma de rangos de Wilcoxon
β	Beta; probabilidad de un error tipo II	G	Número de rachas en la prueba de rachas para detectar aleatoriedad
r	Coeficiente de correlación lineal muestral	μ_G	Media esperada del número de rachas; se utiliza en la prueba de rachas para detectar aleatoriedad
ρ	Rho; coeficiente de correlación lineal poblacional	σ_G	Desviación estándar esperada para el número de rachas; se utiliza en la prueba de rachas para detectar aleatoriedad
r^2	Coeficiente de determinación	$\mu_{\bar{x}}$	Media poblacional de todas las medias muestrales \bar{x} posibles
R^2	Coeficiente de determinación múltiple	$\sigma_{\bar{x}}$	Desviación estándar poblacional de todas las medias muestrales \bar{x} posibles
r_s	Coeficiente de correlación de rangos de Spearman	E	Margen de error del estimado de un parámetro poblacional, o valor esperado
b_1	Estimado puntual de la pendiente de la recta de regresión	Q_1, Q_2, Q_3	Cuartiles
b_0	Estimado puntual del intercepto y de la recta de regresión	D_1, D_2, \dots, D_9	Deciles
\hat{y}	Valor predicho de y	P_1, P_2, \dots, P_{99}	Percentiles
d	Diferencia entre dos valores apareados	x	Valor de un dato
\bar{d}	Media de las diferencias d , calculada a partir de datos muestrales apareados		
s_d	Desviación estándar de las diferencias d , calculada a partir de datos muestrales apareados		
s_e	Error estándar de un estimado		
T	Suma de rangos; se utiliza en la prueba de rangos con signo de Wilcoxon		

Tabla de símbolos

f	Frecuencia con la que ocurre un valor	$z_{\alpha/2}$	Valor crítico de z
Σ	Sigma mayúscula; sumatoria	t	Distribución t
Σx	Suma de los valores	$t_{\alpha/2}$	Valor crítico de t
Σx^2	Suma de los cuadrados de los valores	df	Grados de libertad
$(\Sigma x)^2$	Cuadrado de la suma de todos los valores	F	Distribución F
Σxy	Suma de los productos de cada valor x multiplicado por el valor y correspondiente	χ^2	Distribución chi cuadrada
n	Número de valores en una muestra	χ_R^2	Valor crítico de cola derecha de chi cuadrada
$n!$	n factorial	χ_L^2	Valor crítico de cola izquierda de chi cuadrada
N	Número de valores en una población finita; también se utiliza como el tamaño de todas las muestras combinadas	p	Probabilidad de un suceso o la proporción poblacional
k	Número de muestras o poblaciones o categorías	q	Probabilidad o proporción igual a $1 - p$
\bar{x}	Media de los valores en una muestra	\hat{p}	Proporción muestral
\bar{R}	Media de los rangos muestrales	\hat{q}	Proporción muestral igual a $1 - \hat{p}$
μ	Mu; media de todos los valores en una población	\bar{p}	Proporción obtenida al agrupar dos muestras
s	Desviación estándar de un conjunto de valores muestrales	\bar{q}	Proporción o probabilidad igual a $1 - \bar{p}$
σ	Sigma minúscula; desviación estándar de todos los valores en una población	$P(A)$	Probabilidad del suceso A
s^2	Varianza de un conjunto de datos muestrales	$P(A B)$	Probabilidad del suceso A , suponiendo que el suceso B ya ocurrió
σ^2	Varianza de todos los valores en una población	${}_nP_r$	Número de permutaciones de n elementos seleccionando r a la vez
z	Puntuación estándar	${}_nC_r$	Número de combinaciones de n elementos seleccionando r a la vez

¿NO OLVIDAS ALGO?

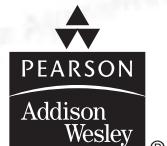
Al comprar este libro de texto, **Pearson Educación** te da acceso a la tecnología más avanzada para complementar tu aprendizaje, dentro y fuera del salón de clases.

Acompañando a este libro, puedes encontrar cuestionarios de autoevaluación, ejercicios interactivos, animaciones, casos de estudio, resúmenes o hasta un curso en línea dentro de nuestra plataforma **CourseCompass***.

Consulta la página Web del libro para conocer los recursos que están disponibles. O pregunta a tu profesor sobre el material que puso a tu disposición para el curso y **entrégale el formulario que está al reverso para solicitar tu código de acceso**.

¡No dejes pasar esta oportunidad y únete a los millones de alumnos que están sacando el máximo provecho de su libro de texto!

***CourseCompass** es una plataforma educativa en línea desarrollada por Blackboard Technologies® exclusivamente para **Pearson Educación**.



SOLICITUD DE CÓDIGO DE ACCESO PARA COURSECOMPASS

DATOS DEL ALUMNO

Nombre completo _____ e-mail _____

DATOS DE LA INSTITUCIÓN

Nombre de la institución _____ Campus o Facultad _____

Dirección _____ Ciudad y estado _____ País _____

Nombre del profesor _____ e-mail del profesor _____

Nombre de la materia _____ Grado (Nº semestre, trimestre, etc.) _____ Nombre de la carrera _____

DATOS DEL LIBRO

Título _____ Edición _____ Autor _____ ISBN _____

¿Es el texto principal? _____ ¿Dónde adquiriste el libro? _____ ¿Consideras adecuado el precio?
Sí No Sí No

¿Cuentas con una computadora propia? _____ ¿Cuentas con acceso a Internet? _____ ¿Cuentas con laboratorio de cómputo en tu escuela?
Sí No Sí No Sí No

¿Has utilizado anteriormente esta u otra plataforma en línea? _____ ¿Ayudó a mejorar tu desempeño?
Sí No Sí No

¿Cuál?

¿Por qué?

¿Utilizas actualmente algún otro libro de Pearson Educación? _____ Sí No

¿Cuáles?
1. Título _____ edición _____ Autor _____

Materia _____ Profesor _____ ¿CourseCompass?
Sí No

2. Título _____ edición _____ Autor _____

Materia _____ Profesor _____ ¿CourseCompass?
Sí No

PARA LLENAR POR EL PROFESOR

(Llenar una sola por grupo y entregar al frente con el resto de las solicitudes)

Clave del curso (Course ID)¹ _____ ISBN del curso¹ _____

Fecha de inicio del curso _____ / _____ / _____ Culminación _____ / _____ / _____ Límite para registro de alumnos² _____ / _____ / _____

Número de códigos solicitados _____ Total de alumnos en el grupo _____ Teléfono de contacto _____

¿Existe el libro en biblioteca? Sí No Fecha de entrega de solicitudes _____ / _____ / _____

¿Le gustaría recibir información sobre otros materiales de Pearson Educación? Sí No

¹ Entre a la sección Course List haciendo clic en la pestaña Courses de CourseCompass. La información aparece debajo de cada curso de su lista.

² En la sección Course List, hacer clic en el botón Course Settings de este curso y luego en la liga Course Dates. La fecha límite para inscripción aparece como Enrollment End Date.

Fórmulas y tablas
 para *Estadística, décima edición*, de Mario Triola
 D.R. © 2006 Pearson Educación de México S.A. de C.V.

Capítulo 3: Estadística descriptiva $\bar{x} = \frac{\sum x}{n} \quad \text{Media}$ $\bar{x} = \frac{\sum f \cdot x}{\sum f} \quad \text{Media (tabla de frecuencias)}$ $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} \quad \text{Desviación estándar}$ $s = \sqrt{\frac{n(\sum x^2) - (\sum x)^2}{n(n-1)}} \quad \text{Desviación estándar (método rápido)}$ $s = \sqrt{\frac{n[\sum(f \cdot x^2)] - [\sum(f \cdot x)]^2}{n(n-1)}} \quad \text{Desviación estándar (tabla de frecuencias)}$ $\text{Varianza} = s^2$	Capítulo 7: Intervalos de confianza (una población) $\hat{p} - E < p < \hat{p} + E \quad \text{Proporción}$ $\text{donde } E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$ <hr/> $\bar{x} - E < \mu < \bar{x} + E \quad \text{Media}$ $\text{donde } E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (\sigma \text{ conocida})$ $\text{o } E = t_{\alpha/2} \frac{s}{\sqrt{n}} \quad (\sigma \text{ desconocida})$ <hr/> $\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2} \quad \text{Varianza}$
Capítulo 4: Probabilidad $P(A \cup B) = P(A) + P(B) \quad \text{si } A \text{ y } B \text{ son mutuamente excluyentes}$ $P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{si } A \text{ y } B \text{ no son mutuamente excluyentes}$ $P(A \cap B) = P(A) \cdot P(B) \quad \text{si } A \text{ y } B \text{ son independientes}$ $P(A \cap B) = P(A) \cdot P(B A) \quad \text{si } A, B \text{ son dependientes}$ $P(\bar{A}) = 1 - P(A) \quad \text{Regla de los sucesos complementarios}$ $nP_r = \frac{n!}{(n-r)!} \quad \text{Permutaciones (sin elementos iguales)}$ $\frac{n!}{n_1! n_2! \dots n_k!} \quad \text{Permutaciones (} n_1 \text{ iguales,...)}$ ${}_nC_r = \frac{n!}{(n-r)! r!} \quad \text{Combinaciones}$	Capítulo 7: Determinación de tamaño de muestra $n = \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2} \quad \text{Proporción}$ $n = \frac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2} \quad \text{Proporción (\hat{p} y \hat{q} conocidas)}$ $n = \left[\frac{z_{\alpha/2} \sigma}{E} \right]^2 \quad \text{Media}$
Capítulo 5: Distribuciones de probabilidad $\mu = \sum x \cdot P(x) \quad \text{Media (distribución de probabilidad)}$ $\sigma = \sqrt{\sum [x - \mu]^2 \cdot P(x)} \quad \text{Desviación estándar (dist. de prob.)}$ $P(x) = \frac{n!}{(n-x)! x!} \cdot p^x \cdot q^{n-x} \quad \text{Probabilidad binomial}$ $\mu = n \cdot p \quad \text{Media (binomial)}$ $\sigma^2 = n \cdot p \cdot q \quad \text{Varianza (binomial)}$ $\sigma = \sqrt{n \cdot p \cdot q} \quad \text{Desviación estándar (binomial)}$ $P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} \quad \text{Distribución de Poisson}$ $\text{donde } e \approx 2.71828$	Capítulo 9: Intervalos de confianza (dos poblaciones) $(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$ $\text{donde } E = z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$ <hr/> $(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E \quad (\text{Indep.})$ $\text{donde } E = t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (\text{gl} = \text{el menor de } n_1 - 1, n_2 - 1) \quad \leftarrow$ <p>(σ_1 y σ_2 desconocidas y se supone que no son iguales)</p> $E = t_{\alpha/2} \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} \quad (\text{gl} = n_1 + n_2 - 2) \quad \leftarrow$ $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$ <p>(σ_1 y σ_2 desconocidas, pero se supone que son iguales)</p> $E = z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \leftarrow$ <p>(σ_1, σ_2 conocidas)</p> <hr/> $\bar{d} - E < \mu_d < \bar{d} + E \quad (\text{Datos apareados})$ $\text{donde } E = t_{\alpha/2} \frac{s_d}{\sqrt{n}} \quad (\text{gl} = n - 1)$
Capítulo 6: Distribución normal $z = \frac{x - \bar{x}}{s} \text{ or } \frac{x - \mu}{\sigma} \quad \text{Puntuación estándar}$ $\mu_{\bar{x}} = \mu \quad \text{Teorema del límite central}$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{Teorema del límite central}$ (error estándar)	

Fórmulas y tablas
 para *Estadística, décima edición*, de Mario Triola
 D.R. © 2006 Pearson Educación de México S.A. de C.V

Capítulo 8: Estadísticos de prueba (una población)

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \quad \text{Proporción: una población}$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \quad \text{Media: una población} \\ (\sigma \text{ conocida})$$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad \text{Media: una población} \\ (\sigma \text{ desconocida})$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \quad \text{Desviación estándar o varianza:} \\ \text{una población}$$

Capítulo 9: Estadísticos de prueba (dos poblaciones)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}} \quad \text{Dos proporciones}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{gl = el menor de} \\ n_1 - 1, n_2 - 1$$

Dos medias: independiente; σ_1 y σ_2 desconocidas y no se supone que sean iguales.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \quad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Dos medias: independiente; σ_1 y σ_2 desconocidas, pero se supone que son iguales.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{Dos medias: independiente;} \\ \sigma_1, \sigma_2 \text{ conocidas.}$$

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} \quad \text{Dos medias: datos apareados} \\ (\text{gl} = n - 1)$$

$$F = \frac{s_1^2}{s_2^2} \quad \text{Desviación estándar o varianza: dos} \\ \text{poblaciones (donde } s_1^2 \geq s_2^2\text{)}$$

Capítulo 11: Multinomiales y tablas de contingencia

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad \text{Multinomial} \\ (\text{gl} = k - 1)$$

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad \text{Tabla de contingencia} \\ [\text{gl} = (r - 1)(c - 1)]$$

$$\text{donde } E = \frac{(\text{total por renglón})(\text{total por columna})}{(\text{gran total})}$$

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c} \quad \text{Prueba de McNemar} \\ \text{para pares apareados} \\ (\text{gl} = 1)$$

Capítulo 10: Correlación lineal/Regresión

$$\text{Correlación } r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

$$b_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b_0 = \bar{y} - b_1\bar{x} \text{ or } b_0 = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$\hat{y} = b_0 + b_1x \quad \text{Ecuación estimada de la recta de regresión}$$

$$r^2 = \frac{\text{variación explicada}}{\text{variación total}}$$

$$s_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}} \circ \sqrt{\frac{\sum y^2 - b_0\sum y - b_1\sum xy}{n-2}}$$

$$\hat{y} - E < y < \hat{y} + E \quad \text{Intervalo de predicción}$$

$$\text{donde } E = t_{\alpha/2}s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\sum x^2) - (\sum x)^2}}$$

Capítulo 12: Análisis de varianza de un factor

Procedimiento para poner a prueba $H_0: \mu_1 = \mu_2 = \mu_3 = \dots$

1. Usar un programa de cómputo o una calculadora para obtener los resultados.
2. Identificar el valor P .
3. Obtener la conclusión:

Si el valor $\leq \alpha$, se rechaza la hipótesis nula de medias iguales.

Si $P > \alpha$, no se rechaza la hipótesis nula de medias iguales.

Capítulo 12: Análisis de varianza de dos factores

Procedimiento:

1. Usar un programa de cómputo o una calculadora para obtener los resultados.
2. Probar H_0 : No hay una interacción entre el factor de renglón y el factor de columna.
3. Detenerse si se rechaza H_0 del paso 2.
 Si no se rechaza H_0 del paso 1 (de manera que al parecer no existe un efecto de interacción), continúe con las siguientes dos pruebas:

Prueba de los efectos del factor de renglón.

Prueba de los efectos del factor de columna.

Fórmulas y tablas
 para *Estadística, décima edición*, de Mario Triola
 D.R. © 2006 Pearson Educación de México S.A. de C.V.

Capítulo 13: Pruebas no paramétricas

$$z = \frac{(x + 0.5) - (n/2)}{\sqrt{n}/2} \quad \text{Prueba del signo para } n > 25$$

$$z = \frac{T - n(n+1)/4}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \quad \begin{array}{l} \text{Prueba de rangos con signo de} \\ \text{Wilcoxon (datos apareados y} \\ n > 30) \end{array}$$

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{R - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}}} \quad \begin{array}{l} \text{Prueba de suma de} \\ \text{rangos de Wilcoxon} \\ (\text{dos muestras} \\ \text{independientes}) \end{array}$$

$$H = \frac{12}{N(N+1)} \left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \dots + \frac{R_k^2}{n_k} \right) - 3(N+1)$$

Prueba de Kruskal-Wallis (chi cuadrada, gl = k - 1)

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \quad \text{Correlación de rangos}$$

$$\left(\text{valor crítico para } n > 30: \frac{\pm z}{\sqrt{n-1}} \right)$$

$$z = \frac{G - \mu_G}{\sigma_G} = \frac{G - \left(\frac{2n_1n_2}{n_1+n_2} + 1 \right)}{\sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}} \quad \begin{array}{l} \text{Prueba de} \\ \text{rachas para} \\ n > 20 \end{array}$$

Capítulo 14: Gráficas de control

Gráfica R: Graficar rangos muestrales

LCS: $D_4 \bar{R}$

Línea central: \bar{R}

LCI: $D_3 \bar{R}$

Gráfica \bar{x} : Graficar medias muestrales

LCS: $\bar{\bar{x}} + A_2 \bar{R}$

Línea central: $\bar{\bar{x}}$

LCI: $\bar{\bar{x}} - A_2 \bar{R}$

Gráfica p : Graficar proporciones muestrales

LCS: $\bar{p} + 3\sqrt{\frac{\bar{p}\bar{q}}{n}}$

Línea central: \bar{p}

LCI: $\bar{p} - 3\sqrt{\frac{\bar{p}\bar{q}}{n}}$

TABLA A-6
Valores críticos del coeficiente
de correlación r de Pearson

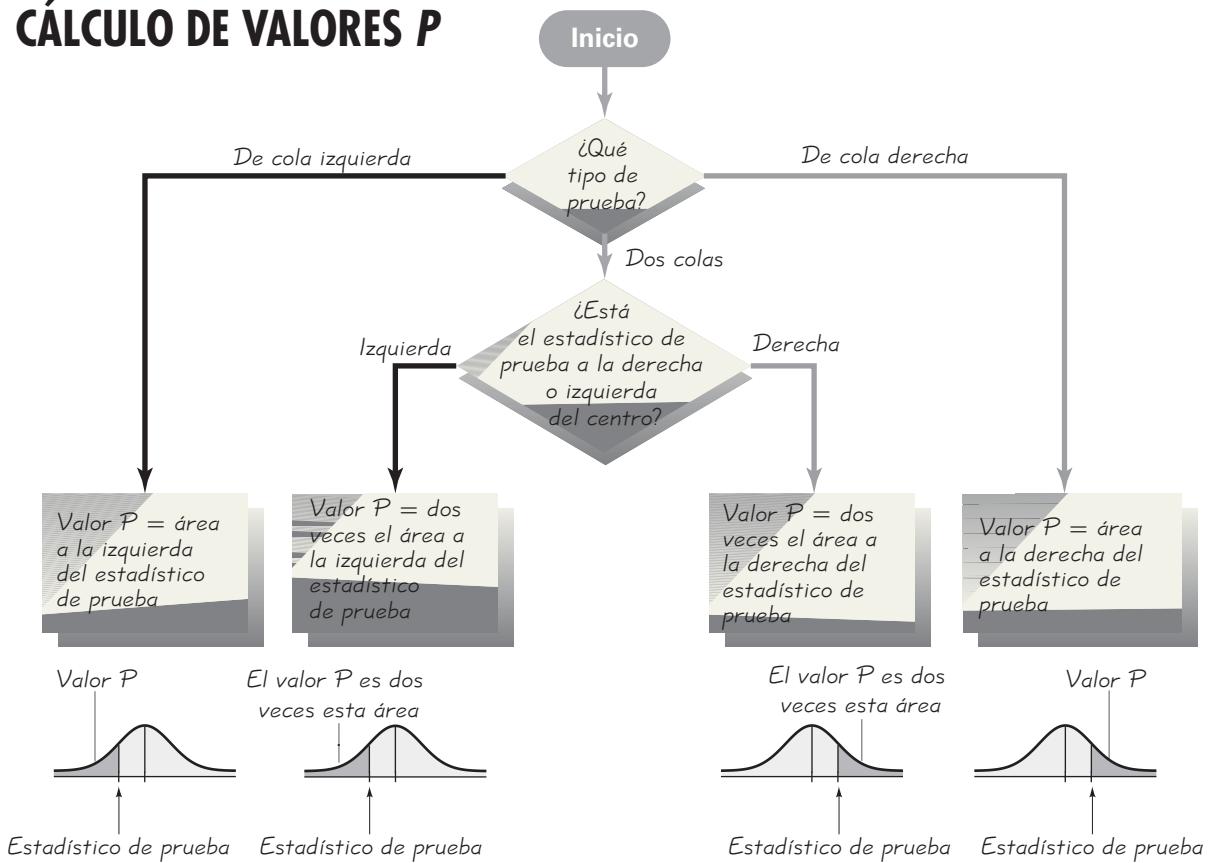
n	$\alpha = .05$	$\alpha = .01$
4	.950	.999
5	.878	.959
6	.811	.917
7	.754	.875
8	.707	.834
9	.666	.798
10	.632	.765
11	.602	.735
12	.576	.708
13	.553	.684
14	.532	.661
15	.514	.641
16	.497	.623
17	.482	.606
18	.468	.590
19	.456	.575
20	.444	.561
25	.396	.505
30	.361	.463
35	.335	.430
40	.312	.402
45	.294	.378
50	.279	.361
60	.254	.330
70	.236	.305
80	.220	.286
90	.207	.269
100	.196	.256

NOTA: Para probar $H_0: \rho = 0$ contra $H_1: \rho \neq 0$, se rechaza H_0 si el valor absoluto de r es mayor que el valor crítico que se indica en la tabla.

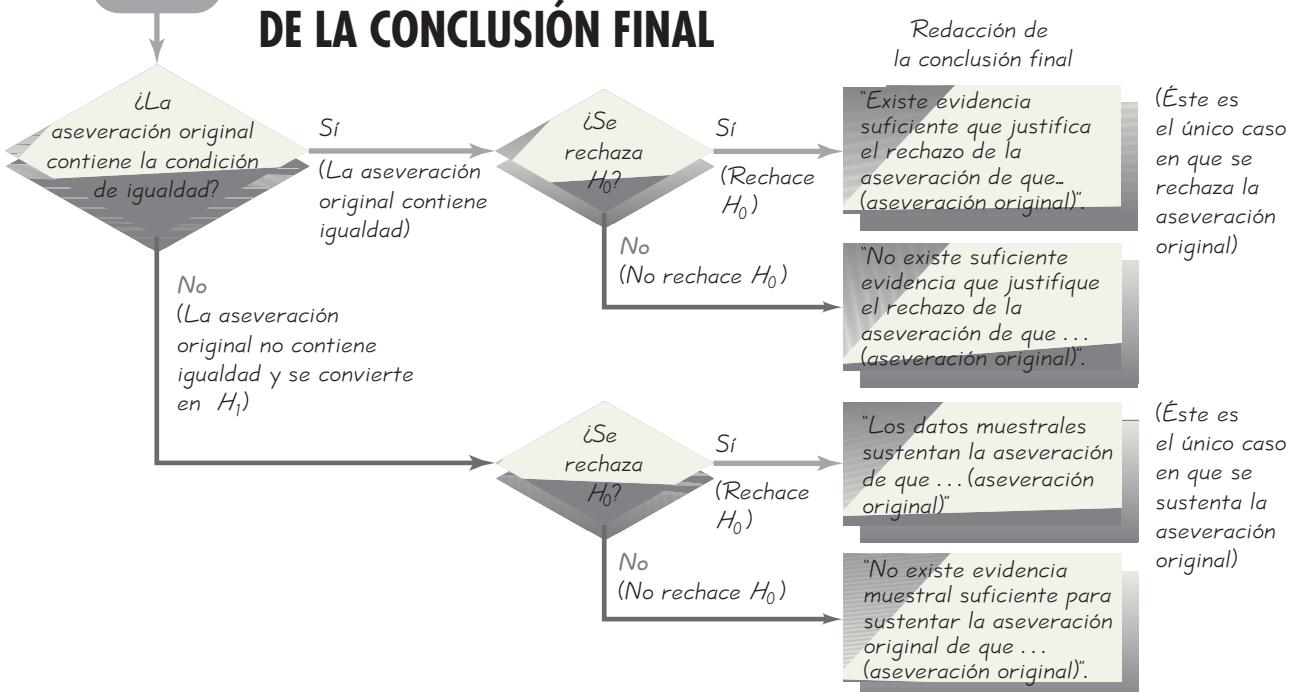
Constantes de una gráfica de control

Tamaño del subgrupo	A_2			
	n	A_2	D_3	D_4
2		1.880	0.000	3.267
3		1.023	0.000	2.574
4		0.729	0.000	2.282
5		0.577	0.000	2.114
6		0.483	0.000	2.004
7		0.419	0.076	1.924

CÁLCULO DE VALORES P



PRUEBA DE HIPÓTESIS: REDACCIÓN DE LA CONCLUSIÓN FINAL

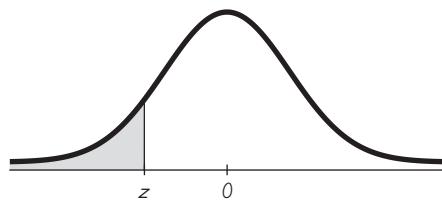


Inferencias acerca de μ : elección entre la distribución t y la distribución normal

Distribución t : Se desconoce σ y la población se distribuye normalmente
o se desconoce σ y $n > 30$

Distribución normal: Se conoce σ y la población se distribuye normalmente
o se conoce σ y $n > 30$

Método no paramétrico o “bootstrap”: La población no se distribuye normalmente y $n \leq 30$



Puntuaciones z NEGATIVAS

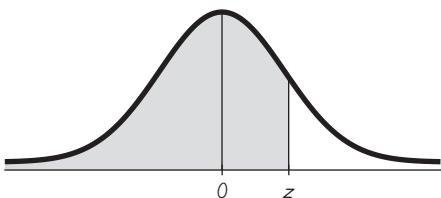
TABLA A-2		Distribución normal estándar (z): Área acumulativa desde la IZQUIERDA									
z		.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.50	y menores	.0001									
-3.4		.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3		.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2		.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1		.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0		.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9		.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8		.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7		.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6		.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5		.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	* .0049	.0048
-2.4		.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	↑ .0066	.0064
-2.3		.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	↑ .0087	.0084
-2.2		.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	↑ .0113	.0110
-2.1		.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	↑ .0146	.0143
-2.0		.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	↑ .0188	.0183
-1.9		.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	↑ .0239	.0233
-1.8		.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	↑ .0301	.0294
-1.7		.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	↑ .0375	.0367
-1.6		.0548	.0537	.0526	.0516	.0505	* .0495	.0485	.0475	↑ .0465	.0455
-1.5		.0668	.0655	.0643	.0630	.0618	↑ .0606	.0594	.0582	↑ .0571	.0559
-1.4		.0808	.0793	.0778	.0764	.0749	↑ .0735	.0721	.0708	↑ .0694	.0681
-1.3		.0968	.0951	.0934	.0918	.0901	↑ .0885	.0869	.0853	↑ .0838	.0823
-1.2		.1151	.1131	.1112	.1093	.1075	↑ .1056	.1038	.1020	↑ .1003	.0985
-1.1		.1357	.1335	.1314	.1292	.1271	↑ .1251	.1230	.1210	↑ .1190	.1170
-1.0		.1587	.1562	.1539	.1515	.1492	↑ .1469	.1446	.1423	↑ .1401	.1379
-0.9		.1841	.1814	.1788	.1762	.1736	↑ .1711	.1685	.1660	↑ .1635	.1611
-0.8		.2119	.2090	.2061	.2033	.2005	↑ .1977	.1949	.1922	↑ .1894	.1867
-0.7		.2420	.2389	.2358	.2327	.2296	↑ .2266	.2236	.2206	↑ .2177	.2148
-0.6		.2743	.2709	.2676	.2643	.2611	↑ .2578	.2546	.2514	↑ .2483	.2451
-0.5		.3085	.3050	.3015	.2981	.2946	↑ .2912	.2877	.2843	↑ .2810	.2776
-0.4		.3446	.3409	.3372	.3336	.3300	↑ .3264	.3228	.3192	↑ .3156	.3121
-0.3		.3821	.3783	.3745	.3707	.3669	↑ .3632	.3594	.3557	↑ .3520	.3483
-0.2		.4207	.4168	.4129	.4090	.4052	↑ .4013	.3974	.3936	↑ .3897	.3859
-0.1		.4602	.4562	.4522	.4483	.4443	↑ .4404	.4364	.4325	↑ .4286	.4247
-0.0		.5000	.4960	.4920	.4880	.4840	↑ .4801	.4761	.4721	↑ .4681	.4641

NOTA: Para valores de z por debajo de -3.49, utilice 0.0001 para el área.

*Utilice estos valores comunes que resultan por interpolación:

Puntuación

z	Área
-1.645	0.0500
-2.575	0.0050



Puntuaciones z POSITIVAS

TABLA A-2

(continuación) Área acumulativa desde la IZQUIERDA

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495 *	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591 ↑	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671 ↑	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738 ↑	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793 ↑	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838 ↑	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875 ↑	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904 ↑	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927 ↑	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945 ↑	.9946	.9948	.9949 *	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959 ↑	.9960	.9961	.9962 ↑	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969 ↑	.9970	.9971	.9972 ↑	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977 ↑	.9978	.9979	.9979 ↑	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984 ↑	.9984	.9985	.9985 ↑	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988 ↑	.9989	.9989	.9989 ↑	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992 ↑	.9992	.9992	.9992 ↑	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994 ↑	.9994	.9994	.9995 ↑	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996 ↑	.9996	.9996	.9996 ↑	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997 ↑	.9997	.9997	.9997 ↑	.9997	.9998
3.50 y mayores	.9999									

NOTA: Para valores de z por encima de 3.49, utilice 0.9999 para el área.

*Utilice estos valores comunes que resultan por interpolación:

Puntuación

<u>z</u>	<u>Área</u>
1.645	0.9500
2.575	0.9950

Valores críticos comunes

Nivel de confianza	Valor crítico
0.90	1.645
0.95	1.96
0.99	2.575

TABLA A-3

Distribución t : Valores críticos t

	Área en una cola				
	0.005	0.01	0.025	0.05	0.10
Grados de libertad	Área en dos colas				
	0.01	0.02	0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316
26	2.779	2.479	2.056	1.706	1.315
27	2.771	2.473	2.052	1.703	1.314
28	2.763	2.467	2.048	1.701	1.313
29	2.756	2.462	2.045	1.699	1.311
30	2.750	2.457	2.042	1.697	1.310
31	2.744	2.453	2.040	1.696	1.309
32	2.738	2.449	2.037	1.694	1.309
34	2.728	2.441	2.032	1.691	1.307
36	2.719	2.434	2.028	1.688	1.306
38	2.712	2.429	2.024	1.686	1.304
40	2.704	2.423	2.021	1.684	1.303
45	2.690	2.412	2.014	1.679	1.301
50	2.678	2.403	2.009	1.676	1.299
55	2.668	2.396	2.004	1.673	1.297
60	2.660	2.390	2.000	1.671	1.296
65	2.654	2.385	1.997	1.669	1.295
70	2.648	2.381	1.994	1.667	1.294
75	2.643	2.377	1.992	1.665	1.293
80	2.639	2.374	1.990	1.664	1.292
90	2.632	2.368	1.987	1.662	1.291
100	2.626	2.364	1.984	1.660	1.290
200	2.601	2.345	1.972	1.653	1.286
300	2.592	2.339	1.968	1.650	1.284
400	2.588	2.336	1.966	1.649	1.284
500	2.586	2.334	1.965	1.648	1.283
750	2.582	2.331	1.963	1.647	1.283
1000	2.581	2.330	1.962	1.646	1.282
2000	2.578	2.328	1.961	1.646	1.282
Grande	2.576	2.326	1.960	1.645	1.282

Fórmulas y tablas
 para *Estadística, décima edición*, de Mario Triola
 D.R. © 2006 Pearson Educación de México S.A. de C.V.

TABLA A-4		Distribución chi cuadrada (χ^2)									
Grados de libertad		Área a la derecha del valor crítico									
		0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879	
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597	
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838	
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860	
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750	
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548	
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278	
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955	
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589	
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188	
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757	
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299	
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819	
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319	
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801	
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267	
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718	
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156	
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582	
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997	
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401	
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796	
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181	
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559	
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928	
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290	
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645	
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993	
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336	
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672	
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766	
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490	
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952	
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215	
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321	
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299	
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169	

De Donald B. Owen, *Handbook of Statistical Tables*, © 1962 Addison-Wesley Publishing Co., Reading, MA. Reproducido con permiso del editor.