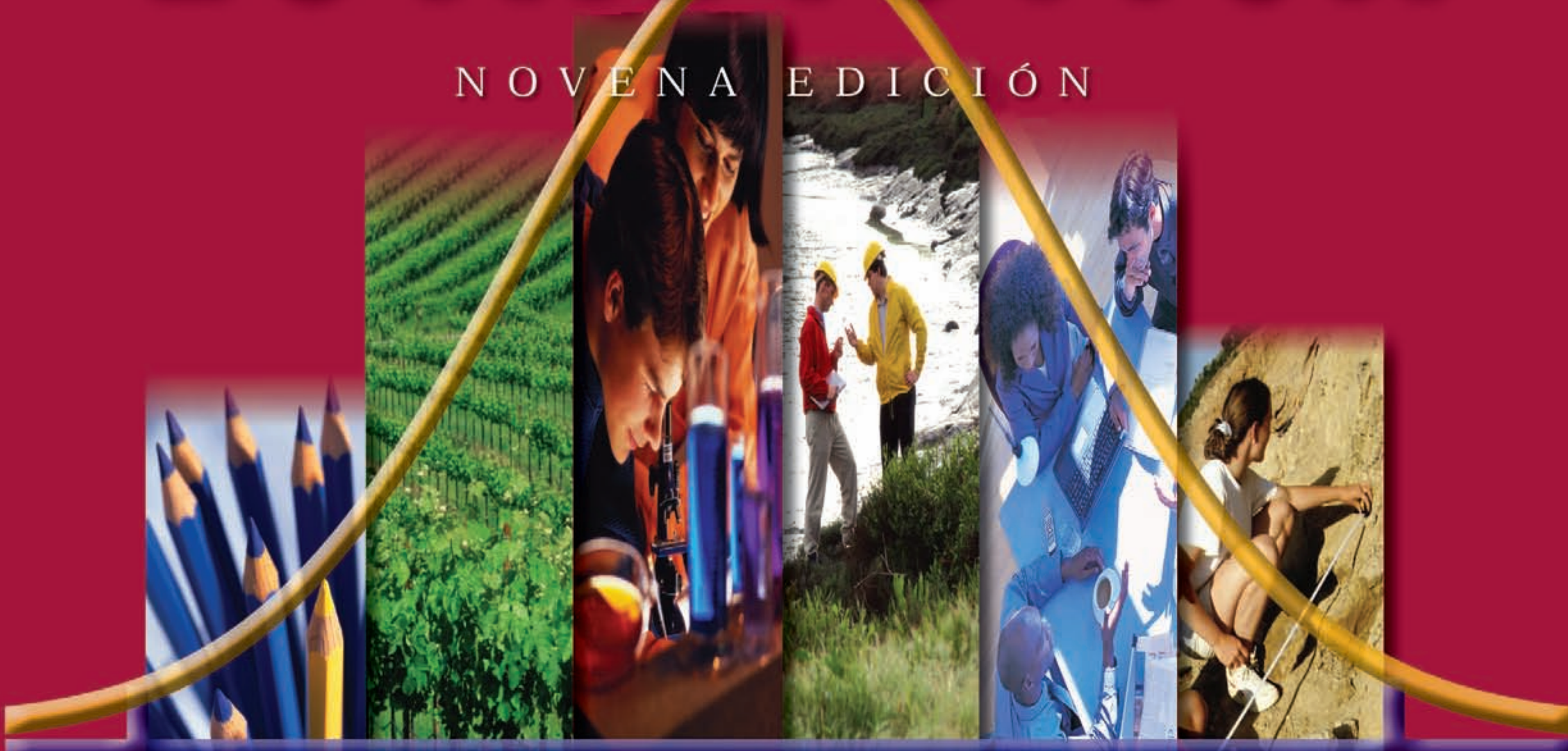
ESTADÍSTICA

NOVENA EDICIÓN



MARIO F. TRIOLA

PEARSON
Addison Wesley®

INCLUYE CD-ROM

ESTADÍSTICA

NOVENA EDICIÓN

ESTADÍSTICA

NOVENA EDICIÓN

MARIO F. TRIOLA

TRADUCCIÓN:
M. Leticia Esther Pineda Ayala
Profesora de Estadística
Universidad Anáhuac

REVISIÓN TÉCNICA:
Mario Alberto González Medina
Profesor Asociado
Universidad de Monterrey

PEARSON Educación ®

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

Datos de catalogación bibliográfica

TRIOLA, MARIO F.
Estadística. Novena edición

PEARSON EDUCACIÓN, México, 2004

ISBN: 970-26-0519-9
Área: Universitarios

Formato: 21 × 27 cm Páginas: 872

ISBN 0-201-77570-0

Edición en español
Editor: Guillemo Trujano Mendoza
e-mail: guillermo.trujano@pearsoned.com
Supervisor de desarrollo: Felipe Hernández Carrasco
Supervisor de producción: José D. Hernández Garduño

Edición en inglés
Publisher: Greg Tobin
Senior Acquisitions Editor: Deirdre Lynch
Executive Project Manager: Christine O'Brien
Editorial Assistant: Keren Blankfeld
Executive Marketing Manager: Yolanda Cossio
Senior Marketing Manager: Pamela Laskey
Managing Editor: Karen Guardino
Senior Production Supervisor: Peggy McMahon
Text and Cover Designer: Barbara T. Atkinson
Cover Photos: Pencils and vineyard © Getty;
People in lab © Stone;
Group at table and men at river bank © Creatas;
Women with tape measure © Corbis.
Associate Media Producer: Jennifer Kerber
Senior Manufacturing Buyer: Evelyn Beaton
Software Development: Kathleen Bowler and Jan Wann
Composition and Production Services: Nesbitt Graphics, Inc.
Printer: VonHoffman Press

NOVENA EDICIÓN, 2004

Atlacomulco 500-5to. piso
Industrial Atoto
53519 Naucalpan de Juárez, Edo. de México
E-mail: editorial.universidades@pearsoned.com

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

ISBN 970-26-0519-9

Impreso en México. *Printed in Mexico.*

1 2 3 4 5 6 7 8 9 0 - 07 06 05 04

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, Mathematics in the Modern World* y *Survey of Mathematics*. Además, es coautor de los títulos *Statistical Reasoning for Everyday Life, Introduction to Technical Mathematics* y *Business Statistics.* También diseñó el programa estadístico de cómputo original STATDISK y ha escrito diversos manuales y libros de trabajo para educación en estadística con soporte tecnológico. Fuera del salón de clases, su trabajo de consultoría incluye el diseño matemático 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 y en el análisis de resultados de elecciones en disputa. Por otro lado, fungió como testigo experto en la Suprema Corte del estado de Nueva York, en un litigio electoral que implicaba a un antiguo alumno. Asimismo, fue miembro del equipo de escritores del proyecto Coalición de la NASA y de la American Mathematics Association of Two-Year Colleges.

Fuera del trabajo, a Marty le gusta viajar, jugar golf, tenis, los ejercicios aeróbicos, el montañismo y cualquier actividad que implique volar. Posee una licencia de piloto comercial con calificación para instrumentos; vuela aviones, helicópteros, planeadores, alas delta y globos aerostáticos. Su pasión por el vuelo incluye saltos en paracaídas, un viaje en un dirigible Goodyear y el uso de un ala paracaídas.

La Text and Academic Authors Association otorgó a Mario F. Triola un *Texty* de Excelencia por su trabajo en el libro *Estadística*.

Prefacio

Acerca de este libro

Aunque se actualizó gran parte de la novena edición de ***Estadística***, el objetivo primordial continúa siendo el mismo: proporcionar el mejor libro de introducción a la estadística, tanto para estudiantes como para profesores. Dicha meta se logra por medio de factores como un estilo ameno de escritura, un contenido que refleja los aspectos importantes de un curso moderno de introducción a la estadística, el uso de las herramientas tecnológicas más recientes, conjuntos reales e interesantes de datos, abundantes componentes pedagógicos y una batería de complementos. El texto sigue las recomendaciones y los 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

Estadística se escribió para estudiantes de cualquier carrera. Aun cuando el uso del álgebra es mínimo, los usuarios deben haber cursado al menos una materia de álgebra elemental en la preparatoria o la universidad. En muchos casos se agregan teorías subyacentes, pero el libro no pone énfasis en el rigor matemático que se adecua más para carreras especializadas en matemáticas. Como la gran cantidad de ejemplos y ejercicios cubren una amplia variedad de aplicaciones estadísticas distintas e interesantes, la obra es propia 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, los negocios, la economía, la ingeniería, las humanidades, las ciencias físicas, el periodismo, las comunicaciones y las artes libres.

Tecnología

Estadística, en su novena edición, puede utilizarse fácilmente sin referencia a tecnología específica alguna. Muchos profesores continúan usando las distintas ediciones con sus alumnos, con la ayuda de una variedad de calculadoras científicas. Sin embargo, para aquellos que deciden complementar el curso con herramientas tecnológicas específicas, éstas se incluyen tanto en el texto como en los materiales complementarios.

Cambios en la organización

- En el capítulo 5, las secciones 5-3 y 5-4 de la octava edición (distribuciones normales no estándar) ahora se combinan en la sección 5-3. El cambio lo motivo el nuevo formato de la tabla A-2, que facilita a los estudiantes el trabajo con las distribuciones normales.
- En el capítulo 5 viene una nueva sección, la 5-4, que describe "las distribuciones y los estimadores de muestreo".
- En los capítulo 6, 7 y 8 los procedimientos para los intervalos de confianza y la comprobación de hipótesis inician con *proporciones*, que los estudiantes suelen considerar más interesantes que las medias. Además, los procedimientos para trabajar con proporciones son más simples, por lo que permiten a los estudiantes concentrarse más en los nuevos métodos de estadística inferencial.
- La sección 6-4 de la octava edición (tamaño de la muestra que se requiere para estimar μ) se incluye en la sección 6-3 (estimación de una media poblacional: σ conocida), junto con los intervalos de confianza que se utilizan para estimar una media poblacional μ.
- Como los profesores incluyen el tema del control estadístico de proceso con menor frecuencia que el tema de la estadística no paramétrica, ambos se modificaron de tal manera que el capítulo 12 cubre los métodos de estadística no paramétrica y el capítulo 13, el control estadístico de procesos.

Cambios en el contenido

- **Procedimientos** En los capítulos 6, 7 y 8 se presenta un cambio de "$n > 30$" a "σ conocida", como criterio clave para elegir entre la distribución normal y la distribución t. Tal cambio refleja la práctica común que utilizan los profesionales, proporciona resultados más precisos y es mejor para los estudiantes que continuarán otros cursos de estadística; además, no es mucho más difícil que el uso del criterio "$n > 30$".
- **Tablas** Ahora hay un nuevo formato para la importante distribución normal en la tabla A-2: las áreas que se acumulan en la izquierda se listan en dos páginas. Por lo general, los estudiantes consideran que dicho formato es más fácil de usar. La tabla A-3 se expandió para incluir tamaños más grandes de muestras para la distribución t de Student.
- **Notación** En la comprobación de hipótesis ya no se utilizan los símbolos $\geq$ y $\leq$ en las expresiones de la hipótesis nula. En el caso de aseveraciones sobre un valor específico de un parámetro, sólo se utiliza el símbolo de igual ($=$). Este cambio refleja la práctica que emplea la inmensa mayoría de profesionales que aplican métodos estadísticos y que reportan hallazgos en revistas científicas.
- **Conjuntos de datos** El Apéndice B comprende 30 conjuntos de datos (en lugar de 20), incluyendo 14 nuevos.
- **Iconos** Los iconos de herramientas tecnológicas T ahora se utilizan para identificar ejercicios que se basan en conjuntos más grandes de datos del Apéndice B, los cuales se realizan mejor usando un programa de computación o una calculadora TI-83 Plus.
- **Interpretación de resultados** A lo largo del libro, ahora se pone mayor énfasis en la *interpretación* de los resultados. En lugar de obtener simplemente las respuestas, se consideran sus implicaciones y consecuencias. Por ejemplo, en el

tema de la probabilidad, en el capítulo 3, en lugar de sólo calcular los valores de probabilidad, los interpretamos estableciendo diferencias entre eventos comunes y eventos extraños. En la comprobación de hipótesis, no sólo finalizamos con una conclusión de rechazo o no rechazo de la hipótesis nula, sino que procedemos a establecer una conclusión práctica que pone énfasis en el resultado real. Se anima a los estudiantes a *pensar* acerca de las implicaciones de los resultados y no a obtener resultados que se parecen a una receta de cocina y carecen de sentido.

Contenido flexible

La organización del libro refleja las preferencias de la mayoría de los profesores de estadística, pero es posible realizar fácilmente dos variaciones con esta novena 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 2. *Las secciones 9-2 (correlación) y 9-3 (regresión) llegan a cubrirse en las primeras etapas.* Sólo hay que omitir el apartado de la sección 9-2, que se identifica con claridad como "Prueba formal de hipótesis" (que requiere el estudio previo del capítulo 7).
- **Variaciones en el tema de probabilidad:** Algunos profesores consideran que el tema de probabilidad debe cubrirse de forma extensa, mientras que otros piensan que la cobertura tiene que ser mínima. Estos últimos llegan a incluir la sección 3-2 y omitir las secciones restantes del capítulo 3, ya que no son esenciales para los capítulos siguientes. Muchos profesores prefieren cubrir sólo los fundamentos de la probabilidad, junto con los aspectos básicos de las reglas de la suma y de la multiplicación; la cobertura de la regla de la multiplicación (secciones 3-4 y 3-5) ofrece dicha flexibilidad ahora.

Ejercicios

Se presentan más de 1500 ejercicios, *¡más del 58 por ciento de ellos nuevos!* En respuesta a las peticiones de los usuarios de la edición previa, ahora se pusieron más ejercicios simples que se basan en conjuntos pequeños de datos. Muchos más de los ejercicios requieren la interpretación de los resultados. Ya que los ejercicios son de gran importancia en cualquier libro de estadística, se tuvo gran cuidado de asegurar su utilidad, relevancia y exactitud. Tres especialistas en estadística leyeron el material con cuidado, en las etapas finales del libro, para verificar la precisión del texto y de las respuestas a los ejercicios. Estos últimos se acomodaron en orden de dificultad creciente para dividirlos en dos grupos: 1. Destrezas y conceptos básicos, y 2. Más allá de lo básico, que se integró con ejercicios que incluyen conceptos más difíciles o que requieren de un mayor acervo matemático. En pocos casos, dichos ejercicios también introducen un concepto nuevo.

Datos reales: *El 64% de los ejercicios utilizan datos reales*. Como el uso de datos reales es tan importante para los estudiantes, se dedicaron cientos de horas para encontrar información real, significativa e interesante. Además de los datos reales que se incluyeron a lo largo del libro, muchos ejercicios se refieren a los 30 conjuntos de datos que se listan en el Apéndice B.

Características distintivas

Más allá del estilo interesante y accesible (y en ocasiones humorístico) de la redacción, se tuvo cuidado en asegurar que cada capítulo de ***Estadística*** ayude a los alumnos a comprender los conceptos que se presentan. Las siguientes características se diseñaron para cumplir ese objetivo:

- **Inicio del capítulo:** Se incluye una lista de secciones que introducen el capítulo al estudiante; un problema que inicia el capítulo, que se basa en datos reales, motiva el material que se introduce; la primera sección implica un panorama general que establece los objetivos del capítulo.
- **Fin del capítulo:**

  Un **repaso del capítulo** resume los conceptos y temas principales;

  los **ejercicios de repaso** proporcionan práctica respecto de los conceptos y procedimientos;

  los **ejercicios de repaso acumulativo** refuerzan el material previo;

  de los DATOS *a la* DECISIÓN

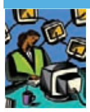
  

  la sección que se denomina **De los datos a la decisión: pensamiento crítico** incluye un problema que requiere de un pensamiento crítico y de un componente de redacción;

  las **actividades de cooperación en equipo** animan en el aprendizaje activo grupal;

  los **proyectos con herramientas tecnológicas** se diseñaron para utilizar STATDISK, Minitab, Excel o la calculadora TI-83 Plus;

  PROYECTO DE INTERNET

  los **proyectos de Internet** ponen en contacto al estudiante con conjuntos de datos de Internet y, en algunos casos, con programas de aplicación.
- **Ensayos al margen:** El texto incluye 120 ensayos al margen, que ilustran los usos y abusos de la estadística en aplicaciones reales, prácticas e interesantes. Incluyen temas tales como "¿Prevalece un género en las familias?", "Precisión del conteo de votos", "Prueba de la terapia de contacto" 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.
- **Programas estadísticos de cómputo:** A lo largo del libro se encuentran instrucciones y resultados de STATDISK, Minitab, Excel y TI-83 Plus.
- **Conjuntos de datos reales:** En todo el libro se utilizan datos reales profusamente. En el Apéndice B se listan 30 conjuntos de datos, 14 de los cuales son nuevos. Dichos conjuntos aparecen de forma impresa en el Apéndice B, así como en forma electrónica en el sitio de Internet y en el disco compacto que se incluye al final del libro. Se agregan temas tan varios como las edades de los polizontes del *Queen Mary*, el uso de alcohol y tabaco en películas infantiles animadas, las erupciones del géiser Old Faithful, las características y el precio de los diamantes, así como datos financieros y de audiencia de películas.
- **Entrevistas:** Cada capítulo incluye una entrevista que realizó el autor a hombres y mujeres profesionales de diversos campos, quienes utilizan la estadística en su trabajo diario.
- **Contraportadas de referencia rápida:** La tabla A-2 (la distribución normal) se reproduce en la segunda de forros y la tabla A-3 (distribución *t*) en la tercera de forros. Al final del libro se incluye una tabla de símbolos que permite consultar con rapidez los símbolos clave.

- **CD-ROM:** El CD-ROM fue elaborado por Mario F. Triola y se incluye en cada nuevo ejemplar del texto, además de los conjuntos de datos del Apéndice B (excepto el conjunto de datos 4). Tales conjuntos se almacenan como archivos de texto, hojas de cálculo de Minitab, archivos de SPSS, SAS, hojas de Excel y aplicaciones de la calculadora TI-83 Plus. El disco compacto también trae programas para la calculadora graficadora TI-83 Plus®, el programa estadístico STATDISK (versión 9-1) y el recurso "Add-Inn" de Excel, que se diseñó para incrementar las capacidades de los programas estadísticos de Excel.

STATDISK

Minitab

Excel

TI-83 Plus

Complementos

Los paquetes complementarios del estudiante y profesor buscan conformar el sistema de aprendizaje más completo y útil disponible para un curso de introducción a la estadística. Los profesores deben contactar a su representante local de ventas de Pearson Educación para recibir copias de los exámenes.

Al profesor

- ***Manual de soluciones para el profesor (disponible en inglés)***, escrito por Mario F. Triola y Milton Loyer; contiene soluciones a todos los ejercicios y un programa muestra del curso. ISBN: 0-321-12212-7.

- **Sistema de evaluación (disponible en inglés):** Se cuidó mucho para asegurar el sistema de evaluación más sólido para la nueva edición de *Estadística*. Además de un banco de exámenes impreso, también hay un generador de exámenes computarizado, el **TestGen4.0** y **Quizmaster3.0**, que permite realizar y editar preguntas del banco de exámenes, transferirlas a otros exámenes y obtener impresiones en diversos formatos. El programa también ofrece muchas opciones para organizar y presentar los bancos de exámenes y los exámenes. Por su capacidad de elaboración aleatoria y su generador de exámenes, el **TestGen-EQ** resulta ideal para crear múltiples versiones de exámenes, ya que ofrece mayor posibilidad de reactivos que las preguntas impresas del banco de exámenes. Sus poderosas funciones de búsqueda y combinación permiten al profesor localizar con facilidad preguntas y presentarlas en el orden que se prefiera. Los usuarios tienen la posibilidad de exportar los exámenes como archivos de texto, de tal modo que éstos pueden leerse en un navegador de Internet. Además, las pruebas que se crearon con TestGen son compatibles con el **QuizMaster**, que permite al estudiante resolver exámenes con la ayuda de una computadora. QuizMaster califica los exámenes de forma automática, almacena los resultados en disco y permite al profesor revisar e imprimir gran diversidad de reportes de estudiantes, clases o cursos. **Printed Testbank** ISBN: 0-321-12214-3; TestGen-EQ para Mac y Windows ISBN: 0-321-12213-5.

- **CD con presentación de conferencia en Power Point®(disponible en inglés):** Gratuito para los clientes que cubran los requisitos, dicho programa de presentación de conferencias para el salón de clase se diseñó específicamente para la secuencia y filosofía de ***Estadística***. Incluye los gráficos clave del libro. Las diapositivas también están disponibles en el sitio Web de Triola www.pearsoneducacion.net/triola. Para Mac y Windows ISBN: 0-321-12215-1.

AL ESTUDIANTE

- **Vídeos (disponible en inglés):** Se diseñaron para complementar muchas secciones del libro; el autor presenta en ellos varios de los temas. Los vídeos ejemplifican todas las herramientas tecnológicas que aparecen en el libro de texto. Los vídeos son un excelente recurso para aquellos estudiantes que han perdido clases o que desean revisar un tema. También es un buen recurso para los profesores que trabajan en programas de aprendizaje a distancia, así como para el estudio individual o los programas de sistemas abiertos. **Cintas de vídeo** ISBN: 0-321-12209-7; **Digital Video Tutor** (versión en CD-ROM) ISBN: 0-321-12231-3.
- **Página en Internet de *Estadística* de Triola:** Se tiene acceso a este sitio en http://www.pearsoneducacion.net/triola. El sitio ofrece proyectos de Internet, relacionados con cada uno de los capítulos del texto, y los conjuntos de datos tal como aparecen en el disco compacto.

Los siguientes manuales tecnológicos incluyen instrucciones y ejemplos del uso de las herramientas tecnológicas. Cada uno se anotó en correspondencia con el libro de texto.

- ***Excel® Manual de laboratorio del estudiante y libro de trabajo (disponible en inglés)***, escrito por Johanna Halsey y Ellena Reda (Dutchess Community College). ISBN: 0-321-12206-2.
- ***Minitab® Manual de laboratorio del estudiante y libro de trabajo***, escrito por Mario F. Triola. ISBN: 0-321-12205-4.
- ***SAS Manual de laboratorio del estudiante y libro de trabajo***, escrito por Joseph Morgan (DePaul University). ISBN: 0-321-12727-7.
- ***SPSS® Manual de laboratorio del estudiante y libro de trabajo***, escrito por Roger Peck (California State University, Bakersfield). ISBN: 0-321-12207-0.
- ***STATDISK Manual de laboratorio del estudiante y libro de trabajo***, escrito por Mario F. Triola. ISBN: 0-321-12216-X.
- ***TI-83 Plus® Compañero de* Estadística**, de Marla Bell (Kennesaw State University). ISBN: 0-321-12208-9.
- ***ActivStats®*, versión de Triola**, que elaboraron Paul Velleman y Data Description, Inc.; ofrece cobertura completa de temas introductorios a la estadística en CD-ROM, con el uso de diversos recursos multimedia. *ActivStat* integra vídeo, simulación, animación, narración, texto, experimentos interactivos, acceso a la Web y Data Desk®, un programa estadístico de cómputo. En el CD-ROM se agregaron problemas y conjuntos de datos para tareas. *ActivStats* para Windows y Macintosh ISBN: 0-201-77139-X. También está disponible en versiones para Excel, JMP, Minitab y SPSS. Consulte a su representante de ventas de Pearson Educación para obtener detalles o consulte el sitio de Internet www.aw.com/activstats.
- La ***Edición del estudiante de Minitab*** **(disponible en inglés)** es una versión condensada del programa estadístico de cómputo profesional de Minitab. Ofrece a los estudiantes la gama completa de métodos estadísticos y capacidades gráficas de Minitab, además de las hojas de cálculo que llegan a contener hasta 5000 puntos de datos. Se acompaña de un manual del usuario, que trae estudios de caso y tutoriales prácticos, mientras que su uso es perfecto para cualquier curso de introducción a la estadística, incluyendo los de ciencias sociales. La versión del estudiante disponible actualmente es la ***Edición del estudiante de Minitab, Release 12 para Windows 95/98 NT.*** ISBN: 0-201-39715-3.

Cualquiera de estos productos puede adquirirse por separado o junto con los libros de texto de Pearson Educación. Para los profesores es posible ponerse en contacto con el representante de ventas local para conocer los detalles sobre la compra de los complementos del libro de texto.

Agradecimientos

El éxito de ***Estadística*** se debe a los esfuerzos de muchos profesionales comprometidos. Quiero expresar mi más sincero agradecimiento a Deirdre Lynch, Christine O'Brien, Greg Tobin, Peggy McMahon, Barbara Atkinson, Brenda Bravener, Yolando Cossio, Karen Guardino, Jen Kerber, Sara Anderson, Keren Blankfeld, Joe Vetere, Joanne Ha, Weslie Lewis y Janet Nuciforo, así como a todo el equipo de Addison-Wesley.

Agradezco también a Paul Lorczak por su trabajo en los proyectos de Internet. A Dale Peterson, de la United States Air Force Academy, por proporcionar excelentes sugerencias para esta nueva edición. A Don Puretz, Heather Carielli y Ashley Mc-Connaughay por ayudar con los conjuntos de datos. Estoy agradecido por las muchas sugerencias y por los comentarios que proporcionaron instructores y estudiantes, así como a los cientos de investigadores que realizaron estudios, llevaron a cabo encuestas, y compilaron datos interesantes y significativos que se usaron en los ejemplos y ejercicios a lo largo de este libro.

También tengo el orgullo de agradecerle a mi esposa, Ginny, y a mis hijos Marc y Scott por su ayuda y estímulo.

Además, me gustaría dar las gracias a las siguientes personas por su ayuda con la novena edición:

Revisores de la precisión del texto
David R. Lund, University of Wisconsin en Eau Claire
Tim Mogill
Kimberly Polly, Parkland College
John Ritschdorff, Marist College
Twin Prime Editorial
Tom Wegleitner

Revisores de la novena edición:
Dan Abbey, Broward Community College
Justine Baker, Peirce College

Donald Barrs, Pellissippi State Technical Community College
Marla Bell, Kennesaw State University
Wayne Ehler, Anne Arundel Community College
Joe Gallegos, Salt Lake Community College
Jim Graziose, Palm Beach Community College
Kristin Hartford, Long Beach City College
John Kozarski, Community College of Baltimore County-Catonsville
Tzong-Yow Lee, University of Maryland
Debra Loeffler, Community College of Baltimore County-Catonsville
Caren McClure, Santa Ana College
Carla Monticelli, Camden County Community College
Rick Moscatello, Southeastern Louisiana University
Nasser Ordoukhani, Barry University
Michael Oriolo, Herkimer County Community College
Kimberly Polly, Parkland College
Diann Reischman, Grand Valley State University
Don Robinson, Illinois State University
Ira Rosenthal, Palm Beach Community College—Eissey Campus
Adele Shapiro, Palm Beach Community College
Lewis Shoemaker, Millersville University
Consuelo Stewart, Howard Community College
David Stewart, Community College of Baltimore County-Dundalk
Sharon Testone, Onondaga Community College
Carol Yin, LeGrange College

Extiendo mis sinceros agradecimientos a los siguientes revisores y usuarios de ediciones previas de este libro, por sus sugerencias:

Mary Abkemeier, Fontbonne College
William A. Ahroon, SUNY en Plattsburgh
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 en Eau Claire
James Baker, Jefferson Community College
Anna Bampton, Christopher Newport University
James Beatty, Burlington County College

Philip M. Beckman, Black Hawk College
Marian Bedee, BGSU, Firelands College
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 Blaueuer, 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 Chambless, 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

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
Bob France, Edmonds Community College
Christine Franklin, University of Georgia
Richard Fritz, Moraine Valley Community College
Maureen Gallagher, Hartwick College
Mahmood Ghamsary, Long Beach City College
Tena Golding, Southeastern Louisiana University
Elizabeth Gray, Southeastern Louisiana University
David Gurney, Southeastern Louisiana University
Francis Hannick, Mankato State University
Joan Harnett, Molloy 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 en 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 en Jacksonville
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

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
Vincent Long, Gaston College
Barbara Loughead, National-Louis University
David Lund, University of Wisconsin en Eau Clair
Rhonda Magel, North Dakota State University—Fargo
Gene Majors, Fullerton College
Hossein Mansouri, Texas State Technical College
Virgil Marco, Eastern New Mexico University
Joseph Mazonec, Delta College
Caren McClure, Rancho Santiago Community 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 en Jacksonville
Kathleen Mittag, University of Texas—San Antonio
Mitra Moassessi, Santa Monica College
Charlene Moeckel, Polk Community College
Theodore Moore, Mohawk Valley Community College
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 en Jacksonville-South
DeWayne Nymann, 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
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
Richard J. Pulskamp, Xavier University
Vance Revennaugh, Northwestern College
C. Richard, Southeastern Michigan College
Sylvester Roebuck, Jr., Olive Harvey College
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
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 Costra College
Richard Stephens, University of Alaska Southeast
W. E. Stephens, McNeese State University
Terry Stephenson, Spartanburg Methodist College
Consuelo Stewart, Howard Community College

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
Gail Wiltse, St. Johns River Community College
Odell Witherspoon, Western Piedmont Community College
Jean Woody, Tulsa Junior College
Thomas Zachariah, Loyola Marymount University
Elyse Zois, Kean College of New Jersey

M.F.T.
LaGrange, Nueva York
Septiembre, 2002

Resumen de contenido

Contenido

ESTADÍSTICA

NOVENA EDICIÓN

1

Introducción a la estadística

PROBLEMA DEL CAPÍTULO

¿Qué podemos aprender de estas encuestas?

A continuación se presentan descripciones breves de cinco encuestas diferentes:

1. A mediados de diciembre de un año reciente, el proveedor de servicios de Internet America Online (AOL) realizó una encuesta entre sus usuarios. La siguiente pregunta se refería a los árboles de Navidad:

   ¿Cuál prefiere?

   - un árbol natural
   - un árbol artificial

   De entre las 7073 respuestas recibidas de los usuarios de Internet, 4650 prefirieron un árbol natural y 2423 un árbol artificial.

2. La revista *Newsweek* hace poco realizó una encuesta acerca del controvertido sitio de Internet llamado Napster, que ofrecía acceso gratuito para copiar discos compactos de música. Se planteó la siguiente pregunta a los lectores:

   ¿Continuaría utilizando Napster si tuviese que pagar?

   Los lectores podían registrar sus respuestas en el sitio de Internet www.newsweek.msnbc.com. De las 1873 respuestas recibidas, el 19% dijo que sí, ya que aun así resultaría más barato que comprar los discos compactos originales. Otro 5% dijo que sí, que se sentiría más cómodo al utilizarlo si lo pagaba.

3. La revista *Good Housekeeping* invitó a mujeres para que visitaran su página en Internet para contestar una encuesta, y se registraron 1500 respuestas. Cuando se les preguntó si preferían tener más dinero o dormir más, el 88% eligió más dinero y sólo el 11%, dormir más.

4. *USA Today* realizó una "Encuesta sobre el cuidado de la salud" de $^3/_4$ de página. A los lectores se les pedía lo siguiente: "Por favor, tómese un momento para llenar esta encuesta y envíenosla". La mayoría de las preguntas se referían a las condiciones de salud, y al consumo de tabaco y de medicamentos de prescripción. La pregunta 17 de la encuesta era: "¿Podríamos establecer contacto nuevamente con usted para que participe en otras encuestas de *USA Today*?".

5. *USA Today* publicó una gráfica de barras con los resultados de una encuesta donde se preguntó a los lectores: "¿Tiene planes para tomar unas vacaciones?". De los 4264 usuarios de Internet que decidieron responder, el 48% dijo que aún no tenía planes, y el 14% contestó que planeaba ir a alguna playa.

¿Qué característica importante tienen en común estas cinco encuestas? Con base en los resultados obtenidos en ellas, ¿cómo se ven afectadas nuestras conclusiones respecto de la población general? ¿Podríamos concluir que la mayoría de los estadounidenses prefieren un árbol de Navidad real que uno artificial? ¿Concluiríamos que la gran mayoría de las mujeres estadounidenses prefieren más dinero que dormir más? ¿O que la gran mayoría de las mujeres lectoras de la revista *Good Housekeeping* prefieren más dinero que dormir más? Las respuestas a tales preguntas son de crucial importancia para evaluar los resultados de las encuestas. El asunto a considerar aquí es el tema más importante de todo este capítulo y podría ser el aspecto más relevante de todo el libro.

En este capítulo estudiaremos temas relevantes sobre la validez de encuestas como las anteriormente descritas. Veremos que con frecuencia sacamos conclusiones relevantes con la simple aplicación del sentido común. Al final de este capítulo, seremos capaces de identificar los aspectos clave que afectan la validez de las cinco encuestas y lograremos una profunda comprensión de los métodos de recolección de datos en general.

1-1 Panorama general

El Problema del capítulo en la página anterior implica a las encuestas. La encuesta es una de muchas herramientas disponibles para recolectar datos. Una meta común de las encuestas es 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 estadística es aprender acerca de un grupo examinando los datos de algunos de sus miembros. En dicho contexto los términos *muestra* y *población* adquieren importancia. Las definiciones formales de éstos y otros términos básicos se presentan a continuación.

Definiciones

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

Estadística es una colección de métodos para planear experimentos, obtener datos, y después organizar, resumir, presentar, analizar, interpretar y llegar a conclusiones basadas en los datos.

Población es la colección completa de todos los elementos (puntuaciones, personas, mediciones, etcétera) a estudiar. Se dice que la colección es completa, pues incluye a todos los sujetos que se estudiarán.

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

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

Por ejemplo, un sondeo de Gallup preguntó a 1087 adultos: "¿Consume bebidas alcohólicas como licor, vino o cerveza o es abstemio?". Los 1087 sujetos de la encuesta constituyen una *muestra* mientras que la *población* consiste en el conjunto de los 202,682,345 estadounidenses adultos. Cada 10 años el gobierno de Estados Unidos intenta obtener un *censo* de cada ciudadano; pero no logra hacerlo porque es imposible localizar a cada uno de ellos. En la actualidad hay polémica en torno al intento de emplear métodos estadísticos acertados para aumentar la exactitud del censo, aunque los aspectos políticos constituyen un factor clave para que los miembros del Congreso se resistan a esta mejoría. Quizás algún día algunos lectores de este texto sean miembros del Congreso y tengan la sabiduría de traer el censo al siglo XXI.

Una función importante de este libro es demostrar cómo utilizar las muestras de datos para llegar a conclusiones respecto de poblaciones. Veremos que es *extremadamente* importante obtener datos muestrales que sean representativos de la población de la que se tomaron. Por ejemplo, si usted encuesta a los estudiantes graduados de su universidad y les pide que anoten sus ingresos anuales y le envíen la respuesta por correo, es probable que los resultados no sean representativos de la población de todo el alumnado. Aquellos con bajos ingresos estarían menos inclinados a responder y quienes respondan pueden mostrar tendencia a exagerar.

Al avanzar en este capítulo debemos enfocarnos en los siguientes conceptos clave:

- **Los datos muestrales deben reunirse de una forma adecuada, como en un proceso de selección *aleatoria*.**
- **Si los datos muestrales no se reúnen de forma adecuada, resultarían tan inútiles que ninguna cantidad de tortura estadística podría salvarlos.**

Ante todo, le pedimos que comience a estudiar estadística con una mente abierta. No considere que el estudio de la estadística es comparable con un procedimiento inflexible. La experiencia del autor es que los estudiantes a menudo se sorprenden por lo interesante que resulta la estadística y también porque realmente llegan a dominar sus principios básicos sin mucha dificultad, incluso si no han sido sobresalientes en otros cursos de matemáticas. Estamos convencidos de que cuando usted termine este curso introductorio, tendrá la firme creencia de que la estadística es una materia rica e interesante con aplicaciones que son extensivas, reales y significativas. También estamos convencidos de que con la asistencia a clases y la dedicación constantes, usted tendrá éxito para 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 compilación 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 negocios se apoyan bastante en datos estadísticos para dirigir sus acciones. Por ejemplo, se compilan datos cuidadosamente y con regularidad para establecer las tasas de desempleo, las tasas de inflación, los índices del consumidor y las tasas de nacimiento y muerte, y los líderes empresariales utilizan los datos resultantes para tomar decisiones que afectan las futuras contrataciones, los niveles de producción y la expansión hacia nuevos mercados.

1-2 Tipos de datos

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

1. **Parámetro:** Cuando Lincoln fue elegido presidente por primera vez, recibió el 39.82% de 1,865,908 votos. Si suponemos que el conjunto de todos esos votos es la población a considerar, entonces el 39.82% es un *parámetro*, no un estadístico.
2. **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, no en la población completa de todos los ejecutivos.

Algunos conjuntos de datos consisten en números (como estaturas 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 alguna característica no numérica.

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, tales como dólares, horas, pies, metros y otras. 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 en *kilogramos*". Ignorar unidades de medida como éstas podría llevar a conclusiones incorrectas. La NASA perdió su Mars Climate Orbiter de 125 millones de dólares cuando la sonda se estrelló, porque la programación 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 *discretos* y *continuos*.

Definiciones

Datos discretos resultan cuando el número de posibles valores es un número finito, o bien, un número que puede contarse. (Es decir, el número de posibles valores es 0, 1, 2, etcétera).

Datos continuos (numéricos) resultan de un infinito de posibles valores que pueden asociarse a puntos de alguna escala continua, cubriendo un rango de valores sin huecos ni interrupciones.

EJEMPLOS

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

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

Otra forma común de clasificación de los datos es el uso de cuatro niveles de medición: nominal, ordinal, de intervalo y de razón. Cuando la estadística se aplica a problemas reales, el nivel de medición de los datos es un factor importante para determinar el procedimiento a usar. (Véase la figura 14.1 en la página 727.) En este libro encontraremos algunas referencias a estos niveles de medición; sin embargo, lo importante aquí es sustentarse en el sentido común: no hay que hacer cálculos ni usar métodos estadísticos con datos que no sean apropiados. Por ejemplo, no tendría sentido calcular un promedio de números del seguro social, ya que estos números son datos que se usan como identificación, y no representan mediciones ni 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

Nivel de medición nominal son los datos consistentes exclusivamente en nombres, etiquetas o categorías que no pueden acomodarse según un esquema de orden (por ejemplo, de bajo a alto).

EJEMPLOS Los ejemplos siguientes ilustran datos muestrales en el nivel de medición nominal.

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

Puesto que los datos nominales carecen de un orden o de un significado numérico, no pueden utilizarse para realizar cálculos. A veces se asignan números a las diferentes categorías (en especial cuando los datos se codifican para el uso de sistemas de cómputo), pero tales números no tienen significado computacional y cualquier promedio que se calcule carece de sentido.

Medición de la desobediencia

¿De qué manera se recolectan datos que parecen imposibles de medir, como el nivel de desobediencia de las personas? El psicólogo Stanley Milgram ideó 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 volts y fue instruido para incrementar los choques en aumentos de 15 volts. El nivel de desobediencia fue el punto donde el sujeto se negaba a incrementar el voltaje. Resultó sorprendente que dos terceras partes de los sujetos obedecieran las órdenes aun cuando el actor gritaba y fingía sufrir un ataque cardiaco.

Apuesta por la ciencia

En ocasiones los datos se recolectan de maneras muy ingeniosas y de fuentes muy extrañas. Un ejemplo es el de ciertos investigadores que estudiaron los cambios climáticos. Ellos se dieron cuenta de que cada primavera, desde 1917, en la pequeña ciudad de Nenana, Alaska, hacían un juego de lotería, en el cual las personas apostaban sobre la hora exacta en que la capa de hielo del río Tanana se rompería (el último premio fue de cerca de 300,000 dólares). Se colocó un tripié en el río congelado y éste se conectó a un reloj. El reloj se detendría cuando el hielo, al quebrarse, moviera el tripié. De esta forma los investigadores supieron el momento preciso en que ocurría la rotura cada año desde 1917, y los datos resultaron muy útiles en el estudio de las tendencias climáticas.

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 Los siguientes son ejemplos de datos muestrales en el nivel de medición ordinal.

1. **Las calificaciones de un curso:** Un profesor universitario asigna calificaciones de A, B, C, D, o E, las cuales pueden acomodarse en orden; sin embargo, no es posible determinar diferencias entre ellas. Por ejemplo, sabemos que A es más alto que B (por lo tanto, existe un orden), pero no podemos restar B de A (por lo tanto, no se calcula la diferencia).
2. **Rangos ordenados:** Con fundamento en varios criterios, una revista clasificó las ciudades de acuerdo con su "calidad de vida"**.** Tales rangos (primero, segundo, tercero, etcétera) determinan un orden; sin embargo, las diferencias entre los rangos ordenados no tienen significado alguno. Por ejemplo, una diferencia de "segundo menos primero" puede sugerir 2 – 1 = 1, pero este resultado de 1 no tiene significado porque no es una cantidad exacta que pueda compararse con otras diferencias del mismo tipo. La diferencia entre la primera ciudad y la segunda no es la misma que la diferencia entre la segunda y la tercera. Utilizando los rangos ordenados de la revista, la *diferencia* entre las ciudades de Nueva York y Boston no puede compararse cuantitativamente con la *diferencia* entre las ciudades de San Luis y Filadelfia.

Los datos ordinales ofrecen información sobre comparaciones relativas, aunque no sobre las magnitudes de las diferencias. Por lo general, los datos ordinales no se usan para cálculos como un promedio, pero esta norma se quebranta en ocasiones (como cuando se usan calificaciones con letras para calcular el punto promedio de calificación).

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 un punto de partida inherente (*natural)* desde cero (donde *nada* de la cantidad esté 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 en este nivel de medición. Tales valores están ordenados, y podemos determinar su diferencia de 0.4°F. Sin embargo, no existe un punto de partida natural. El valor de 0°F quizá parezca un punto de partida, pero es arbitrario y no representa la ausencia total de calor. Como 0°F no es

un punto de partida desde cero natural, es erróneo decir que 50°F es *dos veces* más caliente que 25°F.

2. **Años:** Los años 1000, 2000, 1776 y 1492. (El tiempo no inició en el año 0, así que el año 0 es arbitrario en vez de ser un punto de partida de cero natural, que representaría "ausencia de tiempo").

Definición

El **nivel de medición de razón** se parece al nivel de intervalo, aunque tiene la propiedad adicional de que sí tiene un punto de partida o cero inherente (donde cero indica que *nada* de la cantidad está presente). Para valores en este nivel, tanto las diferencias como las proporciones tienen significado.

EJEMPLOS Los siguientes son ejemplos de datos en el nivel de medición de razón. Observe la presencia del valor cero natural y el uso de proporciones que significan "dos veces" y "tres veces".

1. **Pesos:** Los pesos (en quilates) de anillos engastados con diamante (0 efectivamente representa ausencia de peso y 4 quilates es dos veces el peso de 2 quilates).
2. **Precios:** Los precios de los libros de texto universitarios ($0 efectivamente representa ningún costo y un libro de $90 es tres veces más costoso 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 mayoría de las dificultades surgen con la distinción entre los niveles de intervalo y de razón.

Sugerencia: Para hacer más fácil esta distinción, utilice una sencilla "prueba de razón": considere dos cantidades en las cuales un número es dos veces el otro y pregúntese si "dos veces" se puede usar 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 en la página siguiente, que señala las diferencias entre los cuatro niveles de medición.

1-2 Destrezas y conceptos básicos

En los ejercicios 1 a 4, determine si el valor dado es un estadístico o un parámetro.

1. El Senado actual de Estados Unidos consta de 87 hombres y 13 mujeres.
2. Se selecciona una muestra de estudiantes y el número promedio (media) de libros de texto comprados este semestre es 4.2.
3. Se toma una muestra de estudiantes y el promedio (media) de la cantidad de tiempo de espera en la fila para comprar libros de texto este semestre es 0.65 horas.
4. En un estudio de los 2223 pasajeros del *Titanic*, se encontró que 706 sobrevivieron cuando se hundió.

Tabla 1-1 Niveles de medición de datos

Nivel	Resumen	Ejemplo	
Nominal	Sólo rangos de orden. Los datos no pueden acomodarse en un esquema de orden.	Origen de estudiantes: 5 californianos 20 texanos 40 neoyorquinos	Sólo rangos de orden o nombres.
Ordinal	Rangos de orden que pueden acomodarse, 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 son significativas, pero no hay punto de partida natural y las razones no tienen significado.	Temperaturas del campus: 5°F 20°F 40°F	0°F no es "sin calor". 40°F no es dos veces más caliente que 20°F.
De razón	Hay un punto de partida natural y las razones tienen significado.	Distancias de viaje de estudiantes: 5 km 20 km 40 km	40 km es *dos veces* más lejos que 20 km.

En los ejercicios 5 a 8, determine si los valores dados provienen de un conjunto de datos discreto o continuo.

5. El salario presidencial de George Washington era de 25,000 dólares anuales y el salario presidencial actual es de 400,000 anuales.

6. Un estudiante de estadística obtiene datos muestrales y encuentra que la media del peso de automóviles en la muestra es 3126 libras.

7. En una encuesta de 1059 adultos, se encontró que el 39% de ellos tienen pistolas en sus casas (de acuerdo con una encuesta de Gallup).

8. Cuando se probaron 19,218 máscaras antigás de divisiones de la milicia de Estados Unidos, se encontró que 10,322 estaban defectuosas (de acuerdo con datos de la revista *Time*).

En los ejercicios 9 a 16, determine cuál de los cuatro niveles de medición (nominal, ordinal, de intervalo, de razón) es el más apropiado.

9. Las estaturas de las mujeres que juegan básquetbol en la WNBA.

10. Las calificaciones de fantástico, bueno, promedio, pobre o inaceptable en citas a ciegas.

11. Las temperaturas actuales en los salones de clase en su universidad.

12. Los números en las camisetas de las mujeres que juegan básquetbol en la WNBA.

13. Las calificaciones de la revista *Consumer Reports* de “mejor compra, recomendado, no recomendado”.

14. Los números del seguro social.

15. El número de respuestas “sí” recibidas cuando se les preguntó a 1250 conductores si habían usado alguna vez un teléfono celular mientras conducían.

16. Los códigos postales de la ciudad en que vive.

En los ejercicios 17 a 20, identifique a) *la muestra y* b) *la población. También determine si la muestra parece ser representativa de la población.*

17. Un reportero de *Newsweek* se para en una esquina y pregunta a 10 adultos si creen que el presidente actual está haciendo un buen trabajo.

18. Nielsen Media Research encuesta a 5000 amas de casa seleccionadas al azar y encuentra que el 19% de los televisores encendidos están sintonizados en *60 minutos* (de acuerdo con datos de *USA Today*).

19. En una encuesta de Gallup aplicada a 1059 adultos seleccionados aleatoriamente, el 39% respondió “sí” cuando se le preguntó: “¿Tiene usted una pistola en su casa?”.

20. Una estudiante graduada de la Universidad de Newport realizó un proyecto de investigación acerca de cómo se comunican los adultos estadounidenses. Empezó por una encuesta que envió por correo a 500 de los adultos que ella conocía. Les pidió que le enviaran por correo la respuesta a esta pregunta: “¿Prefiere usted usar el correo electrónico o el correo tortuga (el servicio postal estadounidense)?”. Ella recibió a vuelta de correo 65 respuestas y 42 de ellas indicaron una preferencia por el correo tortuga.

1-2 Más allá de lo básico

21. Interpretación de los incrementos de temperatura En la tira cómica “Born Loser” de Art Sansom, Brutus se alegra por un incremento en la temperatura de 1° a 2°. Cuando alguien 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 está equivocado una vez más.

22. Interpretación de encuesta política Un encuestador aplica una encuesta a 200 personas y les pregunta por el partido político de su preferencia: él codifica las respuestas como 0 (para demócrata), 1 (para republicano), 2 (para independiente) y 3 (para otras respuestas cualesquiera). Entonces calcula el promedio (media) de los números y obtiene 0.95. ¿Cómo se interpreta este valor?

23. Escala para calificar comida Un grupo de estudiantes desarrolló una escala para calificar la calidad de la comida de la cafetería de su escuela, donde 0 representaba “neutral: ni buena ni mala”. Se asignaron números negativos a las comidas malas y números positivos a las comidas buenas; la magnitud del número correspondía a la severidad de lo bueno o lo malo. Las primeras tres comidas se calificaron con 2, 4 y −5. ¿Cuál es el nivel de medición de calificaciones como éstas? Explique su respuesta.

1-3 Pensamiento crítico

El éxito en el curso introductorio de estadística por lo regular requiere de más *sentido común* que destreza matemática (a pesar de la advertencia de Voltaire de que “el sentido común no es muy común”). Ya que ahora tenemos acceso a calculadoras y a computadoras, las aplicaciones modernas de la estadística ya no requieren que dominemos algoritmos complejos de operaciones matemáticas. En su lugar, nos enfocamos en la *interpretación* de los datos y los resultados. Esta sección está diseñada para ilustrar la forma en que se usa el sentido común cuando pensamos de forma crítica acerca de los datos y la estadística.

Hace cerca de un siglo, el estadista Benjamin Disraeli pronunció la famosa frase: "Hay tres clases de mentiras: mentiras, viles mentiras y estadísticas". También se ha dicho que "las cifras no mienten; los mentirosos calculan las cifras". El historiador Andrew Lang dijo que algunas personas utilizan 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 puede usarse para sustentar cualquier cosa, en especial a los estadísticos". En el *Esar's Comic Dictionary* encontramos la definición de que un estadístico es "un especialista que reúne pensamientos y luego los conduce al extravío". Estas afirmaciones se refieren a ejemplos donde los métodos estadísticos se utilizaron de forma errónea, de manera que resultaron engañosos en última instancia. Hay dos fuentes principales de tal engaño: **1.** el intento malintencionado por parte de personas deshonestas, y **2.** los errores de descuido cometidos por personas que no conocen nada mejor. Sin tener en cuenta la fuente, como ciudadanos responsables y como empleados profesionales valiosos, debemos tener una habilidad básica para distinguir entre conclusiones estadísticas que parecen ser válidas de las que son gravemente defectuosas.

Para mantener esta sección en la perspectiva apropiada, hay que saber que éste no es un libro acerca de los malos usos de la estadística. El resto de este libro estará lleno de usos muy importantes de métodos estadísticos válidos. Aprenderemos métodos generales para usar datos muestrales y así poder hacer inferencias relevantes acerca de poblaciones; aprenderemos acerca de encuestas y tamaños de muestra, acerca de mediciones importantes de características fundamentales de los datos. Junto con las explicaciones de estos conceptos generales, veremos muchas aplicaciones específicas reales, tales como los efectos en el fumador pasivo, el predominio del alcohol y el tabaco en las películas de dibujos animados para niños y la calidad de productos de consumo, incluyendo dulces M&M, cereales, Coca Cola y Pepsi. Pero incluso en estas aplicaciones reales y con significado, debemos ser cuidadosos para interpretar correctamente los resultados de métodos estadísticos válidos.

Comenzamos nuestro desarrollo del pensamiento crítico considerando muestras erróneas. Estas muestras son erróneas en el sentido de que el método de muestreo arruina la muestra, de modo que tiene la posibilidad de estar sesgada (es decir, de no ser representativa de la población de la que se obtuvo). La sección siguiente analiza con más detalle los métodos de muestreo y describe la importancia de la *aleatoriedad*. El primer ejemplo sigue un procedimiento de muestreo que tiene una seria carencia de 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

Muestra de respuesta voluntaria (o **muestra autoseleccionada**) es aquella donde los sujetos deciden ser incluidos por sí mismos.

¿Debe creerse en un estudio estadístico?

En la segunda edición del libro *Statistical Reasoning for Everyday Life*, los autores Jeff Bennet, William Briggs y Mario Triola enumeran las siguientes directrices para evaluar de forma crítica un estudio estadístico: **1.** Identifique la meta del estudio, la población considerada y el tipo de estudio. **2.** Considere la fuente, particularmente respecto de la posibilidad de la existencia de prejuicios. **3.** Analice el método de obtención de muestras. **4.** Busque problemas en la definición o medición de variables de interés. **5.** Tenga cuidado con variables confusas que podrían invalidar las conclusiones. **6.** Considere el escenario y la redacción de cualquier encuesta. **7.** Verifique que las gráficas representen los datos con fidelidad y que las conclusiones tengan justificación. **8.** Considere si las conclusiones logran los objetivos del estudio, si tienen sentido y si tienen un significado práctico.

1 Para ver ejemplos, remítase al problema del capítulo. Cuando America Online o alguien más aplica una encuesta por Internet, los individuos por sí mismos deciden si participan o no, así que constituyen una muestra de respuesta voluntaria. Sin embargo,

existen mayores posibilidades de que las personas con opiniones decididas participen, de tal forma que las respuestas no sean representativas de toda la población. A continuación se presentan ejemplos de muestras de respuesta voluntaria que, por su naturaleza, adolecen de una carencia importante, pues no debemos obtener conclusiones sobre una población con base en una muestra sesgada como ésta:

- Las encuestas manejadas a través de Internet, en las que los sujetos deciden si responden o no.
- Las encuestas por correo, donde los sujetos deciden si contestan.
- Las encuestas telefónicas, en las que anuncios en el periódico, la radio, o la televisión, le piden que tome un teléfono voluntariamente y llame a un número especial para registrar su opinión.

Con muestras de respuesta voluntaria como éstas, sólo es posible llegar a conclusiones válidas acerca del grupo específico que decide participar; pero sería una práctica incorrecta común establecer conclusiones acerca de una población más grande. Desde un punto de vista estadístico, una muestra como ésta falla en lo esencial y no debe utilizarse para realizar declaraciones generales acerca de una población mayor.

Muestras pequeñas Las conclusiones no deben basarse en muestras que son sumamente pequeñas. Por ejemplo, el Children's Defense Fund publicó *Children Out of School in America*, donde se reportó que de los estudiantes de escuela secundaria suspendidos en una región, el 67% fueron suspendidos al menos tres veces. ¡Pero esta cifra está basada en una muestra de sólo *tres* estudiantes! Los reportes en los medios de comunicación fallaron al mencionar que el tamaño de la muestra era muy pequeño. (En los capítulos 6 y 7 veremos que *en ocasiones* es posible realizar algunas deducciones valiosas a partir de muestras pequeñas, aunque debemos ser cuidadosos y verificar que se satisfagan los requisitos necesarios).

En ocasiones una muestra puede parecer relativamente grande (como en una encuesta de "2000 adultos estadounidenses seleccionados al azar"), pero si se obtienen conclusiones acerca de los subgrupos, por ejemplo, los republicanos de sexo masculino de 21 años de edad de Pocatello, tales conclusiones estarían basadas en muestras demasiado pequeñas. Si bien es importante tener una muestra que sea suficientemente grande, también lo es el hecho de tener datos muestrales que se recolecten de una forma adecuada, como la selección aleatoria. Aun las muestras grandes llegan a ser muestras erróneas.

Gráficas Las gráficas —como las de barras y las circulares— en ocasiones sirven para exagerar o disfrazar la verdadera naturaleza de los datos. (En el capítulo 2 analizaremos una variedad de gráficas diferentes). Las dos gráficas en la figura 1-1 de la siguiente página representan *los mismos datos* del Bureau of Labor Statistics, aunque el inciso *b*) está diseñado para exagerar la diferencia entre los salarios semanales de hombres y mujeres. Al no iniciar el eje vertical en cero, la gráfica del inciso *b*) tiende a producir una impresión subjetiva engañosa, que hace que los lectores incorrectamente crean que la diferencia es mucho peor de lo que en realidad es. La figura 1-1 enseña una lección importante: para interpretar una gráfica de manera correcta, debemos analizar la información *numérica* dada en ella, para no engañarnos por su forma general. (El término mediana que se utiliza en la figura 1-1 se describirá con claridad en la sección 2-4).

Pictogramas Los dibujos de objetos, llamados pictogramas, también pueden resultar engañosos. Algunos objetos que se usan comúnmente para representar datos incluyen objetos tridimensionales, como bolsas de dinero, pilas de monedas, tanques militares (para gastos militares), barriles (para producción petrolera) y casas

La encuesta del Literary Digest

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

FIGURA 1-1 Salarios semanales de hombres y mujeres de 16 a 24 años

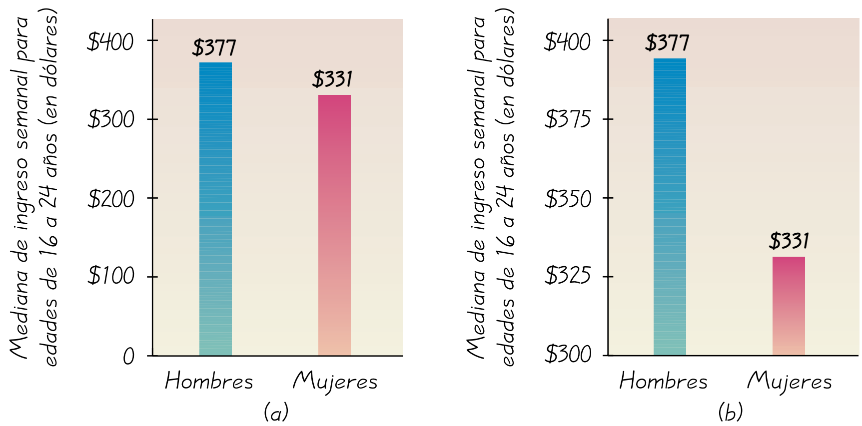


FIGURA 1-2 Pictograma

Duplique el largo, el ancho y la altura de un cubo y el volumen se incrementa por un factor de ocho, como se indica. Si el cubo más pequeño representa los impuestos en un año y el cubo más grande representa el *doble* de los impuestos algún tiempo después, los últimos impuestos parecen ser ocho veces más grande, y no dos, la cantidad original.

(para construcción de viviendas). Al dibujar tales objetos, los artistas llegan a crear impresiones falsas que distorsionan las diferencias. Si duplicamos cada lado de un cuadrado, el área no tan sólo se duplica, sino que aumenta en un factor de cuatro. Si se duplica cada lado de un cubo, el volumen no se duplica simplemente, sino que se incrementa en un factor de ocho, como se observa en la figura 1-2. Si los impuestos se duplican durante una década, un artista podría representar las cantidades de impuestos con una bolsa de dinero para el primer año y otra bolsa de dinero dos veces más ancha, dos veces más alta y dos veces más profunda para el segundo año. En vez de parecer que los impuestos se duplican, parecerá que aumentaron en un factor de ocho y así el dibujo distorsionaría la verdad.

Porcentajes A veces se utilizan porcentajes engañosos o poco claros. Si usted toma el 100% de alguna cantidad, está tomándolo *todo*. (No debería requerir de un 110% de esfuerzo para que la declaración anterior tenga sentido). En referencia a la pérdida de equipaje, la Continental Airlines publicó anuncios afirmando que se trata de “un área en la que ya hemos mejorado un 100% en los últimos seis meses”. En un editorial que criticaba ese dato estadístico, el diario *The New York Times* interpretó correctamente que la cifra de mejora en un 100% significa que ya no se está perdiendo equipaje, 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” significa realmente “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, después multiplique por la cantidad. Este ejemplo muestra que el 6% de 1200 es 72:

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

- **Fracción → Porcentaje:** *Para convertir de una fracción a un porcentaje*, divida el denominador entre el numerador para obtener un número decimal equivalente, y después multiplíquelo por 100 y agregue el símbolo %. Este ejemplo muestra que la fracción 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 demuestra que 0.234 es equivalente a 23.4%:

$$0.234 \rightarrow 0.234 \times 100\% = 23.4\%$$

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

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

Preguntas predispuestas Existen muchos aspectos que afectan las preguntas de una encuesta. Éstas llegan a estar "cargadas" o redactadas intencionalmente de manera que propicien una respuesta deseada. Observe las calificaciones de respuesta "sí" reales para las diferentes redacciones en una pregunta:

- 97% sí: "¿Debe el presidente utilizar su poder de veto para eliminar los desperdicios?".
- 57% sí: "¿Debe el presidente utilizar su poder de veto o no?"

En *The Superpollsters*, David W. Moore describe un experimento donde se preguntó a diferentes sujetos si estaban de acuerdo con las siguientes declaraciones:

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

Aun cuando es el pobre quien recibe 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 cargan de forma no intencional, en virtud de 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 al tránsito y el 27% culpó a la industria; cuando la industria se presentó primero, el 24% culpó al tránsito y el 57% culpó a la industria.

Rechazo Cuando se invita a las personas a contestar una encuesta, algunas se niegan con firmeza a responder. La tasa de rechazo ha crecido en años recientes, en parte porque muchos vendedores persistentes de empresas de telemercadeo buscan vender bienes o servicios comenzando con una inducción de ventas que suena como si fuera parte de una encuesta de opinión. En *Lies, Damn Lies, and Statistics*, el autor Michael Wheeler indica con acierto que "las personas que se niegan a hablar con los entrevistadores parecen ser diferentes de quienes no lo hacen. Algunas quizá tengan miedo a los extraños y otras sean celosas de su privacidad, pero su negativa a hablar demuestra que su visión del mundo circundante es marcadamente diferente de aquellas otras personas que permiten a los entrevistadores entrar a sus hogares".

LA ESTADÍSTICA EN LAS NOTICIAS

La estadística y las minas terrestres

La International Campaign to Ban Land Mines (la Campaña Internacional para Proscribir Minas Terrestres) y el director ejecutivo de la Vietnam Veterans of America Foundation (VVAF) fueron recientemente galardonados con el Premio Nobel de la Paz. Cuando la VVAF pidió ayuda en la recolección de datos acerca de las minas terrestres, se reunió a un equipo de notables estadísticos. En vez de trabajar con datos intangibles, como el valor de la vida humana, ellos trabajaron con datos tangibles en bruto, como es el área que inutiliza un campo minado y el costo de cultivos que no se cosechan. Los datos se incluyeron en el libro *After the Guns Fall Silent: The Enduring Legacy of Landmines*, que vino a ser un recurso clave en las discusiones del tema de las minas terrestres. El *AMSTAT News* citó a uno de los editores del libro: "Este esfuerzo de reunión y análisis de datos es lo que hizo posible presentar el tema ante los legisladores. El trabajo en verdad marcó una diferencia".

Detección de datos falsos

Un maestro de clase asigna la tarea de registrar los resultados de lanzar al aire una moneda 500 veces. Un estudiante deshonesto decide ahorrar tiempo inventando los resultados, en lugar de realmente lanzar la 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 en extremo probable que usted obtenga una serie de seis caras o seis cruces, aunque la gente casi nunca incluye una racha como ésta cuando inventa resultados.

Otra forma de detectar datos "fabricados" consiste en establecer que los resultados violan la ley de Benford: para muchos grupos de datos, los primeros dígitos no están uniformemente distribuidos. Más bien los primeros dígitos de 1, 2, . . . , 9 ocurren con frecuencia de 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).

Correlación y causalidad En el capítulo 9 de este libro analizaremos la asociación estadística entre dos variables, como son la riqueza y el CI. Usaremos el término correlación para indicar que las dos variables están relacionadas. Sin embargo, en el capítulo 9 hacemos esta importante anotación: *la correlación no implica causalidad*. Esto significa que cuando nosotros encontramos una asociación estadística entre dos variables, no podemos concluir que una de las variables es la causa de la otra (o que la afecta directamente). Si encontramos una correlación entre la riqueza y el CI, no podemos concluir que el CI de una persona afecta directamente su riqueza, ni tampoco podemos concluir que la riqueza de la persona afecta directamente su puntuación de CI. En los medios de comunicación es bastante común reportar 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 Algunas veces los estudios reciben el patrocinio de grupos con intereses específicos que buscan promover. Por ejemplo, Kiwi Brands, un fabricante de abrillantador de calzado, encargó un estudio que suscitó esta declaración impresa en algunos periódicos: "De acuerdo con una encuesta nacional realizada a 250 empleadores profesionales, la razón más común del fracaso de un solicitante de trabajo del sexo masculino al dar una buena primera impresión, fue llevar los zapatos desaseados". Debemos ser muy cautos con encuestas como éstas, cuyos resultados generan ganancias económicas para el patrocinador. En los últimos años ha generado preocupación creciente la práctica de las compañías farmacéuticas de financiar a doctores que realizan experimentos clínicos y reportan sus resultados en revistas de prestigio, como *Journal of American Medical Association.*

Números precisos "En la actualidad existen 103,215,027 hogares en Estados Unidos." Puesto que esta cantidad es muy precisa, mucha gente considera erróneamente que también es exacta. En este caso, ese número es un estimado y sería mejor decir que el número de hogares es de alrededor 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 anuncio 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, de modo que la mayoría de esos automóviles que circulaban estaban casi nuevos. La afirmación era técnicamente correcta, aunque muy engañosa, al no presentar 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 ese servicio. Cuando Hertz solicitó información detallada acerca de la encuesta, las respuestas originales de ésta desaparecieron y el coordinador de encuestas de la revista renunció. Hertz demandó a Avis (por publicidad falsa basada en la encuesta) y a la revista; al final las compañías llegaron a un acuerdo.

Además de los casos ya citados, se conocen muchos otros usos incorrectos de la estadística; algunos de estos otros casos se encuentran en libros como el clásico de Darrel Huff, *How to Lie with Statistics*; el de Robert Reichard, *The Figure Finaglers*, y el de Cynthia Crossen, *Tainted Truth*. Comprender tales prácticas resultará extremadamente útil en la evaluación de los datos estadísticos que se encuentran en situaciones cotidianas.

1-3 Destrezas y conceptos básicos

En los ejercicios 1 a 4, utilice el pensamiento crítico para desarrollar una conclusión alternativa. Por ejemplo, considere un reporte de medios de comunicación donde los conductores de BMW gozan de mejor salud que los adultos que no manejan. La conclusión de que los automóviles BMW son la causa de una mejor salud quizás esté equivocada. La siguiente sería una mejor conclusión: Los conductores de BMW tienden a ser más adinerados que los adultos que no manejan y una mayor riqueza se asocia con una mejor atención a la salud.

1. El peso y los camiones Un estudio demostró que los conductores de camiones pesan más que los adultos que no manejan camiones. Conclusión: Los camiones causan que la gente gane peso.

2. Las casas y la longevidad Un estudio concluyó que los propietarios de casas tienden a vivir más tiempo que quienes no habitan viviendas propias. Conclusión: Poseer una casa crea paz y armonía internas que causan que las personas tengan mejor estado de salud y vivan más tiempo.

3. Cumplimiento de las leyes de tránsito Un estudio mostró que en el condado de Orange se expidieron más multas por exceso de velocidad a personas de grupos minoritarios que a los blancos. Conclusión: En el condado de Orange las personas de grupos minoritarios conducen a mayor velocidad que los blancos.

4. Remedio para el resfriado En un estudio de síntomas del resfriado, se encontró que cada uno de los sujetos de estudio con resfriado mejoró dos semanas después de tomar píldoras de jengibre. Conclusión: las píldoras de jengibre curan el resfriado.

En los ejercicios 5 a 16, utilice el pensamiento crítico para señalar el tema principal.

5. El chocolate es un alimento saludable El *New York Times* publicó un artículo que incluyó esta afirmación: “Por fin, el chocolate ocupa el lugar que merece en la pirámide de los alimentos, 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* mostraron que, después de comer chocolate, 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 cardiacas y derrame cerebral. Mars Inc., la compañía de dulces, y la Chocolate Manufacturers Association financiaron gran parte de la investigación”. ¿Qué está equivocado en este estudio?

6. Datos de censo Después de realizado el último censo nacional, el *Poughkeepsie Journal* imprimió este titular de primera página: “281,421,906 en Estados Unidos”. ¿Qué está mal en este titular?

7. Encuesta por correo Cuando la autora Shere Hite escribió *Woman and Love: A Cultural Revolution in Progress*, basó sus conclusiones en 4500 respuestas recibidas después de enviar por correo 100,000 cuestionarios a varios 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 sí o por qué no?

8. 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 demostraron que el 67% de quienes llamaron estuvieron 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, respecto de mantener la sede de las Naciones Unidas en Estados Unidos.

9. Encuestas telefónicas La Hartford Insurance Company lo contrató a usted para encuestar a una muestra de adultos acerca de sus compras de automóviles. ¿Cuál es el error al considerar a las personas cuyos números telefónicos aparecen listados en los directorios como población de la cual se toma la muestra?

10. Crimen y autobuses El *Newport Chronicle* afirma que los paraderos de autobús causan crímenes, porque un estudio concluyó que las tasas de crimen son más altas en las ciudades con paraderos de autobús, que en las zonas rurales que carecen de ellos. ¿Cuál es el error en esta afirmación?

11. Cascos de motocicleta El Senado del estado de Hawai 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 habían resultado 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).

12. La encuesta al 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 incluyó esta declaración: "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". ¿Cuál es el error en esta encuesta?

13. La nicotina de los cigarrillos Remítase al conjunto de datos 5 del Apéndice B y considere el contenido de nicotina de 29 diferentes marcas de cigarrillos. El promedio (media) de esas cantidades es 0.94 mg. ¿Es probable que este resultado sea un buen estimado del promedio (media) de todos los cigarrillos que se han fumado en Estados Unidos? ¿Por qué sí o por qué no?

14. Pregunta incorrecta Una encuesta incluye este reactivo: "Anote su altura en pulgadas". A partir de este dato se espera obtener las estaturas reales de los encuestados y analizarlas, aunque hay dos problemas básicos diferentes en este reactivo; identifíquelos.

15. Longevidad Usted necesita realizar un estudio de longevidad a personas que nacieron después del fin de la Segunda Guerra Mundial en 1945. Si usted visitara los cementerios y utilizara las fechas de nacimiento y muerte indicadas en las lápidas, ¿obtendría buenos resultados? ¿Por qué sí o por qué no?

16. 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 tendido sobre la espalda con la cara hacia arriba. Ella sugirió que esta afirmación es mejor: "Los pediatras aconsejaron la posición supina durante un periodo en que disminuyeron las tasas de SMSI". ¿Qué está equivocado al decir que la posición supina ha disminuido las muertes por SMSI?

En los ejercicios 17 a 22, conteste las preguntas que se hacen en relación con los porcentajes.

17. Porcentajes
- **a.** Convierta la fracción 17/25 a un porcentaje equivalente.
- **b.** Convierta 35.2% a su equivalente decimal.
- **c.** ¿Cuánto es el 57% de 1500?
- **d.** Convierta 0.486 a un porcentaje equivalente.

18. Porcentajes
- **a.** ¿Cuál es el 26% de 950?
- **b.** Convierta 5% en su equivalente decimal.
- **c.** Convierta 0.01 a un porcentaje equivalente.
- **d.** Convierta la fracción 527/1200 a un porcentaje equivalente. Redondee la respuesta a la décima más cercana del porcentaje.

19. Porcentajes en una encuesta de Gallup

a. En una encuesta de Gallup, el 52% de 1038 adultos entrevistados dijo que ser un fumador pasivo es "muy dañino". ¿Cuál es el número real de adultos que dicen que ser un fumador pasivo es "muy dañino"?

b. De entre los 1038 adultos entrevistados, 52 dijeron que ser un fumador pasivo "no es dañino en absoluto". ¿Cuál es el porcentaje de gente que escogió "no es dañino en absoluto"?

20. Porcentajes en un estudio del Lipitor

a. En un estudio del fármaco Lipitor para el colesterol, a 270 pacientes se les administró un placebo; 19 de esos 270 pacientes reportaron dolor de cabeza. ¿Qué porcentaje de este grupo placebo reportó dolor de cabeza?

b. De entre los 270 pacientes del grupo placebo, el 3.0% reportó dolores de espalda. ¿Cuál es el número real de pacientes que reportaron dolores de espalda?

21. Porcentajes delictivos en los planteles universitarios Un estudio de los delitos cometidos por estudiantes bajo la influencia del alcohol o las drogas en los planteles universitarios, se basó en una encuesta por correo a 1875 estudiantes. Un artículo del *USA Today* destacó que "el 8% de los estudiantes, que respondieron de manera anónima, afirmaron haber cometido un delito en el campus. Y el 62% de ese grupo dijo que lo hizo bajo la influencia del alcohol o las 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 del alcohol o las drogas?

22. Porcentajes en los medios de comunicación

a. Un editorial del *New York Times* criticó un gráfico que describía un enjuague bucal que "reduce la placa bacteriana en más del 300%". ¿Que es incorrecto en esta declaración?

b. En el *New York Times Magazine*, un reporte 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 está equivocado en esta declaración?

1-3 Más allá de lo básico

23. Datos falsos 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. Estos valores se dieron para el porcentaje de éxito en cada grupo: 53%, 58%, 63%, 46%, 48%, 67%. ¿Cuál es la principal falla?

24. ¿Qué está mal en el asunto? Trate de identificar cada una de las cuatro fallas principales en lo siguiente. Un diario 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 lectores respondieron y 87% contestó "no", mientras que el 13% dijo "sí".

25. Redacción predispuesta Escriba una pregunta de encuesta que trate sobre un tema de su interés. Primero redacte la pregunta con objetividad, después redáctela para fomentar las respuestas hacia cierta dirección y por tercera vez redáctela para influir en las respuestas en la dirección opuesta.

26. Gráficas Actualmente, las mujeres ganan 74 centavos por cada dólar que ganan los hombres al realizar el mismo trabajo. Dibuje una gráfica que describa esta información de manera objetiva; luego, dibuje una gráfica que exagere la diferencia. (*Sugerencia:* Consulte la figura 1-1.)

1-4 Diseño de experimentos

Si bien esta sección contiene mucha información, existen dos puntos principales que son bastante sencillos. Es necesario entender que el método usado para reunir los datos es extremadamente importante, y debemos reconocer que la *aleatoriedad* resulta importante en particular.

- **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.**
- **La *aleatoriedad* por lo general juega un papel crucial para determinar cuáles son los datos a reunir.**

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**, observamos y medimos características específicas, aunque no intentamos *manipular* a los sujetos que estamos estudiando.

En un **experimento** aplicamos algún *tratamiento* y luego procedemos a observar sus efectos sobre los sujetos.

Una encuesta de Gallup es un buen ejemplo de un estudio observacional, mientras que la prueba clínica del fármaco Lipitor es un buen ejemplo de un experimento. La encuesta de Gallup es observacional en el sentido de que simplemente se observan personas (a menudo por medio de entrevistas) sin modificarlas de ninguna forma. Pero la prueba clínica de Lipitor implica el tratamiento de algunas personas con el fármaco, de manera que se manipula a los sujetos tratados. Hay diferentes tipos de estudios observacionales, como se ilustra en la figura 1-3. Estos términos, que se usan comúnmente en muchas y diferentes revistas profesionales, se definen aquí.

Definiciones

En un **estudio transeccional**, 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 (a través del examen de registros, entrevistas y otros medios).

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

Existe una distinción 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 que nos conciernen, 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

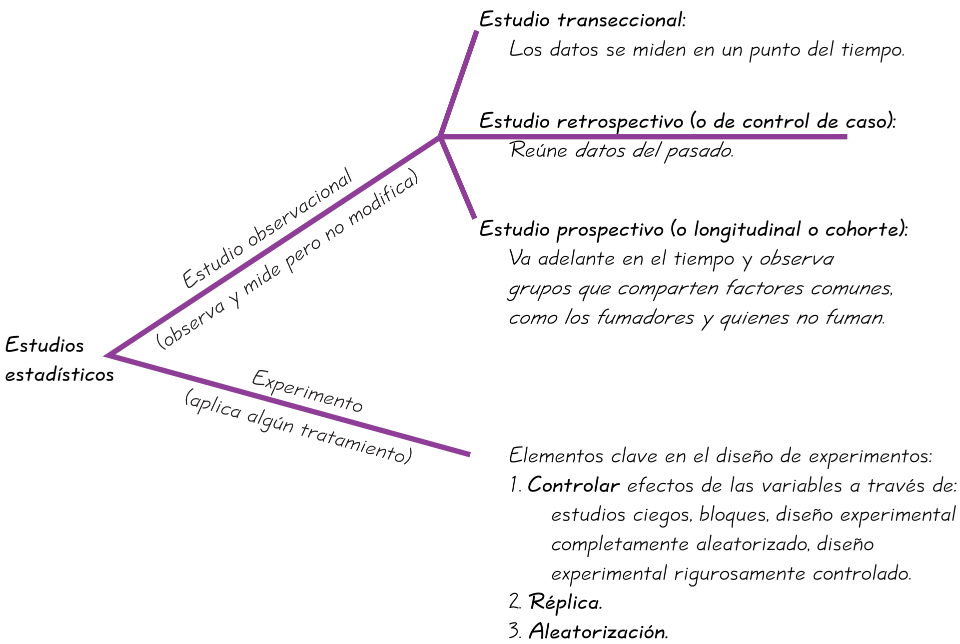


FIGURA 1-3 **Elementos de los estudios estadísticos**

prospectivos vamos adelante en el tiempo siguiendo grupos con un factor causal potencial y grupos que no lo tienen, como un grupo de conductores que utilizan teléfonos celulares y un grupo de conductores que no usan teléfonos celulares.

Las tres definiciones se aplican a los estudios observacionales, aunque por ahora nos enfocaremos en los experimentos. Los resultados de los experimentos algunas veces se empobrecen a causa de la *confusión*.

Definición

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

Intente planear el experimento de manera 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 que carece de nieve y temperaturas muy frías, lo cual en años anteriores obstaculizó la asistencia. Si la asistencia mejora no será posible determinar si la mejoría es atribuible a la nueva política de asistencia o al invierno benigno. Se confunden los efectos de la política de asistencia y del clima.

Control de los efectos de las variables

La figura 1-3 muestra que uno de los elementos clave en el diseño de experimentos es controlar los efectos de las variables. Se adquiere este control utilizando dispositivos como el estudio ciego, los bloques, el diseño experimental completamente aleatorizado o un diseño experimental rigurosamente controlado, que se describen a continuación.

LA ESTADÍSTICA EN LAS NOTICIAS

Pruebas clínicas *vs.* 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 controlados aleatorizados. 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 contraste, los estudios observacionales, en los que los pacientes deciden por sí mismos si toman un fármaco, se consideran menos confiables... Los investigadores manifiestan 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 al empezarlos 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 e informa 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 ser 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 Rosenthall) cuando el investigador o experimentador influye, sin desearlo, en los sujetos mediante factores como la expresión facial, el tono de voz o la actitud.

Estudio ciego En 1954 se diseñó un experimento masivo para probar la efectividad de la vacuna de Salk en la prevención de la poliomielitis que mató o paralizó a miles de niños. En este experimento a un grupo de tratamiento se le administró la vacuna real de Salk, mientras a un segundo grupo se le dio un placebo que no contenía ningún fármaco. En los experimentos que involucran placebos, hay a menudo 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). Este efecto placebo llega a minimizarse o a tomarse en cuenta mediante el uso del **estudio ciego**, una técnica donde el sujeto no sabe si está recibiendo un tratamiento o un placebo. El estudio ciego nos permite determinar si el efecto del tratamiento es significativamente diferente del efecto placebo. El experimento de la poliomielitis fue un **estudio doble ciego**, lo que quiere decir que el estudio ciego ocurrió a dos niveles: **1.** los niños inyectados no sabían si estaban recibiendo la vacuna de Salk o un placebo, y **2.** los doctores que suministraron las inyecciones y evaluaron los resultados tampoco lo sabían.

Bloques Cuando se diseña un experimento para probar la efectividad de uno más tratamientos, es importante poner a los sujetos (con frecuencia llamados *unidades experimentales*) en diferentes grupos (o *bloques*), de manera que estos grupos sean muy similares. Un **bloque** es un grupo de sujetos que son similares en formas que podrían afectar el resultado del experimento.

Cuando realice un experimento con el objetivo de probar uno o más tratamientos diferentes, forme bloques (o grupos) de sujetos con características similares.

Diseño experimental completamente aleatorizado Cuando se decide cómo asignar a los sujetos a los diferentes bloques, se puede utilizar una selección aleatoria o intentar controlar cuidadosamente la asignación, para que los sujetos de cada bloque resulten similares. Una opción consiste en usar un **diseño experimental completamente aleatorizado,** mediante el cual los sujetos se asignan a los diferentes bloques a través de un proceso de *selección aleatoria*. Un ejemplo de un diseño experimental completamente aleatorizado es el experimento de la poliomielitis: los niños fueron asignados al grupo de tratamiento o al grupo placebo a través de un proceso de selección aleatoria (equivalente a lanzar una moneda al aire).

Diseño rigurosamente controlado Otra opción para asignar sujetos a los bloques es el uso del **diseño rigurosamente controlado**, donde los sujetos son *cuidadosamente elegidos* para que quienes formen cada bloque sean similares en las características que sean importantes para el experimento. En un experimento para probar la efectividad de un fármaco para disminuir la presión sanguínea, si el grupo placebo incluye a una persona del sexo masculino de 30 años de edad, con sobrepeso, fumador, con alto consumo de bebidas alcohólicas y con una dieta alta en sal y grasas, el grupo de tratamiento también debe incluir a una persona con características similares (lo cual, en este caso, sería fácil de conseguir).

Réplica y tamaño de muestra

Además de controlar los efectos de las variables, otro elemento clave del diseño experimental es el tamaño de las muestras. Éstas deben ser suficientemente grandes para que el comportamiento errático, que es característico de muestras muy pequeñas, no disfrace los efectos verdaderos de los diferentes tratamientos. La repetición de un experimento se llama **réplica,** la cual se utiliza con efectividad 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 los tamaños de muestra grandes. Sin embargo, una muestra grande no es necesariamente una muestra buena. Aunque es importante tener una muestra que sea suficientemente grande, es más importante tener una muestra en la que los datos se escojan de una forma apropiada, como la selección aleatoria (que se describirá después).

Utilice un tamaño de muestra que sea lo bastante grande para distinguir la verdadera naturaleza de cualquiera de los diferentes efectos, y obtenga la muestra usando un método adecuado, 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 de Salk real, y a otros 200,000 niños se les dio un placebo. Se observó la efectividad de la vacuna porque se usaron tamaños de muestra bastante grandes en el experimento real. No obstante, aunque los grupos de tratamiento y placebo fueran muy grandes, el experimento puede fallar si los sujetos no se asignan a los dos grupos de tal manera que ambos grupos sean similares en las características importantes para el experimento.

Aleatorización y otras estrategias de muestreo

En la estadística, como en la vida, uno de los peores errores es reunir datos en una forma que no sea la adecuada. Insistiremos en este punto muy importante:

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

En la sección 1-3 vimos que una muestra de respuesta voluntaria es aquella donde los sujetos deciden por sí mismos si responden o no. Este tipo de muestras son muy comunes, aunque sus resultados por lo general resultan inútiles para hacer inferencias válidas acerca de poblaciones más grandes.

Ahora definiremos algunos de los métodos de muestreo más comunes.

Definiciones

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

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

EJEMPLO Muestra aleatoria y muestra aleatoria simple 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 tirando un dado y seleccionando la fila correspondiente al resultado. ¿El resultado es una muestra aleatoria? ¿Es una muestra aleatoria simple?

SOLUCIÓN La muestra es una muestra aleatoria porque cada estudiante tiene la misma posibilidad (una posibilidad en seis) de ser elegido. Sin embargo, la muestra no es una muestra 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 seleccionar 10 estudiantes que estén en filas diferentes (aunque hay una posibilidad en seis de seleccionar la muestra que consiste en los 10 estudiantes en la primera fila).

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 define arriba.

Con el muestreo aleatorio se espera que todos los componentes de la población estén (aproximadamente) representados de manera proporcional. Las muestras aleatorias se seleccionan mediante diversos métodos, incluyendo el uso de computadoras para generar números aleatorios. (Antes del uso de las computadoras, las tablas de números aleatorios se utilizaban con frecuencia. 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. El resumen del argumento no está disponible todavía). A diferencia de un muestreo realizado con descuido o por casualidad, el muestreo aleatorio exige una muy cuidadosa planeación y ejecución.

Además del muestreo aleatorio, hay otras técnicas de muestreo en uso, y las más comunes se describen aquí. Observe la figura 1-4, una ilustración que describe los diferentes tipos de muestreo. Tome en cuenta que sólo el muestreo aleatorio y el muestreo aleatorio simple se usarán en el resto de este libro.

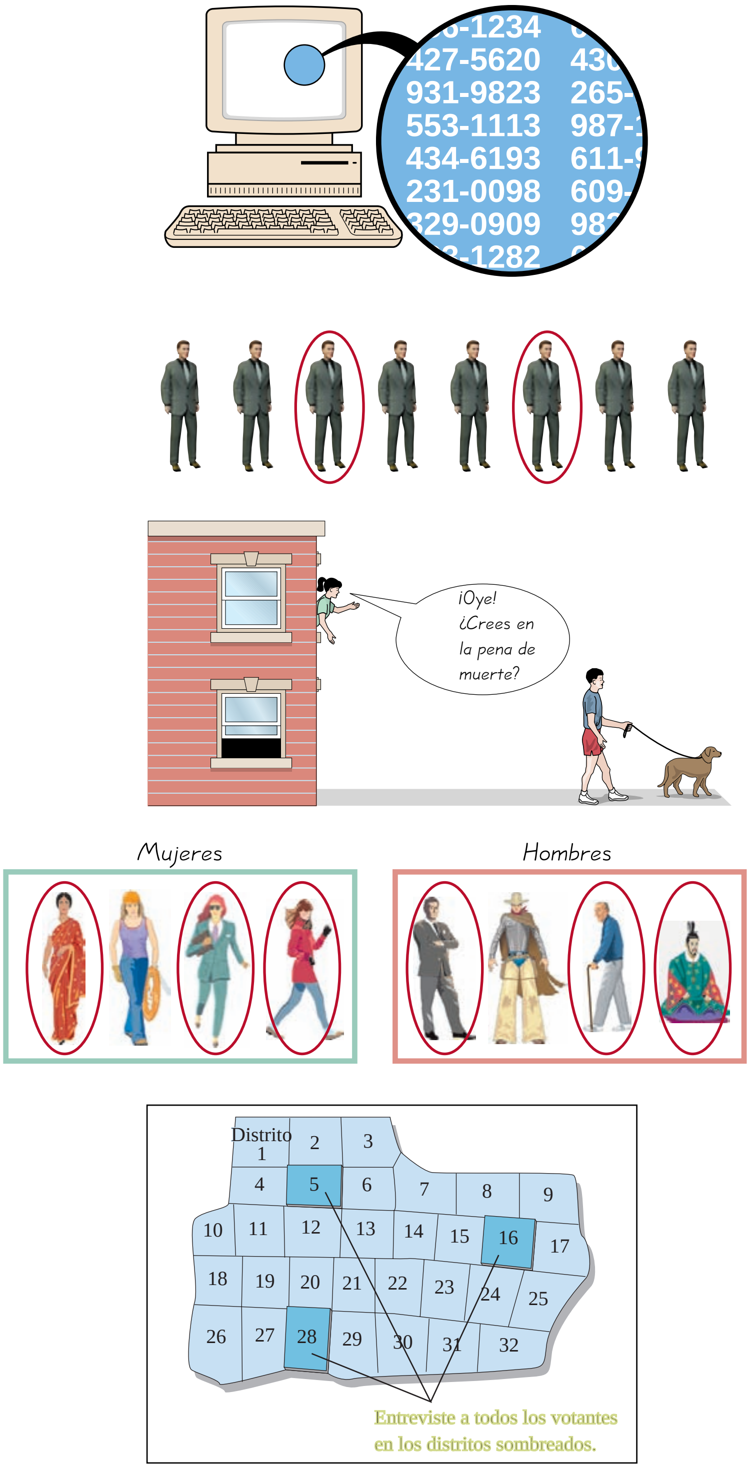
Definiciones

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

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

Con el **muestreo estratificado**, subdividimos la población en al menos dos diferentes subgrupos (o estratos) que comparten las mismas características (por ejemplo, el género o la categoría de edad) y después realizamos un muestreo de cada subgrupo (o estrato).

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



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

Muestreo aleatorio simple:
Se selecciona una muestra de tamaño 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 selecciona cada k-ésimo (por ejemplo, cada quincuagésimo) elemento en la población.

Muestreo de conveniencia:
Se utilizan resultados que son fáciles de obtener.

Muestreo estratificado:
Se subdivide a la población en al menos dos diferentes subgrupos (o estratos) que comparten las mismas características (por ejemplo, el género o categoría de edad), y después se extrae una muestra de cada subgrupo.

Muestreo por racimos:
Se divide el área de la población en secciones (o racimos), se eligen al azar unas cuantas de estas secciones y luego se escogen todos los miembros de los racimos seleccionados.

FIGURA 1-4 Métodos de muestreo comunes

Es fácil confundir el muestreo estratificado y el muestreo por racimos, ya que ambos suponen la formación de subgrupos. Pero el muestreo por racimos usa *todos* los miembros de una *muestra* de racimos, mientras el muestreo estratificado usa una *muestra* de los miembros de *todos* los estratos. Un ejemplo de muestreo por racimos es una encuesta previa a las elecciones, donde se seleccionan aleatoriamente 30 distritos electorales de un número mayor de distritos, luego se encuesta a todas las personas de cada uno de esos distritos escogidos, lo cual es mucho más rápido y mucho menos costoso que seleccionar a una persona de cada uno de los muchos distritos del área de población. Los resultados de la muestra estratificada o por racimos se ajustan o se ponderan para corregir cualquier representación desproporcionada de los grupos.

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

La figura 1-4 ilustra métodos de muestreo comunes. Los profesionales a veces reúnen datos usando cierta combinación de tales métodos. Aquí está un ejemplo típico de lo que se llama un *diseño muestral de etapas múltiples*: primero se selecciona una muestra aleatoria de condados de todos los 50 estados; después se eligen al azar ciudades y pueblos en esos condados; luego aleatoriamente se seleccionan cuadras residenciales en cada ciudad o pueblo; luego se escogen hogares al azar en cada cuadra y, por último, se selecciona al azar a una persona de cada hogar. En este libro no utilizaremos un diseño muestral de este tipo. Hay que recalcar otra vez que los métodos de este libro por lo regular requieren una *muestra aleatoria simple*.

Errores de muestreo

Por muy bien que usted planee y ejecute el proceso de recolección de muestras, es probable que ocurra algún error en los resultados. Por ejemplo, seleccione a 1000 adultos al azar, pregúnteles si se graduaron de bachillerato y registre el porcentaje de respuestas afirmativas en la muestra. Si selecciona 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; tal error es consecuencia de las posibles fluctuaciones de las muestras.

Un **error no de muestreo** ocurre cuando los datos de una muestra se obtienen, registran o analizan de forma incorrecta (como cuando se selecciona una muestra sesgada o predispuesta, cuando se usa un instrumento de medición defectuoso o cuando se cometen errores al copiar los datos).

Si recolectamos con cuidado una muestra que sea representativa de la población, usaremos los métodos de este libro para analizar el error de muestreo, pero debemos tener sumo cuidado para minimizar el error no de muestreo.

Después de leer esta sección, es normal estar un poco abrumado por la variedad de las diferentes definiciones. Sin embargo, recuerde este punto principal: el método usado para reunir datos es sumamente importante y debemos reconocer que la *aleatoriedad* es importante en particular. Si los datos muestrales no se reúnen de manera adecuada, resultarán inútiles por completo, de forma que ninguna cantidad de tortura estadística pueda salvarlos.

1-4 Destrezas y conceptos básicos

En los ejercicios 1 a 4, determine si la distribución dada corresponde a un estudio observacional o a un experimento.

1. Prueba de fármacos A los pacientes se les administra Lipitor para determinar si este fármaco tiene el efecto de disminuir los altos niveles de colesterol.

2. Tratamiento para 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 sífilis.

3. Fraude al consumidor El departamento de pesos y medidas del condado de Dutchess selecciona al azar expendios de gasolina y obtiene un galón de gasolina de cada bomba. La cantidad bombeada se mide para comprobar la precisión.

4. 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 5 a 8, identifique el tipo de estudio observacional (transeccional, retrospectivo o prospectivo).

5. Investigación médica Un investigador de la escuela de medicina de la Universidad de Nueva York obtiene datos acerca de heridas en la cabeza examinando los registros del hospital de los últimos cinco años.

6. Psicología del trauma Un investigador del hospital Monte Sinaí, en la ciudad de Nueva York, planea obtener datos haciendo seguimiento (hasta el año 2010) a los hermanos de las víctimas que perecieron en el ataque terrorista al World Trade Center el 11 de septiembre de 2001.

7. Estadísticas de desempleo El Departamento de Trabajo de Estados Unidos obtiene datos de desempleo reales encuestando a 50,000 personas en este mes.

8. Ganadores de la lotería Un economista reúne datos entrevistando a personas que ganaron la lotería entre los años 1995 y 2000.

En los ejercicios 9 a 20, identifique cuál de estos tipos de muestreo se utiliza: aleatorio, sistemático, por conveniencia, estratificado o por racimos.

9. Noticias televisivas Un reportero de noticias de la empresa de televisión NBC pretende conocer la reacción a una historia triste entrevistando a las personas que van pasando frente a su estudio.

10. Selección de jurado El comisionado de jurados del condado de Dutchess obtiene una lista de 42,763 propietarios de automóviles y compone una junta de jurados seleccionando cada 100-ésimo nombre en esa lista.

11. Encuestas telefónicas En una encuesta de Gallup de 1059 adultos, los sujetos entrevistados fueron seleccionados mediante el uso de una computadora, para generar aleatoriamente los números telefónicos a los que se llamó.

12. Propiedad de automóviles Una investigadora de General Motors dividió todos los automóviles registrados en categorías de subcompacto, compacto, mediano, intermedio y grande. Ella encuesta a 200 propietarios de automóviles de cada categoría.

13. 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 estos grupos.

14. Marketing Una ejecutiva de marketing de General Motors encontró que su departamento de relaciones públicas acababa de imprimir sobres con los nombres y direcciones de todos los propietarios de un Corvette. Ella quiere hacer una prueba piloto de la nueva estrategia de mercadotecnia, así que mezcla cuidadosamente todos los sobres en una urna y obtiene un grupo de muestra sacando 50 de esos sobres.

15. 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).

16. Encuestas de salida La CNN está planeando una encuesta de salida en que se elegirán aleatoriamente 100 casillas electorales y todos los votantes se entrevistarán conforme vayan saliendo de los locales.

17. La educación y el salario Un economista estudia el efecto de la educación en el salario, y realiza una encuesta a 150 trabajadores seleccionados al azar de cada una de estas categorías: estudios menores que la secundaria, grado de escuela secundaria, estudios de mayor grado que la secundaria.

18. Antropometría Un estudiante de estadística obtuvo datos de estatura/peso entrevistando a los miembros de la familia.

19. Investigación médica Un investigador de la Universidad Johns Hopkins encuesta a todos los pacientes cardiacos en cada uno de 30 hospitales seleccionados al azar.

20. Encuesta de MTV Un experto en marketing está planeando una encuesta para MTV, en la cual 500 personas se elegirán aleatoriamente de cada grupo de edades de 10 a 19, 20 a 29, etcétera.

Los ejercicios 21 a 26 se relacionan con muestras aleatorias y muestras aleatorias simples.

21. Muestreo de tabletas de aspirina Un farmacéutico mezcla cuidadosamente un recipiente con 1000 tabletas de Bufferin y luego recoge una muestra de 50 tabletas que se evaluarán para determinar el contenido exacto de aspirina. ¿Este plan de muestreo describe un muestreo aleatorio? ¿Un muestreo aleatorio simple? Explique.

22. Muestreo de estudiantes Un salón de clases consta de 30 estudiantes sentados en cinco filas diferentes, con seis estudiantes en cada fila. El profesor tira un dado y el resultado se utiliza para seleccionar una muestra de los estudiantes de una fila particular. ¿Este plan de muestreo es un muestreo aleatorio? ¿Un muestreo aleatorio simple? Explique.

23. Muestreo de conveniencia Un reportero de noticias se para en la esquina de una calle, obtiene una muestra de residentes de la ciudad seleccionando a cinco adultos que pasan por ahí y les pregunta acerca de sus hábitos de fumar. ¿Este plan de muestreo dará como resultado un muestreo aleatorio? ¿Un muestreo aleatorio simple? Explique.

24. Muestreo sistemático Un ingeniero de control de calidad selecciona cada 100-ésima unidad de fuente de poder de computadora que pasa por una banda transportadora. ¿Resulta este plan de muestreo en un muestreo aleatorio? ¿Un muestreo aleatorio simple? Explique.

25. Muestra estratificada La empresa de alimentos General Foods planea realizar una encuesta de marketing a 100 hombres y 100 mujeres en el condado de Orange, que consiste en un número igual de hombres y mujeres. ¿Resulta este plan de muestreo en un muestreo aleatorio? ¿Un muestreo aleatorio simple? Explique.

26. Muestra por racimos Un investigador de marketing selecciona aleatoriamente 10 cuadras en el pueblo de Newport, luego pregunta a los adultos residentes de las cuadras seleccionadas si tienen un reproductor de DVD. ¿Este plan de muestreo resultará en un muestreo aleatorio? ¿Un muestreo aleatorio simple? Explique.

1-4 Más allá de lo básico

27. Diseño de muestreo La compañía de publicaciones Addison-Wesley le ha comisionado a usted para encuestar a 100 estudiantes usuarios de esta obra. Describa los procedimientos para obtener una muestra de cada tipo: aleatoria, sistemática, de conveniencia, estratificada y por racimos.

28. Confusión Mencione un ejemplo (diferente del que está en el texto) que ilustre la forma en que ocurre la confusión.

29. Selección aleatoria Entre las 50 entidades de Estados Unidos, se elige aleatoriamente una entidad. Después se obtiene el padrón electoral de todo el estado y se selecciona un nombre al azar. ¿Este procedimiento resultará en un votante seleccionado aleatoriamente?

30. 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 común. Los autores resumen el diseño muestral como un “experimento prospectivo, aleatorizado, doble ciego, placebo y controlado”. Describa el significado de cada uno de estos términos en el contexto de este estudio.

31. Conductores con teléfonos celulares ¿Cuáles son los dos problemas principales que pueden encontrarse en un estudio prospectivo, donde algunos conductores no tienen teléfonos celulares mientras que a otros se les pide que usen sus teléfonos celulares mientras conducen?

Repaso

Este capítulo presentó algunos fundamentos importantes, consistentes en definiciones básicas, como las de *muestra* y *población*, junto con algunos principios esenciales. La sección 1-2 analizó los diferentes tipos de datos. La sección 1-3 trató con el uso del pensamiento crítico en el análisis y la evaluación de resultados estadísticos. La sección 1-4 introdujo elementos importantes en el diseño de experimentos. Al terminar el estudio de este capítulo, usted debe ser capaz de:

- Distinguir entre una población y una muestra, y entre un parámetro y un estadístico.
- Identificar el nivel de medición (nominal, ordinal, de intervalo, de razón) de un conjunto de datos.
- Entender la importancia de un buen diseño experimental, incluyendo el control de los efectos de las variables, la réplica y la aleatorización.
- Reconocer la importancia de seguir buenos métodos de muestreo en general y reconocer la importancia de una *muestra aleatoria simple* en particular. Entender que si los datos muestrales no se reúnen de manera adecuada, los datos resultarían tan inútiles que ninguna cantidad de tortura estadística podría salvarlos.

Ejercicios de repaso

1. Muestreo Poco después de que las torres del World Trade Center se colapsaran por los ataques terroristas, America Online aplicó una encuesta a sus suscriptores de Internet y preguntó lo siguiente: "¿Deben reconstruirse las torres del World Trade Center?". De 1,304,240 personas que respondieron, 768,731 respondieron "sí", 286,756 contestaron "no", y 248,753 dijeron que era "demasiado pronto para decidir". Como esta muestra es extremadamente grande, ¿se puede considerar que las respuestas sean representativas de la población de Estados Unidos? Explique.

2. Diseño de muestreo Usted ha sido contratado por Visa para realizar un estudio acerca del uso de tarjeta de crédito 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 racimos.

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 en una muestra de pasajeros de un elevador.
   b. Una clasificación de crítica de cine de *debe verse*, *recomendada*, *no recomendada*, *ni piense en verla*.
   c. Bob, que es distinto en muchas formas, mide el tiempo en días a partir de 0, que corresponde a su fecha de nacimiento. El día anterior a su nacimiento es -1, el día después de su nacimiento es +1, etcétera. 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 La Coca Cola Company tiene 366,000 accionistas y efectúa una encuesta por medio de la selección aleatoria de 30 accionistas de cada uno de los 50 estados de Estados Unidos. Se registra el número de acciones de cada accionista de la muestra.
   a. ¿Los valores obtenidos son discretos o continuos?
   b. Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.
   c. ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?
   d. Si se calcula el número promedio (la media) de acciones, ¿el resultado es un estadístico o un parámetro?
   e. Si usted fuera el ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?
   f. ¿Qué es lo que está incorrecto al evaluar la opinión del accionista enviando un cuestionario por correo que los accionistas podrían llenar y regresar por el mismo medio?

5. Más Coca Cola Identifique el tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) 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 parece resultar en una muestra representativa de la población de los 366,000 accionistas.
   a. Se compila una lista completa de todos los accionistas y se selecciona cada 500-ésimo nombre.
   b. En la junta anual de accionistas, se realiza una encuesta de todos los asistentes.
   c. Se seleccionan al azar 50 diferentes corredores de bolsa y se hace una encuesta a todos sus clientes que tengan acciones de Coca Cola.
   d. Se compila un archivo de computadora de todos los accionistas, de manera que todos ellos se numeran de forma consecutiva y después los números aleatorios generados por computadora se utilizan para seleccionar la muestra de accionistas.
   e. Se reúnen todos los códigos postales de los accionistas y se elige 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. Usará una muestra de sujetos que han sido tratados con el fármaco y otra muestra de sujetos a quienes se les administró un placebo.
   a. ¿Qué es el “estudio ciego” y como puede usarse en este experimento?
   b. ¿Por qué es importante el uso del estudio ciego en este experimento?
   c. ¿Qué es un diseño de bloques completamente aleatorizado?
   d. ¿Qué es un diseño de bloques rigurosamente controlado?
   e. ¿Que es la réplica y por qué es importante?

Ejercicios de repaso acumulativos

Los ejercicios de repaso acumulativos de este libro, están diseñados para incluir temas de capítulos anteriores. Para los capítulos 2 a 13, 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 esta obra. Utilice su calculadora para obtener los valores indicados.

1. Remítase al conjunto de datos 1 del apéndice B y considere sólo los pesos de los primeros 10 varones. ¿Qué valor se obtiene cuando se suman estos 10 pesos y el total se divide entre 10? (Este resultado, llamado media, se analiza en el capítulo 2).

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

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

4. $$\left[\frac{1.96 \cdot 15}{2}\right]^2$$

5. $$\sqrt{\frac{(5-7)^2 + (12-7)^2 + (4-7)^2}{3-1}}$$

6. $$\frac{(183 - 137.09)^2}{137.09} + \frac{(30 - 41.68)^2}{41.68}$$

7. $$\sqrt{\frac{10(513.27) - 71.5^2}{10(10-1)}}$$

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 (o 1.23^5 en algunas calculadoras) puede expresarse como 123,000, y el resultado de 4.65E−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, no en notación científica.

9. 0.95^{500} **10.** 8^{14} **11.** 9^{12} **12.** 0.25^{17}

Actividades de cooperativas en equipo

1. **Actividad en clase** Obtenga 18 popotes de la cafetería. Corte 6 de ellos por la mitad, corte 6 en cuartos y los otros 6 déjelos como están. Ahora debe haber 42 popotes de diferentes longitudes. Póngalos en una bolsa, revuélvalos, luego seleccione un popote, note su longitud y póngalo de nuevo en la bolsa. Repita esto hasta seleccionar 20 popotes. *Importante:* Seleccione los popotes sin mirar al interior de la bolsa y saque el primero que toque. Calcule el promedio (media) de la muestra de 20 popotes. Ahora saque todos los popotes y encuentre la media de la población. ¿La muestra dio un promedio cercano al promedio de la población real? ¿Por qué sí o por qué no?

2. **Actividad en clase** A mediados de diciembre de un año reciente, el proveedor de servicios de Internet America Online (AOL) efectuó una encuesta a sus usuarios. Se les preguntó lo siguiente acerca de los árboles de Navidad: “¿Cuál prefiere usted?”. La respuesta podía ser “un árbol natural” o “un árbol artificial”. Entre las 7073 respuestas recibidas de los usuarios de Internet, 4650 preferían un árbol natural y 2423 un árbol artificial. Ya señalamos que como la muestra es una muestra 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.

3. **Actividad en clase** Identifique los problemas en los siguientes eventos:

   - Un reporte televisado recientemente por *CNN Headline News* incluyó el comentario de que el crimen 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 debían enviarse por correo en sobres que requerían timbres postales.

Proyecto tecnológico

El propósito de este proyecto es introducir los recursos tecnológicos que usted usará en su curso de estadística. Remítase al conjunto de datos 14 en el apéndice B y use las puntuaciones de facilidad de lectura de Flesch para el libro *Harry Potter y la piedra filosofal* de J. K. Rowling. Utilizando su programa de estadística o la calculadora TI-83 Plus, introduzca estos 12 valores, y luego imprima un listado de ellos.

STATDISK Haga clic en **Data (datos)** en la parte superior de la pantalla, después seleccione **Sample Editor (editor de muestra)** y proceda a introducir los datos. Para imprimir seleccione con el ratón **File (archivo)** y luego seleccione **Print (imprimir)**.

Minitab Introduzca los datos en la columna C1, después haga clic en **File (archivo)** y seleccione **Print Worksheet (imprimir hoja de cálculo)**.

Excel Introduzca los datos en la columna A, después haga clic en **File (archivo)** y seleccione **Print (imprimir)**.

TI-83 Plus La impresión de la pantalla de la TI-83 Plus sólo es posible mediante el uso de la conexión a una computadora *Graphlink*.

de los DATOS a la DECISIÓN

Pensamiento crítico

El médico suizo H. C. Lombard en una ocasión compiló datos sobre la longevidad en relación con distintas profesiones. Usó actas de defunción que incluían nombre, edad al morir y profesión. Después procedió a calcular el promedio (media) de la longevidad para las diferentes profesiones, y encontró que los estudiantes eran los más bajos con una media de ¡sólo 20.7 años! (Véase "A Selection of Selection Anomalies" de Wainer, Palmer y Bradlow en *Chance*, vol. 11, núm. 2). Si los mismos datos fueran reunidos el día de hoy en Estados Unidos, se obtendrían resultados similares.

Análisis de los resultados

¿En realidad ser estudiante es más peligroso que ser agente de policía, chofer de taxi o empleado postal? Explique.

PROYECTO DE INTERNET

El sitio Web de *Estadística*/Triola

En esta sección de cada capítulo, se le pedirá que visite la página Web de este libro. Desde ahí usted puede tener acceso a las páginas referentes a todos los proyectos de Internet que vienen en la *novena edición* de *Estadística*/Triola. Vaya a este sitio ahora y familiarícese con todas las características de este libro a las que tiene acceso.

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 Web. 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!

www.pearsoneducacion.net/triola

La estadística @ en el trabajo

"Empleamos la estadística para determinar el grado de aislamiento que existe entre grupos putativos".

Sarah Mesnick

Ecologista conductual y molecular

Sara Mesnick es miembro posdoctorado del National Research Council. Su trabajo como bióloga en mamíferos incluye no sólo investigación en el mar, sino también en el Laboratory of Molecular Ecology. 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 empleamos esta información para crear mejores modelos de manejo para la conservación de ésta y otras especies de mamíferos marinos en peligro de extinción.

¿Qué conceptos de la estadística utiliza?

En la actualidad utilizo 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. Empleamos la chi cuadrada y el estadístico F para determinar la cantidad de poblaciones discretas de ballenas en el Pacífico. Estas poblaciones se manejan como grupos independientes. El análisis de regresión de la relación se utiliza 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, el cachalote y muchas otras especies de mamíferos marinos, aves y tortugas se lastiman o mueren incidentalmente en maniobras 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 estaría amenazada. Empleamos la estadística para determinar el grado de aislamiento que existe entre grupos putativos. Si resultara que los grupos están aislados, usaríamos esta 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 forma enfoca su investigación?

Tratamos de evitar ideas preconcebidas acerca de la forma en que los animales están distribuidos en su medio ambiente. Puesto que los mamíferos marinos en particular son tan difíciles de estudiar, suelen existir ideas aceptadas sobre lo que estos animales hacen, aun cuando esto no se ha investigado de manera profunda. 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 "líder 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.

2

Descripción, exploración y comparación de datos

PROBLEMA DEL CAPÍTULO

¿Realmente son afectadas las personas que no fuman cuando otros sí lo hacen junto a ellas? ¿O es un mito el efecto del *fumador pasivo*?

El conjunto número 6 del Apéndice B incluye algunos de los datos disponibles más recientes del National Institute of Health de Estados Unidos. Los datos, que se reproducen en la tabla 2-1, se obtuvieron como parte del National Health and Nutrition Examination Survey. Los valores de los datos corresponden a los niveles medidos de cotinina sérica (en ng/ml) en personas seleccionadas como sujetos de estudio (los datos se redondearon hacia el entero más cercano, de tal modo que un valor de cero no necesariamente implica la ausencia total de cotinina. De hecho, todos los valores originales fueron mayores que cero). La cotinina es un metabolito de la nicotina, es decir, es una sustancia que se produce cuando el cuerpo absorbe la nicotina. Porque se sabe que la nicotina se absorbe cuando se consumen cigarrillos, hay una forma indirecta de medir la presencia efectiva del humo del tabaco; esto es, por medio de la cotinina.

Existen varios aspectos importantes al respecto: ¿deben preocuparse por su salud las personas que no fuman ante la presencia de fumadores activos? Para preverlo, en los últimos años las autoridades sanitarias han elaborado muchos reglamentos para restringir el tabaquismo en lugares públicos. ¿Son justificadas dichas regulaciones por razones de salud o sólo provocan dificultades innecesarias a los fumadores?

Pensamiento crítico: Una comparación visual de las cifras en los tres grupos de la tabla 2-1 proporciona cierta información. En este capítulo presentamos métodos para lograr una mayor comprensión. Seremos capaces de producir comparaciones productivas e inteligentes; aprenderemos técnicas para describir, explorar y comparar conjuntos de datos, tales como los tres grupos de la tabla 2-1.

Tabla 2-1 Niveles medidos de cotinina en tres grupos

Fumador: Los sujetos reportan su consumo de tabaco.

HTA: Humo de tabaco ambiental). Sujetos que no fuman, pero que están expuestos a humo de tabaco ambiental ("fumadores pasivos"), en su casa o trabajo.

SHTA: (Sin humo de tabaco ambiental): Sujetos que no fuman y que no se exponen a humo de tabaco ambiental en su casa o trabajo. Esto es, no fuman ni son fumadores pasivos.

Fumador:	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:	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:	0	0	0	0	0	0	0	0	0	0
	0	9	0	0	0	0	0	0	244	0
	1	0	0	0	90	1	0	309	0	0
	0	0	0	0	0	0	0	0	0	0

2-1 Panorama general

Este capítulo es sumamente importante, ya que presenta las herramientas básicas para medir y describir diferentes características de un conjunto de datos. Cuando se describen, exploran y comparan conjuntos de datos, las siguientes características suelen ser de enorme importancia.

Características importantes de los datos

1. **Centro:** Valor representativo o promedio 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:** Naturaleza o forma de la distribución de los datos (tales como normales, uniformes o sesgadas).
4. **Datos distantes:** 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: La memorización suele ser ineficaz para recordar información importante. Sin embargo, las cinco características anteriores son tan importantes que deben recordarse con el uso de una técnica mnemónica conocida por las iniciales CVDDT; por ejemplo, "**C**uidado con los **V**irus que **D**estruyen **D**atos y **T**rabajo". Se ha visto que dichas técnicas de memorización son muy efectivas para recordar importantes palabras clave que evocan conceptos básicos.

Pensamiento crítico e interpretación: más allá de las fórmulas

Los profesores de estadística, por lo general, piensan que no es tan importante memorizar fórmulas o realizar cálculos aritméticos complejos a mano. Por el contrario, suelen enfocarse en la obtención de resultados por medio del uso de algún tipo de herramienta tecnológica (calculadoras o programas de cómputo), para después entender, de forma práctica, los resultados a través del pensamiento crítico. Tenga esto en mente conforme avance en el estudio de este capítulo. Por ejemplo, cuando estudie la muy importante *desviación estándar*, en la sección 2-5, trate de observar por qué la fórmula clave funciona como una medida de variación, después aprenda a calcular los valores de las desviaciones estándar, pero trabaje realmente en la *comprensión* y la *interpretación* de los valores de la desviación estándar.

Aun cuando este capítulo incluye, de forma detallada, los casos para procedimientos importantes, no es necesario conocerlos a la perfección en todas las situaciones. No obstante, recomendamos que, en cada caso, realice algunos cálculos manuales antes de utilizar su calculadora o computadora. Lo anterior hará que su comprensión se incremente y podrá apreciar mejor los resultados obtenidos con las herramientas tecnológicas.

Los métodos de este capítulo suelen denominarse métodos de **estadística descriptiva**, porque su objetivo es resumir o *describir* las características importantes de un conjunto de datos. Más adelante, utilizaremos métodos de **estadística inferencial**; lo haremos cuando usemos datos muestrales para hacer inferencias (o generalizaciones) acerca de una población. Con la estadística inferencial realizamos una deducción que va más allá de los datos conocidos. La materia de estadística tiene dos divisiones generales: la descriptiva y la inferencial; este capítulo trata los conceptos básicos de la estadística descriptiva.

2-2 Distribuciones de frecuencias

Cuando se trabaja con conjuntos grandes de datos, con frecuencia es útil organizarlos y resumirlos por medio de la construcción de una tabla que liste los distintos valores posibles de los datos (ya sea de forma individual o por grupos), junto con las frecuencias correspondientes, es decir, el número de veces que ocurren dichos valores.

Definición

Distribución de frecuencias: lista valores de 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 los niveles medidos de cotinina de los 40 fumadores que se muetran en la tabla 2-1. La **frecuencia** de una clase particular es el número de valores originales que caen dentro de esa clase. Por ejemplo, la primera clase de la tabla 2-2 tiene una frecuencia de 11, lo que indica que 11 de los valores originales de los datos están entre 0 y 99, inclusive.

Tabla 2-2
Distribución de frecuencias de los niveles de cotinina de los fumadores

Cotinina	Frecuencia
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2

Para empezar, presentaremos algunos términos estándar utilizados al referirse a la distribución de frecuencia; después describiremos la forma 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 0, 100, 200, 300 y 400).

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 99, 199, 299, 399 y 499).

Las **fronteras de clase** son las cifras utilizadas para separar las clases, aunque sin los espacios creados por los límites de clase. Se obtienen de la siguiente manera: se determina el tamaño del espacio entre el límite de clase superior de una clase y el límite de clase inferior de la siguiente. Se suma la mitad de esa cantidad a cada límite de clase superior, para obtener las fronteras de clase superiores; se resta la mitad de esa cantidad de cada límite de clase inferior, para obtener las fronteras de clase inferiores. (Los espacios de la tabla 2-2 son exactamente de una unidad, de modo que a los límites de clase superiores se les suma 0.5 y a los límites de clase inferiores se les resta 0.5. Las fronteras de la primera clase son -0.5 y 99.5, en tanto que las de la segunda clase son 99.5 y 199.5, y así

continúa

sucesivamente. La lista completa de fronteras utilizadas para todas las clases es la siguiente: −0.5, 99.5, 199.5, 299.5, 399.5 y 499.5.)

Las **marcas de clase** son los puntos medios de las clases. (Las marcas de clase de la tabla 2-2 son 49.5, 149.5, 249.5, 349.5 y 449.5). Cada marca de clase se calcula sumando el límite de clase inferior con el límite de clase superior y dividiendo la suma entre dos.

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 que se utiliza en la tabla 2-2 es igual a 100).

Gráficas de crecimiento actualizadas

Los pediatras acostumbran utilizar gráficas de crecimiento estandarizadas para comparar el peso y la altura de sus pacientes con una muestra de otros niños. Se considera que los pequeños están dentro de un rango normal si su peso y su estatura caen dentro de los percentiles 5o y 95o. Si se encuentran fuera de este rango, les aplicarán pruebas para estar seguros de que no hay problemas médicos graves. Los pediatras se dan cuenta de la existencia de un problema importante gracias a las gráficas; debido a que se basan en niños que vivieron entre 1929 y 1975, se concluyó que las gráficas de crecimiento eran imprecisas. Para rectificar tal problema, las gráficas se actualizaron en el año 2000, con la finalidad de que reflejaran las mediciones 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 que lleva a considerar las características cambiantes de los datos a lo largo del tiempo como un aspecto importante de una población.

Las definiciones de anchura de clase y frontera de clase son engañosas. Hay que tener cuidado para evitar el error común de considerar la amplitud de clase como la diferencia entre el límite de clase inferior y el límite de clase superior. Vea la tabla 2.2 y observe que la anchura de clase es de 100 y no de 99. El proceso para determinar las fronteras de clase se simplifica si se comprende que éstas básicamente llenan los espacios entre clases al dividir la diferencia entre el final de una clase y el inicio de la siguiente.

Procedimiento de construcción de 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 respecto de la naturaleza de los datos, y **3.** se llega a tener un avance para construir gráficas importantes (tales como *histogramas*, que se presentarán en la siguiente sección). Muchas de las herramientas tecnológicas permiten obtener de manera automática las distribuciones de frecuencias, sin necesidad de tenerlas que construir manualmente; no obstante, a continuación se presenta el procedimiento básico:

1. Decida el número de clases que desea tener. Debe ser de entre 5 y 20, y deben utilizarse números enteros o redondeados.

2. Calcule

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

Redondee el 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. Con el uso del límite más bajo de la primera clase y la anchura de clase, proceda a listar los demás límites de clase inferior. (Sume la anchura de clase al punto de partida para obtener el segundo límite de clase inferior. Después, 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 pueden identificarse con facilidad.

6. Ponga una marca en la clase apropiada 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 todos los casos, aun aquellos que tienen una frecuencia de cero. Trate de utilizar la misma anchura para todas las clases, aunque en ocasiones es imposible evitar los intervalos con finales abiertos, como “65 años o mayores”.

PROBLEMA 2 DEL CAPÍTULO

EJEMPLO Niveles de cotinina de fumadores Utilice los 40 niveles de cotinina de los *fumadores* de la tabla 2-1 y siga el procedimiento anterior para crear la distribución de frecuencias que se muestra en la tabla 2-2. Suponga que desea incluir cinco clases.

SOLUCIÓN

Paso 1: Comience seleccionando cinco clases.

Paso 2: Calcule la anchura de clase. En el siguiente cálculo, 98.2 se redondea a 100, ya que es un número más conveniente.

$$\text{anchura de clase} \approx \frac{(\text{valor más alto}) - (\text{valor más bajo})}{\text{número de clases}} = \frac{491 - 0}{5} = 98.2 \approx 100$$

Paso 3: Elija un punto de partida de 0, que es el valor más bajo en la lista y también porque es un número conveniente.

Paso 4: Sume la anchura de clase de 100 al punto de partida de 0 para determinar que el segundo límite inferior de clase es igual a 100. Continúe, y sume la anchura de clase de 100 para obtener los límites inferiores de clase restantes de 200, 300 y 400.

Paso 5: Liste los límites de clase inferiores de forma vertical, como se muestra al margen. Con esta lista se identifican con facilidad los límites de clases superiores correspondientes, tales como 99, 199, 299, 399 y 499.

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

0–
100–
200–
300–
400–

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 lugar 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 reemplazaron con las frecuencias relativas correspondientes, expresadas en porcentajes. La primera clase tiene una frecuencia relativa de $11/40 = 0.275$ o de 27.5%, que se redondea a 28%.

Tabla 2-3 Distribución de frecuencias relativas de los niveles de cotinina en fumadores

Cotinina	Frecuencias relativas
0–99	28%
100–199	30%
200–299	35%
300–399	3%
400–499	5%

Se identifican autores

Entre 1787 y 1788, Alexander Hamilton, John Jay y James Madison publicaron, de forma anónima, el famoso diario *Federalist*, en un intento por convencer a los neoyorquinos de que debían ratificar la Constitución. Se conoció la identidad de la mayoría de los autores de los artículos, pero el autor de 12 de éstos siguió siendo motivo de discusión. Mediante el análisis estadístico del análisis de frecuencias de diversas palabras, ahora concluimos que James Madison es el *probable* autor de esos dos artículos. En muchos de los artículos disputados, la evidencia en favor de la autoría de Madison es abrumadora, al grado de que casi con seguridad afirmamos que estamos en lo correcto.

La segunda clase tiene una frecuencia relativa de $12/40 = 0.3$ o 30%, y así sucesivamente. Si se construye de manera correcta, la suma de las frecuencias relativas debe totalizar 1 (o 100%), con algunas pequeñas discrepancias, que se permiten al redondear los errores. Puesto que 27.5% se redondeó a 28%, y 2.5% se redondeó a 3%, la suma de frecuencias relativas de la tabla 2-3 es de 101%, en lugar de 100%.

Ya que utilizan proporciones simples o porcentajes, las distribuciones de frecuencias nos facilitan la comprensión de la distribución de los datos y nos 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 previas. La tabla 2-4 muestra la distribución de frecuencias acumulativas de la distribución de frecuencias de la tabla 2-2. Con el uso de las frecuencias originales de 11, 12, 14, 1 y 2, sumamos $11 + 12$ para obtener la segunda frecuencia acumulativa de 23; después, sumamos $11 + 12 + 14 = 37$, para obtener la tercera, y así sucesivamente. Vea la tabla 2-4 y observe que, además del uso de frecuencias acumulativas, los límites de clase fueron reemplazados por expresiones como "menor que", las cuales describen el nuevo rango de valores.

Pensamiento crítico: interpretación de las distribuciones de frecuencias

La transformación de datos brutos en una distribución de frecuencias suele ser un medio para un gran fin. Los siguientes ejemplos ilustran la forma en que se utilizan las distribuciones de frecuencias para describir, explorar y comparar conjuntos de datos. (La siguiente sección muestra cómo la elaboración de una distribución de frecuencias suele ser el primer paso en la creación de una gráfica, que presenta la naturaleza de la distribución de forma visual).

EJEMPLO Descripción de los datos Remítase al conjunto de datos 1 en el Apéndice B, que se refieren al pulso de 40 adultos varones que se seleccionaron aleatoriamente. La tabla 2-5 presenta los *últimos dígitos* de tales datos. Si la tasa de pulsaciones se mide contando el número de latidos cardiacos en un minuto, esperamos que los últimos dígitos tengan frecuencias muy similares. Sin embargo, note que la distribución de frecuencias muestra que todos los últimos dígitos son números *pares*; *no* hay números impares. Lo anterior 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*, vemos que cada valor original es un múltiplo de cuatro, lo que sugiere que el número de latidos por minuto se contó durante 15 segundos y que después el resultado se multiplicó por cuatro). Es fascinante aprender el método de recolección de datos con la simple descripción de algunas características de los mismos.

Tabla 2-4 Distribución de frecuencias acumulativas de los niveles de cotinina en fumadores

Cotinina	Frecuencia Relativa
Menos de 100	11
Menos de 200	23
Menos de 300	37
Menos de 400	38
Menos de 500	40

Tabla 2-5 Últimos dígitos de las tasas de pulsaciones de varones

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

EJEMPLO Exploración de datos Para estudiar el comportamiento del géiser Old Faithful, ubicado en el Parque Nacional Yellowstone, los geólogos recolectan datos del tiempo (en minutos) que transcurre entre las erupciones. La tabla 2-6 muestra un resumen de los datos reales obtenidos. Un examen de la distribución de frecuencias reveló un comportamiento inesperado: la distribución del tiempo presenta dos picos distintos. Tal distribución condujo a los geólogos a considerar dos posibles explicaciones.

Tabla 2-6 Tiempo (en minutos) entre las erupciones del géiser Old Faithful

Tiempo	Frecuencia
40–49	8
50–59	44
60–69	23
70–79	6
80–89	107
90–99	11
100–109	1

PROBLEMA 2 DEL CAPÍTULO

EJEMPLO Comparación de conjuntos de datos El problema que abre este capítulo incluye conjuntos de datos que representan los niveles de cotinina que se midieron en fumadores, en no fumadores expuestos al humo del tabaco y en no fumadores sin exposición al

Tabla 2-7 Niveles de cotinina de los tres grupos

Cotinina	Fumadores	No fumadores expuestos al humo	No fumadores sin exposición al humo
0–99	28%	85%	95%
100–199	30%	5%	0%
200–299	35%	3%	3%
300–399	3%	3%	3%
400–499	5%	0%	0%
500–599	0%	5%	0%

continúa

humo del tabaco. La tabla 2-7 presenta las frecuencias relativas de los tres grupos. Al comparar dichas frecuencias relativas, es claro que la distribución de frecuencias de los fumadores es muy diferente de las de los otros dos grupos. Debido a que los dos grupos de no fumadores (expuestos y no expuestos) tienen una frecuencia tan alta de cantidades de la primera clase, sería útil comparar más esos conjuntos de datos examinando los valores con mayor detalle.

2-2 Destrezas y conceptos básicos

En los ejercicios 1 a 4, identifique la anchura de clase, las marcas de clase y las fronteras de clase para las distribuciones de frecuencias dadas, con base en el conjunto de datos 1 del Apéndice B.

1.

Presión sanguínea sistólica de varones	Frecuencia
90–99	1
100–109	4
110–119	17
120–129	12
130–139	5
140–149	0
150–159	1

2.

Presión sanguínea sistólica de mujeres	Frecuencia
80–99	9
100–119	24
120–139	5
140–159	1
160–179	0
180–199	1

3.

Colesterol en varones	Frecuencia
0–199	13
200–399	11
400–599	5
600–799	8
800–999	2
1000–1199	0
1200–1399	1

4.

Índice de masa corporal de mujeres	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

En los ejercicios 5 a 8, elabore la distribución de frecuencias relativas que corresponda a la distribución de frecuencias del ejercicio indicado.

5. Ejercicio 1 **6.** Ejercicio 2 **7.** Ejercicio 3 **8.** Ejercicio 4

En los ejercicios 9 a 12, construya la distribución de frecuencias acumulativas que corresponda a la distribución de frecuencias del ejercicio indicado.

9. Ejercicio 1 **10.** Ejercicio 2 **11.** Ejercicio 3 **12.** Ejercicio 4

13. Dado cargado El autor taladró un hoyo en un dado, lo rellenó con plomo y lo lanzó 200 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 correspondiente y determine si el dado en cuestión difiere significativamente de un dado que no ha sido "cargado".

Tabla del ejercicio 13

Resultado	Frecuencia
1	27
2	31
3	42
4	40
5	28
6	32

14. Lotería La distribución de frecuencias al margen se basa en los números *Win 4* de la lotería del estado de Nueva York, incluidos en el conjunto de datos 26 del Apéndice B. Elabore la distribución de frecuencias relativas correspondiente y determine si los resultados se seleccionaron de tal forma que todos los dígitos sean igualmente probables.

Tabla del ejercicio14

Dígito	Frecuencia
0	18
1	12
2	14
3	9
4	17
5	20
6	21
7	26
8	7
9	16

15. Osos Remítase al conjunto de datos 9 del Apéndice B y construya una distribución de frecuencias con los pesos de los osos. Utilice 11 clases, iniciando con el límite de clase inferior de 0, con una anchura de clase de 50 lb.

16. Temperaturas corporales Remítase al conjunto de datos del Apéndice B; después, construya una distribución de frecuencias de las temperaturas corporales para la medianoche del segundo día. Utilice ocho clases, iniciando con el límite de clase inferior de 96.5, con una anchura de clase de 0.4°F. Describa dos características notables del resultado.

17. Circunferencias de cabezas Remítase al conjunto de datos 3 del Apéndice B. Elabore una distribución de frecuencias con las circunferencias de las cabezas de bebés hombres; luego, construya una distribución de frecuencias separada para las circunferencias de las cabezas de los bebés mujeres. En ambos casos, utilice las clases de 34.0–35.9, 36.0–37.9, etcétera. Después compare los resultados y determine si hay una diferencia significativa entre los dos géneros.

18. Películas de dibujos animados para niños Remítase al conjunto de datos 7 del Apéndice B. Construya una distribución de frecuencias con la duración de las escenas de consumo de tabaco que presentan las películas de dibujos animados para niños; luego, elabore una distribución de frecuencias separada con la duración de las escenas en donde se consume alcohol. En ambos casos, utilice las clases de 0–99, 100–199, etcétera. Compare los resultados y determine si hay una diferencia significativa.

19. Corredores del maratón Remítase al conjunto de datos 8 del Apéndice B. Construya una distribución de frecuencias relativas con las edades de la muestra de hombres que terminaron el maratón de la ciudad de Nueva York; después, elabore una distribución de frecuencias relativas separada con las edades de las mujeres. En ambos casos, inicie la primera clase con el límite de clase inferior de 19, con una anchura de clase de 10. Compare los resultados y determine si hay alguna diferencia notable entre los dos grupos.

20. Coca Cola regular/Coca Cola dietética Remítase al conjunto de datos 17 del Apéndice B. Construya una distribución de frecuencias relativas con los pesos de la Coca Cola regular; inicie 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

21. Interpretación de los efectos de los datos distantes Remítase al conjunto de datos 20 del Apéndice B, de las cargas axiales de latas de aluminio de 0.0111 pulgadas de grosor. A la carga de 504 lb, se le denomina *dato distante*, ya que se encuentra muy lejos del resto de los valores. Construya una distribución de frecuencias que incluya el valor de 504 lb; después, elabore otra sin incluir este valor. En ambos casos, inicie la primera clase en 200 lb y utilice una anchura de clase de 20 lb. Interprete los resultados estableciendo una generalización acerca del efecto que tiene un dato distante en una distribución de frecuencias.

22. Número de clases Los lineamientos de Sturges para la construcción de una distribución de frecuencias sugieren que el número ideal de clases puede aproximarse por medio de $1 + (\log n)/(\log 2)$, donde n es el número de valores de datos. Utilice esta guía para completar la tabla y determine el número ideal de clases.

Tabla del ejercicio 22

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

2-3 Visualización de los datos

Recuerde que el principal objetivo de este capítulo es aprender técnicas importantes para investigar las características "CVDDT" importantes de los conjuntos de datos: centro, variación, distribución, datos distantes y cambios a lo largo del tiempo. En la sección 2-2 se introdujo la distribución de frecuencias como una herramienta para describir, explorar o comparar *distribuciones* de conjuntos de datos. En esta sección continuaremos el estudio de las distribuciones por medio de la introducción de gráficas, que son dibujos de distribuciones. Conforme avance en esta sección, considere que el objetivo no es simplemente la construcción de gráficas, sino más bien aprender algo acerca de los conjuntos de datos, es decir, comprender la naturaleza de sus distribuciones.

Histogramas

Entre los distintos tipos de gráficas que se presentan en esta sección, el histograma es particularmente importante.

Definición

Histograma es una gráfica de barras en 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 espacios entre ellas).

Es posible construir un histograma tras completar una tabla de distribución de frecuencias para un conjunto de datos. En la figura 2-1 se presentan los niveles de cotinina de fumadores, los cuales corresponden, de forma directa, a la distribución de frecuencias de la tabla 2-2, que se presentó en la sección previa. Cada barra del histograma está marcada con su frontera de clase inferior a la izquierda y su frontera

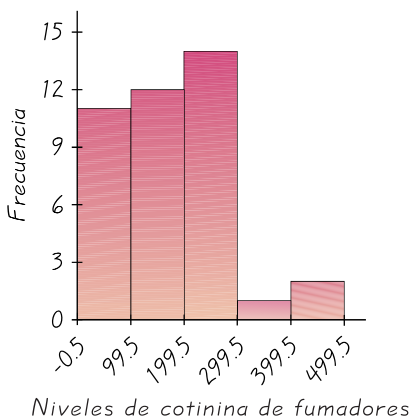


FIGURA 2-1 **Histograma**

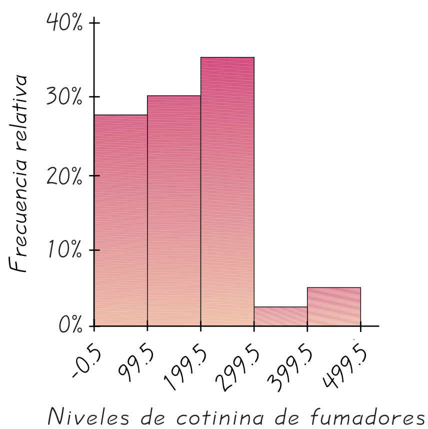


FIGURA 2-2 **Histograma de frecuencias relativas**

de clase superior a la derecha. En lugar de utilizar fronteras de clase a lo largo de la escala horizontal, suele ser más práctico utilizar los valores de las marcas de clase centradas por debajo de sus barras correspondientes. El uso de los valores de las marcas de clase es muy común en los programas de cómputo que generan histogramas de manera automática.

Antes de construir un histograma, a partir de una distribución de frecuencias completa, debemos mencionar algo acerca de las escalas que se utilizan en los ejes vertical y horizontal. La frecuencia máxima (o el siguiente número conveniente más alto) tiene que sugerir un valor para la parte superior de la escala vertical; el cero habrá de colocarse al inicio. En la figura 2-1 se diseñó una escala vertical que va de 0 a 15. La escala horizontal debe subdividirse de modo tal que permita que se ajusten bien todas las clases. De manera ideal, hay que tratar de seguir la regla práctica del intervalo, la cual establece que la altura vertical del histograma debe medir aproximadamente tres cuartas partes de la anchura total. Ambos ejes tienen que etiquetarse de forma clara.

Interpretación de un histograma Recuerde que el objetivo no es la simple construcción de un histograma, sino aprender algo acerca de los datos. Analice el histograma para ver qué es posible aprender acerca de "CVDDT": el centro de los datos, la variación (que se estudiará en la sección 2.5), la forma de la distribución y la existencia o ausencia de datos distantes (valores que se encuentran lejos de los demás). El histograma no es adecuado para determinar si hay cambios a lo largo del tiempo. Al examinar la figura 2-1, se verá que el histograma se centra alrededor del 175, que los valores varían aproximadamente desde 0 hasta 500 y que la distribución está más cargada hacia la izquierda.

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 lugar de las frecuencias reales, tal como sucede en la figura 2-2.

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. Véase la figura 2-3 en la página siguiente, que incluye 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 la izquierda, de manera que la gráfica inicia y termina sobre el eje horizontal.

Ojiva

Una ojiva es una gráfica lineal que representa frecuencias *acumulativas*, de la misma forma que la distribución de frecuencias acumulativas es una lista de éstas (véase la tabla 2-4 en la sección anterior). La figura 2-4 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 empieza con la frontera inferior de la primera clase, en tanto que finaliza con la frontera superior de la última clase. Las ojivas son útiles para determinar el número de valores que se encuentran por debajo de un valor

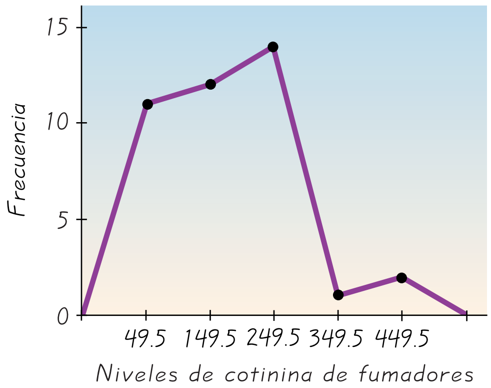


FIGURA 2-3 Polígono de frecuencias

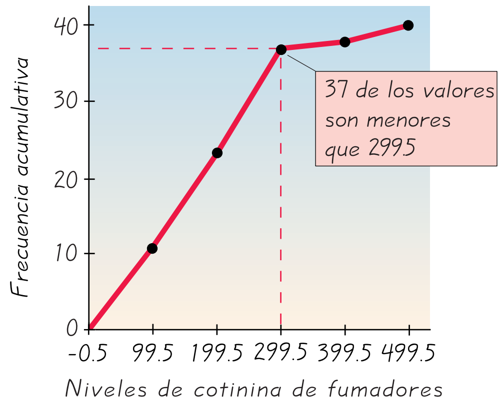


FIGURA 2-4 Ojiva

particular. Por ejemplo, la figura 2-4 muestra que 37 de los valores del nivel de cotinina son menores que 299.5.

Gráficas de puntos

Una **gráfica de puntos** consiste en una gráfica en 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 amontonan. Observe la figura 2-5, que representa la duración de películas de dibujos animados para niños, que se listan en el conjunto de datos 7 del Apéndice B. Por ejemplo, los dos puntos que aparecen a la izquierda representan el valor de 64 minutos, que ocurre dos veces en el conjunto de datos 7. En esta gráfica de puntos vemos que la duración de 120 minutos difiere mucho de las demás.

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 ilustración de la siguiente página muestra una gráfica de tallo y hojas de las mismas duraciones de películas listadas en el conjunto de datos 7 del Apéndice B. Dichas duraciones (en minutos), si se acomodan en orden creciente, son 64, 64, 69, 70, 71, 71, 71, 72, 73, . . . , 120. Es fácil ver cómo el primer valor de 64 se separó en su tallo de 6 y su hoja de 4. Cada uno de los valores restantes, lo hace de una manera similar. Note que las hojas se ordenaron en forma creciente y no en el orden en que aparecen en la lista original.

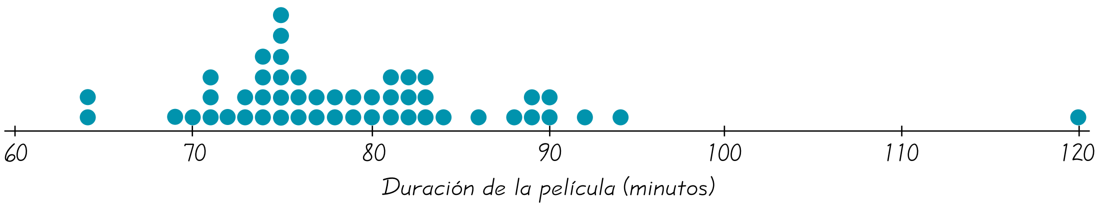


FIGURA 2-5 Gráfica de puntos de la duración de películas para niños

Gráfica de tallo y hojas

Tallo (decenas)	Hojas (unidades)	
6	449	←Los valores son 64, 64, 69.
7	01112334444555555666778899	
8	0011122233346899	
9	0024	
10		
11		
12	0	← El valor es 120.

Si colocamos la página de lado, veremos una distribución de tales datos. Una gran ventaja de la gráfica de tallo y hojas radica en que nos permite ver la distribución de los datos y, al mismo tiempo, retener toda la información de la lista original. En caso de ser necesario, reconstruiríamos la lista original de valores. Otra ventaja es que la construcción de una gráfica de tallo y hojas implica una forma fácil y rápida de *ordenar* datos, y algunos procedimientos estadísticos requieren de un ordenamiento (como el cálculo de una mediana o de los percentiles).

Los renglones de datos de una gráfica de tallo y hojas son similares en naturaleza a las barras de un 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 si redondeamos 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. En nuestro ejemplo, la gráfica de tallo y hojas puede expandirse subdividiendo los renglones en otros con hojas que incluyan dígitos del 0 al 4, así como otros con dígitos del 5 al 9, tal como se muestra en el siguiente diagrama.

Gráfica expandida de tallo y hojas

Tallo	Hojas	
6	44	←Para hojas de 0 a 4
6	9	← Para hojas de 5 a 9
7	01112334444	
7	555555666778899	
8	001112223334	
8	6899	
9	0024	
9		
10		
10		
11		
11		
12	0	

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 la escuela preparatoria "deja muy poco espacio para el estudio de la estadística y la probabilidad. Sin embargo, los estudiantes necesitan fundamentos sobre el análisis de datos". El reportero observó que el cálculo tiene un papel prominente en los estudios universitarios, aun cuando "sólo algunos trabajos, principalmente en áreas técnicas, realmente lo utilizan". Rothstein citó un estudio realizado por el profesor Clifford Konold, de la Universidad de Massachusetts, quien contó el número de desplegados de datos que aparecen en el *New York Times*. En los ejemplares de 1972, el doctor Konold encontró cuatro gráficas o tablas en cada una de las 10 ediciones semanales (sin incluir las secciones de deportes y negocios), pero en 1982 había ocho, en 1992 fueron 44 y "el próximo año, él (el doctor Konold) podría encontrar más de 100". El crecimiento de la estadística como una disciplina se fomenta, en parte, por el uso creciente de dichos desplegados de datos en los medios de comunicación.

Cuando hay necesidad de *reducir* el número de renglones, es posible condensar una gráfica de tallo y hojas al combinar los renglones adyacentes, tal como se indica en la siguiente ilustración. Note que insertamos un asterisco para separar los dígitos en las hojas asociadas con los números en cada tallo. Cada renglón en la gráfica condensada debe incluir exactamente un asterisco, de modo que la forma de la gráfica no se distorsione.

Gráfica condensada de tallo y hojas

Tallo	Hojas	
6-7	449*0111233444555555666778899	← 64, 64, 69, 70, . . . , 79
8-9	00111222333346899*0024	
10-11	*	
12-13	0*	← El valor es 120.

Gráficas de Pareto

La Federal Communications Commission (FCC) verifica la calidad del servicio telefónico en Estados Unidos. Algunas de las quejas en contra de las compañías telefónicas incluyen los cambios, es decir, se cambia de compañía al cliente sin su consentimiento, y el cobro forzoso de 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, 12,478 por cambios sin consentimiento y 1214 por forzamiento. Si usted fuese reportero de un medio impreso, ¿cómo presentaría dicha información? La simple escritura de oraciones con datos numéricos no llevaría a una verdadera comprensión. Un mejor método consiste en utilizar una gráfica conveniente; en este caso, la gráfica de Pareto se adecuaría muy bien.

Una **gráfica de Pareto** es una gráfica de barras para datos cualitativos, donde las barras se ordenan de acuerdo con las frecuencias. Al igual que en los histogramas, las escalas verticales de las gráficas de Pareto representan frecuencias o frecuencias relativas. La barra más alta se coloca a la izquierda y las más pequeñas hacia la derecha. Al ordenar las barras por frecuencias, la gráfica enfoca la atención en las categorías más importantes. La figura 2-6 es una gráfica de Pareto que muestra con claridad que el cambio sin consentimiento es, por mucho, el asunto más grave de las quejas de los clientes respecto de las empresas telefónicas.

Gráficas circulares

Las gráficas circulares también se utilizan para visualizar datos cualitativos. La figura 2-7 es un ejemplo de una **gráfica circular**, que presenta datos cualitativos como si fueran rebanadas de un pastel. La figura 2-7 representa los mismos datos de la figura 2-6. Para construir una gráfica circular, se separa el círculo en las proporciones que se adecuan mejor. La categoría de quejas por cambio sin consentimiento representan un 59% del total, de manera que la porción que representa el cambio sin consentimiento debe abarcar el 59% del total (con un ángulo central de $0.59 \times 360° = 212°$).

La gráfica de Pareto (figura 2-6) y la gráfica circular (figura 2-7) 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, lo cual explica por qué muchas compañías, como Boeing Aircraft, a menudo utilizan las gráficas de Pareto.

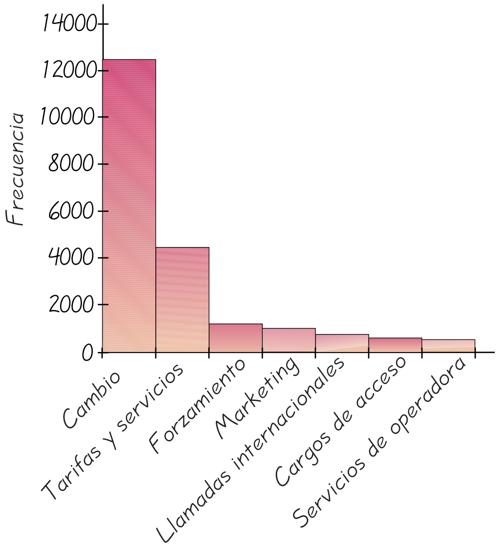


FIGURA 2-6 Gráfica de Pareto de quejas en contra de las compañías telefónicas

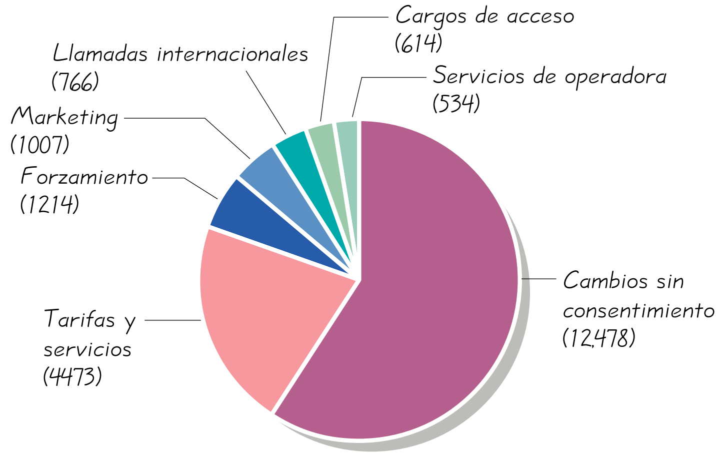
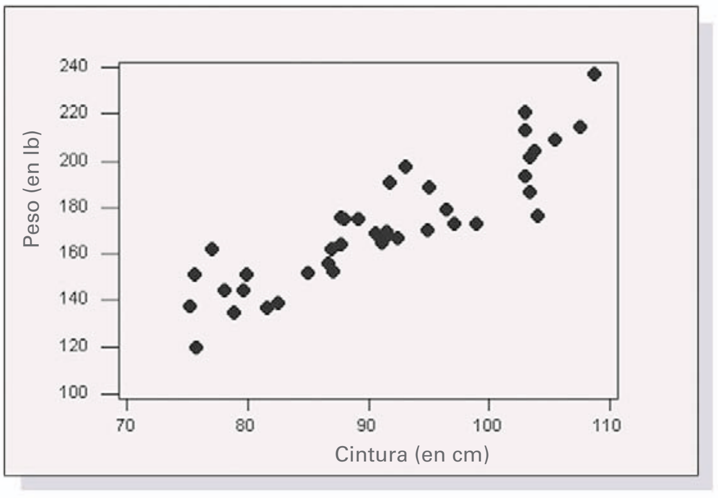


FIGURA 2-7 Gráfica circular de quejas en contra de las compañías telefónicas

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 un diagrama de dispersión manualmente, construya un eje horizontal para los valores de la primera variable y 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 hay alguna relación entre las dos variables. (Este aspecto se estudia a profundidad en el tema de la correlación, en la sección 9-2). Con los datos del peso (en libras) y la circunferencia de la cintura (en cm) de los varones del conjunto de datos 1 del Apéndice B, utilizamos Minitab para generar el diagrama de dispersión que aparece a continuación. Con base en dicha gráfica, parece haber una relación entre el peso y la circunferencia de la cintura, tal como lo muestra el patrón de puntos.

Minitab



Florence Nightingale

A Florence Nightingale (1820-1910) se le reconoce como la fundadora de la profesión de enfermería, aunque también salvó miles de vidas usando la estadística. Cuando encontraba un hospital insalubre y con desabasto, mejoraba dichas condiciones y después utilizaba la estadística para convencer a otros de la necesidad de una reforma médica amplia. Ella diseñó gráficas originales para ilustrar que, durante la guerra de Crimea, murieron más soldados a consecuencia de las condiciones insalubres que en combate. Florence Nightingale fue pionera en el uso de la estadística social y de las técnicas gráficas.

Gráficas de series de tiempo

Los **datos de series de tiempo** son aquellos que se reúnen en diferentes momentos. Por ejemplo, la figura 2-8 muestra el número de pantallas de autocinemas existentes durante un periodo de 14 años (con base en datos de la National Association of Theater Owners). Vemos que durante este tiempo hay 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 la bancarrota porque no verificaban la calidad de sus bienes o servicios; además, de manera incorrecta, creían estar tratando con datos estables. No se dieron cuenta de que sus productos se volvían defectuosos conforme cambiaban importantes características de la población. El capítulo 13 introduce las *gráficas de control*, que son herramientas eficaces para verificar datos de series de tiempo.

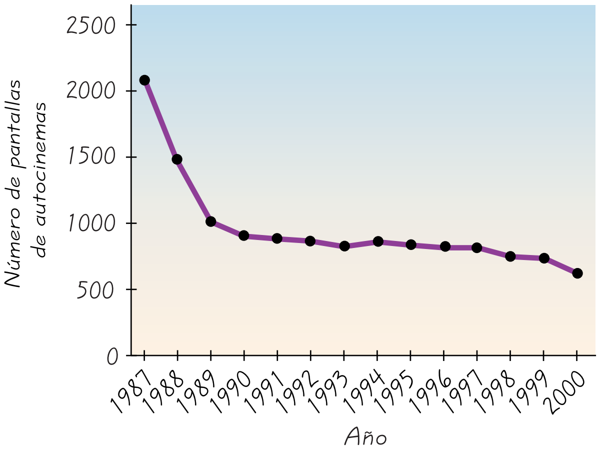
Otras gráficas

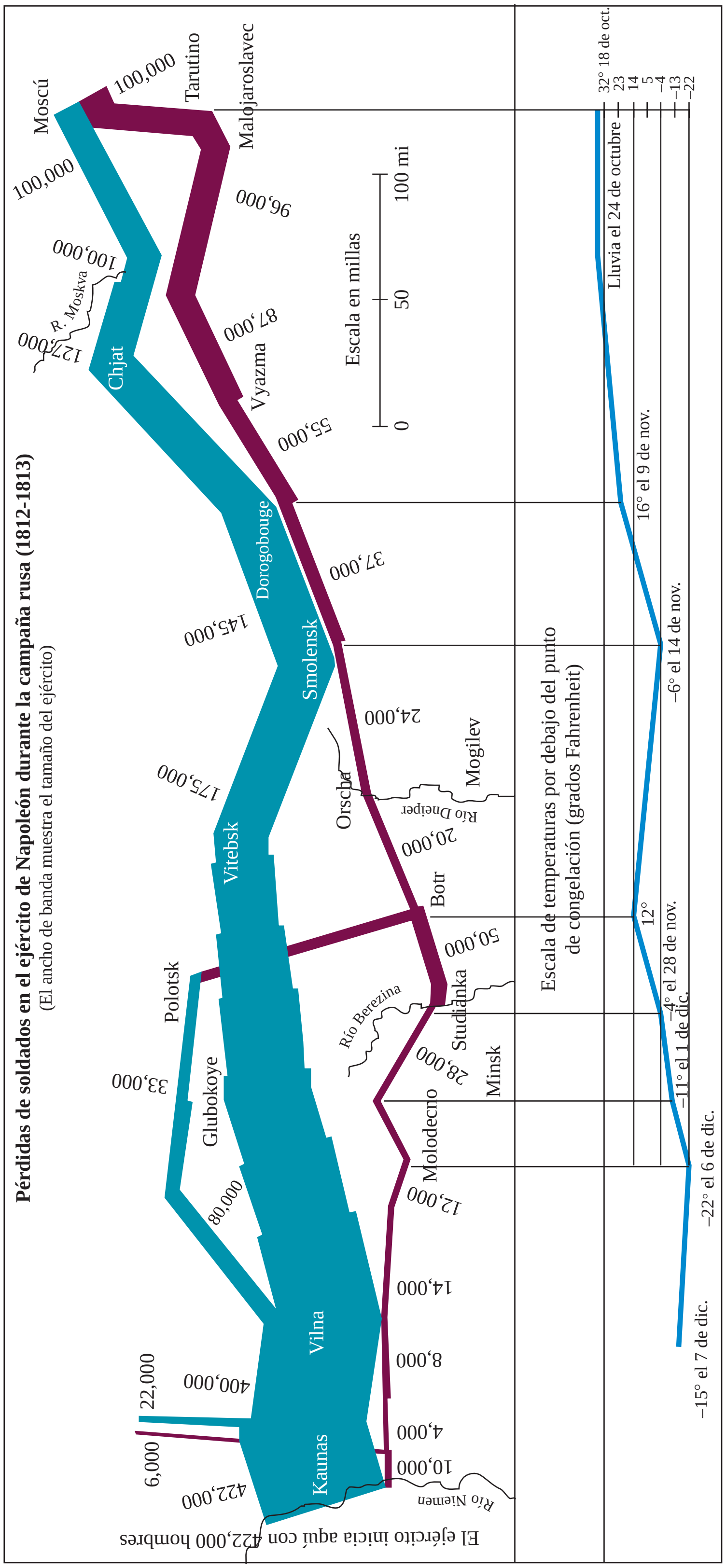
Además de las presentaciones gráficas descritas, hay muchas otras que pueden utilizarse para representar datos de manera llamativa y efectiva. En la sección 2-7 presentaremos las gráficas de cuadro, que son muy útiles para revelar la distribución de los datos. Los pictogramas representan datos con el uso de imágenes de objetos como soldados, tanques, aviones, monedas o bolsas de dinero.

La figura que aparece en la página 53 se ha descrito quizá como "la mejor gráfica estadística que se haya dibujado jamás". Esta figura incluye seis variables diferentes con respecto a la marcha del ejército de Napoleón hacia Moscú en 1812-1813. La banda gruesa a la izquierda representa el tamaño del ejército cuando inició la invasión a Rusia, desde Polonia. 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, incluso 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 2-9 de la página 54, es particularmente interesante porque 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 se asemeja a una gráfica circular, pese a que todos los ángulos centrales

FIGURA 2-8 Datos de series de tiempo: número de pantallas de autocinemas





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

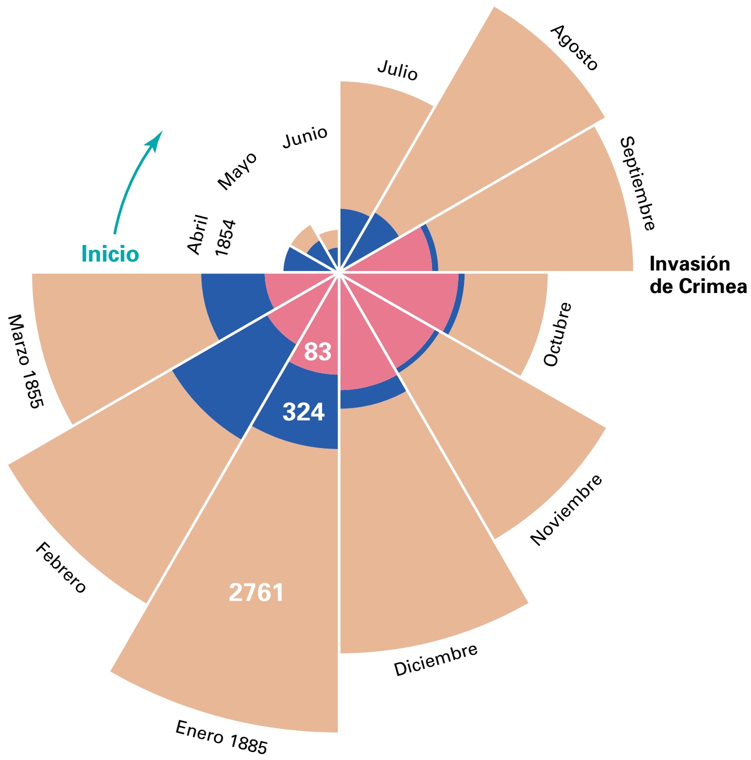


FIGURA 2-9 Muertes en los hospitales militares británicos durante la guerra de Crimea

son iguales y se usan radios diferentes para mostrar los cambios en el número de muertes mensuales. Las regiones externas de la figura 2-9 representan las muertes por enfermedades que pudieron prevenirse; las regiones internas representan muertes por heridas y las regiones centrales, muertes por otras causas.

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 datos distantes (CVDDT, sin el último elemento del tiempo). ¿Cuál es el valor aproximado del centro de la distribución y cuál es el rango aproximado de valores? Considere la forma completa de la distribución. ¿Están los valores distribuidos de manera uniforme? ¿La distribución está sesgada (ladeada) hacia la izquierda o hacia la derecha? ¿Tiene la distribución un pico a la mitad? Identifique cualquier valor extremo y cualquiera 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 2-9 observamos que morían más soldados por cuidados hospitalarios inadecuados que por heridas de 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 debe aparecer a la derecha del polígono de mujeres, mostrando así que los hombres tienen pesos mayores.

Utilizando la tecnología

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

STATDISK Puede generar histogramas y diagramas de dispersión.

Minitab Puede generar todas las gráficas incluidas en esta sección.

Excel Puede generar histogramas, polígonos de frecuencias, gráficas circulares y diagramas de dispersión.

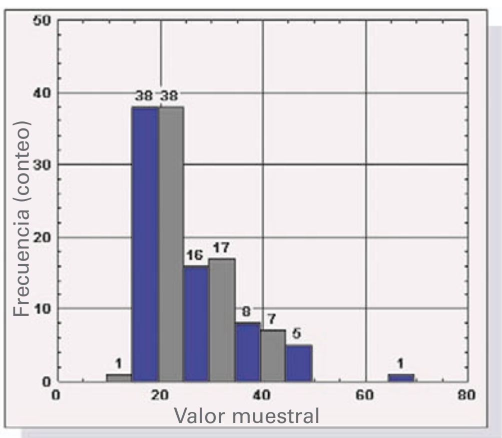
TI-83 Plus Puede generar histogramas y diagramas de dispersión.

2-3 Destrezas y conceptos básicos

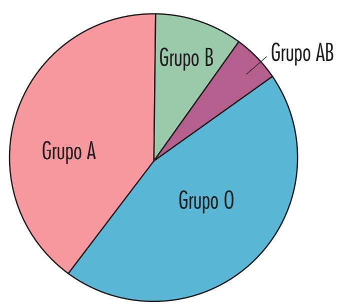
En los ejercicios 1 a 4, conteste las preguntas con respecto al histograma que se genera con STATDISK, el cual representa las edades de todos los polizontes del Queen Mary.

1. Centro ¿Cuál es el valor aproximado del centro? Es decir, ¿qué edad parece estar cerca del centro de todas las edades?
2. Variación ¿Cuáles son las edades más bajas y más altas posibles?
3. Porcentaje ¿Qué porcentaje de los 131 polizontes tenía menos de 30 años de edad?
4. Anchura de clase ¿Cuál es la anchura de clase?

STATDISK



En los ejercicios 5 y 6, remítase a la gráfica circular adjunta referente a los grupos sanguíneos de una muestra grande de personas (con base en datos del Greater New York Blood Program).



5. Interpretación de la gráfica circular ¿Cuál es el porcentaje aproximado de individuos con sangre tipo A? Suponiendo que la gráfica circular se base en una muestra de 500 personas, ¿aproximadamente cuántas de ellas tienen sangre tipo A?
6. Interpretación de la gráfica circular ¿Cuál es el porcentaje aproximado de personas con sangre tipo B? Suponiendo que la gráfica circular se base en una muestra de 500 personas, ¿aproximadamente cuántas de ellas tienen sangre tipo B?

Tabla del ejercicio 7

Edad	Estudiantes	Profesores
0–2	23	30
3–5	33	47
6–8	63	36
9–11	68	30
12–14	19	8
15–17	10	0
18–20	1	0
21–23	0	1

Tabla del ejercicio 8

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

7. Automóviles de estudiantes/profesores Se obtuvieron muestras de automóviles de estudiantes y profesores en la universidad donde trabaja el autor. Sus edades (en años) se resumen en la distribución de frecuencias adyacente. Construya un histograma de frecuencias relativas para los automóviles de los estudiantes y otro histograma de frecuencias relativas para los automóviles de los profesores. Compare ambos. ¿Cuáles son las diferencias más notables?

8. Infracciones por exceso de velocidad La distribución de frecuencias adyacente describe las velocidades de conductores a quienes infraccionó la policía en la ciudad de Poughkeepsie. Estos conductores viajaban en una zona con límite de velocidad de 30 millas/hora en Creek Road, que atraviesa la universidad del autor. Construya un histograma correspondiente a la distribución de frecuencias. ¿Qué sugiere la distribución sobre el límite de velocidad establecido, comparándolo con el límite de velocidad señalado?

9. Osos El ejercicio 15, en la sección 2-2, se refiere al conjunto de datos 9 del Apéndice B. Use la distribución de frecuencias del peso de los osos (con 11 clases, iniciando con un límite de clase inferior de 0 y una anchura de clase de 50 lb), y construya el histograma correspondiente. ¿Cuál es el peso aproximado que se encuentra en el centro?

10. Temperaturas corporales El ejercicio 16, en la sección 2-2, se refiere al conjunto de datos 4 del Apéndice B. Con la distribución de frecuencias de las temperaturas corporales de medianoche del segundo día (con ocho clases, iniciando con el límite de clase inferior de 96.5 y una anchura de clase de 0.4°F), elabore el histograma correspondiente. ¿Qué sugiere la distribución sobre la creencia común de que la temperatura corporal promedio es de 98.6°F? Si se selecciona a los sujetos de forma aleatoria, las temperaturas deberían tener una distribución aproximadamente normal. ¿Es así?

En los ejercicios 11 a 14, realice las comparaciones construyendo las gráficas que se indican.

11. Circunferencia de cabezas El ejercicio 17, en la sección 2-2, se refiere al conjunto de datos 3 del Apéndice B. Utilice la distribución de frecuencias de la circunferencia de la cabeza de los niños y la distribución de frecuencias de la circunferencia de la cabeza de las niñas (con las clases de 34.0-35.9, 36.0-37.9, etcétera.), y construya los dos polígonos de frecuencias correspondientes, utilizando el mismo conjunto de ejes. Compare los resultados y determine si parece haber una diferencia significativa entre los dos géneros.

12. Películas de dibujos animados para niños El ejercicio 18, en la sección 2.2, se refiere al conjunto de datos 7 del Apéndice B. Utilice la distribución de frecuencias de la duración de las películas de dibujos animados para niños que incluyen consumo de tabaco y la distribución de frecuencias de la duración de aquéllas que presentan escenas de consumo de alcohol (con clases de 0-99, 100-199, etcétera.), y construya los dos polígonos de frecuencias correspondientes, usando el mismo conjunto de ejes. Compare los resultados y determine si parece haber una diferencia significativa.

13. Corredores del maratón El ejercicio 19, en la sección 2-2, se refiere al conjunto de datos 8 del Apéndice B. Utilice la distribución de frecuencias relativas de las edades de varones y la distribución de frecuencias relativas de las edades de mujeres (con un límite de clase inferior de 19 y una anchura de clase de 10), y construya los histogramas de frecuencias relativas correspondientes. Compare los resultados y determine si parece haber diferencias notables entre los dos grupos.

14. Coca Cola regular y Coca Cola dietética Remítase al conjunto de datos 17 del Apéndice B y utilice los pesos de la Coca Cola regular y de la Coca Cola dietética. Con las clases de 0.7750-0.7799, 0.7800-0.7849, . . . , 0.8250-0.8299, construya los dos polígonos de frecuencias sobre los mismos ejes. Después, compare los resultados y determine si parece haber una diferencia significativa. ¿Cuál sería una posible explicación de la diferencia?

En los ejercicios 15 y 16, liste los datos originales que se representan con las gráficas de tallo y hojas.

15.

Tallo (decenas)	Hojas (unidades)
20	0005
21	69999
22	2233333
23	
24	1177

16.

Tallo (centenas)	Hojas (decenas y unidades)
50	12 12 12 55
51	
52	00 00 00 00
53	27 27 35
54	72

En los ejercicios 17 y 18, construya la gráfica de puntos con los datos que se representan en la gráfica de tallo y hojas del ejercicio dado.

17. Ejercicio 15

18. Ejercicio 16

En los ejercicios 19 y 20, elabore las gráficas de tallo y hojas para los conjuntos de datos que se indican, que se encuentran en el Apéndice B.

19. Osos La longitud (en pulgadas) de los osos en el conjunto de datos 9. (*Sugerencia*: Primero redondee las longitudes hacia el entero más cercano).

20. Plástico Los pesos (en libras) del plástico que desechan 62 amas de casa: remítase al conjunto de datos 23, e inicie redondeando los pesos hacia el decimal más cercano. (Utilice una gráfica expandida de tallo y hojas, con aproximadamente 11 renglones).

21. Empleos Se realiza un estudio para determinar la manera en que las personas obtienen empleo. La tabla incluye datos de 400 sujetos que se seleccionaron aleatoriamente. Los datos se basan en resultados del National Center for Carrer Strategies. Construya la gráfica de Pareto correspondiente a tales datos. Si alguien deseara obtener un empleo, ¿cuál parece ser el método más efectivo?

Fuentes de empleo de sujetos que se encuestaron	Frecuencia
Anuncios clasificados	56
Empresas que buscan ejecutivos	44
Contactos interpersonales	280
Envíos por correo	20

22. Empleos Remítase a los datos del ejercicio 21 y construya una gráfica circular. Compárela con la gráfica de Pareto. ¿Podría determinar cuál gráfica es más efectiva para mostrar la importancia relativa de las fuentes de empleo?

23. Descarrilamiento de trenes Un análisis del descarrilamiento de trenes mostró que 23 de éstos fueron causados por vías en mal estado, nueve por fallas en el equipo, 12 por errores humanos y seis por otras causas (con base en datos de la Federal Railroad Administration). Construya una gráfica circular que represente tales datos.

24. Descarrilamiento de trenes Remítase a los datos del ejercicio 23 y elabore una gráfica de Pareto. Compare dicha gráfica con la gráfica circular. ¿Podría determinar cuál de las gráficas es más efectiva para mostrar la importancia relativa de las causas de los descarrilamientos de trenes?

En los ejercicios 25 y 26, utilice los datos apareados del apéndice B para construir un diagrama de dispersión.

25. Alquitrán/ CO Para el conjunto de datos 5, ubique el alquitrán en la escala horizontal y el monóxido de carbono en la escala vertical. Determine si parece haber una relación entre el alquitrán y el monóxido de carbono. Si es así, describa dicha relación.

26. Cuello/peso de osos Para el conjunto de datos 9, ubique las medidas del cuello en la escala horizontal y los pesos de los osos en la escala vertical. Con base en los resultados, ¿cuál es la relación existente entre el tamaño del cuello y el peso de los osos?

En los ejercicios 27 y 28, use los datos del Apéndice B para construir una gráfica de series de tiempo.

27. Inversiones en acciones Para el conjunto de datos 25, utilice los valores altos del Dow Jones Industrial Average (DJIA) para construir una gráfica de series de tiempo; después, determine si parece haber alguna tendencia. ¿Cómo podría un inversionista beneficiarse de esta tendencia?

28. Muertes en vehículos automotores En el conjunto de datos 25, utilice los datos de las muertes en vehículos automotores en Estados Unidos para construir una gráfica de series de tiempo; después, determine si parece haber alguna tendencia. Si es así, ofrezca una posible explicación.

En los ejercicios 29 a 32, remítase a la figura de la página 53, que describe la campaña de Napoleón de 1812 hacia Moscú y su retirada. La banda gruesa a la izquierda representa el tamaño del ejército cuando comenzó a invadir Rusia desde Polonia; la banda inferior describe la retirada de Napoleón.

29. Calcule el porcentaje de hombres que sobrevivieron toda la campaña.

30. Calcule el número de hombres y el porcentaje de hombres que murieron durante el cruce del río Berezina.

31. ¿Cuántos hombres murieron durante la retirada de Moscú, en el tiempo cuando la temperatura bajó de 16°F hasta −6°F?

32. De los hombres que lograron llegar a Moscú, ¿cuántos murieron en el viaje de regreso entre Moscú y Botr? (Observe que 33,000 hombres no fueron a Moscú, pero se unieron a los hombres que regresaban).

2-3 Más allá de lo básico

33. **a.** Remítase al conjunto de datos 20 del Apéndice B y elabore un histograma con las cargas axiales de las latas que tienen un grosor de 0.0111 pulgadas. El conjunto de datos incluye un dato distante de 504 lb. (Un dato distante es un valor que aparece muy lejos de los demás valores).

b. Repita el inciso *a*) después de excluir el dato distante de 504 lb.

c. ¿Qué efecto produce un dato distante en la forma del histograma?

34. Los Óscares En el artículo "Ages of Oscar-winning Best Actors and Actresses" (revista *Mathematics Teachers*), escrito por Richard Brown y Gretchen Davis, se utilizan gráficas de tallo y hojas para comparar las edades de los actores y las actrices en el momento que ganaron un Óscar. A continuación, se presentan los resultados de ganadores recientes, para cada categoría.

Actores:	32	37	36	32	51	53	33	61	35	45	55	39
	76	37	42	40	32	60	38	56	48	48	40	43
	62	43	42	44	41	56	39	46	31	47	45	60
	46	40	36									
Actrices:	50	44	35	80	26	28	41	21	61	38	49	33
	74	30	33	41	31	35	41	42	37	26	34	34
	35	26	61	60	34	24	30	37	31	27	39	34
	26	25	33									

a. Construya una gráfica de tallo y hojas, espalda con espalda, con los datos. Las primeras dos edades de cada grupo se insertaron al margen.

b. Utilice los resultados del inciso *a*), compare los dos conjuntos de datos y explique cualquier diferencia.

Edades de los actores (unidades)	Tallo (decenas)	Edades de las actrices (unidades)
	2	
72	3	
	4	4
	5	0
	6	
	7	
	8	

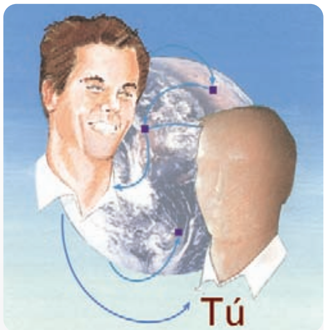
2-4 Medidas de tendencia central

Recuerde que el principal objetivo de este capítulo es lograr manejar las herramientas básicas para medir y describir diferentes características de un conjunto de datos. En la sección 2-1 observamos que, cuando describimos, exploramos y comparamos conjuntos de datos, las siguientes características suelen ser extremadamente importantes: centro, variación, distribución, datos distantes, cambios a través del tiempo. Las siglas CVDDT ("Cuidado con los Virus que Destruyen Datos y Trabajo") son útiles para recordar dichas características. En las secciones 2-2 y 2-3 señalamos que las distribuciones de frecuencias y las gráficas, así como los histogramas, sirven para investigar la distribución. En esta sección trataremos las características del *centro*.

Definición

Medida de tendencia central: 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 tanto, tenemos diferentes definiciones de las medidas de tendencia central, incluyendo media, mediana, moda y mitad del rango. Comenzaremos con la media.



Seis grados de separación

Los psicólogos sociales, historiadores, científicos políticos y especialistas en comunicaciones se interesan en el "problema del mundo pequeño": dadas dos personas cualesquiera en el mundo, ¿cuántos vínculos intermedios son necesarios para conectarlas? El psicólogo social Stanley Milgram realizó un experimento donde algunos sujetos intentaron ponerse en contacto con otras personas en específico, enviando por correo un archivo de información a un conocido que ellos pensaban estaba más cerca de la persona que buscaban. De las 160 cadenas de este tipo que se iniciaron, sólo 44 se completaron. El número de conocidos intermediarios varió de entre 2 y 10, con una mediana de 5 (o "seis grados de separación"). El experimento fue criticado por incluir sujetos muy sociales y por no hacer ajustes a las muchas conexiones perdidas de personas con ingresos más bajos. Otro estudio matemático mostró que si las cadenas perdidas se hubiesen completado, la mediana sería ligeramente mayor que 5.

Media

La media (aritmética) generalmente es la más importante de todas las medidas numéricas utilizadas para describir datos; constituye lo que la mayoría de la gente denomina *promedio*.

Definición

Media aritmética (de un conjunto de puntajes): medida de tendencia central que se obtiene sumando los puntajes y dividiendo el total entre el número de puntajes. Tal medida de tendencia central se utilizará de manera frecuente a lo largo del libro; además, nos referiremos a ella simplemente como la **media**.

Esta definición se expresa como la fórmula 2-1, que utiliza la letra griega Σ (sigma mayúscula) para indicar que los valores de los datos deben sumarse. Esto es, Σx representa la sumatoria de todos los valores de los datos. El símbolo n denota el **tamaño de la muestra**, que es el número de puntajes en el conjunto de datos.

Fórmula 2-1

$$\text{media} = \frac{\Sigma x}{n}$$

La media se denota como $\bar{x}$ (se denomina "x barra"), si el conjunto de datos es una *muestra* de una población más grande; si se utilizan todos los puntajes de la población, entonces la media se simboliza con μ (mu minúscula). (Los estadísticos de una muestra generalmente se representan con letras inglesas, tales como $\bar{x}$, y los parámetros de la población con letras griegas, tales como μ).

Notación

Σ	denota la *sumatoria* de un conjunto de valores.
x	es la *variable* que suele utilizarse para representar los valores de datos individuales.
n	representa el *número de valores* de una *muestra*.
N	representa el *número de valores* de una *población*.
$\bar{x} = \frac{\Sigma x}{n}$	es la media de un conjunto de valores *muestrales*.
$\mu = \frac{\Sigma x}{N}$	es la media de todos los valores de una *población*.

EJEMPLO Verificación del plomo en el aire A continuación se presentan cantidades de plomo medidas (en microgramos por metro cúbico o $\mu g/m^3$) en el aire. La Environmental Protection Agency estableció un estándar de calidad del aire respecto del plomo: 1.5 $\mu g/m^3$. Las mediciones que se presentan más adelante 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. Después del colapso de los dos edificios del World Trade Center, surgió una gran preocupación respecto de la calidad del aire. 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 2-1. Primero se suman los puntajes y después se dividen entre el número de ellos:

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

La media del nivel de plomo es 1.538 $\mu g/m^3$. 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 "dato distante" como éste. En tal caso, el nivel de plomo de 5.40 $\mu g/m^3$ se midió un día después del colapso de las torres del World Trade Center y los niveles de polvo y humo ofrecieron una explicación razonable para un valor tan extremo.

Una desventaja de la media es su sensibilidad a cada valor, de modo que un puntaje excepcional puede afectarla de manera drástica. La mediana resuelve, en gran parte, esa desventaja.

Mediana

Definición

Mediana (de un conjunto de datos): medida de tendencia central que implica el valor que está en medio, cuando los valores originales de los datos se presentan en orden de magnitud creciente (o decreciente). La mediana suele denotarse con $\tilde{x}$ (se pronuncia "*x* con tilde").

Para calcular la mediana, primero clasifique los valores (acomódelos en orden), luego siga uno de estos 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.

La figura 2-10 demuestra este procedimiento para el cálculo de la mediana.

EJEMPLO Verificación del plomo en el aire A continuación se presentan cantidades de plomo medidas (en $\mu g/m^3$) 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

continúa

Paradoja del tamaño del grupo escolar

Hay al menos dos formas de obtener la media de un grupo escolar, y pueden dar resultados muy distintos. En una universidad, si consideramos las cantidades de estudiantes en 737 grupos, obtenemos una media de 40 estudiantes. Pero si compilamos una lista de los tamaños de grupo para cada estudiante y utilizamos dicha lista, obtendríamos una media de grupo de 147. Esta gran discrepancia es por el hecho de que hay muchos estudiantes en los grupos grandes, mientras que en los grupos pequeños hay muy pocos. Sin modificar el número de grupos ni de profesores, reduciríamos el tamaño medio del grupo escolar que los estudiantes experimentan, haciendo que todos los grupos tengan aproximadamente el mismo tamaño. Lo anterior también incrementaría la asistencia, que es mayor en los grupos escolares pequeños.

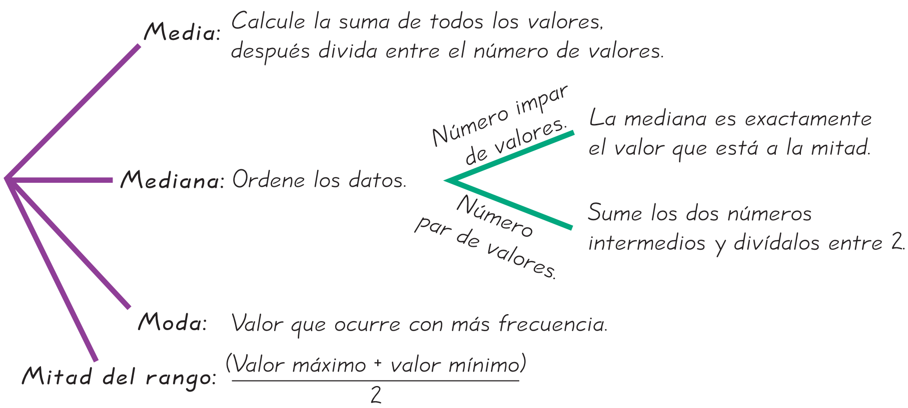


FIGURA 2-10 **Procedimientos para calcular las medidas de tendencia central**

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 es el número que se encuentra exactamente a la mitad de la lista ordenada; por lo tanto, la mediana es 0.915 $\mu g/m^3$. Note que la mediana es muy diferente de la media de 1.538 $\mu g/m^3$, que se obtuvo del mismo conjunto de datos muestrales del ejemplo anterior. La razón de esa gran discrepancia es el efecto que el puntaje 5.40 tuvo en la media. Si este valor extremo se redujera a 1.20, la media caería de 1.538 $\mu g/m^3$ hasta 0.838 $\mu g/m^3$, pero la mediana no cambiaría.

EJEMPLO Verificación del plomo en el aire Repita el ejemplo anterior, después de incluir la medición de 0.66 $\mu g/m^3$, que se registró otro día. Es decir, calcule la mediana de estas mediciones del 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 exactamente el valor a la mitad de la lista ordenada: 0.73 $\mu g/m^3$.

Después de estudiar los ejemplos anteriores, debe quedar claro el procedimiento para obtener la mediana; además, que la media se ve afectada de manera drástica por valores extremos, mientras que la mediana no. Puesto que la mediana no es tan sensible a los valores extremos, con frecuencia se utiliza para conjuntos de datos que tienen un número relativamente pequeño de datos distantes. Por ejemplo, recientemente la oficina de censos de Estados Unidos reportó que la *mediana* del ingreso familiar es de 36,078 dólares anuales. Se usó la mediana ya que existen pocas familias con ingresos realmente altos.

Moda

Definiciones

Moda (de un conjunto de datos, que suele denotarse como *M*): valor que ocurre con mayor frecuencia.

- Cuando dos valores ocurren 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 ocurren 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.

a. 5.40 1.10 0.42 0.73 0.48 1.10

b. 27 27 27 55 55 55 88 88 99

c. 1 2 3 6 7 8 9 10

SOLUCIÓN

a. El número 1.10 es la moda, ya que es el valor que ocurre con mayor frecuencia.

b. Los números 27 y 55 son modas, ya que ambos ocurren con la frecuencia más alta. Este conjunto de datos es bimodal, porque tiene dos modas.

c. No hay moda, ya que ningún valor se repite.

En realidad, la moda no se utiliza mucho con datos numéricos. Sin embargo, entre las distintas medidas de tendencia central que consideramos, la moda es la única que puede usarse con datos de nivel nominal de medición. (Recuerde que el nivel nominal de medición se aplica a datos que consisten únicamente en nombres, etiquetas o categorías). Por ejemplo, una encuesta que se aplicó a estudiantes universitarios mostró que el 84% tiene aparato de televisión; el 76%, videocasetera; el 60%, reproductor de discos compactos portátil; el 39%, sistema de videojuegos y el 35%, reproductor de DVD (con base en datos del National Center for Education Statistics). En tanto que el televisor es el aparato más frecuente, es posible afirmar que la moda es el televisor. No podemos calcular una media o mediana para datos como éstos, a nivel nominal.

LA ESTADÍSTICA EN LAS NOTICIAS

Un hombre promedio

La revista *Men´s Health* publicó estadísticas que describen al "hombre promedio", que tiene 34.4 años de edad, pesa 175 libras, mide cerca de 5 pies 10 pulgadas y se llama Mike Smith. La edad, el peso y la estatura son valores medios, pero el nombre de Mike Smith es la moda que corresponde al nombre y apellido más comunes. Otra estadística notable es la siguiente: el hombre promedio duerme aproximadamente 6.9 horas por noche, bebe cerca de 3.3 tasas de café al día y consume 1.2 bebidas alcohólicas diariamente; además, gana alrededor de 36,100 dólares anuales, debe 2,563 dólares en las tarjetas de crédito y tiene 3,100 dólares ahorrados en el banco.

Mitad del rango

Definición

Mitad del rango: medida de tendencia central que constituye el valor que está a medio camino, entre el puntaje más alto y el más bajo, en el conjunto original de datos. Se calcula sumando el valor máximo con el mínimo y luego dividiendo dicha suma entre 2, como en la siguiente fórmula.

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

LA ESTADÍSTICA EN LAS NOTICIAS

Maniquís $\neq$ realidad

La revista *Health* comparó las medidas de los maniquís 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. Cintura de los maniquíes: 23 pulgadas; cintura de las mujeres: 29 pulgadas. Tamaño de la cadera de los maniquíes: 34 pulgadas; tamaño de la 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.

La mitad del rango se utiliza en pocas ocasiones. Puesto que sólo utiliza los valores máximo y mínimo es demasiado sensible a dichos extremos. Sin embargo, la mitad del rango posee tres características positivas: **1.** es fácil de calcular; **2.** ayuda a reforzar el hecho importante de que existen diferentes formas para definir el centro de un conjunto de datos; **3.** en ocasiones se utiliza de manera incorrecta como si fuese la mediana, de manera que es posible disminuir la confusión al definir con claridad la mitad del rango con respecto a la mediana.

EJEMPLO Verificación del plomo en el aire A continuación se presentan medidas de las cantidades de plomo (μg/m^3) en el aire, en el lugar donde estaba el World Trade Center, días después del 11 de septiembre del 2001. Calcule la mitad del rango para 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 μg/m^3.

Desafortunadamente, el término *promedio* en ocasiones se utiliza para cualquier medida de tendencia central y, en ocasiones, para implicar la media. Por esta ambigüedad, no debemos usar el término *promedio* cuando nos referimos a una medida de tendencia central en particular. En su lugar, habrá que aplicar el término específico, tal como media, mediana, moda o mitad del rango. Cuando nos encontremos un valor reportado como *promedio*, tendremos que saber que el valor puede ser el resultado de cualquiera de las distintas definiciones.

Con la idea de describir, explorar y comparar datos, incluimos la tabla 2-8, que resume las distintas medidas de tendencia central para los niveles de cotinina que se presentan en la tabla 2-1, en el problema del capítulo. Recuerde que la cotinina es un metabolito de la nicotina, de modo que cuando el cuerpo absorbe la nicotina se produce la cotinina. Una comparación de las medidas de tendencia central sugiere que los niveles de cotinina son más altos en los fumadores. Además, los niveles de cotinina de los individuos que no fuman, pero están expuestos al humo del tabaco, son más altos que los de personas que tampoco fuman y no están expuestas al humo. Lo anterior sugiere que "los fumadores pasivos" sí se ven afectados. Se

Tabla 2-8 Comparación de los niveles de cotinina de fumadores, de no fumadores expuestos al humo ambiental del tabaco (HAT) y de no fumadores no expuestos al humo ambiental del tabaco (SHAT).

	Fumadores	HAT	SHAT
Media	172.5	60.6	16.4
Mediana	170.0	1.5	0.0
Moda	1 y 173	1	0
Mitad del rango	245.5	275.5	154.5

dispone de métodos para determinar si estas aparentes diferencias son estadísticamente significativas. Más adelante, consideraremos algunos de estos métodos.

Regla de redondeo

Una regla sencilla para redondear respuestas es la siguiente:

> **Aumente una posición decimal más a las que están presentes 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 3.333333 . . . , que se redondea a 3.3. Como los valores originales son números enteros, redondeamos al décimo 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).

Media de una distribución de frecuencias

Cuando los datos se resumen en una distribución de frecuencias es probable que no conozcamos los valores exactos de una clase en particular. Para hacer que los cálculos sean posibles, pretendemos que todos los valores muestrales sean iguales a la marca de clase. Ya que cada marca de clase se repite un número de veces igual a la frecuencia de clase, la sumatoria de todos los valores muestrales es $\Sigma(f \cdot x)$, donde f denota la frecuencia y x representa la marca de clase. El número total de valores muestrales es la sumatoria de frecuencias Σf. La fórmula 2-2 se utiliza para calcular la media cuando los datos muestrales se resumen en una distribución de frecuencias. La fórmula 2-2 en realidad no es un concepto nuevo, sino una variación de la fórmula 2-1.

Primero multiplique cada frecuencia y marca de clase, después sume los productos
↓

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

↑
sumatoria de las frecuencias

Por ejemplo, observe la tabla 2-9 en la siguiente página. Las primeras dos columnas son iguales a la distribución de frecuencias (tabla 2-2) de los niveles de cotinina de fumadores. La tabla 2-9 ilustra el procedimiento que se sigue para aplicar la fórmula 2-2 cuando se calcula la media de datos resumidos en una distribución de frecuencias. En realidad, por lo general se utilizan programas de cómputo o calculadoras, en lugar del cálculo manual. El resultado de la tabla 2-9 es $\bar{x} = 177.0$, aunque obtenemos $\bar{x} = 172.5$ si utilizamos la lista original con 40 valores. Recuerde, la distribución de frecuencias produce una aproximación de $\bar{x}$, ya que no se basa en la lista original exacta de valores muestrales.

Nadie 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. Sugieren ponderar los resultados con base en la frecuencia en que la gente no se encuentra en su casa. Por ejemplo, alguien 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 dicha persona por primera vez, sus resultados se ponderan de modo que se cuenten tres veces, respecto de un sujeto 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. Tan inteligente solución se presentó inicialmente en 1949.

Tabla 2-9 Cálculo de la media de una distribución de frecuencias

Nivel de cotinina	Frecuencia f	Marca de clase x	$f \cdot x$
0–99	11	49.5	544.5
100–199	12	149.5	1794.0
200–299	14	249.5	3493.0
300–399	1	349.5	349.5
400–499	2	449.5	899.0
Totales:	$\Sigma f = 40$		$\Sigma(f \cdot x) = 7080.0$

$$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f} = \frac{7080}{40} = 177.0$$

Media ponderada

En algunos casos los valores varían su grado de importancia, de modo que es posible que queramos acomodarlos de acuerdo con ello. Después, será posible proceder al cálculo de una **media ponderada**, que es una media que se obtiene asignando distintos pesos a los valores, tal como se muestra en la fórmula 2-3.

Fórmula 2-3 media ponderada: $$\bar{x} = \frac{\Sigma(w \cdot x)}{\Sigma w}$$

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 2-3, como sigue:

$$\bar{x} = \frac{\Sigma(w \cdot x)}{\Sigma w}$$
$$= \frac{(20 \times 85) + (30 \times 90) + (50 \times 75)}{20 + 30 + 50} = \frac{8150}{100} = 81.5$$

Otro ejemplo son los promedios universitarios (que utilizan letras), que pueden calcularse si asignamos a cada calificación con letras el número adecuado de puntos (A = 4, B = 3, etcétera), y después asignamos a cada puntaje un peso igual al número de horas crédito. Nuevamente, se utiliza la fórmula 2-3 para calcular el promedio de calificaciones.

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? Desafortunadamente, 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 resumen en la tabla 2-10.

Tabla 2-10 Comparación de la media, mediana, moda y mitad del rango

Medida de tendencia central	Definición	¿Qué tan común es?	Existencia	¿Toma en cuenta cada valor?	¿Se ve afectada por valores extremos?	Ventajas y desventajas
Media	$\bar{x} = \dfrac{\Sigma x}{n}$	"promedio" más conocido	siempre existe	sí	sí	se usa a lo largo de este libro; funciona bien con muchos métodos estadísticos
Mediana	valor en medio	de uso común	siempre existe	no	no	suele ser una buena opción si hay algunos valores extremos
Moda	valor más frecuente	se usa en ocasiones	podría no existir; podría haber más de una	no	no	apropiada para datos en el nivel nominal
Mitad del rango	$\dfrac{(\text{máx} + \text{mín})}{2}$	poco usada	siempre existe	no	sí	muy sensible a los valores extremos

Comentarios generales:

- En el caso de una colección de datos que es aproximadamente simétrica con una moda, la media, la mediana, la moda y la mitad del rango tienden a ser iguales.
- En el caso de una colección de datos obviamente asimétrica, sería bueno reportar tanto la media como la mediana.
- La media es relativamente *confiable*. Es decir, cuando las muestras se extraen de la misma población, las medidas muestrales tienden a ser más consistentes que las demás medidas de tendencia central (consistentes en el sentido de que las medias muestrales, extraídas de la misma población, no varían tanto como las otras medidas de tendencia central).

Una ventaja importante de la media es que toma en cuenta cada valor; una desventaja notoria es que en ocasiones se ve afectada de manera drástica por unos cuantos valores extremos. Tal desventaja se supera usando una media recortada, como se describe en el ejercicio 21.

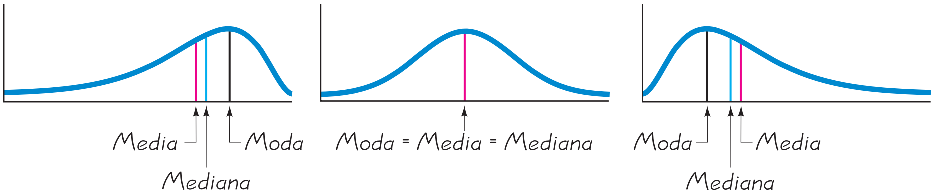
Sesgo

Una comparación de la media, la mediana y la moda puede revelar información acerca de la característica de sesgo, que se define a continuación y se ilustra en la figura 2-11.

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).

FIGURA 2-11 Sesgo



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

Los datos **sesgados a la izquierda** (que también se denomina como *sesgo negativo*) poseen una cola izquierda más larga, en tanto que 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 2-11*a*). Los datos **sesgados a la derecha** (lo que también se denomina *sesgo positivo*) poseen una cola derecha más larga, mientras que la media y la mediana se encuentran a la derecha de la moda. Nuevamente, aunque no siempre es posible predecirlo, en los datos sesgados a la derecha, la media suele estar a la derecha de la mediana, como en la figura 2-11c).

Si examinamos el histograma de la figura 2-1, para los niveles de cotinina de fumadores, observaremos una gráfica sesgada hacia la derecha. En la práctica, muchas distribuciones de datos son simétricas y carecen de sesgo. Las distribuciones sesgadas hacia la derecha son más comunes que aquéllas sesgadas hacia la izquierda, porque con frecuencia es más fácil obtener valores excepcionalmente grandes que valores excepcionalmente pequeños. En el caso de los ingresos anuales, por ejemplo, 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, los ingresos anuales tienden a estar sesgados hacia la derecha, como en la figura 2-11*c*).

Utilizando 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 para muestras, que se engloban en la *estadística descriptiva*. (Véase la sección 2-6, donde se incluyen representaciones visuales de muestras que se obtienen con el STATDISK, Minitab, Excel y la calculadora TI-83 Plus.) A continuación, se incluyen algunos de los procedimientos para la obtención de dichas representaciones visuales.

STATDISK Elija **Data** del menú principal y utilice el **Sample Editor** para ingresar los datos. Presione **Copy** y luego presione **Data** nuevamente, pero ahora elija la opción **Descriptive Statistics**. Haga clic en **Paste**, para recuperar el conjunto de datos que usted alimentó. Ahora haga clic en **Evaluate** para obtener los estadísticos descriptivos, incluyendo la media, la mediana, la mitad del rango y otros que se discutirán en las próximas secciones.

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:A40 para 40 valores en la columna A), haga clic en **Summary Statistics** y después haga clic en **OK**. (Si Data Analysis no aparece en el menú Tools, deberá instalarlo haciendo clic en **Tools** y seleccionando **Add-Ins.**)

TI-83 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 **1-Var Stats** y finalmente **ENTER** dos veces. La representación visual incluirá la media, 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.

2-4 Destrezas y conceptos básicos

En los ejercicios 1 a 8 calcule a) la media, b) la mediana, c) la moda y d) la mitad del rango, de los datos muestrales dados.

1. Consumo de tabaco en películas infantiles En el artículo "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), se registró la duración (en segundos) de escenas de películas animadas, de los Universal Studios, que muestran consumo de tabaco. A continuación se presentan los primeros seis valores, que están incluidos en el conjunto de datos 7 del Apéndice B. ¿Hay algún problema al incluir escenas de consumo de tabaco en películas infantiles de dibujos animados?

0	223	0	176	0	548

2. Harry Potter En un intento por medir el nivel de lectura de un libro, se obtuvieron los puntajes de la facilidad de lectura de Flesch Reading de 12 páginas, seleccionadas aleatoriamente, de la obra *Harry Potter y la piedra filosofal*, de J. K. Rowling. Dichos puntajes, que se incluyen en el conjunto de datos 14 del Apéndice B, se listan a continuación. En tanto que estos puntajes se basan en 12 páginas seleccionadas aleatoriamente, ¿será la media de esta muestra un estimado razonable del nivel medio de lectura de todo libro?

85.3	84.3	79.5	82.5	80.2	84.6
79.2	70.9	78.6	86.2	74.0	83.7

3. Cereal Un nutriólogo obtiene las cantidades de azúcar (en gramos) de un gramo de 16 cereales diferentes, incluyendo Cheerios, Corn Flakes, Fruit Loops, Trix y 12 más. Estos valores, que se incluyen en el conjunto de datos 16 del Apéndice B, se listan a continuación. ¿Será la media de estos valores un buen estimado de la cantidad media de azúcar que hay en cada gramo del cereal consumido por la población de todos los estadounidenses que lo comen? ¿Por qué?

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

4. Índice de masa corporal Como parte del examen nacional de salud (National Health Examination) en Estados Unidos, se mide el índice de masa corporal en una muestra aleatoria de mujeres. Algunos de los valores, que se incluyen en el conjunto de datos 1 del Apéndice B, se listan a continuación. ¿Estará la media de tal muestra razonablemente cerca de la media de 25.74, que es la media de las 40 mujeres incluidas en el conjunto de datos 1?

19.6	23.8	19.6	29.1	25.2	21.4	22.0	27.5
33.5	20.6	29.9	17.7	24.0	28.9	37.7	

5. Conductores alcoholizados Más adelante se incluyen las concentraciones de alcohol en la sangre de conductores que se vieron envueltos en accidentes fatales y que después fueron sentenciados a prisión (con base en datos del Departamento de Justicia de Estados Unidos). Si las leyes estatales actuales prohíben conducir con niveles por encima de 0.08 o 0.10, ¿están estos niveles significativamente por arriba del máximo permitido?

0.27	0.17	0.17	0.16	0.13	0.24	0.29	0.24
0.14	0.16	0.12	0.16	0.21	0.17	0.18	

6. Muertes en motocicletas A continuación se presentan las edades de motociclistas que se accidentaron mortalmente en accidentes de tránsito (con base en datos del Departamento del Transporte estadounidense). ¿Apoyan los resultados la creencia común de que una mayor proporción de conductores jóvenes se ven implicados en tales accidentes?

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

7. Tiempos de reacción El autor visitó el Museo de Ciencias Reuben H. Fleet en San Diego y repitió un experimento de tiempos de reacción. Se obtuvieron los siguientes tiempos (en centésimas de segundo). ¿Qué tan consistentes son estos resultados y de qué forma afecta la consistencia el uso de la media muestral, como un estimado de la media poblacional?

19	20	17	21	21	21	19	18	19	19
17	17	15	17	18	17	18	18	18	17

8. Tabletas de Bufferin A continuación se listan los pesos medidos (en miligramos) de una muestra de tabletas de aspirina Bufferin. ¿Cuál sería la consecuencia grave de tener pesos que varían tanto?

672.2	679.2	669.8	672.6	672.2	662.2
662.7	661.3	654.2	667.4	667.0	670.7

En los ejercicios 9 a 12 calcule media, mediana, moda y mitad del rango para cada una de las dos muestras; luego, compare los dos conjuntos de resultados.

9. 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):

Jefferson Valley:	6.5	6.6	6.7	6.8	7.1	7.3	7.4	7.7	7.7	7.7
Providence:	4.2	5.4	5.8	6.2	6.7	7.7	7.7	8.5	9.3	10.0

Interprete los resultados y determine si hay una diferencia entre los dos conjuntos de datos, que no sea aparente cuando se comparen las medidas de tendencia central. Si es así, ¿cuál es?

10. 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:

Regular:	0.8192	0.8150	0.8163	0.8211	0.8181	0.8247
De dieta:	0.7773	0.7758	0.7896	0.7868	0.7844	0.7861

¿Parece haber una diferencia significativa entre los dos conjuntos de datos? ¿Cómo se explicaría una diferencia como ésa?

11. Mickey D *vs.* Jack Al investigar los tiempos que se requieren en el servicio en automóvil (en segundos), se obtuvieron los siguientes resultados (con base en datos del QSR Drive-Thru Time Study).

McDonald’s:	287	128	92	267	176	240	192	118	153	254	193	136
Jack in the Box:	190	229	74	377	300	481	428	255	328	270	109	109

¿Cuál de los dos gigantes de comida rápida parece ser más veloz? ¿La diferencia parece ser significativa?

12. Anchura de cráneos Las anchuras máximas de muestras de cráneos egipcios de varones, que datan del 4000 a.C. y del 150 d.C. (de acuerdo con los datos de *Ancient Races of the Thebaid*, de Thomson y Randall-Maciver) se muestran a continuación:

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 del tamaño de las cabezas a través del tiempo sugieren la mezcla con personas de otras regiones. ¿Parece haber cambiado el tamaño de las cabezas del 4000 a.C. al 150 d.C.?

En los ejercicios 13 a 16, remítase a los conjuntos de datos del Apéndice B. Utilice un programa de cómputo o una calculadora para obtener las medias y las medianas, luego compare los resultados, tal como se indica.

T **13.** Circunferencia de cabezas Para diagnosticar de forma correcta la hidrocefalia, un pediatra investiga la circunferencia de las cabezas de niños y niñas de dos años de edad. Utilice los resultados muestrales listados en el conjunto de datos 3. ¿Parece haber alguna diferencia entre los dos géneros?

T **14.** Clancy, Rowling, Tolstoi Un psicólogo infantil investiga las diferencias en cuanto a la facilidad de la lectura; obtiene datos 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. Remítase al conjunto de datos 14 del Apéndice B y utilice las puntuaciones de la calificación de Flesch-Kincaid de las 12 páginas seleccionadas aleatoriamente, para cada uno de los tres libros. ¿Los datos parecen ser diferentes?

T **15.** Lluvia en el fin de semana Utilice el conjunto de datos 11 del Apéndice B para calcular la media y la mediana de las cantidades de lluvia que caen en Boston los jueves; calcule también, la media y la mediana de las cantidades de lluvia que caen en Boston los domingos. Los reportes de los medios de comunicación afirmaron que llueve más durante los fines de semana que entre semana. ¿Estos resultados apoyan dicha afirmación?

T **16.** Consumo de tabaco/alcohol en películas infantiles En el artículo "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), se registraron las duraciones (en segundos) de escenas que muestran consumo de tabaco y alcohol en películas infantiles de dibujos animados. Remítase al conjunto de datos del Apéndice B, luego calcule la media y la mediana de las duraciones de escenas con tabaco así como la media y la mediana de las escenas con alcohol. ¿Parece haber una diferencia entre tales duraciones? ¿Cuál parece ser el problema mayor: las escenas que presentan consumo de tabaco o aquellas que muestran consumo de alcohol?

En los ejercicios 17 a 20 calcule la media de los datos que se resumen en la distribución de frecuencias dada.

17. Old Faithful Los visitantes del Parque Nacional Yellowstone consideran que una erupción del géiser Old Faithful es una gran atracción que uno no debe perderse. La distribución de frecuencias dada resume una muestra de los tiempos (en minutos) entre las erupciones.

18. Dado cargado El autor taladró un hoyo en un dado y lo relleno con plomo, después procedió a lanzarlo 200 veces. Los resultados se presentan en la distribución de frecuencias al margen. ¿El resultado parece ser muy diferente del resultado esperado con un dado inalterado?

19. Infracciones de tránsito La distribución de frecuencias describe las velocidades de conductores infraccionados por la policía en la ciudad de Poughkeepsie. Los conductores viajaban por una zona con límite de velocidad de 30 millas/hora en Creek Road, que pasa por la universidad del autor. ¿Cómo se compara la media con el límite de velocidad de 30 mi/h?

20. Temperaturas corporales La distribución de frecuencias al margen resume una muestra de temperaturas corporales humanas. (Véanse las temperaturas de medianoche del segundo día, listadas en el conjunto de datos 4 del Apéndice B.) ¿Cómo se compara la media con el valor de 98.6°F, que es el valor que la mayoría de la gente supone como la media?

Tabla del ejercicio 17

Tiempo	Frecuencia
40–49	8
50–59	44
60–69	23
70–79	6
80–89	107
90–99	11
100–109	1

Tabla del ejercicio 18

Resultado	Frecuencia
1	27
2	31
3	42
4	40
5	28
6	32

Tabla del ejercicio 19

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

Tabla del ejercicio 20

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

2-4 Más allá de lo básico

T **21.** 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; finalmente, se calcula la media de los valores restantes. Para los pesos de los osos, en el conjunto de datos 9 del Apéndice B, calcule *a*) la media; *b*) la media recortada del 10%; *c*) la media recortada del 20%. ¿Cómo se comparan los resultados?

22. Media de medias Con el uso de un almanaque, un investigador obtiene la media del salario de maestros de cada estado de Estados Unidos. Suma los 50 valores y luego los divide entre 50 para obtener la media. ¿Será el resultado igual a la media nacional del salario de maestros? ¿Por qué?

23. 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.
a. Calcule el décimo valor.
b. Necesitamos crear una lista de n valores que contenga 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?

24. Datos censurados Se realizó un experimento para probar la vida de baterías de automóviles. El experimento se llevó a cabo durante un tiempo 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 la batería aún funcionaba al final del experimento). ¿Qué se concluye acerca de la vida media de las baterías?

25. Media ponderada Kelly Bell obtiene calificaciones parciales de 65, 83, 80 y 90. En el examen final recibe una calificación de 92. Calcule la media ponderada, si cada uno de los exámenes parciales cuenta el 15% y el examen final cuenta el 40% de la calificación total.

26. 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.
a. La misma constante k se suma a cada valor del conjunto de datos.
b. Cada valor del conjunto de datos se multiplica por la misma constante k.

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

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

(Ningún valor puede ser cero). Cuatro estudiantes conducen desde Nueva York hasta Florida (1,200 millas), a una velocidad de 40 mi/h (¡sí, como no!). Como necesitan llegar a su clase de estadística a tiempo, viajan de regreso a una velocidad de 60 mi/h. ¿Cuál es la velocidad promedio del viaje completo? (La media armónica se utiliza para promediar velocidades).

28. 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%, el 8%, el 9%, el 12% y el 7% se obtiene determinando la media geométrica de 1.10, 1.08, 1.09, 1.12 y 1.07. Calcule el factor de crecimiento promedio.

29. La **media cuadrática** (o **cuadrado medio de raíz**, o **CMR**) suele utilizarse en aplicaciones físicas. Por ejemplo, en los sistemas de distribución de energía, los montajes y las corrientes suelen referirse 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 el número de valores n y después sacando la raíz cuadrada del resultado, el cual se expresa como

$$\text{media cuadrática} = \sqrt{\frac{\Sigma x^2}{n}}$$

Calcule el CMR de estas fuentes de poder (en volts): 110, 0, −60, 12.

30. Mediana Cuando los datos se resumen en una distribución de frecuencias, la mediana puede calcularse si primero se identifica la *clase de la mediana* (la clase que contiene a la mediana). Entonces, suponemos que los valores en esa clase se distribuyen uniformemente y podemos interpolar. Este proceso se describe de la siguiente forma:

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

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. ¿Cómo se compara este resultado con la mediana de la lista original de datos, que es de 170? ¿Cuál valor de la mediana es mejor: el que se calculó para la tabla de frecuencias o el de 170?

2-5 Medidas de variación

Sugerencia: Ya que la sección introduce el concepto de variación, que es muy relevante en la estadística, es una de las más importantes de todo el libro. Primero lea esta sección en forma rápida y obtenga una comprensión general de las características de variación. Después, aprenda a calcular las medidas de variación, en especial la desviación estándar. Finalmente, trate de comprender el razonamiento que subyace a la fórmula de la desviación estándar, pero no gaste demasiado tiempo memorizando fórmulas o haciendo cálculos aritméticos. En lugar de ello, dé prioridad a aprender a *interpretar* los valores de la desviación estándar.

En la figura 2-12, de la página 74, se presenta un ejemplo visual de variación, el cual incluye muestras de tornillos de dos compañías diferentes. Puesto que dichos tornillos se utilizan para unir las alas al fuselaje, su calidad es muy importante. Si sólo tomamos en consideración la media, no reconoceríamos cualquier diferencia entre dos muestras, ya que ambas tienen una media de $\bar{x} = 2{,}000$ pulgadas. Sin embargo, debe ser evidente que las muestras difieren mucho con respecto a las *variaciones* de las longitudes de los tornillos. Los tornillos fabricados por Precision Bolt Co. parecen tener longitudes muy similares, mientras que las longitudes de los tornillos de Ruff Bolt Co. varían mucho. En muchos procesos de fabricación, este mismo aspecto tiene una gran importancia. Se logra una mejor calidad a través de una variación menor. En esta sección queremos que desarrolle la habilidad para medir y comprender la variación.

Otra situación ideal, que ilustra la importancia de la variación, se percibe en las filas de espera en los bancos. En el pasado, muchos bancos requerían que sus clientes esperaran en filas separadas, frente a cada una de las ventanillas; sin embargo, ahora la mayoría utiliza una sola fila de espera. ¿Por qué hicieron este cambio? El tiempo medio de espera no cambió, ya que la configuración de la fila de

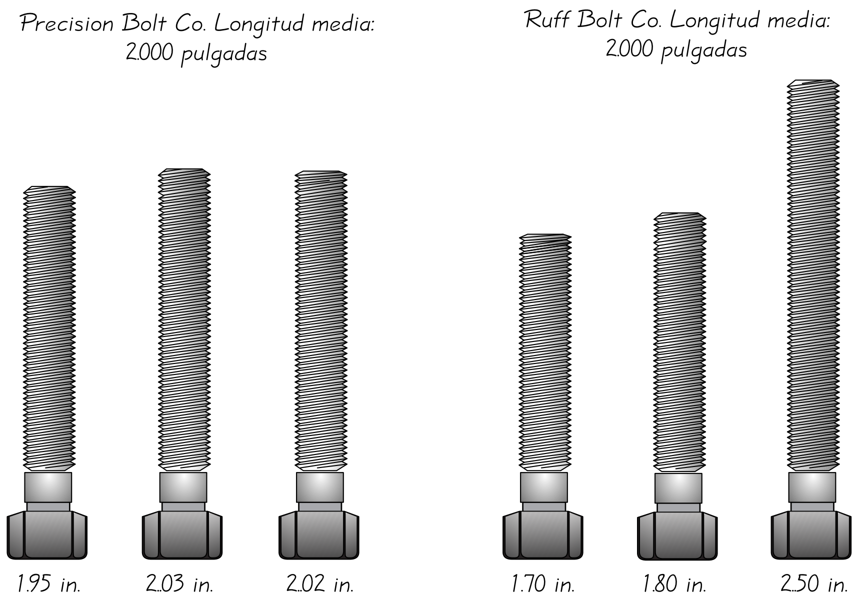


FIGURA 2-12 Tornillos fabricados por dos compañías diferentes

espera no afecta la eficiencia de los cajeros. El cambio a una sola fila se hizo porque los clientes prefieren esperar periodos que sean más *consistentes*, con menos variación. Miles de bancos hicieron un cambio que resultó en una menor variación (y clientes más contentos), aun cuando no se afectó la media. Consideremos algunos tiempos de espera específicos (en minutos) de clientes bancarios.

Banco Salem (una fila de espera)	4	7	7
Banco Mulberry (múltiples filas de espera)	1	3	14

Es fácil calcular que $\bar{x} = 6.0$ para ambos conjuntos de datos. También es fácil notar, mediante una inspección visual, que los tiempos de espera de 4, 7, 7 varían mucho menos que los tiempos de espera de 1, 3, 14. Procedamos ahora a desarrollar algunas formas específicas de medición real de la variación, de modo que sea posible utilizar números específicos en lugar de juicios subjetivos. Comencemos por el rango.

Rango

Definición

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

$$\textbf{Rango} = \textbf{(valor máximo)} - \textbf{(valor mínimo)}$$

Para calcular el rango, sólo se resta el valor mínimo del valor máximo. Para los clientes del Banco Salem, el rango es $7 - 4 = 3$ min. El Banco Mulberry tiene tiempos de espera con un rango de 13 min y este valor más grande sugiere una mayor variación.

Es muy fácil calcular el rango, pues depende únicamente de los valores máximo y mínimo, pero no es tan útil como otras medidas de variación que usan cada valor. (Véase el ejercicio 35 como un ejemplo en el que el rango causó confusión).

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 (véase la página 81).

Definición

Desviación estándar (de un conjunto de valores muestrales): 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 2-4 o 2-5.

Fórmula 2-4 $$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$ desviación estándar de la muestra

Fórmula 2-5 $$s = \sqrt{\frac{n\Sigma(x^2) - (\Sigma x)^2}{n(n - 1)}}$$ fórmula abreviada de la desviación estándar de la muestra

Más adelante, en esta sección, analizaremos los fundamentos de estas fórmulas, aunque por ahora le recomendamos que utilice la fórmula 2-4, para resolver algunos ejemplos, y que después aprenda a calcular los valores de la desviación estándar por medio del uso de su calculadora y de un programa de cómputo. (La mayoría de las calculadoras científicas se diseñaron de tal modo que permiten que se introduzca una lista de valores y se obtenga de forma automática la desviación estándar). 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 suele ser positivo. Sólo es igual a cero cuando todos los valores de los datos son el mismo número. Además, valores grandes de s indican mayores cantidades de variación.
- El valor de la desviación estándar s se puede incrementar de manera drástica con la inclusión de uno o más datos distantes (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.

Más acciones, menos riesgo

En su libro *Investments*, 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 de la cartera 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 poca cantidad de acciones implicadas. Con más de 30 acciones, hay muy poco riesgo específico de una empresa; en su lugar, casi todo el riesgo es "riesgo de mercado", atribuible al mercado de acciones global. Además, señalan que estos principios son "sólo una aplicación de la bien conocida ley de promedios".

Procedimiento para calcular la desviación estándar con la fórmula 2-4

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

Paso 2: Reste la media de cada valor individual para tener 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 $\Sigma(x - \bar{x})^2$.

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 2-4 Use la fórmula 2-4 para calcular la desviación estándar de los tiempos de espera de los clientes del Banco Mulberry. Dichos tiempos (en minutos) son 1, 3, 14.

SOLUCIÓN Utilizaremos los seis pasos en este proceso. Remítase a dichos pasos y a la tabla 2-11, 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 que se obtuvo en el paso 2 para lograr 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; dichas interpretaciones se analizarán más tarde en esta sección.

Tabla 2-11 Cálculo de la desviación estándar de los tiempos de espera de los clientes del Banco Mulberry

x	$x - \bar{x}$	$(x - \bar{x})^2$
1	−5	25
3	−3	9
14	8	64
Totales: 18		98

$$\bar{x} = \frac{18}{3} = 6.0 \text{ min} \qquad s = \sqrt{\frac{98}{3-1}} = \sqrt{49} = 7.0 \text{ min}$$

¿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. ¿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 2-5 En el ejemplo anterior se utilizó la fórmula 2-4 para calcular la desviación estándar de los tiempos de espera de los clientes del Banco Mulberry. Con el mismo conjunto de datos, calcule la desviación estándar con la fórmula 2-5.

SOLUCIÓN La fórmula 2-5 requiere que primero encontremos valores para n, Σx y Σx^2.

$n = 3$ (ya que existen tres valores en la muestra)

$\Sigma x = 18$ (se obtiene al sumar los tres valores muestrales)

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

Si usamos la fórmula 2-5, obtendremos

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

Una actividad adecuada es detenerse aquí y calcular la desviación estándar de los tiempos de espera del Banco Salem. Siga los mismos procedimientos de los dos ejemplos anteriores y verifique que, para el Banco Salem, $s = 1.7$ 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 las compararemos para darnos cuenta de que la desviación estándar de los tiempos de espera del Banco Salem (1.7 min) es mucho menor que la desviación estándar del Banco Mulberry (7.0 min). Esto apoya nuestra conclusión subjetiva de que los tiempos de espera del Banco Salem tienen una variación mucho menor que los tiempos del Banco Mulberry.

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 lugar de dividir entre $n - 1$, se hace entre el tamaño N de la población, como en la siguiente expresión:

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

Ya que generalmente usamos datos muestrales, a menudo utilizaremos la fórmula 2-4, en la cual dividimos entre $n - 1$. Muchas calculadoras dan tanto la desviación estándar muestral como la desviación estándar poblacional, pero con una gran variedad de notaciones diferentes. Asegúrese de identificar la notación de su calculadora, de modo que obtenga el resultado correcto. (La TI-83 Plus utiliza Sx para la desviación estándar muestral y σx para la desviación estándar poblacional).

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 lugar de *variación*). El término *varianza* se refiere a una definición específica.

Definiciones

Varianza (de un conjunto de valores): medida de variación igual al cuadrado de la desviación estándar.

Varianza muestral: cuadrado de la desviación estándar s.

Varianza poblacional: 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 hacerlo de manera sistemática, a sobreestimar o subestimar σ^2. (Véase el ejercicio 41).

EJEMPLO Cálculo de la varianza En el ejemplo anterior, empleamos los tiempos de espera de los clientes del Banco Mulberry para descubrir que la desviación estándar está dada por $s = 7.0$ 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 muestran abajo. Note que las unidades de los valores de los datos están dadas en minutos y que la desviación estándar es de 7.0 minutos; 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 en el capítulo 11. Para nuestros propósitos presentes, la varianza tiene la siguiente gran desventaja:

las unidades de la varianza son diferentes a las unidades del conjunto original de datos. Por ejemplo, si los tiempos originales de espera de los clientes están dados en minutos, las unidades de varianza estarán dadas en minutos cuadrados (min^2). ¿Qué es un minuto cuadrado? (Diviértase elaborando una respuesta creativa a dicha pregunta). Ya que la varianza utiliza unidades distintas, es sumamente difícil comprender la varianza si la se relaciona con el conjunto original de datos. Por esta propiedad, nos enfocaremos en la desviación estándar, mientras tratamos de comprender la variación.

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 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 2-4:

Aumentar una posición decimal a la que había en los datos originales.

Redondee sólo 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).

Comparación de la variación en diferentes poblaciones

Anteriormente afirmamos que, ya que 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 que se tomaron de distintas poblaciones. El *coeficiente de variación* resuelve tal desventaja.

Definición

Coeficiente de variación o **CV** de un conjunto de datos muestrales o poblacionales, expresado como porcentaje, describe la desviación estándar relativa a la media, y está dada de la siguiente forma:

Muestra

$$CV = \frac{s}{\bar{x}} \cdot 100\%$$

Población

$$CV = \frac{\sigma}{\mu} \cdot 100\%$$

EJEMPLO **Estatura y peso de hombres** Si utilizamos los datos muestrales estatura y peso de 40 hombres, incluidos en el 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 el coeficiente de variación de los pesos; finalmente, compare los dos resultados.

	Media ($\bar{x}$)	Desviación estándar (s)
Estatura	68.34 in	3.02 in
Peso	172.55 lb	26.33 lb

SOLUCIÓN Debido a que tenemos estadísticos muestrales, los dos coeficientes de variación se obtienen de la siguiente manera:

Estaturas: $$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{3.02 \text{ in}}{68.34 \text{ in}} \cdot 100\% = 4.42\%$$

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. Se observa 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 adulto que mida el doble que otro, pero es mucho más común ver a uno que pese el doble que otro.

Cálculo de la desviación estándar a partir de una distribución de frecuencias

En ocasiones necesitamos calcular la desviación estándar de un conjunto de datos que se resume en una distribución de frecuencias, como en la tabla 2-2 de la sección 2-2. Si se dispone de la lista original de valores muestrales, se utiliza la fórmula 2-4 o la 2-5, de modo que el resultado es más exacto. Si los datos originales no están disponibles, se utiliza uno de los dos métodos siguientes:

1. Si el número total de valores no es demasiado grande, trabaje con su calculadora o programa de cómputo e introduzca cada marca de clase tantas veces como el número de la frecuencia de clase.
2. Calcule la desviación estándar con la fórmula 2-6.

Fórmula 2-6

$$s = \sqrt{\frac{n\left[\Sigma(f \cdot x^2)\right] - \left[\Sigma(f \cdot x)\right]^2}{n(n-1)}}$$

desviación estándar para una distribución de frecuencias

Tabla 2-12 Cálculo de la desviación estándar a partir de una distribución de frecuencias

Cotinina	Frecuencia f	Marca de clase, x	$f \cdot x$	$f \cdot x^2$
0–99	11	49.5	544.5	26952.75
100–199	12	149.5	1794.0	268203.00
200–299	14	249.5	3493.0	871503.50
300–399	1	349.5	349.5	122150.25
400–499	2	449.5	899.0	404100.50
Totales:	$\Sigma f = 40$		$\Sigma(f \cdot x) = 7080$	$\Sigma(f \cdot x^2) = 1692910$

EJEMPLO Niveles de cotinina de fumadores Calcule la desviación estándar de los 40 valores que se resumen en la distribución de frecuencias de la tabla 2.2, considerando que no se dispone del conjunto original de datos.

SOLUCIÓN

Método 1: La tabla 2-12 tiene marcas de clase de 49.5, 149.5, 249.5, 349.5 y 449.5. Con una calculadora o un programa de cómputo introduzca el valor de 49.5, 11 veces (ya que la frecuencia de la primera clase es 11); introduzca 149.5, 12 veces y así sucesivamente. Obtenga la desviación estándar de este conjunto de 40 marcas de clase. El resultado debe ser 106.2.

Método 2: Utilice la fórmula 2-6. La aplicación de la fórmula 2-6 requiere que primero obtengamos los valores de n, $\Sigma(f \cdot x)$ y $\Sigma(f \cdot x^2)$. Después de obtener estos valores de la tabla 2-12, apliquemos la fórmula 2-6 de la siguiente manera:

$$s = \sqrt{\frac{n\left[\Sigma(f \cdot x^2)\right] - \left[\Sigma(f \cdot x)\right]^2}{n(n-1)}} = \sqrt{\frac{40[1692910] - [7080]^2}{40(40-1)}}$$

$$= \sqrt{\frac{17{,}590{,}000}{1560}} = \sqrt{11275.64103} = 106.2$$

Calculadora TI-83 Plus A diferencia de la mayoría de las calculadoras, la TI-83 Plus calcula la desviación estándar de valores resumidos en una distribución de frecuencias. Primero, introduzca las marcas de clase en la lista L1, después introduzca las frecuencias en la lista L2. Ahora presione **STAT**, seleccione **CALC**, luego seleccione **1-VarStats** e introduzca L1, L2 para obtener resultados que incluyan la media y la desviación estándar. Nuevamente, la desviación estándar muestral se identifica con Sx y la desviación estándar poblacional con σx.

Interpretación y comprensión de la desviación estándar

Este apartado es sumamente importante, puesto que ahora trataremos de que la desviación estándar tenga sentido. 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.

Ya que la variación es un concepto tan importante y que la desviación estándar es una herramienta tan útil para medir la variación, consideraremos tres formas diferentes para lograr una apreciación de los valores de las desviaciones estándar. La primera 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 de la media. (Es posible mejorar la precisión de tal regla si tomamos en cuenta factores como el tamaño de la muestra y la naturaleza de la distribución, aunque preferimos sacrificar precisión en aras de la sencillez. Además, podríamos usar tres o, incluso, cuatro desviaciones estándar en lugar de 2, lo cual constituye una decisión un poco 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 burdo 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 burdos de los valores muestrales mínimos y máximos "comunes" por medio de

$$\text{valor mínimo "común"} \approx (\text{media}) - 2 \times (\text{desviación estándar})$$

$$\text{valor máximo "común"} \approx (\text{media}) + 2 \times (\text{desviación estándar})$$

Cuando calcule una desviación estándar por medio de las fórmulas 2-4 o 2-5, la regla práctica del intervalo resulta útil para verificar el resultado, pero debe estar consciente de que, aun cuando la aproximación nos acerca a la respuesta, puede tener un error considerable.

EJEMPLO Niveles de cotinina de fumadores Utilice la regla práctica del intervalo para calcular un estimado burdo de la desviación estándar de la muestra de 40 niveles de cotinina de fumadores, como se observa en la tabla 2-1.

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 los niveles de cotinina, notaremos que el mínimo es 0 y el máximo 491; por lo tanto, el rango es de 491. La desviación estándar s se estima de la siguiente manera:

$$s \approx \frac{\text{rango}}{4} = \frac{491}{4} = 122.75 \approx 123$$

INTERPRETACIÓN Este resultado es muy cercano al valor correcto de 119.5, que se obtiene al calcular el valor exacto de la desviación estándar con las fórmulas 2-4 o 2-5. No espere que la regla práctica del intervalo funcione tan bien en otros casos.

El siguiente ejemplo es particularmente importante como ilustración de una forma de *interpretar* el valor de una desviación estándar.

EJEMPLO Circunferencias de la cabeza de niñas Resultados anteriores del National Health Survey sugieren que las circunferencias de las cabezas de niñas de dos meses de edad tienen una media de 40.05 cm y una desviación estándar de 1.64 cm. Utilice la regla práctica del intervalo para calcular el mínimo y el máximo "comunes" de las circunferencias de las cabezas. (Estos resultados serían prácticos para un médico al que le interese identificar circunferencias "infrecuentes", que serían el resultado de un trastorno como la hidrocefalia). Después, determine si una circunferencia de 42.6 cm sería considerada "infrecuente".

SOLUCIÓN Con una media de 40.05 cm y una desviación estándar de 1.64 cm, empleamos la regla práctica del intervalo para calcular las circunferencias mínima y máxima comunes, de la siguiente manera:

$$\begin{aligned}\text{mínimo} &\approx (\text{media}) - 2 \times (\text{desviación estándar})\\ &= 40.05 - 2(1.64) = 36.77 \text{ cm}\end{aligned}$$

$$\begin{aligned}\text{máximo} &\approx (\text{media}) + 2 \times (\text{desviación estándar})\\ &= 40.05 + 2(1.64) = 43.33 \text{ cm}\end{aligned}$$

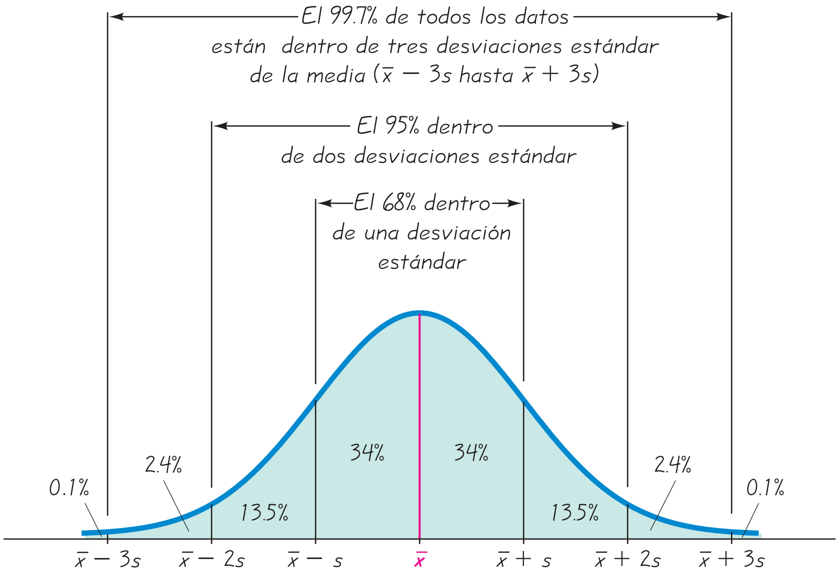
INTERPRETACIÓN Con base en estos resultados, esperamos que, generalmente, las niñas de dos meses de edad tengan una cabeza cuya circunferencia mida entre 36.77 cm y 43.33 cm. Como 42.6 cm está dentro de estos límites, se consideraría una niña normal.

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 que tienen una distribución aproximadamente normal*. (Véase la figura 2-13).

- 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 dos desviaciones estándar de la media.
- Aproximadamente el 99.7% de todos los valores están dentro de tres desviaciones estándar de la media.

FIGURA 2-13 La regla empírica



EJEMPLO Puntuaciones de CI Las puntuaciones de adultos normales en la prueba Weschler tienen una distribución normal, con una media de 100 y una desviación estándar de 15. ¿Qué porcentaje de adultos tienen puntuaciones de CI entre 55 y 145?

SOLUCIÓN La clave para resolver el problema es reconocer que 55 y 145 están exactamente a tres desviaciones estándar de la media de 100, como se indica abajo.

$$3 \text{ desviaciones estándar} = 3s = 3(15) = 45$$

Por lo tanto, tres desviaciones estándar de la media son igual a

$$100 - 45 = 55$$

o

$$100 + 45 = 145$$

La regla empírica nos indica que aproximadamente el 99.7% de todos los valores están dentro de tres desviaciones estándar de la media; por lo tanto, el 99.7% de todas las puntuaciones de CI se encuentran entre 55 y 145.

Sugerencia: Las dificultades para aplicar la regla empírica suelen surgir de la confusión al interpretar frases tales como “dentro de 3 desviaciones estándar de la media”. Deténgase aquí, revise el ejemplo anterior hasta que el significado de dicha frase esté claro. Además, observe las siguientes interpretaciones generales de esa frase.

Frase	Significado
Dentro de una desviación estándar de la media	Entre $(\bar{x} - s)$ y $(\bar{x} + s)$
Dentro de dos desviaciones estándar de la media	Entre $(\bar{x} - 2s)$ y $(\bar{x} + 2s)$
Dentro de tres desviaciones estándar de la media	Entre $(\bar{x} - 3s)$ y $(\bar{x} + 3s)$

Un tercer concepto útil para comprender 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 lugar de limitarse a conjuntos de datos con distribuciones normales se aplica a cualquier conjunto de datos, pero sus resultados son muy aproximados.

Teorema de Chebyshev

La proporción (o fracción) de cualquier conjunto de datos que está dentro de K desviaciones estándar de la media es siempre *al menos* $1 - 1/K^2$, donde K es cualquier número positivo mayor que 1. Para $K = 2$ y $K = 3$, tenemos los siguientes enunciados:

- Al menos $3/4$ (o 75%) de todos los valores están dentro de dos desviaciones estándar de la media.
- Al menos $8/9$ (u 89%) de todos los valores están dentro de tres desviaciones estándar de la media.

EJEMPLO Puntuaciones de CI Las puntuaciones de CI de adultos normales tomadas de la prueba Weschler tienen una media de 100 y una desviación estándar de 15. ¿Qué podemos concluir a partir del teorema de Chebyshev?

SOLUCIÓN Al aplicar el teorema de Chebyshev, con una media de 100 y una desviación estándar de 15, llegamos a las siguientes conclusiones:

- Por lo menos $3/4$ (o 75%) de todos los adultos tienen puntuaciones de CI que están dentro de dos desviaciones estándar de la media (entre 70 y 130).
- Al menos $8/9$ (u 89%) de todos los adultos tienen puntuaciones de CI que están dentro de tres desviaciones estándar de la media (entre 55 y 145).

Cuando intentemos darle un significado a un 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 2-4, que es la base de su definición. (La fórmula 2-5 es sencillamente otra versión de la fórmula 2-4, derivada de modo que los cálculos aritméticos pueden simplificarse).

Fundamentos de la fórmula 2-4

La desviación estándar de un conjunto de datos muestrales se define con las fórmulas 2-4 y 2-5, las cuales son equivalentes en el sentido de que siempre producen el mismo resultado. La fórmula 2-4 tiene la ventaja de reforzar el concepto de que la desviación estándar es un tipo de desviación promedio. La fórmula 2-5, la de ser más fácil de usar cuando hay que calcular desviaciones estándar por nuestra cuenta. La fórmula 2-5 también elimina los errores de redondeo intermedios que se introducen en la fórmula 2-4 cuando no se utiliza el valor exacto de la media. La fórmula 2-5 se aplica en calculadoras y programas, ya que requiere sólo de tres lugares de memoria (para n, $\sum x$ y $\sum x^2$), 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 2.4? 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 del Banco Mulberry de 1, 3, 14, la media es 6.0, de modo que las desviaciones de la media son -5, -3 y 8. Sería bueno combinar, de alguna forma, dichas 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, necesitamos evitar la cancelación de números positivos y negativos. Un método consiste en sumar valores absolutos, como en $\Sigma|x - \bar{x}|$. Si calculamos la media de esta suma, obtendremos la **desviación media absoluta** (o **DMA**), que es la distancia media de los datos con respecto a la media.

$$\text{desviación media absoluta} = \frac{\Sigma|x - \bar{x}|}{n}$$

Ya que los tiempos de espera del Banco Mulberry de 1, 3, 14 tienen desviaciones de -5, -3 y 8, la desviación media absoluta es $(5 + 3 + 8)/3 = 6/3 = 5.3$.

¿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 8.3 presentó un método para hacer inferencias acerca de las medias de dos poblaciones; dicho 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 se selecciona aleatoriamente un valor de cada población y se suman, dichas 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 presentan en el capítulo 9, y el análisis de varianza que se introduce en el capítulo 11. Además, el ejercicio 42 demuestra que el valor de la media absoluta presenta un sesgo, lo cual 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 sólo 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 usan en álgebra. Hay muchos ejemplos en los cuales un procedimiento estadístico se encuentra sesgado en una suma de cuadrados similar. Por lo tanto, en lugar de emplear valores absolutos, obtenemos una mejor medida de variación si logramos que todas las desviaciones $(x - \bar{x})$ no sean negativas, lo cual haremos elevándolas al cuadrado; este método conduce a la desviación estándar. Por tales razones, las calculadoras científicas suelen incluir una función para la desviación estándar, pero casi nunca la desviación media absoluta.

¿Por qué dividir entre $n - 1$? Después 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 hay solamente $n - 1$ valores independientes. Es decir, con una media dada, sólo a $n - 1$ valores se le puede asignar un número con libertad, antes de que se determine el último valor. Véase el ejercicio 41, que proporciona números concretos que ilustran como tal división entre $n - 1$ es mejor que la división entre n. Este ejercicio muestra 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 incrementando su valor general, haciendo que su denominador sea más pequeño (usando $n - 1$ en lugar de n). El ejercicio 41 demuestra cómo la división entre $n - 1$ provoca que la varianza muestral s^2 iguale el valor de la varianza poblacional σ^2; en tanto que la división entre n causa que la varianza muestral s^2 subestime el valor de la varianza poblacional σ^2.

El paso 6 de la fórmula 2-4, para el cálculo de una 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 se da en minutos, la desviación estándar de dichos tiempos también estará 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.

Después de estudiar dicha sección, usted debe comprender que la desviación estándar es una medida de variación entre valores. Al tener datos muestrales, será capaz de calcular el valor de la desviación estándar y de interpretar los valores de las desviaciones estándar que obtuvo. Debe saber que, para un conjunto de datos común, es raro que un valor difiera de la media por más de 2 o 3 desviaciones estándar.

2-5 Destrezas y conceptos básicos

En los ejercicios 1 a 8 calcule el rango, la varianza y de desviación estándar para los datos muestrales dados. (En la sección 2-4 se utilizaron los mismos datos para calcular medidas de tendencia central. Aquí calculamos medidas de variación).

1. Consumo de tabaco en películas infantiles En el artículo "Tobbaco 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), se registró la duración (en segundos) de escenas de películas de dibujos animados, de los Universal Studios, que muestran consumo de tabaco. A continuación se presentan los primeros seis valores, que se incluyen en el conjunto de datos 7 del Apéndice B. ¿Parecen tales duraciones ser consistentes o varían ampliamente?

0 223 0 176 0 548

2. Harry Potter En un intento por medir el nivel de lectura de un libro, se obtuvieron los puntajes de la facilidad de lectura de Flesch de 12 páginas que se seleccionaron aleatoriamente, de la obra *Harry Potter y la piedra filosofal*, de J. K. Rowling. Dichos puntajes, que se encuentran en el conjunto de datos 14 del Apéndice B, se listan a continuación. Debido a que tales puntajes se basan en 12 páginas que se seleccionaron aleatoriamente, ¿es probable que la desviación estándar de esta muestra sea un estimado razonable de la desviación estándar de los niveles de lectura de todas las páginas del libro?

85.3	84.3	79.5	82.5	80.2	84.6
79.2	70.9	78.6	86.2	74.0	83.7

3. Cereal Un nutriólogo obtiene las cantidades de azúcar (en gramos) de un gramo de 16 cereales diferentes, incluyendo Cheerios, Corn Flakes, Fruit Loops, Trix y 12 más. Estos valores, que se incluyen en el conjunto de datos 16 del Apéndice B, se listan a continuación. ¿Será la desviación estándar de dichos valores un buen estimado de la desviación estándar de la cantidad de azúcar en cada gramo del cereal consumido por la población de todos los estadounidenses? ¿Por qué?

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

4. Índice de masa corporal Como parte del National Health Examination se mide el índice de masa corporal en una muestra aleatoria de mujeres. Algunos de los valores, que se anexan en el conjunto de datos 1 del Apéndice B, se listan a continuación. ¿Estará la desviación estándar de la muestra razonablemente cerca de la desviación estándar de 6.17, que es la desviación estándar de las 40 mujeres que se incluyen en el conjunto de datos 1?

19.6	23.8	19.6	29.1	25.2	21.4	22.0	27.5
33.5	20.6	29.9	17.7	24.0	28.9	37.7	

5. Conductores alcoholizados Abajo se listan las concentraciones de alcohol en la sangre de conductores que se vieron envueltos en accidentes fatales y que después fueron sentenciados a prisión (de acuerdo con datos del Departamento de Justicia de Estados Unidos). Cuando un estado lanza una campaña para "reducir el número de conductores alcoholizados", ¿es la intención de la campaña disminuir la desviación estándar?

0.27	0.17	0.17	0.16	0.13	0.24	0.29	0.24
0.14	0.16	0.12	0.16	0.21	0.17	0.18	

6. Muertes en motocicleta A continuación se presentan las edades de motociclistas cuando se accidentaron fatalmente en accidentes de tránsito (de acuerdo con datos del Departamento del Transporte de Estados Unidos). ¿De qué manera se compara la variación de estas edades con la variación de las edades de conductores con licencia en la población general?

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

7. Tiempos de reacción El autor visitó el Museo de Ciencias Reuben H. Fleet, en San Diego, y repitió un experimento de tiempos de reacción. Se obtuvieron los siguientes tiempos (en centésimas de segundo). ¿De que manera las medidas de variación reflejan el hecho de que tales tiempos parezcan muy consistentes?

19	20	17	21	21	21	19	18	19	19
17	17	15	17	18	17	18	18	18	17

8. Tabletas de Bufferin A continuación se listan los pesos medidos (en miligramos) de una muestra de tabletas de aspirina Bufferin. Como este medicamento debe fabricarse de forma consistente para que las dosis se controlen, ¿las medidas de variación parecen indicar que la variación tiene un nivel aceptable?

672.2	679.2	669.8	672.6	672.2	662.2
662.7	661.3	654.2	667.4	667.0	670.7

En los ejercicios 9 a 12 calcule el rango, la varianza y la desviación estándar para cada una de las dos muestras; luego, compare los dos conjuntos de resultados. (En la sección 2.4 se utilizaron los mismos datos).

9. 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):

Jefferson Valley:	6.5	6.6	6.7	6.8	7.1	7.3	7.4	7.7	7.7	7.7
Providence:	4.2	5.4	5.8	6.2	6.7	7.7	7.7	8.5	9.3	10.0

10. 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:

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

11. Mickey D *vs.* Jack Al investigar los tiempos que se requieren en el servicio para automóvil (en segundos), se obtienen los siguientes resultados (con base en datos del QSR Drive-Thru Time Study).

McDonald's:	287	128	92	267	176	240	192	118	153	254	193	136
Jack in the Box:	190	229	74	377	300	481	428	255	328	270	109	109

12. Anchura de cráneos Las anchuras máximas de muestras de cráneos egipcios de varones que datan del 4000 a.C. y del 150 d.C (de acuerdo con datos de *Ancient Races of the Thebaid*, de Thomson y Randall-Maciver) se muestran a continuación:

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

En los ejercicios 13 a 16 remítase a los conjuntos de datos del Apéndice B. Utilice un programa de cómputo o una calculadora para obtener las desviaciones estándar; luego, compare los resultados.

T **13.** Circunferencia de cabezas Para diagnosticar de forma correcta el trastorno de hidrocefalia, un pediatra investiga la circunferencia de las cabezas de niños y niñas de dos años de edad. Utilice los resultados muestrales listados en el conjunto de datos 3. ¿Hay alguna diferencia entre los dos géneros?

T **14.** Clancy, Rowling, Tolstoi Un psicólogo infantil investiga las diferencias en la facilidad de lectura; obtiene datos con *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. Remítase al conjunto de datos 14 del Apéndice B y utilice las puntuaciones de la calificación de Flesch-Kincaid de las 12 páginas seleccionadas aleatoriamente para cada uno de los tres libros.

T **15.** Lluvia el fin de semana Utilice el conjunto de datos 11 del Apéndice B, sobre las cantidades de lluvia que caen en Boston los jueves y las que caen los domingos.

T **16.** Consumo de tabaco/alcohol en películas infantiles En el artículo "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), se registraron las duraciones (en segundos) de escenas que muestran consumo de tabaco y alcohol en películas infantiles de dibujos animados. En el conjunto de datos 7 del Apéndice B, utilice las duraciones de escenas con tabaco y después las escenas con alcohol.

En los ejercicios 17 a 20 calcule la desviación estándar de los datos que se resumen en la distribución de frecuencias dada. (En la sección 2.4 se utilizaron las mismas distribuciones de frecuencias.)

17. Old Faithful Los visitantes del Parque Nacional Yellowstone consideran que una erupción del géiser Old Faithful es una gran atracción que uno no debe perderse. La distribución de frecuencias dada resume una muestra de los tiempos (en minutos) entre las erupciones.

18. Dado cargado El autor taladró un hoyo en un dado y lo rellenó con plomo; después, procedió a lanzarlo 200 veces. Los resultados se presentan en la distribución de frecuencias al margen.

19. Infracciones de tránsito La distribución de frecuencias describe las velocidades de conductores infraccionados por la policía en la ciudad de Poughkeepsie. Los conductores viajaban a través de una zona con límite de velocidad de 30 millas/hora en Creek Road, que pasa por la universidad del autor.

Tabla del ejercicio 17

Tiempo	Frecuencia
40–49	8
50–59	44
60–69	23
70–79	6
80–89	107
90–99	11
100–109	1

Tabla del ejercicio 18

Resultado	Frecuencia
1	27
2	31
3	42
4	40
5	28
6	32

Tabla del ejercicio 19

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

Tabla del ejercicio 20

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

20. Temperaturas corporales La distribución de frecuencias al margen resume una muestra de temperaturas corporales humanas. (Véanse las temperaturas de medianoche del segundo día, listadas en el conjunto de datos 4 del Apéndice B).

21. Edades de profesores Utilice la regla práctica del intervalo para estimar la desviación estándar de las edades de todos los profesores de su universidad.

22. Calificaciones de prueba Con la regla práctica del intervalo estime la desviación estándar de las calificaciones del primer examen de estadística de su clase.

23. Longitudes de piernas En los datos muestrales del conjunto de datos 1 del Apéndice B, las longitudes del muslo de la muestra de 40 mujeres tienen una media de 38.86 centímetros y una desviación estándar de 3.78 centímetros. Use la regla práctica del intervalo para estimar las longitudes mínima y máxima "comunes" de los muslos de las mujeres. En dicho contexto, ¿una longitud de 47.0 centímetros sería considerada infrecuente?

24. Estaturas de mujeres La media de las estaturas de las mujeres es de 63.6 pulgadas, con una desviación estándar de 2.5 pulgadas (con base en datos del National Health Survey). Utilice la regla práctica del intervalo para estimar las estaturas mínima y máxima "comunes" de las mujeres. En tal contexto, ¿es poco común que una mujer mida seis pies?

25. Estaturas de mujeres Las estaturas de las mujeres tienen una distribución normal, con una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas. Utilice la regla empírica para determinar el porcentaje aproximado de mujeres que están entre
a. 61.1 y 66.1 pulgadas
b. 56.1 y 71.1 pulgadas

26. Pesos de la Coca Cola regular Con los pesos de la Coca Cola regular listados en el conjunto de datos 17 del Apéndice B, encontramos que la media es de 0.81682 libras, la desviación estándar es de 0.00751 libras y la distribución es aproximadamente normal. Aplique la regla empírica y determine el porcentaje aproximado de latas de Coca Cola regular que tienen pesos entre
a. 0.80931 y 0.82433 libras
b. 0.80180 y 0.83184 libras

27. Estaturas de mujeres Si las estaturas de mujeres tienen una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas, ¿qué se concluye a partir del teorema de Chebyshev acerca del porcentaje de mujeres que están entre 58.6 pulgadas y 68.6 pulgadas?

28. Pesos de la Coca Cola regular Utilizando los pesos de la Coca Cola regular listados en el conjunto de datos 17 del Apéndice B, encontramos que la media es de 0.81682 libras y la desviación estándar es de 0.00751 libras. ¿Qué concluye a partir del teorema de Chebyshev acerca del porcentaje de latas de Coca Cola regular con pesos que están entre 0.79429 y 0.83935 libras?

T **29.** Coeficiente de variación del cereal Remítase al conjunto de datos 17 del Apéndice B. Calcule el coeficiente de variación de las calorías y el coeficiente de variación de los gramos de azúcar por gramo de cereal. Compare los resultados.

T **30.** Coeficiente de variación de Coca Cola y de Pepsi Remítase al conjunto de datos 17 del Apéndice B. Calcule el coeficiente de variación de los pesos de la Coca Cola regular y después el coeficiente de variación de los pesos de la Pepsi regular. Compare los resultados. ¿Alguna de las dos compañías parece tener pesos significativamente más consistentes?

31. Igualdad para todos ¿Qué sabe usted acerca de los valores en un conjunto de datos con una desviación estándar $s = 0$?

32. Comprensión de las unidades de medición Si un conjunto de datos consiste en multas por exceso de velocidad (en dólares), ¿qué unidades se utilizan para la desviación estándar? ¿Qué unidades se utilizan para la varianza?

33. Comparación de baterías para automóviles Las marcas de baterías para automóviles Everlast y Endurance aseguran en su etiqueta una duración de 48 meses. En realidad, ambas tienen una vida media de 50 meses, pero las baterías Everlast tienen una desviación estándar de dos meses, mientras que la de las baterías Endurance es de seis meses. ¿Cuál de las marcas sería una mejor opción? ¿Por qué?

34. Interpretación de datos distantes Un conjunto de datos consta de 20 valores, bastante cercanos entre sí. Se incluye otro valor, pero éste es un dato distante (muy lejos de los demás). ¿De qué manera se ve afectada la desviación estándar por el dato distante? ¿No genera efecto alguno? ¿Tiene un efecto pequeño? ¿Un efecto grande?

2-5 Más allá de lo básico

35. Comparación de conjuntos de datos Dos secciones diferentes de un curso de estadística resuelven el mismo examen, cuyas calificaciones se muestran abajo. Calcule el rango y la desviación estándar de cada sección. ¿Qué se concluye acerca de la variación en las dos secciones, a partir de los valores del rango? ¿Por qué el rango causa confusión en este caso? ¿Qué se concluye acerca de la variación en las dos secciones con respecto a los valores de la desviación estándar?

Sección 1:	1	20	20	20	20	20	20	20	20	20	20
Sección 2:	2	3	4	5	6	14	15	16	17	18	19

36. Transformación de datos Describa de qué forma se afectan el rango y la desviación estándar de un conjunto de datos en los siguientes casos:

a. Se suma la misma constante K a cada valor del conjunto de datos.

b. Cada valor del conjunto de datos se multiplica por la misma constante K.

c. Para los datos de temperaturas corporales listados en el conjunto de datos 4 del Apéndice B (12 a. m. del día 2), $\bar{x} = 98.20°F$ y $s = 0.62°F$. Calcule los valores de y s después de convertir cada temperatura a grados Celsius. [*Considere que* $C = 5(F - 32)/9$].

37. Genichi Taguchi creó un método para mejorar la calidad y reducir los costos de fabricación, por medio de una combinación de ingeniería y estadística. Una herramienta básica en el método de Taguchi es el **cociente señal-ruido**. La forma más simple para calcular tal cociente es dividir la media entre la desviación estándar. Calcule el cociente señal-ruido de los niveles de cotinina de fumadores listados en la tabla 2-1.

38. Sesgo En la sección 2-4, introdujimos el concepto general de sesgo. El sesgo puede medirse por medio del **índice de sesgo de Pearson**:

$$I = \frac{3(\bar{x} - \text{mediana})}{s}$$

Si $I \geq 1.00$ o $I \leq -1.00$, los datos se consideran *significativamente sesgados*. Calcule el índice de sesgo de Pearson de los niveles de cotinina de fumadores, listados en la tabla 2-1, y determine si hay un sesgo significativo.

39. Comprensión de la desviación estándar Una muestra consiste en 10 calificaciones de pruebas, que caen entre 70 y 100, inclusive. ¿Cuál es la desviación estándar más grande posible?

40. ¿Datos falsos? Para cualquier conjunto de datos de n valores, con desviación estándar $s\sqrt{n-1}$, cada valor debe estar dentro de la media. Una profesora de estadística reporta que las calificaciones de una prueba que se aplicó a 17 estudiantes de su clase tuvo una media de 75.0 y una desviación estándar de 5.0. Kelly, a quien se considera la mejor estudiante de la clase, afirma haber recibido una calificación de 97. ¿Podría Kelly estar diciendo la verdad?

41. ¿Por qué dividir entre $n-1$? Sea que una población consista en los valores 3, 6, 9. Suponga que muestras de los valores se seleccionan aleatoriamente *con reemplazo*.

a. Calcule la varianza σ^2 de la población {3, 6, 9}.

b. Liste las nueve muestras diferentes posibles de los valores seleccionados con reemplazo; luego, calcule la varianza muestral s^2 (que incluye la división entre $n-1$) de cada una de ellas. Si se seleccionan de forma repetida dos valores muestrales, ¿cuál es el valor medio de la varianza muestral s^2?

c. Para cada una de las nueve muestras, calcule la varianza tratando cada muestra como si fuese una población. (Asegúrese de utilizar la fórmula de la varianza poblacional, que incluye la división entre n). Si selecciona de forma repetida dos valores muestrales, ¿cuál es el valor medio de las varianzas poblacionales?

d. ¿Con qué método se obtienen mejores estimados de σ^2: el inciso *b*) o el inciso *c*)? ¿Por qué? Al calcular las varianzas muestrales, ¿debe utilizarse la división entre n o $n-1$?

e. Los incisos anteriores muestran que s^2 es un estimador sin sesgo de σ^2. ¿Será s un estimador sin sesgo de σ?

42. ¿Por qué no utilizar la DMA? El ejercicio 41 demuestra que la varianza muestral s^2 es un estimador sin sesgo de σ^2. Haga lo que se le pide con la misma población de {3, 6, 9}, para demostrar que la desviación media absoluta de una muestra es un estimador sesgado de la desviación media absoluta de una población.

a. Calcule la desviación media absoluta de la población {3, 6, 9}.

b. Liste las nueve muestras diferentes posibles de dos valores seleccionadas con reemplazo, después calcule la desviación media absoluta de cada una de ellas. Si se seleccionan de forma repetida dos valores muestrales, ¿cuál es el valor medio de las desviaciones medias absolutas?

c. Con base en los resultados de los incisos *a*) y *b*), ¿la desviación media absoluta de una muestra tiende a igualar la desviación media absoluta de una población? ¿La división entre $n-1$, en lugar de la división entre n, convierte a la desviación media absoluta en un estimado sin sesgo de la desviación media absoluta de la población?

2-6 Medidas de posición relativa

Esta sección incluye medidas que pueden utilizarse para comparar valores de diferentes conjuntos de datos o para comparar valores dentro del mismo conjunto de datos. Aquí introducimos las puntuaciones z (para comparar valores de distintos conjuntos de datos), así como los cuartiles y percentiles (para comparar valores dentro del mismo conjunto de datos).

Puntuaciones *z*

Una puntuación z (o puntuación estándar) se calcula convirtiendo un valor a una escala estandarizada, como se establece en la siguiente definición. Utilizaremos ampliamente las puntuaciones z en el capítulo 5 y en capítulos posteriores, ya que son muy importantes.

Definición

Puntuación estándar o **puntuación z:** número de desviaciones estándar que un valor x se encuentra por arriba o por debajo de la media. Se calcula utilizando las siguientes expresiones:

$$\text{Muestra} \qquad z = \frac{x - \bar{x}}{s} \qquad \text{o} \qquad \text{Población} \qquad z = \frac{x - \mu}{\sigma}$$

(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 La superestrella de la NBA Michael Jordan mide 78 pulgadas, en tanto que la jugadora de basquetbol de la WNBA Rebecca Lobo mide 76 pulgadas. En efecto, Jordan es más alto por dos pulgadas, pero ¿cuál de los jugadores es *relativamente* más alto? ¿La estatura de Jordan, entre los hombres, excede la estatura de Lobo entre las mujeres? Los hombres tienen estaturas con una media de 69.0 pulgadas, con una desviación estándar de 2.8 pulgadas; las mujeres tienen estaturas con una media de 63.6 pulgadas, con una desviación estándar de 2.5 pulgadas (datos basados en el National Health Survey).

SOLUCIÓN Para comparar las estaturas de Michael Jordan y Rebecca Lobo, en relación con las poblaciones de hombres y mujeres, necesitamos estandarizar dichas estaturas convirtiéndolas en puntuaciones z.

$$\text{Jordan:} \qquad z = \frac{x - \mu}{\sigma} = \frac{78 - 69.0}{2.8} = 3.21$$

$$\text{Lobo:} \qquad z = \frac{x - \mu}{\sigma} = \frac{76 - 63.6}{2.5} = 4.96$$

INTERPRETACIÓN La estatura de Michael Jordan está a 3.21 desviaciones estándar por arriba de la media, pero la estatura de Rebecca Lobo está a 4.96 desviaciones estándar por arriba de la media. La estatura de Rebecca Lobo entre las mujeres es relativamente mayor que la estatura de Michael Jordan entre los hombres.

Puntuaciones *z* y valores infrecuentes

En la sección 2-5 utilizamos la regla práctica del intervalo para concluir que un valor es "infrecuente" o poco común 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 2-14 en la página 94). Si aplicamos este criterio, tanto Michael Jordan como Rebecca Lobo tienen estaturas infrecuentes, ya que ambos cuentan con estaturas con puntuaciones z mayores que 2.

Si consideramos a jugadores profesionales de basquetbol con estaturas excepcionales, tomemos en cuenta a otro jugador, Mugsy Bogues, que alcanzó el éxito aun cuando sólo mide 5 pies y 3 pulgadas. (Nuevamente usamos el hecho de que

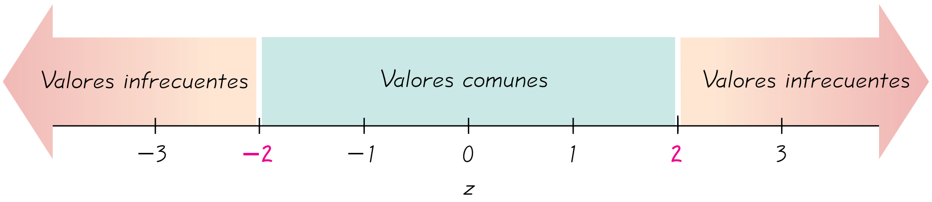


FIGURA 2-14
Interpretación de las puntuaciones *z*
Los valores infrecuentes son aquellos con puntuaciones *z* menores que −2.00 o mayores que 2.00.

los hombres tienen estaturas con una media de 69.0 pulgadas, con una desviación estándar de 2.8 pulgadas). Después de convertir 5 pies y 3 pulgadas a 63 pulgadas, convertimos su estatura en una puntuación *z* de la siguiente manera:

$$\text{Bogues: } z = \frac{x - \mu}{\sigma} = \frac{63 - 69.0}{2.8} = -2.14$$

Demos gracias a Mugsy Bogues por sus muchos años de juego inspirado y por ilustrar este principio:

> **Siempre que un valor sea menor que la media, su puntuación *z* correspondiente será negativa**
>
> **Valores comunes:** $\mathbf{-2 \le z}$ **puntuación** $\mathbf{\le 2}$
>
> **Valores infrecuentes:** ***z*** **puntuación** $\mathbf{< -2}$ *o* ***z*** **puntuación** $\mathbf{> 2}$

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 encima* 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*; son útiles para comparar valores dentro del mismo conjunto de datos o entre distintos conjuntos de datos.

Cuartiles y percentiles

De la sección 2-4, recuerde que la mediana de un conjunto de datos es el valor que está a la mitad, de modo que 50% de los valores son iguales o menores a la mediana y el 50% de los valores son mayores o iguales a la mediana. Tal 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 se acomodan en orden).

He aquí 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, y con frecuencia los distintos programas de cómputo producen resultados diferentes. Por ejemplo, si usted utiliza un 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
Excel	5.25	12.5	22.75
TI-83 Plus	4.5	12.5	24.5

Para este conjunto de datos, STATDISK y la calculadora TI-83 Plus coinciden, aunque esto no siempre sucede. Si 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 hay tres cuartiles que separan un conjunto de datos en cuatro partes, también se tienen 99 **percentiles**, que se denotan $P_1, P_2, \ldots, P_{99}$, 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 casi el mismo número de valores).

El proceso para calcular percentiles, que corresponde a un valor particular x, es bastante sencillo, tal como se indica en la siguiente expresión:

$$\text{percentil del valor } x = \frac{\text{número de valores menores que } x}{\text{número total de valores}} \cdot 100$$

EJEMPLO Niveles de cotinina de fumadores La tabla 2-13 lista los 40 niveles ordenados de cotinina de fumadores que se incluyen en la tabla 2-1. Calcule el percentil correspondiente al nivel de cotinina de 112.

SOLUCIÓN A partir de la tabla 2-13 se ve que hay dos valores menores que 112; por lo tanto,

$$\text{percentil de } 112 = \frac{12}{40} \cdot 100 = 30$$

INTERPRETACIÓN El nivel de cotinina de 112 es el percentil 30o.

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 usaremos se resume en la figura 2-15, que emplea la notación que viene inmediatamente después.

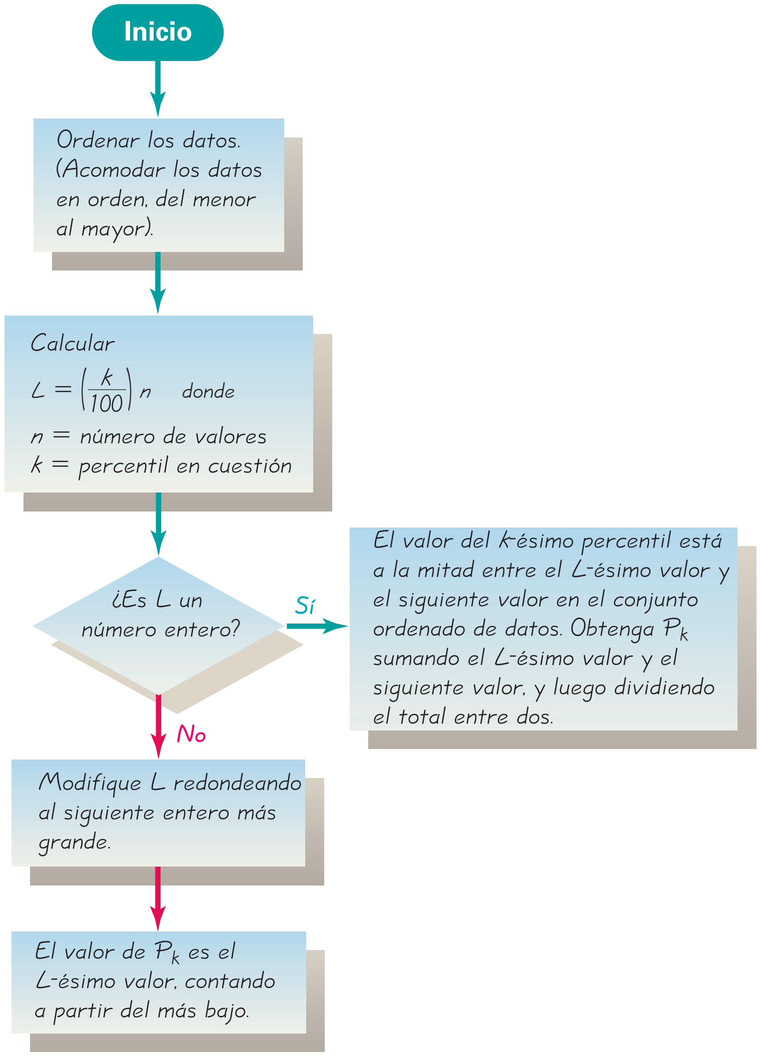
Tabla 2-13 *Niveles ordenados de cotinina de 40 fumadores*

0	1	1	3	17	32	35	44	48	86
87	103	112	121	123	130	131	149	164	167
173	173	198	208	210	222	227	234	245	250
253	265	266	277	284	289	290	313	477	491

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, tales como el ICR y el IPC, nos permiten comparar el valor de alguna variable con su valor en un periodo base. El valor de un número índice es el valor real dividido entre el valor base, multiplicado por 100.

FIGURA 2-15 Conversión del *k*-ésimo percentil al valor del dato correspondiente



Notación

n = número total de valores en el conjunto de datos
k = percentil utilizado (ejemplo: para el percentil 25o, $k = 25$).
L = localizador que da la *posición* de un valor (ejemplo: para el valor 12o en la lista ordenada, $L = 12$).
P_k = percentil k-ésimo (ejemplo: P_{25} es el percentil 25o).

PROBLEMA 2 DEL CAPÍTULO

EJEMPLO Niveles de cotinina de fumadores Remítase a los niveles ordenados de cotinina de fumadores de la tabla 2-13, y utilice la figura 2-15 para calcular el valor del percentil 68o, P_{68}.

SOLUCIÓN Si nos referimos a la figura 2-15, veremos que los datos muestrales ya están ordenados, de modo que es posible proceder al cálculo del valor del localizador L. En el cálculo utilizamos $k = 68$, ya que estamos tratando de obtener el valor del percentil 68o. Usamos $n = 40$, porque tenemos 40 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{68}{100} \cdot 40 = 27.2$$

Después, nos preguntamos si L es un número entero; respondemos que no. Por lo tanto, procederemos al siguiente recuadro hacia abajo y modificamos L, redondeándola de 27.2 a 28. (En este libro solemos redondear de la forma común, pero es uno de los casos donde redondeamos hacia *arriba* y no hacia el entero *más cercano*). Por último, el recuadro final muestra que el valor de P_{68} es el valor 28o, contando hacia arriba, desde el mínimo. En la tabla 2-13, el valor 28o es 234. Es decir, $P_{68} = 234$.

PROBLEMA 2 DEL CAPÍTULO

EJEMPLO Niveles de cotinina de fumadores Remítase a la muestra de niveles de cotinina de fumadores que aparece en la tabla 2-13. Utilice la figura 2-15 para calcular el valor de Q_1, que es el primer cuartil.

SOLUCIÓN Primero observamos que Q_1 es igual que P_{25}, por lo que procedemos a calcular el valor del percentil 25o. Si nos referimos a la figura 2-15, veremos que los datos muestrales ya se ordenaron, de manera que procedemos a calcular el valor del localizador L. En este cálculo utilizamos $k = 25$, ya que tratamos de obtener el valor del percentil 25o y usamos $n = 40$, porque tenemos 40 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{25}{100} \cdot 40 = 10$$

Después, nos preguntamos si L es un número entero; respondemos que sí. Por lo tanto, vamos al recuadro que se ubica a la derecha. Vemos que el valor del percentil k-ésimo (25o) está a la mitad entre el valor L-ésimo (10o) y el siguiente valor en el conjunto original de datos. Es decir, el valor del percentil 25o se ubica a la mitad, entre el 10o valor y el 11o valor. El 10o valor es 86 y el 11o valor es 87; por lo tanto, el valor a la mitad de ellos es 86.5. Concluimos que el percentil 25o es $P_{25} = 86.5$. El valor del primer cuartil Q_1 es también 86.5.

El ejemplo anterior demuestra que al calcular un valor cuartilar (como Q_1), es posible utilizar el valor del percentil equivalente (como P_{25}) en su lugar. Al margen, se indican las relaciones equivalentes entre cuartiles y percentiles.

$Q_1 = P_{25}$
$Q_2 = P_{50}$
$Q_3 = P_{75}$

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:

$$\text{rango intercuartilar (o RIC)} = Q_3 - Q_1$$

$$\text{rango semiintercuartilar} = \frac{Q_3 - Q_1}{2}$$

$$\text{cuartil medio} = \frac{Q_3 + Q_1}{2}$$

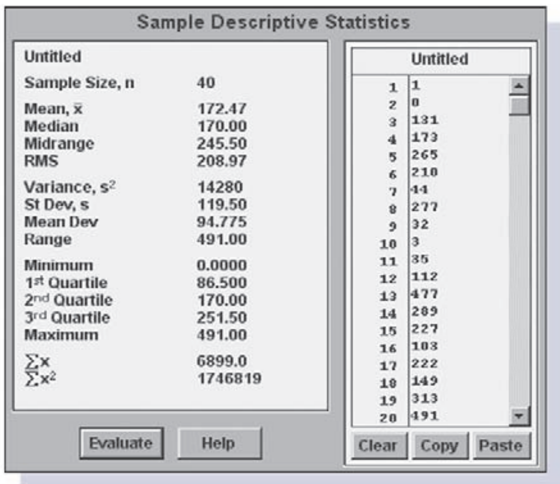
$$\text{rango de percentiles 10—90} = P_{90} - P_{10}$$

Después de completar esta sección, usted debe ser capaz de convertir un valor en su puntuación z (o puntuación estándar) correspondiente, de manera que sea posible compararlo con otros valores que provienen de diferentes conjuntos de datos. También tendrá que ser capaz de convertir un valor en su valor percentil 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 $Q_3 = P_{75}$).

Utilizando la tecnología

Se puede utilizar una variedad de programas de cómputo y calculadoras diferentes para calcular muchos de los estadísticos estudiados hasta ahora en este capítulo. En la sección 2-4 dimos instrucciones específicas para el uso de STATDISK, Minitab, Excel y la calculadora TI-83 Plus. Señalamos que en ocasiones es posible introducir un conjunto de datos y utilizar 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 los niveles de cotinina de *fumadores* que se presentaron en la tabla 2-1, en el problema del capítulo. Los resultados de la calculadora TI-83 Plus se muestran en dos pantallas, ya que no caben en una sola.

STATDISK



Minitab

Variable	N	Mean	Median	TrMean	StDev	SE Mean
SMOKER	40	172.5	170.0	164.7	119.5	18.9

Variable	Minimum	Maximum	Q1	Q3
SMOKER	0.0	491.0	86.3	252.3

Excel

Column1	
Mean	172.475
Standard Error	18.89434
Median	170
Mode	1
Standard Deviation	119.4983
Sample Variance	14279.85
Kurtosis	0.519621
Skewness	0.587929
Range	491
Minimum	0
Maximum	491
Sum	6899
Count	40

TI-83 Plus

```
1-Var Stats
 x̄=172.475
 Σx=6899
 Σx²=1746819
 Sx=119.4983076
 σx=117.9951244
↓n=40
```

```
1-Var Stats
↑n=40
 minX=0
 Q1=86.5
 Med=170
 Q3=251.5
 maxX=491
```

2-6 Destrezas y conceptos básicos

En los ejercicios 1 a 4 exprese todas las puntuaciones z con dos decimales.

1. Puntuaciones de CI Las puntuaciones de CI de la prueba Stanford Binet tienen una media de 100 y una desviación estándar de 16. Albert Einstein obtuvo un CI de 160.

a. ¿Cuál es la diferencia entre el CI de Einstein y la media?

b. ¿Cuántas desviaciones estándar implica esto [la diferencia obtenida en el inciso *a*)]?

c. Convierta la puntuación de CI de Einstein a puntuación z.

d. Si consideramos que las puntuaciones de CI "comunes" son aquellas que, convertidas en puntuaciones z, caen entre -2 y 2, ¿es el CI de Einstein común o infrecuente?

2. Pulso de adultos Suponga que los adultos tienen pulsos (latidos por minuto) con una media de 72.9 y una desviación estándar de 12.3 (con base en datos del National Health Examination). Cuando escribió este ejercicio, el autor tenía un pulso de 48.

a. ¿Cuál es la diferencia entre el pulso del autor y la media?

b. ¿Cuántas desviaciones estándar implica esto [la diferencia obtenida en el inciso *a*)]?

c. Convierta el pulso de 48 a puntuación z.

d. Si consideramos que los pulsos "comunes" son aquellos que, convertidos en puntuaciones z, caen entre -2 y 2, ¿es el pulso de 48 común o infrecuente? ¿Podría explicar por qué un pulso sería inusualmente bajo? (La razón de este pulso tan bajo *no* es que los autores de libros de estadística estén comúnmente en un estado que se describiría como comatoso).

3. Estaturas de hombres Las estaturas de hombres adultos tienen una media de 69.0 pulgadas y una desviación estándar de 2.8 pulgadas. Calcule las puntuaciones z que corresponden a los siguientes individuos:

a. El actor Danny DeVito, que mide 5 pies.

b. El jugador de basquetbol de la NBA, Shaquille O'Neal, que mide 7 pies 1 pulgada.

c. El autor, quien es un "jugador" de golf y tenis, que mide 69.72 pulgadas.

4. Temperaturas corporales La temperatura corporal humana tiene una media de 98.20°F y una desviación estándar de 0.62°F. Convierta las temperaturas dadas a puntuaciones z.
a. 100° **b.** 96.96° **c.** 98.20°

En los ejercicios 5 a 8 exprese todas las puntuaciones z con dos decimales. Considere una puntuación como infrecuente si es menor que −2.00 o mayor que 2.00.

5. Estaturas de mujeres El Club Beanstalk es sólo para mujeres y hombres muy altos. La estatura mínima que se requiere en las mujeres es de 70 pulgadas. Las estaturas de las mujeres tienen 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; después, determine si dicha estatura es infrecuente.

6. Duración del embarazo Una mujer, que escribió a *Dear Abby*, afirmó que dio a luz 308 días después de una visita de su esposo, que estaba en la marina. La duración del embarazo tiene una media de 268 días y una desviación estándar de 15 días. Calcule la puntuación z de 308 días. ¿Es infrecuente la duración? ¿Qué concluye?

7. Temperatura corporal La temperatura corporal humana tiene una media de 98.20°F y una desviación estándar de 0.62°F. Se descubre que un paciente de urgencias tiene una temperatura de 101°F. Convierta 101° en puntuación z. ¿Es la temperatura inusualmente alta? ¿Qué sugiere esto?

8. Niveles de colesterol Para hombres de entre 18 y 24 años de edad, los niveles séricos de colesterol (en mg/100ml) tienen una media de 178.1 y una desviación estándar de 40.7 (con base en datos del National Health Survey). Calcule la puntuación z correspondiente de un hombre de entre 18 y 24 años, quien presenta un nivel sérico de colesterol de 259.0 mg/100ml. ¿Es este nivel inusualmente alto?

9. Comparación de calificaciones de una prueba ¿Cuál es relativamente mejor: una calificación de 85 en una prueba de psicología o una calificación de 45 en una prueba de economía? Las calificaciones en la prueba de psicología tienen una media de 90 y una desviación estándar de 10. Las calificaciones en la prueba de economía tienen una media de 55 y una desviación estándar de 5.

10. Comparación de calificaciones Tres estudiantes resuelven pruebas equivalentes del sentido del humor; una vez que la risa disminuye, se calculan sus calificaciones. ¿Cuál es la calificación relativa más alta?
a. Una calificación de 144 en una prueba que tiene una media de 128 y una desviación estándar de 34.
b. Una calificación de 90 en una prueba que tiene una media de 86 y una desviación estándar de 18.
c. Una calificación de 18 en una prueba que tiene una media de 15 y una desviación estándar de 5.

T **11.** Peso de Coca Cola Remítase en el conjunto de datos 17 del Apéndice B, a la muestra de 36 pesos de Coca Cola regular. Convierta en peso de 0.7901 en puntuación z. ¿Es 0.7901 un peso inusual de la Coca Cola regular?

T **12.** M&M verdes Remítase en el conjunto de datos 19 del apéndice B, a la muestra de pesos de dulces M&M verdes. Convierta el peso del M&M verde más pesado en puntuación z. ¿Es infrecuente el peso del M&M verde más pesado en estos dulces?

En los ejercicios 13 a 16 utilice los 40 niveles ordenados de cotinina de fumadores, listados en la tabla 2.13. Calcule el percentil correspondiente a los niveles de cotinina dados.

13. 149 **14.** 210 **15.** 35 **16.** 250

En los ejercicios 17 a 24 utilice los 40 niveles ordenados de cotinina de fumadores, listados en la tabla 2.13. Calcule el percentil o cuartil indicado.

17. P_{20} **18.** Q_3 **19.** P_{75} **20.** Q_2
21. P_{33} **22.** P_{21} **23.** P_1 **24.** P_{85}

En los ejercicios 25 a 28 utilice los niveles de colesterol de mujeres listados en el conjunto de datos 1 del Apéndice B. Calcule el percentil correspondiente al nivel de colesterol dado.

25. 123 **26.** 309 **27.** 271 **28.** 126

En los ejercicios 29 a 36 utilice los niveles de colesterol de mujeres listados en el conjunto de datos 1 del Apéndice B. Calcule el percentil o cuartil indicado.

29. P_{85} **30.** P_{35} **31.** Q_1 **32.** Q_3

33. P_{18} **34.** P_{36} **35.** P_{58} **36.** P_{96}

2-6 Más allá de lo básico

37. Unidades de medición Cuando se calcula una puntuación z para la estatura de un jugador de basquetbol de la NBA, ¿de qué manera se afecta el resultado si, en lugar de utilizar pulgadas, todas las estaturas se expresan en centímetros? En general, ¿de qué manera se afectan las puntuaciones z por la unidad particular de medición que se utiliza?

38. Conversión de una puntuación z La estatura de las mujeres tiene una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas.

a. Julia Roberts, que es una de las actrices más exitosas de los últimos años, tiene una estatura que, convertida a puntuación *z, es* de 2.16. ¿Qué tan alta es (en pulgadas)?

b. La cantante de rap Lil'Kim tiene una estatura que, convertida a puntuación *z, es* de -1.84. ¿Qué tan alta es (en pulgadas)?

39. Distribución de puntuaciones z

a. Un conjunto de datos tiene una distribución uniforme. Si todos los valores se convierten a puntuaciones z, ¿cuál es la forma de la distribución de las puntuaciones z?

b. Un conjunto de datos tiene una distribución normal. Si todos los valores se convierten a puntuaciones z, ¿cuál es la forma de la distribución de las puntuaciones z?

c. En general, ¿cómo se ve afectada la forma de una distribución si todos los valores se convierten en puntuaciones z?

40. Secuencia de Fibonacci Éstos son los primeros de muchos términos de la famosa secuencia de Fibonacci: 1, 1, 2, 3, 5, 8, 13.

a. Calcule la media $\bar{x}$ y la desviación estándar s; después, convierta cada valor a puntuación z. No redondee las puntuaciones z. Ocupe tantos datos como su calculadora pueda manejar.

b. Calcule la media y la desviación estándar de las puntuaciones z que se obtuvieron en el inciso *a*).

c. Si utilizara cualquier otro conjunto de datos, ¿obtendría los mismos resultados que en el inciso *b*)?

41. Niveles de cotinina de fumadores Utilice los niveles ordenados de cotinina de fumadores que se listan en la tabla 2-3.

a. Calcule el rango intercuartilar.

b. Calcule el cuartil medio.

c. Calcule el rango de percentiles 10-90.

d. ¿Es $P_{50} = Q_2$? Si es así, ¿es P_{50} *siempre* igual a Q_2?

e. ¿Es $Q_2 = (Q_1 + Q_3)/2$? Si es así, ¿es Q_2 *siempre* igual a $(Q_1 + Q_3)/2$?

42. Interpolación Cuando se calculan percentiles con el uso de la figura 2-15, si el localizador L no es un número entero, lo redondeamos hacia el siguiente número entero mayor. Una alternativa para este procedimiento es interpolar, de modo que un localizador de 23.75 conduce a un valor que está a 0.75 (o 3/4) del camino entre los valores 23o y 24o. Utilice este método de interpolación para calcular P_{35} y Q_1 para los pesos de los osos que se listan en el conjunto de datos 9 del Apéndice B.

43. Deciles y cuartiles En un conjunto de datos hay nueve **deciles**, que se denotan con D_1, $D_2, \ldots, D_9$, que dividen los datos ordenados en 10 grupos, con aproximadamente 10% de los valores en cada grupo. También existen cuatro **quintiles**, que dividen los datos ordenados en cinco grupos, con aproximadamente 20% de los valores en cada grupo. (Note la diferencia entre los quintiles y los cuartiles, que ya describimos en esta sección).
- **a.** ¿Qué percentil es equivalente a D_1? ¿A D_5? ¿A D_8?
- **b.** Utilice los niveles ordenados de cotinina de fumadores de la tabla 2-13 y calcule los nueve deciles.
- **c.** Utilice los niveles ordenados de cotinina de fumadores de la tabla 2-13 y calcule los cuatro quintiles.

2-7 Análisis exploratorio de datos (AED)

El presente capítulo presenta las herramientas básicas para describir, explorar y comparar datos; esta sección se enfoca en la exploración de datos. Iniciamos definiendo el análisis exploratorio de datos; después, introduciremos los datos distantes, el resumen de 5 números y las gráficas de cuadro.

Definición

Análisis exploratorio de datos: proceso para 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, y que iniciamos con **1.** *el centro*, **2.** *la variación*, y **3.** la naturaleza de la *distribución*. Tales características pueden investigarse calculando los valores de la media y la desviación estándar, así como por medio de la construcción de un histograma. Por lo general, es importante investigar más el conjunto de datos para identificar cualquier particularidad notable, en especial aquélla que llegue a afectar de forma importante los resultados y las conclusiones. Una de estas características es la presencia de datos distantes.

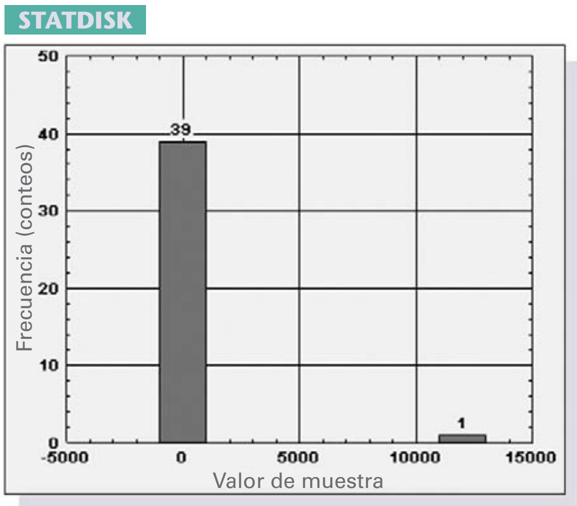
Datos distantes

Dato distante: valor que está muy alejado de la mayoría de los demás valores. Un dato distante es un valor *extremo* en relación con los otros datos. Cuando se explora un conjunto de datos, se deben considerar los datos distantes, 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 un dato distante; aunque no todos los datos distantes son errores, algunos de ellos son valores correctos.

PROBLEMA DEL CAPÍTULO 2

EJEMPLO **Niveles de cotinina de fumadores** Cuando se utiliza un programa de cómputo o una calculadora, es muy fácil cometer errores con los dedos. Remítase a los niveles de cotinina de fumadores que se listan en la tabla 2-1, en el problema del capítulo; suponga que el primer dato de 1 se introduce de manera incorrecta como 11111, porque usted estaba distraído viendo un meteorito que aterrizaba en su jardín. El dato incorrecto de 11111 es un dato distante, ya que se localiza muy lejos de los demás valores. ¿De qué manera afecta ese dato distante a la media, a la desviación estándar y al histograma?

SOLUCIÓN Cuando el dato 1 se reemplaza con el valor distante de 11111, la media cambia de 172.5 a 450.2, de modo que el efecto del dato distante es muy grande. El dato incorrecto de 11111 causa que la desviación estándar cambie de 119.5 a 1732.7, por lo que el efecto del dato distante también es muy grande. La figura 2-1, en la sección 2-3, muestra el histograma con los valores correctos de los niveles de cotinina de fumadores de la tabla 2-1, pero la representación visual del STATDISK que se muestra aquí, contiene el histograma que resulta del uso de los mismos datos con el valor de 1, reemplazado por el valor incorrecto de 11111. Compare el histograma del STATDISK con la figura 2-1 y verá fácilmente que la presencia del dato distante afecta de manera drástica la forma de la distribución.



El ejemplo anterior ilustra estos principios importantes:

1. **Un dato distante puede tener un efecto importante sobre la media.**
2. **Un dato distante puede tener un efecto importante sobre la desviación estándar.**
3. **Un dato distante puede tener un efecto importante sobre la escala del histograma, de modo que la verdadera naturaleza de la distribución se oculta totalmente.**

Un procedimiento sencillo para encontrar datos distantes es el examen de una lista *ordenada* de los datos. En particular, observe los valores mínimo y máximo muestrales; luego, determine si se alejan mucho de los demás valores. Algunos datos distantes son valores correctos y algunos son errores, como en el ejemplo anterior. Si estamos seguros de que un dato distante es un error, debemos corregirlo o eliminarlo. Si incluimos un dato distante, 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 no incluyan los datos distantes.

Una propina extrema

Es importante tomar en cuenta los datos distantes ya que, en muchos casos, un valor extremo puede tener un efecto muy importante en los estadísticos y en las conclusiones que se derivan de ellos. A veces un dato distante es un error que debe ser corregido o eliminado. En otros, un dato distante es un valor válido que debe investigarse para obtener información importante. Algunos alumnos del autor, al reunir datos consistentes de facturas y propinas de restaurantes, no encontraron datos distantes sobresalientes en esos valores muestrales. Sin embargo, un dato distante es la propina de 16,000 dólares que se dio por una cuenta de 8,899.78 dólares 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 ya antes había atendido al cliente y que "él siempre es generoso, pero nunca antes de esta forma. Tengo que hablarle a mi hermana acerca de él".

Gráficas de cuadro

Además de las gráficas presentadas en la sección 2.3, 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 presencia de datos distantes. La construcción de una gráfica de cuadro requiere que primero se obtenga el valor mínimo, el valor máximo y los cuartiles, tal como se define en el resumen de los cinco números.

Definiciones

Para un conjunto de datos, el **resumen de los cinco 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.

Gráfica de cuadro (o **diagrama de cuadro y bigotes**): gráfica de un conjunto de datos que consiste en una línea que se extiende desde el valor mínimo hasta el valor máximo, así como una caja con líneas trazadas en el primer cuartil, Q_1; la mediana y el tercer cuartil, Q_3. (Véase la figura 2-16).

Procedimiento para construir una gráfica de cuadro

1. Elabore el resumen de los cinco números, consistente 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 afuera 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 podría no ser 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 manera que sea posible realizar comparaciones correctas.

PROBLEMA DEL CAPÍTULO 2

EJEMPLO Niveles de cotinina de fumadores Remítase a los 40 niveles de cotinina de fumadores en la tabla 2-1 (sin el error de 11111 utilizado en lugar del 1 en el ejemplo anterior).

a. Obtenga los valores que constituyen el resumen de los cinco números.

b. Construya una gráfica de cuadro.

SOLUCIÓN

a. El resumen de los cinco números consta del valor mínimo, Q_1, la mediana, Q_3, y el valor máximo. Para obtener dichos valores, primero ordene los datos (acomódelos en orden del más bajo al más alto). El mínimo de 0 y el máximo de 491 son fáciles de identificar en la lista ordenada. Ahora, proceda a calcular

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 algunas permiten que los estudiantes se gradúen sin entrenamiento alguno en números. ¿Cómo pueden estos reporteros escribir con sensibilidad sobre 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". Este escritor 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 cubrir el déficit presupuestal de 2.7 mil millones de dólares; en el artículo nunca se menciona el tamaño *total* del presupuesto, de modo que la cifra de 2.7 mil millones de dólares está fuera de contexto.

los cuartiles. Si usamos el diagrama de flujo de la figura 2-15, obtendremos $Q_1 = P_{25} = 86.5$, que se sitúa al calcular el localizador $L = (25/100)40 = 10$, y al encontrar el valor que está a la mitad entre el 10o y el 11o valores en la lista ordenada. La mediana es 170, que es el valor que está a la mitad entre los valores 20o y 21o. También encontramos que $Q_3 = 251.5$, al utilizar la figura 2.15 para el percentil 75o. Por lo tanto, el resumen de los cinco números es 0, 86.5, 170, 251.5 y 491.

b. En la figura 2-16 creamos la gráfica de cuadro para los datos. Usamos el valor mínimo (0) y el valor máximo (491) para determinar la escala de valores; después, graficamos los valores del resumen de los cinco números, como se indica a continuación.

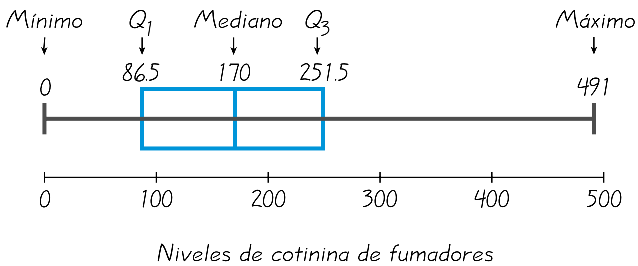


FIGURA 2-16

Gráfica de cuadro

En la figura 2-17, presentamos algunas gráficas de cuadro genéricas, junto con formas comunes de distribución. Parece ser que los niveles de cotinina de fumadores tienen una distribución sesgada.

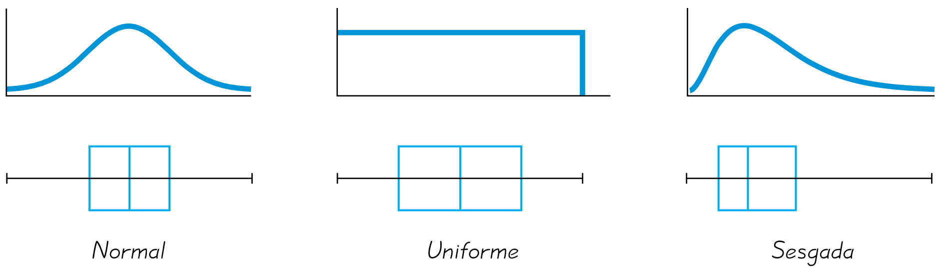
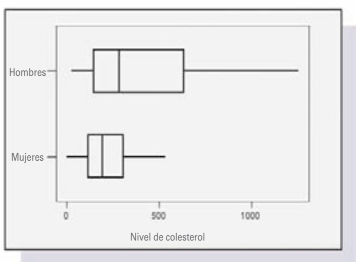


FIGURA 2-17 Gráficas de cuadro que corresponden a distribuciones normales, uniformes y sesgadas

Para ilustrar el uso de gráficas de cuadro que permiten comparar conjuntos de datos, véase la representación visual de Minitab de los niveles de colesterol para una muestra de hombres y una muestra de mujeres, con base en datos del National Health Examination, que se incluyen en el conjunto de datos 1 del Apéndice B. De acuerdo con el conjunto de datos, parece que los hombres tienen niveles de colesterol generalmente más altos que las mujeres, y que los niveles de colesterol de los hombres varían más que los de las mujeres.



LA ESTADÍSTICA EN LAS NOTICIAS

“Mejores” universidades

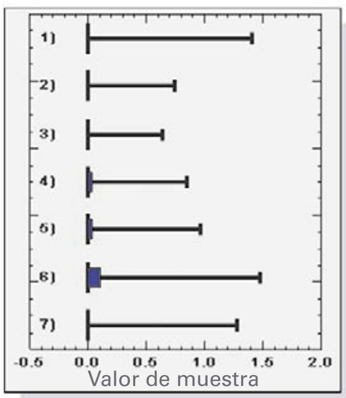
Cada año, el *U.S. News and World Report* publica un número con una lista de “las mejores universidades de Estados Unidos”. Generalmente las ventas de ese ejemplar aumentan hasta un 40%. Existen críticos de la lista que argumentan en contra de los criterios y el método de recolección de datos. Las quejas más comunes son: que se da demasiada importancia a los criterios de riqueza, la reputación, las calificaciones del consejo universitario, las donaciones de los alumnos y las opiniones de los presidentes universitarios; que se da poca importancia a la satisfacción de los estudiantes y a las prácticas educativas efectivas. El *New York Times* entrevistó a Kenneth Auchincloss, que es editor de la obra *How to Get into College* (de Kaplan/Newsweek), quien respondió que “nunca nos hemos sentido cómodos tratando de cuantificar en términos numéricos los diversos criterios empleados al calificar a una universidad como buena o menos buena, y no queremos dedicar los recursos a realizar un análisis estadístico elaborado que, con franqueza, no pensamos que sea válido”.

EJEMPLO ¿Llueve más durante los fines de semana? Remítase al conjunto de datos 11 del Apéndice B, que incluye una lista de las cantidades de lluvia (en pulgadas) que cayeron en Boston todos los días de un año reciente. La reunión de este conjunto de datos se inspiró con reportes de los medios de comunicación acerca de que llueve más durante los fines de semana (sábado y domingo) que entre semana. Más adelante, en este libro, describiremos métodos estadísticos importantes que permitan probar, de manera formal, dicha aseveración; por ahora, exploremos el conjunto de datos, para ver qué puede aprenderse. (Aun cuando sepamos aplicar estos métodos estadísticos formales, primero habrá que explorar los datos, antes de proceder con el análisis formal.)

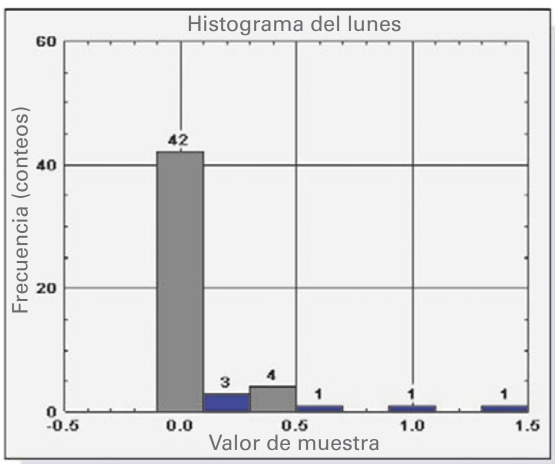
SOLUCIÓN Comencemos investigando los elementos clave del centro, la variación, la distribución, los datos distantes y las características en el tiempo (la misma lista “CVDDT” que se introdujo en la sección 2-1). Abajo se presentan medidas de tendencia central (media), medidas de variación (desviación estándar) y el resumen de los cinco números para las cantidades de lluvia que caen cada día de la semana. La representación visual del STATDISK muestra gráficas de cuadro de cada uno de los siete días de la semana, iniciando con el lunes en la parte superior. Debido a que los histogramas de los siete días son muy similares, únicamente mostramos el histograma de las cantidades de lluvia del lunes.

	Media	Desviación estándar	Mínimo	Q_1	Mediana	Q_3	Máximo
Lunes	0.100	0.263	0.000	0.000	0.000	0.010	1.410
Martes	0.058	0.157	0.000	0.000	0.000	0.015	0.740
Miércoles	0.051	0.135	0.000	0.000	0.000	0.010	0.640
Jueves	0.069	0.167	0.000	0.000	0.000	0.040	0.850
Viernes	0.095	0.228	0.000	0.000	0.000	0.040	0.960
Sábado	0.143	0.290	0.000	0.000	0.000	0.100	1.480
Domingo	0.068	0.200	0.000	0.000	0.000	0.010	1.280

STATDISK



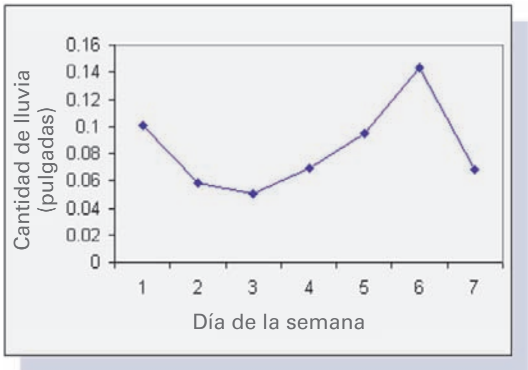
STATDISK



INTERPRETACIÓN Al examinar y comparar los estadísticos y las gráficas, hicimos las siguientes observaciones importantes:

- *Medias*: Las medias varían desde un mínimo de 0.051 pulgadas hasta un máximo de 0.143 pulgadas. Las siete medias varían en cantidades considerables. En capítulos siguientes presentaremos métodos para determinar si tales diferencias son *significativas*. (Métodos posteriores mostrarán que las medias no difieren en cantidades significativas). Si colocamos las medias en orden de menor a mayor, obtendremos la siguiente secuencia de días: miércoles, martes, domingo, jueves, viernes, lunes, sábado. No parece haber un patrón de mayor cantidad de lluvia durante los fines de semana (aunque la media más alta corresponde al sábado). Además, observe la gráfica de Excel de las siete medias, en donde la media del lunes se graficó primero. La gráfica de Excel no apoya la aseveración de mayor cantidad de lluvia durante los fines de semana (aunque podría argumentarse que llueve más los sábados).

Excel



- *Variación*: Las siete desviaciones estándar varían de 0.135 pulgadas a 0.290 pulgadas, pero estos valores no son muy diferentes. No parece haber algo infrecuente en las cantidades de variación.
- Los *mínimos*, *primeros cuartiles* y *medianas* son todos iguales a 0.00 para cada uno de los siete días. Lo anterior se explica por el hecho de que por cada día de la semana hay muchos días en los que no llueve. La abundancia de ceros también se observa en las gráficas de cuadro y en los histogramas, los cuales muestran que los datos tienen distribuciones cargadas hacia el extremo de los mínimos (sesgo derecho).
- *Datos distantes*: No aparecen datos distantes o valores inusuales. En el extremo de los mínimos hay muchas cantidades de lluvia iguales a cero. En el extremo de los máximos, la lista en que se ordenan las 365 cantidades de lluvia termina con los valores máximos de 0.92, 0.96, 1.28, 1.41 y 1.48.
- *Distribuciones*: Las distribuciones de las cantidades de lluvia están sesgadas hacia la derecha. No son normales, como esperaríamos. Si el uso de un método particular de estadística requiere poblaciones distribuidas normalmente (en forma de campana), este requisito no se satisface en las cantidades de lluvia.

Ahora comprendemos en gran medida la naturaleza de las cantidades de lluvia que caen en Boston durante distintos días de la semana. Con base en nuestra exploración, concluimos que en Boston no cae más lluvia durante los fines de semana que los demás días (aunque podríamos argumentar que llueve más los sábados).

Pensamiento crítico

Si nos armamos con una lista de herramientas para investigar el centro, la variación, la distribución, los datos distantes y las características de los datos a través del tiempo, tendríamos la tentación de desarrollar un procedimiento descuidado, por lo que el pensamiento crítico es sumamente importante. Además de utilizar las herramientas que se presentan en este capítulo, deberemos considerar cualesquiera otros factores que puedan ser cruciales para las conclusiones que elaboremos. En tal caso, plantearíamos preguntas como las siguientes: ¿Es posible que la muestra sea representativa de la población o está sesgada de alguna manera? ¿Cuál es la fuente de los datos? ¿Sería posible que la fuente fuera alguien con intereses que puedan afectar la calidad de los datos? Suponga, por ejemplo, que deseamos estimar el ingreso medio de estudiantes universitarios. También, suponga que enviamos por correo cuestionarios a 500 estudiantes y que recibimos 20 respuestas. Podríamos calcular la media y la desviación estándar, así como construir gráficas, identificar datos distantes, etcétera, pero los resultados serán lo que los estadísticos llaman desperdicios. La muestra es de respuesta voluntaria, por lo que no tiene posibilidades de ser representativa de la población de todos los estudiantes universitarios. Además de las herramientas estadísticas específicas presentadas en este capítulo, ¡también debemos *pensar*!

Utilizando la tecnología

Esta sección introdujo los datos distantes, los resúmenes de los cinco números y las gráficas de cuadro. Para encontrar datos distantes, se acomodan los datos en orden de menor a mayor; después, se examinan los valores máximo y mínimo para determinar si están muy lejos de los otros valores muestrales. El STATDISK, Minitab, Excel y la calculadora TI-83 Plus proporcionan valores de cuartiles, de modo que es fácil elaborar el resumen de los cinco números. El STATDISK, Minitab, Excel y la calculadora TI-83 Plus pueden utilizarse para crear gráficas de cuadro. Ahora describiremos los distintos procedimientos. (*Precaución*: Recuerde que los valores cuartilares calculados por medio de Minitab y la calculadora TI-83 Plus pueden diferir ligeramente de los calculados a partir de la figura 2-15, por lo que tal vez las gráficas de cuadro también difieran ligeramente).

STATDISK Elija el elemento **Data** del menú principal, y utilice el **Sample Editor** para introducir los datos; después, haga clic en **COPY**. Ahora seleccione **Data**, luego **Boxplot,** y haga clic en **PASTE** y en **Evaluate**.

Minitab Introduzca los datos en la columna C1; luego, seleccione **Graph** y **Boxplot**. Introduzca C1 en la primera celda, debajo de la columna Y; luego, haga clic en **OK**.

Excel Aunque Excel no se diseñó para generar gráficas de cuadro, éstas pueden crearse utilizando el Data Desk XL add-in, que complementa este libro. Primero introduzca los datos en la columna A. Haga clic en **DDXL** y seleccione **Charts y Plots.** Estando en la función Type, elija la opción de **Boxplot**. En el cuadro de diálogo, haga clic en el icono del lápiz e introduzca el rango de datos, como A1:A40, si usted tiene 40 valores listados en la columna A. Haga clic en **OK**. El resultado es una gráfica de cuadro modificada, tal como se describe en el ejercicio 13. También se muestran los valores del resumen de los cinco números.

TI-83 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 **ON** y elija el tipo de gráfica de cuadro que se ubica a la mitad el segundo renglón. Xlist debe indicar L1 y el valor Freq tiene que ser 1. Ahora presione la tecla **ZOOM** y elija la opción 9 para **ZoomStat**. Presione la tecla **ENTER**; debe aparecer la gráfica de cuadro. Puede utilizar las teclas con flechas para moverse hacia la derecha o hacia la izquierda, de manera que le sea posible leer los valores desde la escala horizontal.

2-7 Destrezas y conceptos básicos

1. Lotería Remítase al conjunto de datos 26 y utilice sólo los 40 dígitos en la primera columna de los resultados *Win 4* de la lotería del estado de Nueva York (9, 7, 0, etcétera). Encuentre el resumen de los cinco números y construya una gráfica de cuadro. ¿Qué características de la gráfica de cuadro sugieren que los dígitos fueron seleccionados con un procedimiento aleatorio y justo?

2. Presupuestos de películas Remítase al conjunto de datos 21 del Apéndice B, con los montos de presupuesto de las 15 películas con clasificación R. Elabore el resumen de los cinco números y construya una gráfica de cuadro. Determine si los valores muestrales son representativos de las películas realizadas este año.

3. Calorías de cereales Remítase al conjunto de datos 16 del Apéndice B de los 16 valores de las calorías por gramo de cereales. Elabore el resumen de los cinco números y construya una gráfica de cuadro. Determine si los valores muestrales serían representativos de los cereales consumidos por la población en general.

4. Nicotina en cigarrillos Remítase al conjunto de datos 5 de las 29 cantidades de nicotina (en miligramos por cigarrillo). Elabore el resumen de los cinco números y construya una gráfica de cuadro. ¿Podrían los valores muestrales ser representativos de los cigarrillos fumados por un consumidor individual?

5. M&M rojos Remítase al conjunto de datos 9 de los 21 pesos (en gramos) de los dulces M&M rojos. Elabore el resumen de los cinco números y construya una gráfica de cuadro. ¿Podrían los valores muestrales ser representativos de los dulces M&M de todos los colores?

T **6.** Longitudes de osos Remítase al conjunto de datos 9 de las longitudes (en pulgadas) de los 54 osos que anestesiamos y medimos. Elabore el resumen de los cinco números y construya una gráfica de cuadro. ¿Podría la distribución de longitudes ser simétrica? ¿O está sesgada?

T **7.** Alcohol en películas infantiles Remítase al conjunto de datos 7 de las 50 duraciones (en segundos) de escenas que presentan consumo de alcohol en películas infantiles de dibujos animados. Elabore el resumen de los cinco números y construya una gráfica de cuadro. Con base en la gráfica de cuadro, ¿la distribución parece simétrica o está sesgada?

T **8.** Temperaturas corporales Remítase al conjunto de datos 4 del Apéndice B de las 106 temperaturas corporales a las 12 a. m. del día 2. Elabore el resumen de los cinco números y construya una gráfica de cuadro; después, determine si los valores muestrales apoyan la creencia común de que la temperatura corporal media es de 98.6°F.

En los ejercicios 9 a 12 elabore los resúmenes de los cinco números, construya gráficas de cuadro y compare los conjuntos de datos.

9. Premios Óscar En el artículo “Ages of Oscar Winning Best Actors and Actresses” (revista *Mathematics Teacher*), de Richard Brown y Gretchen Davis, los autores comparan las edades de actores y actrices en el momento de ganar el Óscar. En la siguiente tabla se presentan los resultados de los ganadores de ambas categorías. Utilice gráficas de cuadro para comparar los dos conjuntos de datos.

Actores:	32	37	36	32	51	53	33	61	35	45	55
	39	76	37	42	40	32	60	38	56	48	48
	40	43	62	43	42	44	41	56	39	46	31
	47	45	60	46	40	36					
Actrices:	50	44	35	80	26	28	41	21	61	38	49
	33	74	30	33	41	31	35	41	42	37	26
	34	34	35	26	61	60	34	24	30	37	31
	27	39	34	26	25	33					

T **10.** Coca Cola regular/Coca Cola dietética Remítase al conjunto de datos 17 del Apéndice B; utilice los pesos de la Coca Cola regular y los pesos de la Coca Cola dietética. ¿Parece haber una diferencia significativa? Si es así, ¿qué explicación le encuentra?

11. Niveles de cotinina Remítase a la tabla 2-1 ubicada en el problema del capítulo. Ya calculamos que el resumen de los cinco números para los niveles de cotinina de fumadores es 0, 86.5, 170, 251.5 y 491. Elabore los resúmenes de los cinco números para los otros dos grupos; después, construya las tres gráficas de cuadro utilizando la misma escala. ¿Existe alguna diferencia aparente?

12. Clancy, Rowling, Tolstoi Remítase al conjunto de datos 14 del Apéndice B y utilice las puntuaciones de la escala de facilidad de lectura de Flesch para las páginas muestra de las obras *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. (Las puntuaciones más altas indican una lectura más fácil). ¿Parece haber alguna diferencia en la facilidad de lectura? ¿Son los resultados consistentes con sus expectativas?

2-7 Más allá de lo básico

13. Las gráficas de cuadro introducidas en esta sección suelen denominarse gráficas de cuadro *de esqueleto* (o *regulares*). Las **gráficas de cuadro modificadas** se construyen de la siguiente forma:

a. Calcule el *RIC*, que denota el rango intercuartilar, definido por $RIC = Q_3 - Q_1$.

b. Dibuje el cuadro con la mediana y los cuartiles como siempre; pero, cuando trace las líneas a la derecha e izquierda del cuadro, dibújelas sólo hasta los puntos que corresponden a los valores máximo y mínimo, que están dentro de 1.5 *RIC* del cuadro.

c. Los **datos ligeramente distantes**, que se grafican como puntos *sólidos*, son valores que están por debajo de Q_1 o por arriba de Q_3, por una cantidad mayor que 1.5 *RIC*, pero no mayor que 3 *RIC*. Es decir, los datos ligeramente distantes son valores *x*, tales que

$$Q_1 - 3\,RIC \leq x < Q_1 - 1.5\,RIC$$

o

$$Q_3 + 1.5\,RIC < x \leq Q_3 + 3\,RIC$$

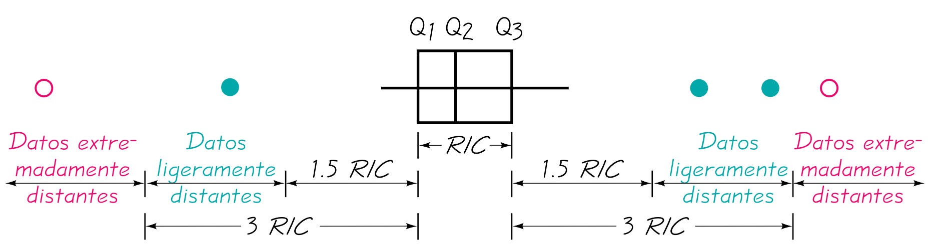
d. Los **datos extremadamente distantes**, que se grafican como pequeños círculos *vacíos*, son valores que están por debajo de Q_1 por más de 3 *RIC* o por encima de Q_3 por más de 3 *RIC*. Es decir, los datos extremadamente distantes son valores *x*, tales que

$$x < Q_1 - 3\,RIC$$

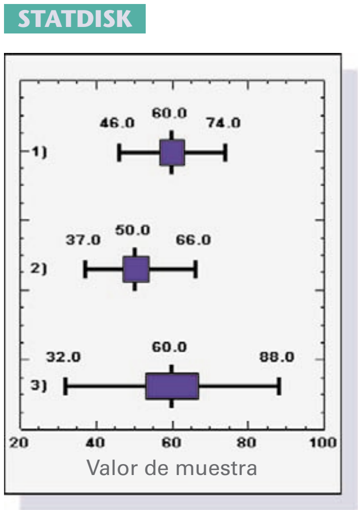
o

$$x > Q_3 + 3\,RIC$$

La figura anexa es un ejemplo de una gráfica de cuadro modificada. Remítase a los niveles de cotinina de fumadores de la tabla 2-1, incluida en el problema del capítulo. Ya sabemos que este conjunto de datos tiene un resumen de los cinco números de 0, 86.5, 170, 251.5 y 491. Identifique el valor de *RIC*, identifique los rangos de valores utilizados para localizar sus datos ligeramente distantes y extremadamente distantes; después, identifique cualesquiera datos ligeramente distantes y extremadamente distantes reales.



14. Remítase a la representación visual de STATDISK, de las tres gráficas de cuadro que representan la medida de longevidad (en meses) de muestras de tres distintas baterías para automóvil. Si usted es el encargado de una flotilla de automóviles y debe seleccionar una de las tres marcas, ¿cuál gráfica de cuadro representa la marca que debería elegir? ¿Por qué?



Repaso

En este capítulo consideramos métodos para describir, explorar y comparar conjuntos de datos. Cuando se investigan conjuntos de datos, las siguientes características son, por lo general, 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. *Datos distantes*: Valores muestrales que se ubican muy lejos de la mayoría de valores muestrales.
5. *Tiempo*: Características cambiantes de los datos a través del tiempo.

Después de completar este capítulo, usted será capaz de hacer lo siguiente:

- Resumir datos por medio de la construcción de una distribución de frecuencias o de una distribución de frecuencias relativas (sección 2-2).
- Representar visualmente la naturaleza de la distribución por medio de la construcción de un histograma, una gráfica de puntos, una gráfica de tallo y hojas, una gráfica circular o una gráfica de Pareto (sección 2-3).
- Calcular medidas de tendencia central como la media, la mediana, la moda y la mitad del rango (sección 2-4).
- Calcular medidas de variación como la desviación estándar, la varianza y el rango (sección 2-5).
- Comparar valores individuales utilizando puntuaciones *z*, cuartiles o percentiles (sección 2-6).
- 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 2-7).

Además de crear dichas tablas, gráficas y medidas, usted será capaz de comprender e interpretar los resultados. Por ejemplo, debe entender con claridad que la desviación estándar es una medida acerca de qué tanto varían los datos, y saber utilizar la desviación estándar para distinguir entre valores frecuentes e infrecuentes.

Ejercicios de repaso

1. Edades de presidentes La senadora Hayes está considerando competir por la presidencia de Estados Unidos, pero sólo tiene 35 años de edad, que es la edad mínima requerida. Al investigar este tema, descubre las edades de presidentes anteriores en el momento de tomar el cargo; dichas edades se listan abajo. Utilice las edades 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; y *j*) P_{10}.

57	61	57	57	58	57	61	54	68	51	49	64	50	48
65	52	56	46	54	49	51	47	55	55	54	42	51	56
55	51	54	51	60	62	43	55	56	61	52	69	64	46
54													

2. **a.** John F. Kennedy tenía 43 años de edad cuando tomó posesión. Utilice los resultados del ejercicio 1 y convierta esta edad a puntuación *z*.
b. ¿Será la edad de 43 años de Kennedy "infrecuente"? ¿Por qué?
c. Aplique la regla práctica del intervalo para identificar otras edades de la lista que sean infrecuentes.
d. Aun cuando la lista de edades no incluye la de 35 años, ¿sería esa edad infrecuente? ¿Es probable que un candidato a la presidencia, de 35 años de edad, descubra que su edad sea un tema importante de campaña?

3. Distribución de frecuencias Utilice la misma lista de edades del ejercicio 1 y construya una distribución de frecuencias. Use seis clases, con 40 como el límite inferior de la primera clase, y una anchura de clase de 5.

4. Histograma Utilice la distribución de frecuencias del ejercicio 3, construya un histograma e identifique la naturaleza general de la distribución (ya sea uniforme, normal, sesgada).

5. Gráfica de cuadro Utilice las mismas edades de la lista en el problema 1, construya una gráfica de cuadro e identifique los valores que constituyen el resumen de los cinco números.

6. Regla empírica Suponga que las edades de presidentes pasados, presentes y futuros tienen una distribución normal, con una media de 54.8 años y una desviación estándar de 6.2 años.
a. ¿Qué dice la regla empírica acerca del porcentaje de edades entre 48.6 años y 61.0 años (o dentro de una desviación estándar de la media)?
b. ¿Qué dice la regla empírica acerca del porcentaje de edades entre 42.4 y 67.2 años?

7. Comparación de puntuaciones Un psicólogo industrial de la Citation Corporation crea dos pruebas diferentes para medir la satisfacción laboral. ¿Cuál puntuación es mejor: una de 72 en la prueba de administración, la cual tiene una media de 80 y una desviación estándar de 12, o una de 19 en la prueba de producción de empleados, con una media de 20 y una desviación estándar de 5? Explique.

8. **a.** Estime la media de la edad de los automóviles que conducen los estudiantes de su universidad.
b. Utilice la regla práctica del intervalo para hacer un estimado de la desviación estándar de la edad de los automóviles que conducen los estudiantes de su universidad.

9. Transformación de datos Una profesora de estadística encontró que el tiempo que emplean los estudiantes para resolver su examen final tiene una media de 135 minutos y una desviación estándar de 15 minutos. Planea añadir una nueva pregunta, la cual requerirá cinco minutos adicionales de cada estudiante.
   a. ¿Cuál es la media después de incluir la nueva pregunta?
   b. ¿Cuál es la desviación estándar después de incluir la nueva pregunta?
   c. ¿Cuál es la varianza después de incluir la nueva pregunta?

10. Quejas de pasajeros aéreos En un año reciente hubo 23,000 quejas por parte de pasajeros aéreos. Las categorías y frecuencias de las quejas, proporcionadas por el Departamento de Transporte de Estados Unidos, son las siguientes: atención al cliente (4370); problemas con el vuelo (9200); reservaciones, boletaje y abordaje (1610); equipaje (3450); obtención de reembolsos (1150), y otras razones (3220). Construya una gráfica de Pareto que resuma los datos.

Ejercicios de repaso acumulativos

1. Errores del reloj de pulso Como parte de un proyecto para la clase de estadística, un estudiante reúne datos sobre la precisión de relojes de pulso; obtiene los siguientes errores del tiempo (en segundos). (Los valores positivos representan relojes que se adelantan; los valores negativos representan relojes que se atrasan).

   140 −125 105 −241 −85 41 186 −151 325 80 27 20 20 30 −65

   a. Calcule la media, la mediana, la moda y la mitad del rango.
   b. Calcule la desviación estándar, la varianza y el rango.
   c. ¿Provendrán los tiempos dados de una población discreta o de una continua?
   d. ¿Cuál es el nivel de medición de estos valores (nominal, ordinario, de intervalo, de razón)?

2. **a.** Un conjunto de datos tiene un nivel nominal de medición y usted desea obtener un dato representativo. ¿Cuál de los siguientes es el más apropiado: la media, la mediana, la moda o la mitad del rango? ¿Por qué?
   b. Se obtiene una muestra al llamar por teléfono a las primeras 250 personas listadas en el directorio telefónico local. ¿Qué tipo de muestreo se utilizó (aleatorio, estratificado, sistemático, de racimo, de conveniencia)?
   c. Se lleva a cabo una encuesta de salida, consistente en encuestar a cada persona que sale de la casilla electoral en 50 distritos seleccionados aleatoriamente. ¿Qué tipo de muestreo se utilizó (aleatorio, estratificado, sistemático, de racimo, de conveniencia)?
   d. Un fabricante recarga cartuchos para impresoras de computadora. Un gerente descubre que la cantidad de tinta vertida en el recipiente no es muy consistente, de tal forma que algunos cartuchos duran más de lo esperado, en tanto que otros se agotan demasiado pronto. Él desea mejorar la calidad haciendo que la cantidad de tinta en los cartuchos tenga más consistencia. Cuando se analizan las cantidades de tinta, ¿cuál de los siguientes estadísticos es más importante: la media, la mediana, la moda, la mitad del rango, la desviación estándar, el primer cuartil, el tercer cuartil? ¿Debería elevarse, disminuirse o dejar sin cambio el valor de dicho estadístico?

3. Consumo de energía Cada año, el Departamento de Energía de Estados Unidos publica un *Annual Energy Review*, que incluye el consumo de energía *per capita* (en millones de BTU) de cada uno de sus 50 estados. Si se calcula la media de estos 50 valores, ¿el resultado será la media de consumo de energía *per capita* de la población de los 50 estados combinados? Si no es así, explique cómo calcularía la media del consumo de energía *per capita* de los 50 estados combinados.

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*, 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 pueden incluirse en los avalúos de bienes raíces, así como en las estimaciones del valor de un automóvil y en las de la posibilidad de una guerra nuclear.

Realice un experimento para probar dicha teoría. Seleccione a algunos sujetos y pídales que estimen rápidamente el valor de

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

Después, seleccione otros sujetos y pídales 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 modo que las condiciones sean uniformes y que los dos grupos muestrales se seleccionen de forma que se minimice cualquier sesgo. No describa la teoría a los sujetos hasta después de proporcionar 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 detallados utilizados, el método de análisis, las gráficas y/o los estadísticos relevantes, así como las conclusiones. Incluya una crítica dando razones por las que los resultados podrían ser incorrectos y describa formas para mejorar el experimento.

Una variante del experimento anterior consiste en encuestar personas acerca de su conocimiento sobre la población de Kenia. Primero pregúntele a la mitad de los sujetos si piensa que la población es mayor o menor que cinco millones; después, pídales que estimen la población dando un número real. Pregunte a la otra mitad de los sujetos si cree que la población es mayor o menor que 80 millones; después, pídales que estimen la población. (La población de Kenia es de 28 millones). Compare los dos conjuntos de resultados e identifique el efecto “de anclaje” en la cifra inicial que dieron los sujetos encuestados.

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. Proporcione lo siguiente: *a*) una lista de valores muestrales, *b*) resultados de computadora impresos con estadísticos descriptivos y gráficas, y *c*) una descripción por escrito de la naturaleza de los datos, el método de recolección y las características importantes.

3. Actividad en clase A continuación se indican las edades que un grupo de motociclistas tenía cuando se hirió fatalmente en accidentes de tránsito (según datos del Departamento del Transporte de Estados Unidos). Si su objetivo es dramatizar el peligro que constituyen las motocicletas para la gente joven, ¿cuál de los siguientes sería el más efectivo: el histograma, la gráfica de Pareto, la gráfica circular, la gráfica de puntos, la media, la mediana? Construya la gráfica y encuentre el estadístico que cumple mejor el objetivo. ¿Es correcto distorsionar los datos de manera deliberada si el objetivo es 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	

4. Actividad fuera de clase Pida a cada equipo, que se formó con tres o cuatro estudiantes, que seleccione uno de los siguientes reactivos y que construya una gráfica que sirva para poner énfasis en la pregunta:

a. ¿Habrá una diferencia entre los valores del índice de masa corporal (IMC) de los hombres y el de las mujeres? (Véase el conjunto de datos 1 del Apéndice B).

b. ¿Existirá alguna relación entre la estatura de los hijos (o hijas) y la estatura de sus padres (o madres)? (Véase el conjunto de datos 2 del Apéndice B).

c. ¿Parecen los dígitos generados por la lotería *Win 4* del estado de Nueva York haber sido seleccionados aleatoriamente? ¿O estarán sesgados? (Véase el conjunto de datos 26 del Apéndice B).

Proyecto tecnológico

Cuando se manejan conjuntos grandes de datos, la introducción manual de éstos suele ser tediosa y requerir de mucho tiempo. Hay mejores actividades que hacer con su tiempo, como aprender los principios aerodinámicos de un *frisbee*. Remítase al conjunto de datos 30 del Apéndice B, que incluye las distancias de los jonrones “”de tres jugadores de beisbol excepcionales: Barry Bonds (temporada 2001), Mark McGwire (temporada 1998) y Sammy Sosa (temporada 1998). En lugar de introducir manualmente las 209 distancias de los tres conjuntos de datos, utilice la calculadora TI-83 Plus, el STATDISK, el Minitab o Excel para cargar el conjunto de datos, los cuales están disponibles en el CD incluido en este libro. Proceda a generar histogramas, obtendrá estadísticos apropiados que le permitirán comparar los tres conjuntos de datos. ¿Hay algunas diferencias significativas? ¿Existen algunos datos distantes? ¿Parece que los jugadores que golpean más lejos hacen más jonrones””? ¿Por qué? Analice los últimos dígitos de las distancias y determine si los valores parecen ser estimaciones o mediciones. Escriba un reporte breve que contenga sus conclusiones y gráficas de apoyo.

Conjunto de datos 30: Distancias de jonrones

Las distancias de los jonrones de Mark McGwire (1998), Sammy Sosa (1998) y Barry Bonds (2001) están en pies.

Nombre de los archivos de STATDISK y de texto: MCGWR, SOSA, BONDS.
Minitab: el nombre de la hoja de cálculo es HOMERUNS.MTW.
Excel: el nombre del libro es HOMERUNS.XLS.
TI-83 Plus: el nombre es HOMERUNS, y los nombres de los archivos son los mismos de STATDISK y de los archivos de texto.

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							

de los DATOS a la DECISIÓN

Pensamiento crítico

Las muertes por choques de automóviles son devastadoras para las familias de las víctimas y con frecuencia implican procesos legales y pagos de seguro costosos. A continuación se presentan las edades de 100 conductores que murieron en choques de automóviles, seleccionados aleatoriamente. También se incluye una distribución de frecuencias, por edades, de conductores con licencia.

Edad (en años) de conductores muertos en choques de automóviles

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

Edad	Conductores con licencia (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

Análisis

Convierta la distribución de frecuencias a una distribución de frecuencias relativas; después, elabore una distribución de frecuencias relativas con las edades de los conductores que murieron en choques de automóviles. Compare las dos distribuciones de frecuencias relativas. ¿Cuáles categorías de edad parecen tener proporciones sustancialmente mayores de muertes que las proporciones de los conductores con licencia? Si usted fuese el responsable de establecer las tasas de seguros de automóviles, ¿a qué categorías de edad les asignaría las tasas más altas? Construya una gráfica que permita identificar las categorías de edad más propensas a accidentes automovilísticos fatales.

PROYECTO DE INTERNET

Datos en Internet

Internet posee una enorme cantidad de información, de la cual, gran parte, se presenta en forma de datos brutos que pueden estudiarse y resumirse con el uso de los estadísticos presentados en este capítulo. Por ejemplo, encontramos la siguiente información con tan sólo unos cuantos clics:

- El valor de la acción de Walt Disney Corporation tiende a subir durante los meses de invierno, pero varía el resto del año. En el 2001, los precios máximo y mínimo de la acción variaron más de 17 puntos.
- Ichiro Suzuki, jardinero lateral del equipo de beisbol de los Marineros de Seattle, tuvo un porcentaje de “hits” de 0.457 durante la temporada 2001.
- Las poblaciones de California, Nueva York y Texas representan más del 25% del total de la de Estados Unidos.

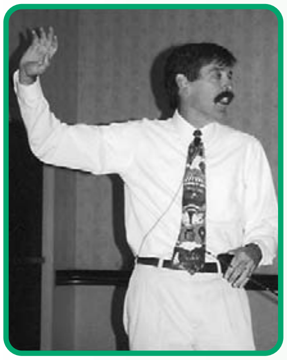
El proyecto de Internet para este capítulo, que se encuentra en el sitio de Internet de *Estadística elemental*, lo conducirá a conjuntos de datos en las áreas de deportes, finanzas y clima. Una vez que haya armado un conjunto de datos, aplicará los métodos de este capítulo para resumir y clasificar los datos.

El sitio Web para este capítulo se encuentra en

www.pearsoneducacion.net/triola

La estadística @ en el trabajo

"El periodista debe ser capaz de tener una visión crítica hacia la investigación, y comprender el verdadero contexto y el significado del trabajo".



Mark Fenton

Editor de la Revista Large Walking Magazine

Mark Fenton también es defensor de la marcha y campeón de este deporte. Perteneció al equipo nacional de marcha de Estados Unidos en cinco ocasiones, y ha representado a este país en numerosas competencias internacionales. Estudió biomecánica y fisiología del deporte en el Olympic Training Center's Sports Science Laboratory, ubicado en Colorado Springs, Colorado.

¿Cuáles conceptos de estadística utiliza?

Debo estar familiarizado con todas las herramientas comunes que se utilizan en los análisis estadísticos de las ciencias del deporte y de la investigación en salud pública; desde las medias y las desviaciones estándar, hasta la significancia o diferencias estadísticas, los intervalos de confianza y el análisis de varianza, entre otros.

¿De qué manera utiliza la estadística en su trabajo?

Suelo leer revistas de investigaciones médicas (*Medicine and Science in Sports and Exercise* y el *Journal of the American Medical Association* son los más relevantes) sobre fisiología del deporte, salud pública e investigación epidemiológica. Debo mantenerme informado sobre la manera en que se controlan y analizan los estudios, así como comprender la potencia o el valor relativo del resultado de un estudio. Esta información la utilizo en los artículos y conferencias que elaboro sobre los hallazgos de dicho trabajo.

Por favor, describa un ejemplo específico que ilustre la forma en que el uso de la estadística haya logrado mejorar un producto o servicio.

Por lo general, leo artículos de investigación y debo estar atento a los vacíos de la evidencia, tales como tamaños de muestra demasiado pequeños; significancia estadística con diferencias absolutas pequeñas entre grupos; desviaciones estándar extremadamente grandes; todo lo cual puede oscurecer la relevancia de un hallazgo para una persona promedio.

Un ejemplo simple es el problema que existe al afirmar que el pulso máximo **promedio** de una mujer de 36 años es de 190 latidos por minuto (226 menos la edad). Esto es útil para calcular el pulso que se busca, hasta que uno encuentra que la desviación es demasiado alta y la cifra resultante puede alejarse tanto como 10 latidos por minuto para casi 1/3 de la población, un error suficientemente grande para causar problemas al hacer ejercicio. El estudio que señaló esto alteró la forma en que recomiendo la intensidad del ejercicio.

¿Se estará incrementando o disminuyendo el uso de la probabilidad y la estadística, o permanece estable?

Debido a mi trabajo cada vez más intenso en salud pública y al incremento en la complejidad de las herramientas estadísticas que se utilizan en los análisis de poblaciones grandes, mis conocimientos en este campo **deben** seguir acrecentándose si quiero mantenerme al día.

¿Recomienda el uso de la estadística a los estudiantes universitarios de hoy?

Recomiendo al menos un curso de estadística, ya que es una herramienta útil para evaluar la información con que se nos bombardea cada día, mucha de la cual carece de contexto. Cualquier persona que se interesa en la ciencia del periodismo debe tomar al menos un curso introductorio extenso de estadística. El periodista debe ser capaz de tener una visión crítica hacia la investigación, y comprender el verdadero contexto y el significado del trabajo.

3

Probabilidad

PROBLEMA DEL CAPÍTULO

Falsos positivos y falsos negativos

En diferentes etapas de nuestra vida, todos nos sometemos a una variedad de exámenes médicos. Algunos exámenes médicos son tan simples, como aquellos donde se usa un termómetro para establecer si la temperatura corporal es muy alta o muy baja, o como un esfigmomanómetro con el que se determina si la presión sanguínea es muy alta o muy baja. Otros exámenes clínicos incluyenn el análisis de muestras de sangre para identificar la presencia de alguna enfermedad. En este problema del capítulo consideramos los resultados obtenidos en un estudio clínico consistente en una prueba de embarazo. Para una mujer es importante saber si está embarazada con la finalidad de interrumpir prácticas que serían potencialmente dañinas para el bebé, como las actividades físicas, la medicación, la exposición a tóxicos en el trabajo, el tabaquismo o el consumo de alcohol. Las pruebas de embarazo, como casi todas las pruebas médicas, arrojan resultados que distan de ser 100% precisos. Los resultados mostrados en la tabla 3-1 se obtuvieron con la prueba de embarazo de Abbot, a partir de muestras de sangre (según datos de "Specificity and Detection Limit of Ten Pregnancy Tests", de Tiitinen y Stenman, *Scandinavian Journal of Clinical Laboratory Investigation*, vol. 53, suplemento 216). Existen factores, como el avance del embarazo, que afectan la precisión de dichos exámenes. Las pruebas de embarazo son, por lo regular, más confiables cuando se aplican al menos dos semanas después de la concepción. Otras pruebas ofrecen resultados más confiables que los de la tabla 3-1. Por ejemplo, la de Abbot Testpack Plus es una prueba de orina con una tasa de falso positivo de 0.2% y una tasa de falso negativo de 0.6%. Los términos *falso positivo* y *falso negativo* se incluyen entre los siguientes términos, que se usan comúnmente en las pruebas médicas o en procedimientos de vigilancia:

- *Falso positivo*: La prueba indica incorrectamente embarazo cuando la mujer no está embarazada.
- *Falso negativo*: La prueba indica incorrectamente que la mujer no está embarazada cuando en realidad lo está.
- *Verdadero positivo*: La prueba indica correctamente que la mujer está embarazada cuando en realidad lo está.
- *Verdadero negativo*: La prueba indica correctamente que la mujer no está embarazada cuando no lo está.
- *Sensibilidad de la prueba*: La probabilidad de un verdadero positivo.
- *Especificidad de la prueba*: La probabilidad de un verdadero negativo.

Con base en los resultados de la tabla 3-1, ¿cuál es la probabilidad de que una mujer esté embarazada si la prueba indica un resultado negativo? ¿Cuál es la probabilidad de un falso positivo? En este capítulo resolveremos estas preguntas.

Tabla 3-1 Resultados de prueba de embarazo

	Resultado de prueba positivo (indicó embarazo)	Resultado de prueba negativo (no indicó embarazo)
La mujer está embarazada	80	5
La mujer no está embarazada	3	11

3-1 Panorama general

La probabilidad es la base sobre la que se construyen los métodos importantes de la estadística inferencial. Como un sencillo ejemplo, suponga que usted hubiera ganado el premio mayor de la lotería nacional cinco veces seguidas. Habría acusaciones de que usted hizo trampa de alguna forma. Las personas saben que aun cuando existe la posibilidad de que alguien gane cinco veces consecutivas, por pura suerte, la posibilidad es tan increíblemente baja, que rechazarían la suerte como una explicación razonable. Ésta es precisamente la forma de pensar de los estadísticos: las personas rechazan las explicaciones basadas en probabilidades muy bajas. Los estadísticos usan la *regla del suceso infrecuente*.

Regla del suceso infrecuente para estadística inferencial

Si, bajo un supuesto dado (como un juego de lotería justo), la probabilidad de un suceso particular observado (como ganar cinco veces consecutivas) es extremadamente pequeña, concluimos que el supuesto probablemente es incorrecto.

El objetivo principal de este capítulo es desarrollar una comprensión válida de los valores de probabilidad, la cual se usará 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.

3-2 Fundamentos

Al considerar la probabilidad, tratamos con procedimientos (como tirar un dado, contestar una pregunta de opción múltiple en un examen o aplicar una prueba de embarazo) que producen resultados.

Definiciones

Suceso: 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 se forma con todos los resultados que ya no es posible desglosar más.

EJEMPLOS

Procedimiento	Ejemplo de suceso	Espacio muestral
Tirar un dado	5 (suceso simple)	{1, 2, 3, 4, 5, 6}
Tirar dos dados	7 (suceso no simple)	{1-1, 1-2, . . . , 6-6}

Cuando se tira un dado, el resultado 5 es un *suceso simple* porque no es posible desglosarlo en otros. Cuando se tiran dos dados, el resultado 7 *no es un suceso simple*, porque esto puede todavía desglosarse en eventos más simples, tales como 3-4 o 6-1. Cuando tiramos dos dados, el resultado de 3-4 se considera un suceso simple, porque no es posible desglosarlo más. Tal vez pensaríamos de forma incorrecta que 3-4 se desglosaría en los resultados individuales de 3 y 4, pero, cuando se tiran dos dados, 3 y 4 no son resultados individuales. Cuando se tiran dos dados, existen exactamente 36 resultados que son sucesos simples: 1-1, 1-2, . . . , 6-6.

Hay diferentes formas para definir la probabilidad de un suceso; 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 ocasiones que el suceso A ocurre en realidad. Con base en estos resultados reales, $P(A)$ se estima de la siguiente forma:

$$P(A) = \frac{\text{número de veces que ocurre } A}{\text{número de veces que se repitió el ensayo}}$$

Regla 2: Método clásico de la probabilidad (requiere resultados igualmente probables)

Suponga que un procedimiento dado tiene n sucesos simples distintos, cada uno de los cuales 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 obtiene simplemente suponiendo o estimando su valor con base en el conocimiento de las circunstancias relevantes.

LA ESTADÍSTICA EN LAS NOTICIAS

Asteroides asesinos

La probabilidad de que nuestra civilización sea destruida por un asteroide que choque con nuestro planeta tiene una importancia evidente para casi todos nosotros. En junio de 2002 un artículo del *New York Times* reportó que un asteroide "suficientemente grande como para arrasar una gran ciudad se acercó hasta 75,000 millas (unos 120,000 kilómetros) de la Tierra... pero no fue detectado hasta pasados tres días". Cuando se intenta determinar esa probabilidad, el método de las frecuencias relativas no se aplica, porque es imposible realizar ensayos y no hay datos históricos de una destrucción de ese tipo. El método clásico no se aplica porque los resultados posibles no tienen la misma probabilidad de ocurrir. Sólo puede aplicarse el método de la probabilidad subjetiva.

Con base en observaciones más actuales, los astrónomos estiman que hay 700,000 asteroides lo suficientemente grandes y cercanos como para destruirnos. Al usar este número y el conocimiento de las órbitas de los asteroides, los astrónomos desarrollaron la probabilidad subjetiva de que a nuestra civilización la destruya una colisión con un asteroide en algún momento en los próximos 100 años: la probabilidad es aproximadamente de 1/5000.

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 notó que una concepción en la noche de Año Nuevo podría resultar probablemente en 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.

(a) (b) (c)

FIGURA 3-1 Tres métodos para calcular la probabilidad

a) Método de las frecuencias relativas (regla 1). Cuando se trata de determinar: *P* (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.

$$P(2) = \frac{\text{número de formas en que 2 puede ocurrir}}{\text{número total de sucesos simples}}$$

$$= \frac{1}{6}$$

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.

Es muy importante notar 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 3-1 ilustra los tres métodos.

Al calcular probabilidades con el método de frecuencias relativas (regla 1), obtenemos un *estimado* en lugar de un valor exacto. Conforme el número total de observaciones se incrementa, los estimados correspondientes tienden a acercarse a la probabilidad real. Tal propiedad se enuncia en forma de teorema, al que se conoce comúnmente como la *ley de los grandes números*.

Ley de los grandes números

Conforme un procedimiento se repite una y otra vez, la probabilidad de frecuencias relativas (regla 1) de un suceso, tiende a aproximarse a la probabilidad real.

La ley de los grandes números 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: un estimado de probabilidad basado en sólo unos cuantos ensayos puede desviarse en cantidades sustanciales; pero, con un número muy grande de ensayos, el estimado tiende a ser mucho más preciso. Por ejemplo, es muy fácil que una encuesta de opinión entre sólo una docena de personas seleccionadas al azar resulte errónea en gran medida, pero si se aplica a miles de personas seleccionadas al azar, puede acercarse bastante a los valores reales de la población.

EJEMPLO Volando alto Calcule la probabilidad de que un adulto que se seleccionó aletoriamente haya volado en una línea aérea comercial.

SOLUCIÓN El espacio muestral consiste en dos sucesos simples: la persona que se seleccionó ya voló en una línea aérea comercial o no lo ha hecho. Puesto que el espacio muestral consiste en sucesos que no son igualmente probables, no es posible utilizar el método clásico (regla 2). Con estos resultados de una encuesta de Gallup, podemos usar el método de frecuencias relativas (regla 1): de 855 adultos que se seleccionaron al azar, 710 indicaron que ya volaron en líneas aéreas comerciales. Obtenemos el resultado siguiente:

$$P(\text{haber volado en una línea aérea comercial}) = \frac{710}{855} \approx 0.830$$

EJEMPLO Ruleta Usted planea apostar al número 13 en el próximo giro de una ruleta. ¿Cuál es la probabilidad de que pierda?

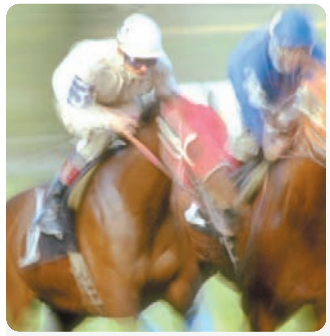
SOLUCIÓN Una ruleta tiene 38 ranuras distintas y sólo una corresponde al número 13. La ruleta se diseñó de manera que las 38 ranuras sean igualmente probables de resultar. De las 38 ranuras, 37 resultan en una pérdida. Ya que el espacio muestral incluye resultados igualmente probables, usamos el método clásico (regla 2) para obtener

$$P(\text{pérdida}) = \frac{37}{38}$$

EJEMPLO Choque de meteoritos ¿Cuál es la probabilidad de que su automóvil sea impactado por un meteorito este año?

SOLUCIÓN La ausencia de datos históricos de meteoritos que chocan contra automóviles impide usar el método de frecuencias relativas de la regla 1. Hay dos posibles resultados (chocar o no chocar), pero no son igualmente probables, de tal forma 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 tal caso, todos sabemos que la probabilidad en cuestión es muy, muy pequeña. Estimemos que sea, digamos, de 0.000000000001 (equivalente a una en un billón). Este estimado subjetivo, que se basa en nuestro conocimiento general, puede encontrarse en el campo general de la probabilidad real.

En problemas de probabilidad básicos del tipo de los que estamos considerando, es muy importante examinar con cuidado la información disponible e identificar correctamente el número total de posibles resultados. En algunos casos, el número total de resultados posibles está dado, pero en otros tiene que calcularse, como en el siguiente ejemplo, que requiere que calculemos el número total de resultados posibles.



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 que se apostaron, en tanto que las probabilidades reales se calcularon a partir de los resultados reales de las carreras.

EJEMPLO Pena de muerte Se seleccionan adultos al azar para una encuesta de Gallup; a ellos se les pregunta si están a favor de la pena de muerte para una persona convicta por homicidio. Las respuestas incluyen a 319 personas que están a favor de la pena de muerte, 133 personas que están en contra y 39 que no tienen una opinión al respecto. Con base en tales resultados, estime la probabilidad de que una persona seleccionada aleatoriamente esté a favor de la pena de muerte.

SOLUCIÓN *Sugerencia:* En lugar de tratar de formular una respuesta directamente del extracto escrito, resuma la información dada en un formato que le permita comprenderla mejor. Por ejemplo:

319	a favor de la pena de muerte
133	en contra de la pena de muerte
39	sin opinión
491	total

Ahora utilicemos el método de frecuencias relativas (regla 1) como sigue:

$$P\text{(personas a favor de la pena de muerte)}$$
$$= \frac{\text{número de personas a favor de la pena de muerte}}{\text{total}} = \frac{319}{491} = 0.650$$

Estimamos que hay una probabilidad de 0.650 de que cuando un adulto se selecciona al azar, él o ella estén a favor de la pena de muerte para alguien convicto por homicidio. Como sucede con todas las encuestas, la precisión de tal resultado depende de la calidad del método de muestreo y del procedimiento de la encuesta. Ya que la encuesta fue realizada por la organización Gallup, es probable que los resultados sean razonablemente precisos. El capítulo 6 incluirá procedimientos más avanzados para analizar resultados de encuesta como éstos.

	1ro 2do 3ro
	niño–niño–niño
exactamente 2 niños →	niño–niño–niña
→	niño–niña–niño
	niño–niña–niña
→	niña–niño–niño
	niña–niño–niña
	niña–niña–niño
	niña–niña–niña

EJEMPLO Género de hijos Determine la probabilidad de que una pareja con tres hijos tenga exactamente dos niños. 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 del otro.

SOLUCIÓN El mayor obstáculo 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 ocho diferentes formas en que los tres hijos pueden presentarse; las listamos al margen. Como los ocho resultados son igualmente probables, utilizamos la regla 2. De los ocho posibles resultados, tres corresponden exactamente a dos niños, así que

$$P(2\text{ niños en 3 nacimientos}) = \frac{3}{8} = 0.375$$

INTERPRETACIÓN Existe una probabilidad de 0.375 de que si un matrimonio 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) resulta impráctico. En el juego del Solitario, por ejemplo, los resultados (del reparto) son todos igualmente probables, pero es muy frustrante tratar de utilizar la regla 2 para calcular la probabilidad de ganar. En estos casos es posible obtener buenos estimados con mayor facilidad usando el método de frecuencias relativas (regla 1). Es muy común que las simulaciones sean útiles cuando se emplea este método. (Una **simulación** de un procedimiento es un proceso que se comporta en las mismas formas que el procedimiento mismo; por lo tanto, produce resultados similares). 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.

EJEMPLO Día de Acción de Gracias Si se selecciona un año al azar, calcule la probabilidad de que el Día de Acción de Gracias sea un *a*) miércoles, *b*) jueves.

SOLUCIÓN

a. El Día de Acción de Gracias se celebra siempre el cuarto jueves de noviembre. Por lo tanto, es imposible que un Día de Acción de Gracias caiga en miércoles. Cuando un suceso es imposible, decimos que su probabilidad es 0.

b. Es cierto que el Día de Acción de Gracias es un jueves. Cuando es seguro que un suceso ocurra, decimos que su probabilidad es 1.

Ya que cualquier suceso imaginable es imposible, o cierto, o está en alguna parte intermedia, se deduce que la probabilidad matemática de cualquier suceso es 0, 1, o un número entre 0 y 1 (véase figura 3-2).

- **La probabilidad de un suceso imposible es 0.**
- **La probabilidad de un suceso que ocurrirá con certeza es 1.**
- $0 \leq P(A) \leq 1$ **para cualquier suceso *A*.**

En la figura 3-2, la escala de 0 hasta 1 se muestra a la izquierda, mientras que las expresiones más comunes y familiares de probabilidad se indican a la derecha.

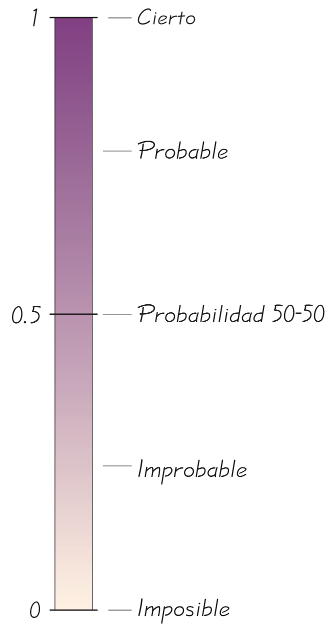


FIGURA 3-2 Valores posibles para probabilidades

Sucesos complementarios

Algunas veces necesitamos calcular la probabilidad de que un suceso *A* *no* ocurra.

Definición

Complemento de un suceso *A*, denotado por $\overline{A}$: 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 que el bebé *no* sea un niño?

continúa

¿Qué tan probable?

¿Cómo interpretamos términos como probable, improbable o extremadamente probable? La Federal Aviation Administration (FAA) interpreta estos términos de la manera siguiente: *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 aeroplano. *Improbable:* una probabilidad del orden de 0.00001 o menor. Sucesos de esta clase no se espera que ocurran durante toda la vida operacional de un solo aeroplano de un tipo en particular, aunque pueden ocurrir durante toda la vida operacional de todos los aeroplanos de un tipo en particular. *Extremadamente improbable:* una probabilidad en el orden de 0.000000001 o menor. Eventos como éste son tan improbables que no es necesario considerar su ocurrencia.

SOLUCIÓN Ya que 105 de los 205 bebés son niños, se deduce que 100 de ellos son niñas, entonces

$$P(\text{no seleccionar a un niño}) = P(\overline{\text{niño}}) = P(\text{niña}) = \frac{100}{205} = 0.488$$

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 texto.

Redondeo de probabilidades

Cuando se expresa el valor de una probabilidad, hay que dar la fracción o el número decimal *exactos*, o 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 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 relevantes (21491), por lo cual puede redondearse a 0.0215, con tres dígitos relevantes.
- La probabilidad de $1/3$ puede permanecer como fracción o redondearse a 0.333. *No* redondee a 0.3.
- La probabilidad de caras en un lanzamiento de una moneda puede expresarse como $1/2$ o 0.5; ya que 0.5 es exacto, no hay necesidad de expresarlo como 0.500.
- La fracción $432/7842$ es exacta, pero su valor no es evidente. 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; la usaremos de aquí en adelante en. Por ejemplo, un resultado de computadora típico puede incluir una expresión “valor *P*” como “nivel de significancia 0.001”. Más tarde analizaremos el significado del valor *P*, pero tales valores son esencialmente probabilidades del tipo que se analizó 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 un promedio de sólo una vez en mil ensayos.

Posibilidades

Las expresiones de probabilidad a veces se proponen como *posibilidades*, como 50:1 (o “50 a 1”). Una grave desventaja de las posibilidades es que hacen que muchos cálculos sean extremadamente difíciles. Por ello, los estadísticos, los matemáticos y los científicos prefieren usar probabilidades. La ventaja de las posibilidades es que facilitan el manejo de las transacciones de dinero asociadas con los juegos de azar, por lo cual 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 pistas de carreras y los casinos están en el negocio con la finalidad 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 son el cociente $P(\overline{A})/P(A)$, 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)

EJEMPLO Si usted apuesta \$5 al número 13 en la ruleta, su probabilidad de ganar es $1/38$, en tanto que las posibilidades de pago están dadas por el casino como 35:1.

a. Calcule las posibilidades reales en contra del resultado de 13.

b. ¿Cuánta ganancia neta podría usted obtener si gana apostando al 13?

c. Si el casino estuviera funcionando solamente por diversión y las posibilidades de pago fueran cambiadas para igualar las posibilidades reales en contra del 13, ¿cuanto ganaría usted si el resultado fuera 13?

SOLUCIÓN

a. Con $P(13) = 1/38$ y $P(\text{no } 13) = 37/38$, tenemos

$$\text{posibilidades reales en contra del 13} = \frac{P(\text{no } 13)}{P(13)} = \frac{37/38}{1/38} = \frac{37}{1} \text{ o } 37{:}1$$

b. 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 que se apuesta. Para una apuesta de \$5, la ganancia neta es de \$175. El apostador que gane podría recoger \$175 más la apuesta original de \$5.

c. 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 se cambiaran de 35:1 a 37:1, usted obtendría una ganancia neta de \$37 por cada \$1 que apostara. Si usted apuesta \$5, su ganancia neta sería de \$185. (El casino logra su ganancia pagando sólo \$175, en lugar de los \$185 que se pagarían con un juego de ruleta justo, en lugar de uno que favorece al casino).

Puedes apostarlo

En una lotería estatal típica, la "casa" tiene una ventaja del 65 o el 70%, ya que sólo entre el 35 y el 40% del dinero que se apuesta 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, del 5.9% para el 21, del 1.4% para los dados y del 3 al 22% para las máquinas tragamonedas. Algunos jugadores profesionales pueden ganar sistemáticamente en el 21, usando complicadas técnicas de conteo de cartas. Ellos saben cuándo 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.

3-2 Destrezas y conceptos básicos

En los ejercicios 1 y 2 exprese el grado indicado de probabilidad como un valor de probabilidad.

1. Identificación de valores de probabilidad
a. "Usted tiene una probabilidad de 50-50 de escoger el camino correcto".
b. "Hay un 20% de probabilidad de que llueva mañana".
c. "Usted tiene una probabilidad de un pelo de rana de casarse con mi hija".

2. Identificación de valores de probabilidad
a. "Hay un 90% de probabilidad de que mañana nieve".
b. "Definitivamente, por la noche oscurecerá".
c. "Usted tiene una probabilidad en diez de estar en lo correcto".

3. Identificación de valores de probabilidad ¿Cuáles de los siguientes valores *no pueden* ser probabilidades?

$$0, \quad 1, \quad -1, \quad 2, \quad 0.0123, \quad 3/5, \quad 5/3, \quad \sqrt{2}$$

4. 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?

5. Género de hijos En esta sección, dimos un ejemplo que incluye una lista de los ocho resultados 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.

6. Teléfonos celulares y cáncer cerebral En un estudio de 420,000 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 que se seleccionó al azar desarrolle un cáncer de este tipo. Se encontró que la probabilidad para la población general es de 0.000340; ¿es el resultado muy diferente de éste? ¿Qué sugiere el resultado acerca de los teléfonos celulares como causantes de cáncer de este tipo, como ya se afirmó?

7. Probabilidad de un jonrón El jugador de beisbol Barry Bonds rompió un récord importante cuando dio 73 jonrones en la temporada 2001. Durante esa temporada, estuvo al bat 476 veces. Si se selecciona al azar una de las ocasiones que estuvo al bat, calcule la probabilidad de que sea una de las veces que pegó un jonrón. ¿Difiere mucho el resultado de la probabilidad de 0.0715 que resulta de sus 567 jonrones en 7932 ocasiones que estuvo al bat?

8. 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 a una persona que se selecciona al azar en Estados Unidos sea alcanzada por un rayo este año.

Uso de la probabilidad para identificar sucesos infrecuentes. *En los ejercicios 9 a 16 considere un suceso como "infrecuente" si su probabilidad es igual o menor que 0.05. (Esto equivale al mismo criterio que se usa comúnmente en estadística inferencial, pero*

Probabilidades que desafían a 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 César exhaló en el último aliento al morir. En este mismo ánimo morboso y poco intuitivo, si la fatal taza con cicuta que mató a Sócrates hubiera contenido 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 verificar: en grupos de 25 estudiantes, la probabilidad de que al menos dos cumplan años el mismo día es de más del 50%.

como el valor de 0.05 no es absolutamente rígido, algunas veces se emplean otros valores en su lugar; por ejemplo, 0.01).

9. Probabilidad de un resultado equivocado La tabla 3-1 muestra que, de 85 mujeres embarazadas, la prueba de embarazo arrojó un resultado equivocado cinco veces.
a. Con base en los resultados disponibles, calcule la probabilidad de obtener un resultado equivocado de la prueba para una mujer embarazada.
b. Para mujeres no embarazadas, ¿será "infrecuente" que el resultado de la prueba esté equivocado?

10. Probabilidad de un resultado equivocado La tabla 3-1 muestra que de 14 mujeres que *no* están embarazadas, la prueba de embarazo produjo un resultado equivocado tres veces.
a. Con base en los resultados disponibles, calcule la probabilidad de obtener un resultado de prueba equivocado para una mujer no embarazada.
b. Para mujeres no embarazadas ¿será "infrecuente" que el resultado de la prueba esté equivocado?

11. Encuesta de tabaquismo En una encuesta de Gallup, se interrogó a 1038 adultos acerca de los efectos del tabaquismo pasivo; 52 de ellos indicaron que tales efectos "no son dañinos en absoluto".
a. Si usted selecciona al azar a uno de los adultos que se encuestaron, ¿cuál es la probabilidad de seleccionar a alguien que opine que ser fumador pasivo no es dañino en absoluto?
b. ¿Es "infrecuente" que alguien opine que ser fumador pasivo no es dañino en absoluto?

12. Fármaco reductor del colesterol En un ensayo clínico de Lipitor, un fármaco común que se usa para disminuir el colesterol, a un grupo de pacientes se les administró un tratamiento de tabletas de Atorvastatin de 10 miligramos. En dicho grupo, 19 pacientes sufrieron síntomas de gripe y 844 no los sufrieron (según datos de Pfizer, Inc.).
a. Estimar la probabilidad de que un paciente que toma el fármaco sufra síntomas de gripe.
b. ¿Es "infrecuente" que un paciente que toma el fármaco sufra síntomas de gripe?

13. Pasajeros de líneas aéreas "rebotados" En un año reciente, a 2624 pasajeros de American Airlines se les impidió abordar sus vuelos contra su voluntad, en tanto que hubo otros 168,262 que fueron "rebotados" voluntariamente a cambio de efectivo o vales.
a. Estime la probabilidad de que un pasajero rebotado de American Airlines, que se selecciona al azar, sea uno de los que fueron rebotados contra su voluntad.
b. ¿Es "infrecuente" que alguien sea "rebotado" en contra de su voluntad?

14. Llegadas de vuelo a tiempo Un estudio de 150 vuelos de American Airlines, seleccionados aleatoriamente, mostró que 108 llegaron a tiempo (según datos del Departamento del Transporte de Estados Unidos).
a. ¿Cuál es la probabilidad estimada de que un vuelo de American Airlines llegue *retrasado*?
b. ¿Es "infrecuente" que un vuelo de American Airlines llegue retrasado?

15. 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 estaría convencida 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?

16. Lotería En la antigua lotería del estado de Nueva York, usted tenía que escoger seis números entre 1 y 54, inclusive. Había 25,827,165 diferentes combinaciones de seis números posibles y se tenía que seleccionar la combinación correcta de los seis números para ganar el premio mayor. Para una apuesta de $1, usted escogía dos distintas combinaciones de seis números. (No era posible seleccionar sólo una combinación de seis números, usted tenía que seleccionar dos).

a. Si usted apostaba $1 y seleccionaba dos combinaciones diferentes de seis números, ¿cuál era la probabilidad de ganar el premio mayor?

b. ¿Era infrecuente ganar el premio mayor?

17. Probabilidad de un cumpleaños

a. Si se selecciona 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.

b. Si se selecciona a una persona aleatoriamente, calcule la probabilidad de que su cumpleaños sea cualquier otro día. Ignore los años bisiestos.

c. Seleccione al azar a una persona y calcule la probabilidad de que naciera un día de la semana que termine con la letra *s* o con la letra *o*.

18. Probabilidad de 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 dichos resultados para estimar la probabilidad de que un consumidor que se selecciona al azar reconozca la sopa Campbell's.

b. *Estime* la probabilidad de que un consumidor adulto estadounidense, seleccionado al azar, conozca el nombre de la marca McDonald's, la cadena de restaurantes de comida rápida más famosa en Estados Unidos.

c. *Estime* la probabilidad de que un consumidor adulto estadounidense que se selecciona aleatoriamente reconozca el nombre de la marca Veeco Instruments, un fabricante de productos de microelectrónica.

19. Encuesta del pastel de frutas (*fruitcake*) En una encuesta de Bruskin-Goldring Research, se preguntó cómo debía usarse un pastel de frutas. Ciento treinta y dos personas respondieron que como tope para una puerta y 880 citaron otros fines, incluyendo como comida para pájaros, relleno para terrenos y regalo. Si se selecciona al azar a una de estas personas, ¿cuál es la probabilidad de que sea alguien que usaría el pastel de frutas como tope para una puerta?

20. Probabilidad de un accidente automovilístico De entre 400 conductores aleatoriamente seleccionados, en el rango de edades de 20 a 24 años, 136 sufrieron un accidente automovilístico durante el año anterior (de acuerdo con datos del Consejo de Seguridad Nacional de Estados Unuidos). 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 año próximo? ¿Será el valor resultante suficientemente alto como para preocupar a los individuos de 20 a 24 años de edad?

21. Probabilidad de ganar en el Solitario Remítase al conjunto de datos 27 del Apéndice B y suponga que se juega el mismo Solitario de Microsoft.

a. Estime la probabilidad de ganar cuando se juega una partida.

b. Estime la probabilidad de ganar $208 agotando todas las cartas.

22. Probabilidad de reacción adversa a un fármaco Cuando el fármaco Viagra se probó clínicamente, 117 pacientes reportaron dolor de cabeza y 617 no lo hicieron (según datos de Pfizer, Inc.). Use esta muestra para estimar la probabilidad de que un usuario de Viagra sufra dolor de cabeza. ¿La probabilidad es suficientemente alta como para preocupar a los usuarios de Viagra?

23. Género de hijos: construcción de espacio muestral La sección 3-2 incluye una tabla que resume los resultados de género para una pareja que planea tener tres hijos.

a. Construya una tabla similar para una pareja que planea tener dos hijos.

b. Suponiendo que los resultados que se listan en el inciso *a*) sean igualmente probables, calcule la probabilidad de tener dos niñas.
c. Calcule la probabilidad de tener exactamente un hijo de cada género.

24. Genética: construcción de espacio muestral Ambos padres tienen el par de genes de color de ojos café/azul, y cada uno transmite un gene a un hijo. Suponga que si el hijo tiene al menos un gene café, ese color dominará y los ojos serán cafés. (La determinación real del color de los ojos es un tanto más complicada).
a. Haga una lista de los posibles resultados diferentes. Suponga que dichos resultados son igualmente probables.
b. ¿Cuál es la probabilidad de que un hijo de estos padres tenga el par de genes azul/azul?
c. ¿Cuál es la probabilidad de que el hijo tenga ojos cafés?

25. Posibilidades en el Derby de Kentucky Cuando el caballo Monarchos ganó el 127º Derby de Kentucky, una apuesta de $2 a que Monarchos ganaría resultó en un reintegro de $23.
a. ¿Qué ganancia neta hubo al ganar con una apuesta de $2 a Monarchos?
b. ¿Cuáles fueron las posibilidades de pago en contra de que Monarchos ganara?
c. Con base en el paseo preliminar a la carrera, los apostadores colectivamente creyeron que Monarchos tenía una probabilidad de ganar de 1/15. Suponiendo que 1/15 era la probabilidad real de la victoria de Monarchos, ¿cuáles fueron las posibilidades reales en contra?
d. Si las posibilidades de pago fueran iguales a las posibilidades reales que se calcularon en el inciso *c*), ¿cuanto valdría un boleto de $2 después de que Monarchos ganó?

26. Cálculo de posibilidades en la ruleta Una rueda de ruleta tiene 38 ranuras, una ranura es 0, otra es 00 y las otras están numeradas del 1 hasta el 36. Usted apuesta a número impar.
a. ¿Cuál es su probabilidad de ganar?
b. ¿Cuáles son las posibilidades reales en contra?
c. Cuando se apuesta a número impar, las posibilidades de pago son 1:1. ¿Qué ganancia obtendría usted si apostara $18, si pudiera, de alguna manera, convencer al casino de cambiar sus posibilidades de pago para que fueran las mismas que las posibilidades reales en contra? (*Sugerencia:* No trate realmente de convencer a ningún casino de esto; su sentido del humor está ausente por completo cuando se trata de asuntos como éste).

3-2 Más allá de lo básico

27. Interpretación de la eficacia Se diseña un experimento doble ciego para probar la eficacia del fármaco Estadisticzeno, como tratamiento para la ceguera ante los números. Cuando a los sujetos se les trata con Estadisticzeno, parecen mostrar mejoría. Los investigadores calculan que hay una probabilidad de 0.04 de que el grupo de tratamiento muestre mejoría si el fármaco no surte efecto. ¿Qué concluye acerca de la eficacia del Estadisticzeno?

28. Determinación de un jurado aleatorio Un abogado defiende a un cliente a quien se le acusó de no cumplir con sus obligaciones de pensión alimenticia. La junta de jurados potenciales consta de 20 mujeres, por lo que el abogado calcula que hay una probabilidad de 1/1,048,576 de que 20 personas seleccionadas aleatoriamente sean todas mujeres. ¿Hay fundamento para argumentar que la junta de jurados es injusta para su cliente?

29. Cálculo de probabilidad a partir de posibilidades Puesto que las posibilidades reales en contra de un suceso A sean $a{:}b$, entonces $P(A) = b/(a + b)$. Calcule la probabilidad de que Millenium gane su próxima carrera, ya que las posibilidades reales en contra son de 3:5.

30. Riesgo relativo y cociente de posibilidades En un ensayo clínico de 734 sujetos, a quienes se les trató con Viagra, 117 reportaron dolor de cabeza. En un grupo control con 725 sujetos sin dicho tratamiento, 29 reportaron dolores 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 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 el **cociente de posibilidades**, 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 control, el cual se calcula evaluando lo siguiente:

$$\frac{p_t/(1-p_t)}{p_c/(1-p_c)}$$

El riesgo relativo y el cociente de posibilidades a menudo se utilizan en estudios médicos y epidemiológicos. Calcule el riesgo relativo y el cociente de posibilidades para los datos del dolor de cabeza.

31. Años bisiestos y adivinación de cumpleaños En el inciso *a*) del ejercicio 17, para encontrar la probabilidad de que una persona que se selecciona al azar cumpla años el 18 de octubre, se ignoraron los años bisiestos.

a. Recalcule tal probabilidad considerando que un año bisiesto ocurre cada cuatro años. (Exprese su respuesta en forma de fracción exacta).

b. Los años bisiestos ocurren en años divisibles entre cuatro, salvo en los años centesimales (años que terminan en 00), en los que ocurre un año bisiesto en una serie de cuatro. Por ejemplo, los años 1700, 1800, y 1900 no fueron años bisiestos, mientras que el 2000 sí lo fue. Calcule la probabilidad exacta para el caso que se menciona, incluyendo los años bisiestos, y exprésela como una fracción exacta.

32. Moscas en una naranja Si dos moscas se paran en una naranja, calcule la probabilidad de que ambas se localicen en puntos pertenecientes al mismo hemisferio.

33. Puntos en un palo En un palo recto se seleccionan al azar dos puntos longitudinales. 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. (Quizás éste sea el ejercicio más difícil del libro).

3-3 Regla de la suma

El objetivo principal de esta sección es introducir la *regla de la suma* como un método para calcular probabilidades que pueden expresarse de la forma $P(A \text{ o } B)$, o sea, 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. La palabra clave que debemos recordar es *o*. En todo este texto usaremos la expresión *o inclusive*, que significa: uno o el otro, o ambos. (Con la excepción del ejercicio 27, no consideramos el *o exclusivo*, que significa que o bien uno o bien el otro, pero no ambos).

En la sección anterior presentamos aspectos fundamentales de la probabilidad y estudiamos sucesos calificados como *simples*. En esta sección y en la siguiente, introducimos *sucesos compuestos*.

Definición

Suceso compuesto: cualquier suceso que combina dos o más sucesos simples.

Notación de la regla de la suma

$P(A \text{ o } B) = P(\text{ocurre el suceso } A \text{ u ocurre el suceso } B \text{ o ambos ocurren})$.

Las probabilidades juegan un papel prominente en la genética. Observe la figura 3-3, que presenta una muestra de chícharos, como los que usó Mendel en sus famosos experimentos de hibridación. Los chícharos que se muestran tienen vainas verdes o amarillas y flores moradas o blancas. En esta muestra de 14 chícharos, ¿cuántos de ellos tienen "vainas verdes o flores moradas"? (Recuerde, "vaina verde o flor morada" significa realmente "vaina verde, o flor morada, o ambas"). La revisión de la figura 3-3 debe indicar que, en total, 12 chícharos tienen vainas verdes o flores moradas. (*Nota importante: Es erróneo* sumar los ocho chícharos con vainas verdes a los nueve chícharos con flores moradas, ya que el total de 17 toma en cuenta dos veces a cinco de los chícharos, pero éstos son individuos, por lo cual deben contarse una vez cada uno). Puesto que 12 de los 14 chícharos tienen "o vainas verdes o flores moradas", la probabilidad de seleccionar un chícharo al azar con una vaina verde o una flor morada se expresa como P(vaina verde o flor morada) $= 12/14 = 6/7$.

El 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 método 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 cualquier traslape entre estos dos conjuntos, compensar restando el número de resultados que se contaron dos veces, como se hace en la regla siguiente:

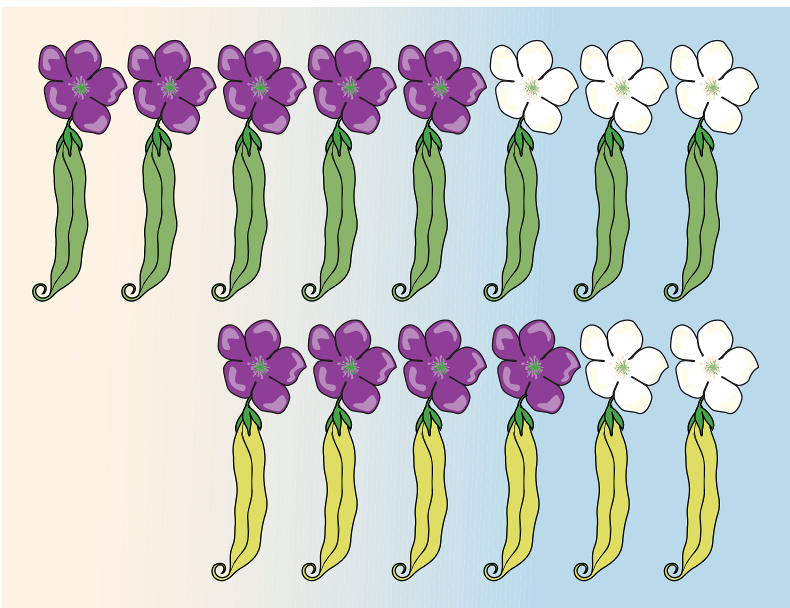


FIGURA 3-3 Chícharos utilizados en un estudio genético

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 de ensayo o de procedimiento.

Aunque la regla formal de la suma se presenta como una fórmula, en general es mejor entender el espíritu de la regla y aplicarlo intuitivamente de la siguiente forma:

Regla intuitiva de la suma

Para obtener $P(\text{A o } B)$, 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 modo que cada resultado se cuente sólo una vez*. $P(\text{A o } B)$ es igual a esa suma dividida entre el número total de resultados en el espacio muestral.

En la figura 3-4 se muestra un diagrama de Venn, que nos ofrece una comprensión visual de la regla formal de la suma. En esta figura se observa 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 futbol americano). La 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 sustenta la regla de la suma. Por la relación entre la regla de la suma y el diagrama de Venn que se muestra en la figura 3-4, es común el uso de la notación $P(A \cup \text{B})$ en lugar de $P(A \text{ o } B)$. Asimismo, se utiliza con frecuencia la notación $P(A \cap B)$ en lugar de $P(A \text{ y } B)$, de modo 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 si A y B no pueden ocurrir simultáneamente, porque $P(A \text{ y } B)$ se convierte en cero. La figura 3-5 ilustra que si A y B no se traslapan, tenemos $P(A \text{ o } B) = P(A) + P(B)$. La definición siguiente formaliza la ausencia de intercepto que se muestra en la figura 3-5.

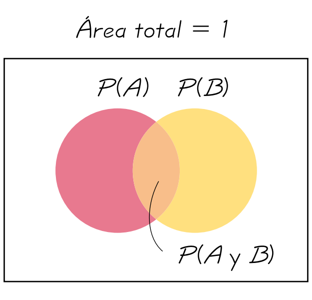


FIGURA 3-4 Diagrama de Venn que muestra sucesos traslapados

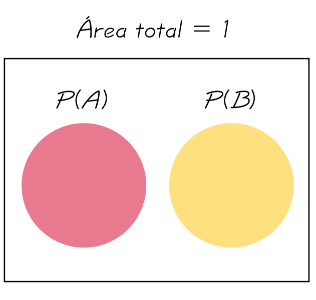


FIGURA 3-5 Diagrama de Venn que muestra sucesos no traslapados

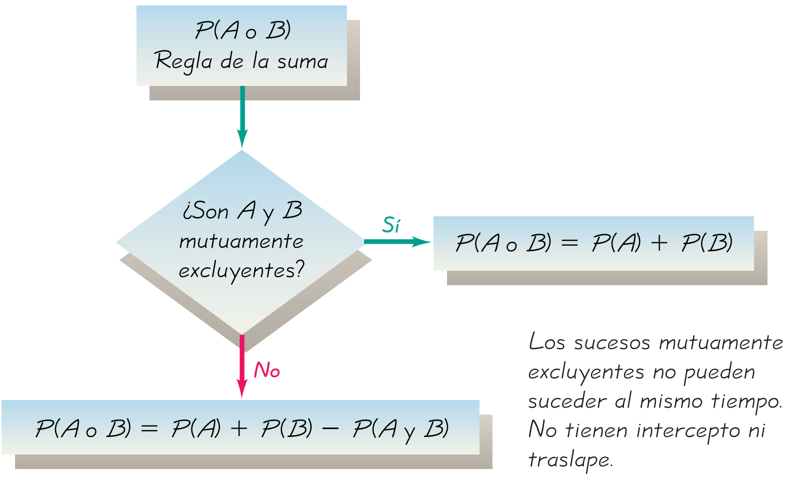


FIGURA 3-6 Aplicación de la regla de la suma

Definición

Los sucesos A y B son **desarticulados** (o **mutuamente excluyentes**) cuando ambos no pueden ocurrir juntos.

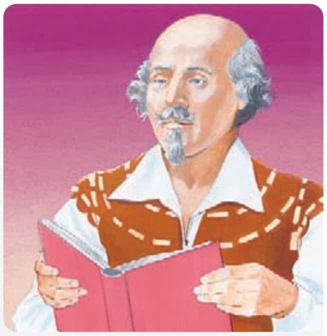
El diagrama de flujo de la figura 3-6 muestra cómo los sucesos mutuamente excluyentes afectan la regla de la suma.

EJEMPLO Ensayos clínicos de prueba de embarazo Remítase a la tabla 3-1, que se reproduce aquí para su conveniencia. Suponiendo que de entre 99 mujeres incluidas en el estudio, se selecciona una al azar, aplique la regla de la suma para calcular la probabilidad de seleccionar a una mujer que está embarazada o que tuvo un resultado de prueba positivo.

Tabla 3-1 Resultados de prueba de embarazo

	Resultado de prueba positivo (indicó embarazo)	Resultado de prueba negativo (no indicó embarazo)
La mujer está embarazada	80	5
La mujer no está embarazada	3	11

SOLUCIÓN Revisando la tabla, fácilmente advertimos que hay 88 mujeres embarazadas o que marcaron positivo. Obtenemos este total de 88 sumando las mujeres embarazadas a las mujeres que marcaron positivo, teniendo cuidado de contar a las 80 mujeres embarazadas que marcaron positivo sólo una vez. Sería erróneo sumar las 85 mujeres embarazadas con las 83 mujeres que marcaron positivo, ya que el total de 168 incluiría a algunas mujeres dos veces; hay que considerar que son individuos que deben contarse sólo una vez. Dividiendo el total correcto de 88 entre el total general de 99, obtenemos este resultado: $P(\text{embarazada o positiva}) = 88/99 = 8/9$ o 0.889.



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 se presenta aquí es otra aplicación interesante. (Véase "Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?" en *Biometrika*, vol. 63, núm. 3).

Hay varias estrategias que usted podría usar en el ejemplo anterior, para contar a las mujeres que estaban embarazadas o que marcaron positivo. Cualquiera de los siguientes funcionaría:

- Resalte con color las casillas que representan mujeres que estaban embarazadas o que marcaron positivo; luego, sume los números en estas casillas, teniendo cuidado de sumar cada número sólo una vez. Este método nos da

$$3 + 80 + 5 = 88$$

- Sume las 85 mujeres embarazadas a las 83 mujeres que marcaron positivo, pero compense el doble conteo restando las 80 mujeres embarazadas que marcaron positivo. Este método da un resultado de

$$85 + 83 - 80 = 88$$

- Comience con el total de 85 mujeres embarazadas; después, sume aquellas que marcaron positivo y que no se incluyeron aún en ese total, para obtener un resultado de

$$85 + 3 = 88$$

Estudie cuidadosamente el ejemplo anterior, que clarifica esta característica esencial de la regla de la suma: "o" implica suma, y la suma debe hacerse sin doble conteo.

Los puntos clave de esta sección se resumen como sigue:

1. Para calcular $P(A \text{ o } B)$, empiece por asociar "o" con la suma.
2. Considere si los sucesos A y B son mutuamente excluyentes, es decir, ¿pueden ocurrir al mismo tiempo? Si no son mutuamente excluyentes (es decir, pueden ocurrir al mismo tiempo), asegúrese de evitar (o por lo menos compensar) el doble conteo cuando sume las probabilidades relevantes. Entendiendo la importancia de no contar por duplicado cuando calcule $P(A \text{ o } B)$, no será necesario tener que determinar el valor de $P(A) + P(B) - P(A \text{ y } B)$.

Los errores que se cometen al aplicar la regla de la suma a menudo implican doble conteo, es decir, sucesos que no son mutuamente excluyentes se tratan como si lo fueran. Una indicación de semejante error es una probabilidad total mayor que 1; sin embargo, los errores que se relacionan con la regla de la suma no siempre hacen que la probabilidad total rebase 1.

Sucesos complementarios

En la sección 3-2, definimos el complemento del suceso A y lo denotamos como $\overline{A}$. Decimos que consiste en todos los resultados donde el suceso A no puede ocurrir. Los sucesos A y $\overline{A}$ son por implicación mutuamente excluyentes, ya que es imposible que un suceso y su complemento ocurran al mismo tiempo. Además, es posible estar absolutamente seguros de que A ocurre, o bien, que no ocurre, lo que implica que debe ocurrir A o $\overline{A}$. Estas observaciones nos permiten aplicar la regla de la suma para sucesos mutuamente excluyentes, así que:

$$P(A \text{ o } \overline{A}) = P(A) + P(\overline{A}) = 1$$

Justificamos $P(A \text{ o } \overline{A}) = P(A) + P(\overline{A})$ señalando que A y $\overline{A}$ son mutuamente excluyentes; 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 expresiones equivalentes.

Regla de los sucesos complementarios

$$P(A) + P(\overline{A}) = 1$$
$$P(\overline{A}) = 1 - P(A)$$
$$P(A) = 1 - P(\overline{A})$$

La figura 3-7 es una representación visual de la relación entre $P(A)$ y $P(\overline{A})$.

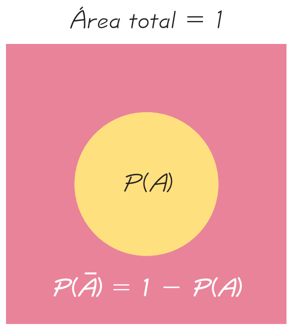


FIGURA 3-7 Diagrama de Venn para el complemento del suceso *A*

EJEMPLO En la realidad, cuando nace un bebé, $P(\text{niño}) = 0.5121$. Calcule $P(\overline{\text{niño}})$.

SOLUCIÓN Usando la regla de los sucesos complementarios, tenemos

$$P(\overline{\text{niño}}) = 1 - P(\text{niño}) = 1 - 0.5121 = 0.4879.$$

Es decir, la probabilidad de no tener un niño, que es la misma que la de tener una niña, es de 0.4879.

La principal ventaja de la regla de los sucesos complementarios es que permite simplificar ciertos problemas considerablemente. Ilustraremos esta ventaja en la sección 3-5.

3-3 Destrezas y conceptos básicos

Determine si los sucesos son mutuamente excluyentes. *Para cada inciso de los ejercicios 1 y 2, ¿son los dos eventos mutuamente excluyentes en un mismo experimento?* (Sugerencia: *Considere "mutuamente excluyentes" como equivalente a "separados" o "sin traslape".*)

1. a. Seleccionar aleatoriamente a un cirujano cardiaco.
Seleccionar aleatoriamente a un médico de sexo femenino.
b. Seleccionar aleatoriamente a una estudiante universitaria.
Seleccionar aleatoriamente a un estudiante universitario que conduzca motocicleta.
c. Seleccionar aleatoriamente a una persona que se trata con el fármaco para reducir el colesterol Lipitor.
Seleccionar aleatoriamente a una persona de un grupo control que no recibió tratamiento.

2. a. Seleccionar aleatoriamente a un padre de familia que esta noche, a las 8:15, se encuentre viendo la cadena NBC en la televisión.
Seleccionar aleatoriamente a un padre de familia que esta noche, a las 8:15, se encuentre viendo la cadena CBS en la televisión.
b. Recibir una llamada telefónica en respuesta a una encuesta voluntaria, de una persona que se opone a todos los impuestos gubernamentales.
Recibir una llamada telefónica en respuesta a una encuesta voluntaria, de una persona que aprueba todos los impuestos gubernamentales.
c. Seleccionar aleatoriamente a un senador de Estados Unidos que ocupe actualmente un puesto en el gobierno.
Seleccionar aleatoriamente a una mujer funcionaria electa.

3. Cálculo de complementos
a. Si $P(A) = 0.05$, calcule $P(\overline{A})$
b. Según datos de la oficina de censos de Estados Unidos, cuando se selecciona al azar a una mujer de 25 años de edad, hay una probabilidad de 0.218 de que tenga

una licenciatura. Si se selecciona al azar a una mujer de 25 años, calcule la probabilidad de que no tenga licenciatura.

4. Cálculo de complementos

a. Calcule $P(\overline{A})$ puesto que $P(A) = 0.0175$.

b. Una encuesta de Reuters/Zogby mostró que el 61% de los estadounidenses creen que existe vida en otro sitio en la galaxia. ¿Cuál es la probabilidad de seleccionar al azar a algún estadounidense que no crea esto?

5. Uso de la regla de la suma Remítase a la figura 3-3 y calcule la probabilidad de que al seleccionar aleatoriamente uno de los chícharos, obtenga uno con vaina verde o flor blanca.

6. Uso de la regla de la suma Remítase a la figura 3-3 y calcule la probabilidad de que al seleccionar aleatoriamente uno de los chícharos, obtenga uno con vaina amarilla o flor morada.

7. Día nacional de la estadística Si se selecciona una persona al azar, calcule la probabilidad de que su cumpleaños no sea el 18 de octubre, que es el Día Nacional de la Estadística en Japón. Ignore los años bisiestos.

8. Cumpleaños y complemento Si selecciona a una persona al azar, calcule la probabilidad de que su cumpleaños no sea en octubre. Ignore años bisiestos.

En los ejercicios 9 a 12 utilice los datos de la siguiente tabla, que resume resultados del hundimiento del Titanic.

	Hombres	Mujeres	Niños	Niñas
Sobrevivientes	332	318	29	27
Muertos	1360	104	35	18

9. Pasajeros del *Titanic* Si selecciona al azar a uno de los pasajeros del *Titanic*, calcule la probabilidad de que sea una mujer o una niña.

10. Pasajeros del *Titanic* Si selecciona al azar a uno de los pasajeros del *Titanic*, calcule la probabilidad de que sea un hombre o una persona que sobrevivió al hundimiento.

11. Pasajeros del *Titanic* Si selecciona al azar a uno de los pasajeros del *Titanic*, calcule la probabilidad de que sea un niño o un sobreviviente.

12. Pasajeros del *Titanic* Si selecciona al azar a uno de los pasajeros del Titanic, calcule la probabilidad de que sea una mujer o alguna persona que no sobrevivió al hundimiento.

Uso de la regla de la suma con grupos sanguíneos. *En los ejercicios 13 a 20 remítase a la gráfica adjunta, que describe los grupos sanguíneos y los tipos de Rh de 100 personas (según datos del Greater New York Blood Program). En cada caso, suponga que se selecciona uno de los 100 sujetos aleatoriamente; calcule la probabilidad que se indica.*

13. P(no grupo A)

14. P(tipo Rh^-)

15. P(grupo A o tipo Rh^-)

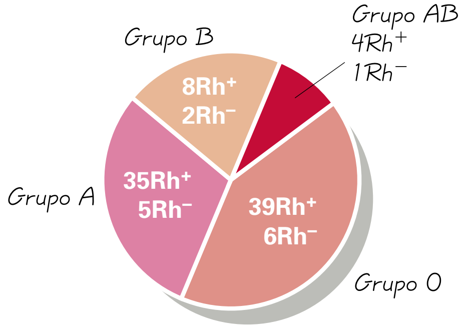
16. P(grupo A o grupo B)

17. P(no tipo Rh^+)

18. P(grupo B o tipo Rh^+)

19. P(grupo AB o tipo Rh^+)

20. P(grupo A u O o tipo Rh^+)



21. Datos de automóviles Remítase al conjunto de datos 22 del Apéndice B. Si uno de los 20 automóviles se selecciona al azar, calcule la probabilidad de que tenga transmisión manual o seis cilindros.

22. Tabaquismo y género Remítase al conjunto de datos 4 del Apéndice B. Si uno de los 107 sujetos de estudio se selecciona al azar, calcule la probabilidad de tener un varón o un fumador.

23. Resistencia a la encuesta Las empresas que realizan encuestas se interesan en los niveles decrecientes de cooperación de las personas que se contactan para que las encuesten. Un encuestador contacta a 84 individuos de entre 18 y 21 años, y descubre que 73 responden y 11 se rehúsan a hacerlo. Cuando se contacta a 275 personas de entre 22 y 29 años, 255 responden y 20 se rehúsan (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 las 359 personas. Calcule la probabilidad de que sea una persona en el rango de edad de 18 a 21 años o alguien que rechaza responder.

24. Resistencia a la encuesta Remítase al mismo conjunto de datos que se utilizó en el ejercicio 23. Suponga que se selecciona al azar a 1 de las 359 personas, calcule la probabilidad de que sea una persona en el rango de 18 a 21 años o que sí respondió.

3-3 Más allá de lo básico

25. Determine si los sucesos son mutuamente excluyentes

a. Si $P(A) = 3/11$, $P(B) = 4/11$ y $P(A \text{ o } B) = 7/11$, ¿qué puede inferir acerca de los sucesos A y B?

a. Si $P(A) = 5/18$, $P(B) = 11/18$, y $P(A \text{ o } B) = 13/18$, ¿qué puede inferir acerca de los sucesos A y B?

26. Sucesos mutuamente excluyentes Si los sucesos A y B son mutuamente excluyentes y los sucesos B y C también son mutuamente excluyentes, ¿tienen que ser mutuamente excluyentes los sucesos A y C? Dé un ejemplo que fundamente su respuesta.

27. *O* exclusivo ¿Como se transformaría la regla de la suma si el *o exclusivo* se usara en lugar del *o inclusivo*? En esta sección se explicó que el *o exclusivo* significa uno u otro, pero no ambos.

28. Extensión de la regla de la suma La regla formal de la suma, que se incluye 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).

3-4 Regla de la multiplicación: fundamentos

En la sección 3-3 presentamos la regla de la suma para calcular $P(A \text{ o } B)$, es decir, la probabilidad de que un ensayo tenga un resultado de A o B o ambos. El objetivo de esta sección es desarrollar una regla para calcular $P(A \text{ y } B)$, esto es, la probabilidad de que el suceso A ocurra en un primer ensayo y el suceso B ocurra en un segundo ensayo.

Sentenciados por probabilidad

Un testigo describió a una asaltante de Los Ángeles como una mujer de raza caucásica con pelo rubio, peinado en cola de caballo, que escapó en un automóvil amarillo que conducía un hombre afroamericano que usaba barba y bigote. Janet y Malcom Collins se ajustaban a tal descripción y se les condenó con fundamento en el testimonio de que hay aproximadamente una posibilidad en 12 millones de que cualquier pareja tenga tales características. Se estimó que la probabilidad de poseer 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 aleatoriamente, se cometió un error grave al no considerar la probabilidad de que hubiera *otras* parejas en la misma región con las mismas características.

Notación

$P(A \text{ y } B) = P($el suceso A ocurre en un primer ensayo y el suceso B ocurre en un segundo ensayo$)$

En la sección 3-3 asociamos *o* con sumar; en esta sección asociaremos *y* con multiplicar. Veremos que $P(A \text{ y } B)$ implica la multiplicación de probabilidades, y que en ocasiones deberemos ajustar la probabilidad del suceso B para discernir el resultado del suceso A.

La teoría de la probabilidad se utiliza extensamente en el análisis y el diseño de pruebas estandarizadas, como SAT, ACT, LSAT (para leyes) y MCAT (para medicina). Para facilitar la calificación, las pruebas de este tipo suelen incluir preguntas del tipo verdadero/falso o de opción múltiple. Supongamos que el primer reactivo de un examen es del tipo verdadero/falso, y que el segundo es de opción múltiple con cinco respuestas posibles (*a*, *b*, *c*, *d*, *e*). Usemos los dos reactivos siguientes. ¡Intente responderlos!

1. Verdadero o falso: Una libra de plumas pesa más que una libra de oro.

2. Entre lo siguiente, ¿qué es lo que tiene la mayor influencia en la sociedad moderna?

- **a.** El control remoto
- **b.** Este libro
- **c.** Las computadoras
- **d.** Los tenis con luces en el tacón
- **e.** Las recepcionistas coquetas

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 sea también 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 conjeturas al azar, tenemos que los 10 posibles resultados son igualmente probables, entonces

$$P(\text{ambas correctas}) = P(T \text{ y } c) = \frac{1}{10} = 0.1$$

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 modo que

$$P(T \text{ y } c) = P(T) \times P(c)$$

Esto sugiere que, en términos generales, $P(A \text{ y } B) = P(A) \cdot P(B)$, pero antes de hacer dicha generalización, consideremos otro ejemplo.

Por lo pronto, notamos que los diagramas de árbol suelen utilizarse 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, que se muestran como líneas que emanan de un punto de partida. Estos diagramas son útiles para calcular el número de resultados posibles, cuando el número de posibilidades no es demasiado grande. El diagrama de árbol de la figura 3-8 muestra los resultados de los reactivos de verdadero/falso y opción múltiple. En la figura 3-8 vemos que si las dos respuestas son conjeturas 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 cinco dos veces, o sea, 10. El diagrama de árbol de la figura 3-8 ilustra la razón del uso de la multiplicación.

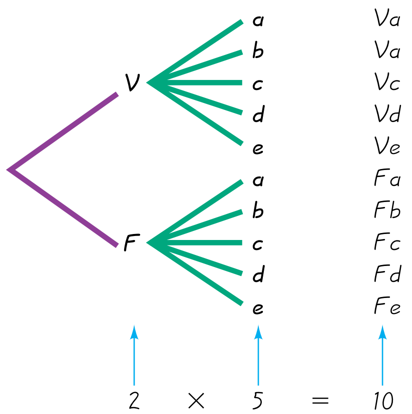


FIGURA 3-8 Diagrama de árbol de reactivos de examen

Nuestro primer ejemplo, el de los reactivos de verdadero/falso y opción múltiple, sugiere que $P(A \text{ y } B) = P(A) \cdot P(B)$; el siguiente ejemplo introducirá otro elemento importante.

EJEMPLO Experimento de genética En los famosos experimentos de hibridación de Mendel se emplearon chícharos, como los que se muestran en la figura 3-3, que se incluye en la sección 3-3 y se reproduce en la página siguiente. Si dos de los chícharos que se observan en la figura 3-3 se seleccionan al azar sin reemplazo, calcule la probabilidad de que la primera selección tenga una vaina verde y la segunda una vaina amarilla. (Es posible ignorar los colores de las flores en la parte superior).

SOLUCIÓN

Primera selección:	(porque hay 14 chícharos, ocho
$P(\text{vaina verde}) = 8/14$	de los cuales tienen vainas verdes)
Segunda selección:	(hay 13 chícharos sobrantes, seis
$P(\text{vaina amarilla}) = 6/13$	de los cuales tienen vainas amarillas)

continúa

Calificación perfecta en el sat

Si se selecciona al azar un sujeto que responde el 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? Se trata de dos preguntas muy diferentes.

En un año reciente, aproximadamente 1.3 millones de personas respondieron el SAT, y sólo 587 recibieron calificaciones perfectas de 1,600, entonces hay una probabilidad de 587 ÷ 1.3 millones, o alrededor de 0.000452, de seleccionar aleatoriamente a uno de los sujetos de la prueba y a una persona con una calificación perfecta. Sólo una parte del SAT incluye 35 preguntas de opción múltiple, y la probabilidad de responder a todas ellas correctamente adivinando es de $(1/5)^{35}$, cantidad tan pequeña que, cuando se escribe como un decimal, resultan 24 ceros después del punto decimal.

Recomendación para la lotería

Un columnista del diario *New York Daily News*, Stephen Allensworth, hace poco dio recomendaciones para seleccionar números en el juego *New York State's Daily Numbers*. 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, se están dejando engañar y perderán más dinero, ya que creen erróneamente que sus probabilidades de ganar mejoran.

Con P(primer chícharo con vaina verde) $= 8/14$ y P(segundo chícharo con vaina amarilla) $= 6/13$, tenemos

$$P(\text{primer chícharo con vaina verde y segundo chícharo con vaina amarilla}) = \frac{8}{14} \cdot \frac{6}{13} \approx 0.264$$

El punto clave es que se tiene que ajustar la probabilidad del segundo suceso para reflejar el resultado del primer suceso. Ya que el segundo chícharo se selecciona sin reemplazar el primero, la segunda probabilidad debe tomar en cuenta el resultado de la primera selección de un chícharo con vaina verde. Después de que se ha seleccionado un chícharo con vaina verde en el primer ensayo, sólo quedan 13 chícharos y seis de ellos tienen vainas amarillas, entonces la segunda selección nos da: P(chícharo con vaina amarilla) $= 6/13$.

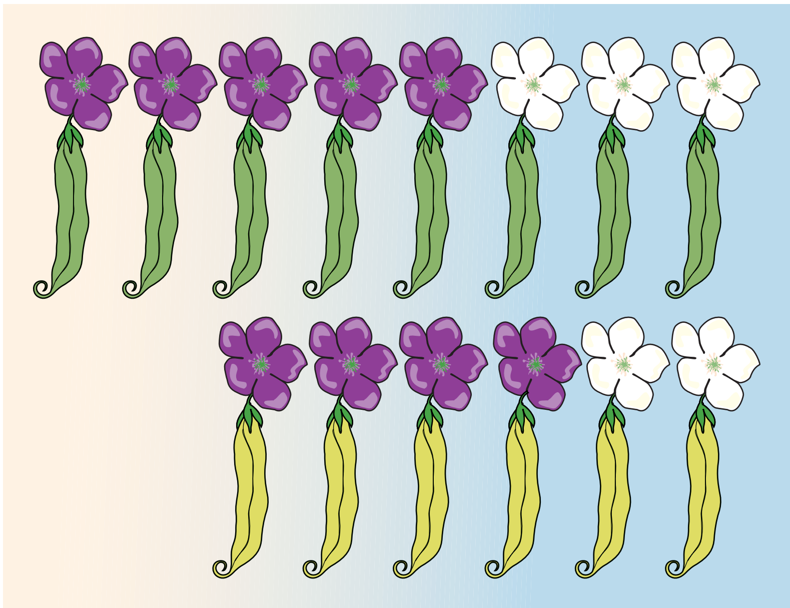


FIGURA 3-3 Chícharos usados en un estudio de genética

Este ejemplo manifiesta 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 notación siguiente.

Notación para la probabilidad condicional

$P(B \mid A)$ representa la probabilidad de que un suceso B ocurra después de admitir que el suceso A ya ocurrió. (Es posible leer $B \mid A$ como "B dado A").

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, algunos 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**.

Por ejemplo, jugar a la lotería de California y después a la lotería de Nueva York son sucesos *independientes*, porque el resultado de la lotería de California no surte efecto alguno 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.

Con la notación y las definiciones anteriores, junto con los principios ilustrados en los ejemplos, resumimos el concepto clave de la sección como la siguiente *regla formal de la multiplicación*, pero se recomienda que usted trabaje con la *regla intuitiva de la multiplicación*, que tiene más probabilidades de manifestar comprensión que el uso a ciegas de una fórmula.

Regla formal de la multiplicación

$$P(A \text{ y } B) = P(A) \cdot P(B|A)$$

Si A y B son sucesos independientes, $P(B|A)$ es realmente lo mismo que $P(B)$. (Para hacer un estudio más amplio y determinar si los sucesos son independientes o dependientes, véase el apartado "Prueba de independencia", en la sección 3-5. Por lo pronto, trate de entender el concepto básico de independencia y la forma en que afecta las probabilidades calculadas). Observe la siguiente *regla intuitiva de la multiplicación*. (Véase también la figura 3-9).

Regla intuitiva de la multiplicación

Cuando se trata de calcular 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 tome en cuenta la ocurrencia previa del suceso A.

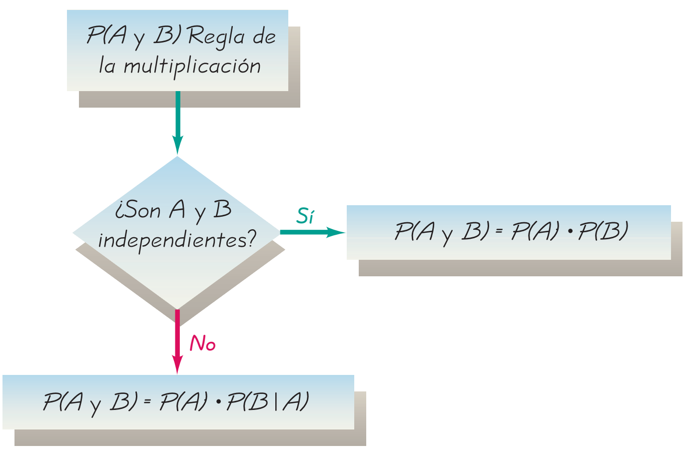


FIGURA 3-9 **Aplicación de la regla de la multiplicación**

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 luego falló el último motor que funcionaba. El jet descendió sin propulsión desde 13,000 pies hasta 4,000 pies, por lo que la tripulación logró arrancar un motor y la aeronave, con las 172 personas a bordo, que aterrizaron con seguridad. Con motores a reacción independientes, la probabilidad de que los tres fallen es de sólo 0.0001^3, 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.

EJEMPLO Bienes dañados Telektronics fabrica computadoras, televisores, reproductores de CD y otros productos electrónicos. Cuando los artículos que se envían se dañan, las causas del daño se clasifican como agua (A), compresión (C), perforación (P) o marcas en la caja (M). Abajo se encuentra una lista de las causas codificadas de cinco artículos que se dañaron. Una analista de control de calidad quiere seleccionar aleatoriamente dos artículos para elaborar una investigación más amplia. Calcule la probabilidad de que el primer artículo fuese dañado por compresión (C) y el segundo también por lo mismo (C). Suponga que las selecciones se hacen *a*) con reemplazo; *b*) sin reemplazo.

A C C P M

SOLUCIÓN

a. Si los dos artículos se seleccionan con reemplazo, las dos selecciones son independientes, ya que al segundo suceso no le afecta el primer resultado. En cada una de las dos selecciones hay dos artículos que se dañaron por compresión (C) entre los cinco; entonces, tenemos

$$P(\text{el primer artículo es C y el segundo artículo es C}) = \frac{2}{5} \cdot \frac{2}{5} = \frac{4}{25} \text{ o } 0.16$$

b. Si los dos artículos se seleccionan sin reemplazo, las dos selecciones son dependientes porque el segundo suceso se afectó por el primer resultado. En la primera selección, a dos de los cinco artículos los daño la compresión (C). Después de seleccionar un artículo dañado por compresión, estamos dejando cuatro artículos incluyendo a uno al que dañó también la compresión. Por lo tanto, tenemos

$$P(\text{el primer artículo es C y el segundo artículo es C}) = \frac{2}{5} \cdot \frac{1}{4} = \frac{2}{20} = \frac{1}{10} \text{ o } 0.1$$

Nótese que en este caso ajustamos la segunda probabilidad para tomar en cuenta la selección de un artículo al que dañó la compresión (C) en el primer resultado. Después de seleccionar C la primera vez, había sólo un C entre los cuatro artículos que quedaban.

Hasta aquí ya analizamos 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 caras es de $0.5 \cdot 0.5 \cdot 0.5 = 0.125$. También es posible extender la regla de la multiplicación para aplicarla a varios sucesos dependientes; simplemente hay que ajustar las probabilidades conforme se avanza. Por ejemplo, la probabilidad de sacar cuatro cartas diferentes (sin reemplazo) de un mazo revuelto y que todas sean ases es

$$\frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} \cdot \frac{1}{49} = 0.00000369$$

El inciso *b* del último ejemplo implicó la selección de artículos sin reemplazo; por lo tanto, tratamos los sucesos como dependientes. Sin embargo, es común tratar sucesos como independientes cuando se toman *muestras pequeñas* de *poblaciones grandes*. En tales casos, es raro seleccionar 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 modo que sean técnicamente dependientes).

Los encuestadores usan este lineamiento cuando encuestan apenas a 1,000 adultos de poblaciones de millones. Ellos suponen independencia, aunque toman la muestra sin reemplazo. El siguiente ejemplo es otra ilustración de dicho lineamiento. También ejemplifica cómo se utiliza la probabilidad para probar una afirmación hecha acerca de una población. Nos da idea del importante procedimiento de *prueba de hipótesis* que se estudiará en el capítulo 7.

EJEMPLO Control de calidad Una gerente de producción de Telektronics afirma que su nuevo proceso de manufactura de DVD es mejor porque la tasa de defectos es menor del 3%, que fue la tasa de defectos en el pasado. Para fundamentar su afirmación, ella fabrica un lote de 5,000 DVD; después, selecciona al azar 200 de ellos para probarlos, con el resultado de que no hay defectos en ninguno de los 200 DVD que se seleccionaron. Suponiendo que el nuevo método tuviera la misma tasa de defectos del 3% como en el pasado, calcule la probabilidad de que no haya defectos en los 200 DVD. Con base en el resultado, ¿hay suficiente evidencia para fundamentar la afirmación de la gerente de que su nuevo proceso es mejor?

SOLUCIÓN La probabilidad de que no tengan defectos es la misma que la probabilidad de que los 200 DVD estén en buen estado. Por lo tanto, queremos encontrar P(todos los 200 DVD en buen estado). También suponemos que la tasa de defectos es del 3% para observar si el resultado de cero defectos, por su probabilidad, llegaría a ocurrir fácilmente y poder así compararlo con el antiguo proceso de fabricación. Si la tasa de defectos es del 3%, tenemos P(DVD en buen estado) $=$ 0.97. Los DVD que se seleccionaron fueron escogidos sin reemplazo, pero la muestra de 200 DVD es menor al 5% de la población de 5,000. Entonces, trataremos los sucesos como si fueran independientes. Obtenemos este resultado:

P(1o en buen estado y 2o en buen estado y 3o en buen estado . . . y 200o en buen estado)

$= P$(DVD en buen estado) $\cdot$ P(DVD buen estado) $\cdot$ . . . $\cdot$ P(DVD en buen estado)

$= 0.97 \cdot 0.97 \cdot \ldots \cdot 0.97$

$= 0.97^{200} = 0.00226$

La baja probabilidad de 0.00226 indica que en lugar de obtener un resultado poco común, con una tasa de defecto del 3%, una explicación más razonable es que no ocurrieron defectos porque la tasa de defectos es realmente menor que el 3%. Debido a que hay una probabilidad tan pequeña (0.00226) de producir todos los DVD en buen estado, con un tamaño de muestra de 200 y una tasa de defectos del 3%, tenemos suficiente evidencia para concluir que el nuevo método es mejor.

Los fundamentos de las reglas de la suma y de la multiplicación se resumen como sigue:

- En la regla de la suma, la palabra "o" en $P(A$ o $B)$ sugiere una suma. Sume $P(A)$ y $P(B)$, siendo cuidadoso para hacerlo de forma que cada resultado se cuente sólo una vez.

LA ESTADÍSTICA EN LAS NOTICIAS

Redundancia

Es posible mejorar considerablemente la confiabilidad de los sistemas con la redundancia de componentes críticos. Los automóviles de carreras de las series de la NASCAR Winston Cup tienen dos sistemas de ignición para que, si uno falla, haya 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* sobre los aviones antirradar: "Un avión construido en buena parte con fibra de carbono fue el Lear Fan 2100, que debía 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 la 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.

- En la regla de la multiplicación, la palabra "y" en $P(A \text{ y } B)$ sugiere una multiplicación. Multiplique $P(A)$ y $P(B)$, pero asegúrese de que la probabilidad del suceso B toma en cuenta la ocurrencia previa del suceso A.

3-4 Destrezas y conceptos básicos

Identificación de sucesos como independientes o dependientes*. En los ejercicios 1 y 2 clasifique cada par de sucesos como independientes o dependientes.*

1. **a.** Tirar un dado y obtener un 5.
Lanzar una moneda y obtener cara.
b. Seleccionar aleatoriamente a un televidente que ve *Monday Night Football.*
Seleccionar aleatoriamente a un segundo televidente que ve *Monday Night Football.*
c. Usar pantalones cortos a cuadros con calcetines negros y sandalias.
Pedir a alguien una cita y tener una respuesta positiva.

2. **a.** Descubrir que su calculadora no funciona.
Descubrir que su refrigerador no funciona.
b. Descubrir que la luz de su cocina no funciona.
Descubrir que su refrigerador no funciona.
c. Beber hasta deteriorar su capacidad de conducir.
Verse involucrado en un accidente automovilístico.

3. Moneda y dado Calcule la probabilidad de que al lanzar una moneda y tirar un dado, los resultados sean cruz y 3.

4. Letra y dígito La propietaria de una computadora nueva crea una contraseña que consta de dos caracteres. Ella selecciona al azar una letra del alfabeto para el primer carácter y un dígito (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) para el segundo. ¿Cuál es la probabilidad de que su contraseña sea "K9"? ¿Sería eficaz esta contraseña como obstáculo contra alguien que trate de tener acceso a su computadora?

5. Aplicación de la regla de la multiplicación Si dos de los elementos que se muestran abajo se seleccionan al azar, calcule la probabilidad de que ambos elementos sean de color verde. Estos elementos se utilizan en pruebas de percepción.

rojo amarillo verde rojo azul amarillo

a. Suponga que el primer elemento se reemplaza antes de seleccionar el segundo.
b. Suponga que el primer elemento no se reemplaza antes de seleccionar el segundo.

6. Aplicación de la regla de la multiplicación Usando los mismos seis elementos del ejercicio 5, calcule la probabilidad de seleccionar al azar tres elementos y obtener uno de color rojo en la primera selección, uno de color verde en la segunda y un elemento azul en la tercera.
a. Suponga que cada elemento se reemplaza antes de que se seleccione el siguiente.
b. Suponga que ninguno de los elementos que se seleccionaron se reemplaza antes de que los otros sean seleccionados.

7. Máscaras antigás defectuosas La revista *Time* reportó que cuando se probaron 19,218 máscaras antigás en divisiones de la milicia de Estados Unidos, se encontró que 10,322 estaban defectuosas (según datos de la Organización Mundial de la Salud). Si una investigación más a fondo comienza por la selección aleatoria de dos máscaras antigás de esta población, calcule la probabilidad de que ambas estén defectuosas.

a. Suponga que la primera máscara antigás se reemplaza antes de seleccionar la segunda.

b. Suponga que la primera máscara antigás no se reemplaza antes de seleccionar la segunda.

c. Compare los resultados que se obtuvieron en *a*) y *b*).

d. Dada la alternativa entre seleccionar con reemplazo y sin reemplazo, ¿cual alternativa es más lógica para tal situación? ¿Por qué?

8. Uso de ropa naranja de cazador Un estudio de heridas de caza en relación con el uso de ropa naranja "de cazador" mostró que de 123 cazadores heridos al confundirlos con presas, seis usaban ropa naranja (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 que se seleccionaron usaran ropa naranja.

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?

9. Probabilidad y adivinar Una profesora de psicología hace un examen sorpresa que consta en 10 reactivos verdadero/falso; establece que para aprobar se requieren al menos siete respuestas correctas. Suponga que un estudiante que no se preparó adopta la cuestionable estrategia de adivinar cada respuesta.

a. Calcule la probabilidad de que las primeras siete respuestas sean correctas y las últimas tres sean incorrectas.

b. ¿La probabilidad del inciso *a*) es igual a la probabilidad de aprobar? ¿Por qué?

10. Selección de senadores en Estados Unidos En el 107o Congreso, el Senado consta de 13 mujeres y 87 hombres. Si un cabildero de la industria del tabaco selecciona al azar a tres diferentes senadores, ¿cuál es la probabilidad de que sean mujeres? ¿Sería probable que un cabildero usara la selección aleatoria en esta situación?

11. Cumpleaños coincidentes

a. El autor nació el 27 de noviembre. ¿Cuál es la probabilidad de que otras dos personas que se seleccionen al azar nacieran también el 27 de noviembre? (Ignore los años bisiestos).

b. ¿Cuál es la probabilidad de que dos personas que se seleccionaron al azar tengan la misma fecha de cumpleaños? (Ignore los años bisiestos).

12. Cumpleaños coincidentes

a. Una pareja atrajo la atención de los medios de comunicación, ya que sus tres hijos, que nacieron en años diferentes, llegaron al mundo el 4 de julio. Ignorando los años bisiestos, calcule la probabilidad de que tres personas seleccionadas al azar nacieran el 4 de julio. ¿Es la probabilidad lo suficientemente baja como para que un suceso como éste no tenga probabilidades de ocurrir, en algún lugar de Estados Unidos, en el transcurso de varios años?

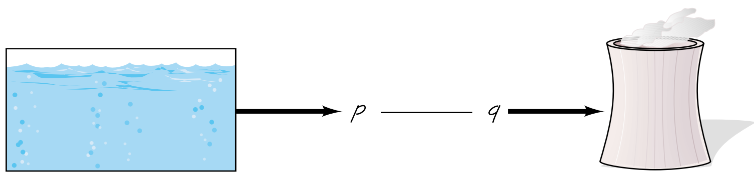
b. Ignore los años bisiestos y calcule la probabilidad de que tres personas que se seleccionen al azar tengan todas la misma fecha de cumpleaños.

13. Muestreo de aceptación Con cierto método de un procedimiento que se llama *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 Elctronics Company acaba de fabricar 5,000 CD, de los cuales el 3% están defectuosos. Si se seleccionan al azar 12 de estos CD para probarlos, ¿cuál es la probabilidad de que se acepte el lote completo?

14. Nivel de confianza de una encuesta En las encuestas de opinión pública es común manejar un "nivel de confianza" del 95%, lo que quiere decir que hay un 0.95 de probabilidad de que los resultados de la encuesta sean precisos dentro de los márgenes de error que se consideró. Si cinco organizaciones diferentes realizan encuestas independientes, ¿cuál es la probabilidad de que las cinco sean precisas dentro de los márgenes de error que se consideraron? ¿Sugiere el resultado que con un nivel de confianza del 95% es posible esperar que casi todas las encuestas estén dentro del margen de error que se consideró?

15. Prueba de efectividad de un método de selección de género Descubrimientos recientes parecen hacer posible que las parejas incrementen, de forma impresionante, la posibilidad de tener un hijo con el género de su elección. En una prueba de un método de selección del género, 10 parejas desean tener niñas. Si este método de selección del género no tuviera efecto, ¿cuál es la probabilidad de que 10 bebés sean todos niñas? Si en realidad resultan 10 niñas en 10 hijos, ¿parece ser efectivo este método de selección de género? ¿Por qué?

16. Confiabilidad de un reactor nuclear En un reactor nuclear se utilizan sensores remotos para controlar cada una de dos válvulas separadas e independientes, que se abren para abastecer agua para enfriamiento en caso de emergencia, las cuales se denotan por p y q. Cada válvula tiene un 0.9968 de probabilidad de abrirse cuando se le dispara. Para la configuración dada, calcule la probabilidad de que cuando ambos sensores se disparen, el agua fluya a través del sistema y ocurra enfriamiento. ¿El resultado es suficientemente alto para considerarse seguro?



17. La excusa de la llanta que se reventó Cuatro estudiantes que perdieron un examen ofrecen una excusa clásica, afirman que se le reventó una llanta al automóvil en el que los cuatro viajaban. En la reposición del examen, el maestro pide a cada uno de los estudiantes que identifique la llanta en particular que se reventó. Si ellos en realidad no tuvieron ninguna avería en los neumáticos, pero seleccionan al azar una llanta que supuestamente se reventó, ¿cuál es la probabilidad de que todos ellos escojan la misma llanta?

18. Identificación de la voz de un criminal En un caso legal en Riverhead, Nueva York, nueve víctimas 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 hubiesen hecho al azar, calcule la probabilidad de que las nueve víctimas seleccionaran a la misma persona. ¿Constituye esto una duda razonable?

19. Control de calidad Una gerente de producción de Telektronics afirma que su nuevo proceso de fabricación de reproductores de CD es mejor porque su tasa de defectos es más baja que el 2%, la tasa de defectos en el pasado. Para fundamentar su afirmación, ella fabrica un lote de 5,000 reproductores de CD, luego selecciona aleatoriamente 15 de ellos para probarlos, con el resultado de que no hay defectos en los 15 reproductores de CD que se seleccionaron. Con base en el resultado, ¿hay suficiente evidencia para fundamentar la afirmación de la gerente de que su nuevo proceso es mejor?

20. Redundancia El principio de la 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.

continúa

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, hay un 0.975 de probabilidad de ser despertados. ¿Cuál es la probabilidad de ser despertado si estamos usando dos relojes despertadores?

Resultados de prueba de embarazo *En los ejercicios 21 a 24 utilice los datos de la tabla 3-1, que se reproduce aquí.*

Tabla 3-1 Resultados de prueba de embarazo

	Resultado de prueba positivo (indicó embarazo)	Resultado de prueba negativo (no indicó embarazo)
La mujer está embarazada	80	5
La mujer no está embarazada	3	11

21. Resultado de prueba positivo Si se seleccionan al azar dos mujeres, calcule la probabilidad de que ambas pruebas den resultado positivo.

22. Embarazo Si se selecciona al azar una de las mujeres, calcule la probabilidad de elegir una que probó negativo o una que no está embarazada.

23. Embarazo Si se seleccionan al azar dos diferentes mujeres, calcule la probabilidad de que ambas estén embarazadas.

24. Resultado de prueba negativo Si se selecciona al azar tres mujeres, calcule la probabilidad de que todas obtuvieran un resultado negativo.

3-4 Más allá de lo básico

25. Las mismas fechas de cumpleaños Calcule la probabilidad de que dos personas no tengan la misma fecha de cumpleaños, cuando el número de personas que se seleccionó al azar es

a. 3 **b.** 5 **c.** 25

26. Género de hijos Si una pareja planea tener ocho hijos, calcule la probabilidad de que todos sean del mismo género.

27. 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.

28. Complementos y la regla de la suma
a. Desarrolle una fórmula para la probabilidad de no obtener A o B, ni ninguno de los dos, en un mismo ensayo. Esto es, calcule una expresión para $P(\overline{A \text{ o } B})$.
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 $P(\overline{A} \text{ o } \overline{B})$.
c. Compare los resultados de los incisos *a* y *b*. ¿Es $P(\overline{A \text{ o } B}) = P(\overline{A} \text{ o } \overline{B})$?

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 que se mezclaron, los reclutas con sífilis se identificaban 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 pareando especímenes 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. Dicha técnica se conoce como *muestreo compuesto*.

3-5 Regla de la multiplicación: complementos y probabilidad condicional

La sección 3-4 introdujo el concepto básico de la regla de la multiplicación; en esta sección extenderemos el uso de tal regla a dos aplicaciones especiales. Comencemos con situaciones en las cuales queremos identificar la probabilidad de que, entre varios ensayos, *uno al menos* dé un resultado que se especifica.

Complementos: La probabilidad de "uno al menos"

La regla de la multiplicación y la regla de los complementos pueden utilizarse juntas para simplificar en gran medida la solución a este tipo de problema: calcular la probabilidad de que entre varios ensayos, uno al menos dé algún resultado que se especificó. En casos como éste, es esencial que el significado del lenguaje se comprenda con claridad:

- "Uno al menos" equivale a "uno o más".
- El complemento de obtener uno al menos, de los elementos de un tipo en particular, es que usted no reciba elementos de ese tipo.

Suponga que una pareja planea tener tres hijos y quiere conocer la probabilidad de al menos una sea niña. Véanse las interpretaciones siguientes:

Una niña al menos entre tres hijos = 1 o más niñas.

El complemento de "una niña al menos" = no niñas = los tres hijos son niños.

Calcularíamos esta probabilidad con facilidad realizando una lista del espacio muestral completo de ocho resultados, pero queremos ilustrar el uso de los complementos, ya que son útiles en muchos otros problemas que no se resuelven tan fácilmente.

EJEMPLO Género de hijos Calcule la probabilidad de que una pareja tenga al menos una niña entre tres hijos. Suponga que los niños y las niñas son igualmente probables, así como 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 uno de los tres hijos es una niña.

Paso 2: Identifique el suceso que es el complemento de A.

$$\begin{aligned}\overline{A} &= \textit{no} \text{ tener al menos una niña entre tres hijos}\\ &= \text{todos los tres hijos son niños}\\ &= \text{niño y niño y niño}\end{aligned}$$

Paso 3: Calcule la probabilidad del complemento.

$$\begin{aligned}P(\overline{A}) &= P(\text{niño y niño y niño})\\ &= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8}\end{aligned}$$

Paso 4: Calcule $P(A)$ evaluando $1 - P(\bar{A})$.

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{1}{8} = \frac{7}{8}$$

INTERPRETACIÓN Existen 7/8 de probabilidad de que si una pareja tiene tres hijos, al menos uno sea una niña.

El principio que se utiliza en este ejemplo se resume como sigue:

Para calcular la probabilidad de *uno al menos* de algo, calcule la probabilidad de *ninguno*, y reste el resultado de 1. Esto es,

$$\boldsymbol{P(\text{uno al menos}) = 1 - P(\text{ninguno}).}$$

Probabilidad condicional

Ahora consideraremos el siguiente punto principal de esta sección, que 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 hombre es de 0.5, pero si ya sabe que la persona a seleccionar cambia con frecuencia los canales de la televisión, con un control remoto, la probabilidad es de 0.999 (bueno, tal vez sea una pequeña exageración). Una *probabilidad condicional* de un suceso ocurre cuando la probabilidad se afecta por el conocimiento de otras circunstancias. La probabilidad condicional de que el suceso B ocurra, puesto que el suceso A ya ocurrió, se calcula usando la regla de la multiplicación [$P(A \text{ y } B) = P(A) \cdot P(B|A)$] y resolviendo para $P(B|A)$, así como dividiendo ambos lados de la ecuación entre $P(A)$.

Definición

Una **probabilidad condicional** de un suceso es una probabilidad que se obtiene con la información adicional de algún otro suceso que ya ocurrió. $P(B | A)$ denota la probabilidad condicional de que el suceso B ocurra, puesto que el suceso A ya ocurrió, y se calcula dividiendo la probabilidad de que ambos sucesos A y B ocurran entre la probabilidad del suceso A:

$$P(B | A) = \frac{P(A \text{ y } B)}{P(A)}$$

Esta fórmula es una expresión formal de la probabilidad condicional, pero recomendamos el siguiente método intuitivo.

Método intuitivo para la probabilidad condicional

La probabilidad condicional de B dado A, se calcula suponiendo que el suceso A ya ocurrió; bajo ese supuesto, se calcula la probabilidad de que el suceso B ocurra.

¿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 fue asesinado en un automóvil Lincoln hecho por la Ford Motor Company. Los sucesores a la presidencia, tanto de Lincoln como de Kennedy, fueron vicepresidentes con apellido Johnson. Catorce años *antes* del naufragio del *Titanic*, una novela describió el hundimiento del *Titán*, un barco que chocó con un iceberg; véase *The Wreck of the Titanic Foretold?*, de Martin Gardner. Gardner señala: “En casi todos los casos de coincidencias desconcertantes, es imposible hacer aunque sea una estimación burda de su probabilidad”.

EJEMPLO Ensayos clínicos de prueba de embarazo Remítase a la tabla 3-1, que se reproduce aquí para su conveniencia.

a. Si se elige aleatoriamente uno de los 99 sujetos, encuentre la probabilidad de que esa mujer pruebe positivo, ya que está embarazada.

b. Si se elige aleatoriamente uno de los 99 sujetos, encuentre la probabilidad de que ella esté embarazada, ya que la prueba resultó positiva.

Tabla 3-1 Resultados de prueba de embarazo

	Resultado de prueba positivo (indicó embarazo)	Resultado de prueba negativo (no indicó embarazo)
La mujer está embarazada	80	5
La mujer no está embarazada	3	11

SOLUCIÓN **a.** Queremos $P(\text{positivo} \mid \text{embarazada})$, la probabilidad de elegir a alguna mujer en quien la prueba fue positiva, *puesto que la mujer que se seleccionó estaba embarazada*. Aquí está el punto relevante: si suponemos que la mujer que se seleccionó estaba embarazada, estamos tratando sólo con las 85 mujeres del primer renglón de la tabla 3-1. De entre estas 85 mujeres, 80 dieron positivo; entonces,

$$P(\text{positivo} \mid \text{embarazada}) = \frac{80}{85} = 0.941$$

Puede calcularse el mismo resultado con la fórmula que se dio con la definición de probabilidad condicional. En los siguientes cálculos, utilizamos el hecho de que 80 de las 99 mujeres estaban embarazadas y dieron positivo. Además, 85 de las 99 mujeres estaban embarazadas. Tenemos

$$\begin{aligned} P(\text{positivo} \mid \text{embarazada}) &= \frac{P(\text{embarazada y positivo})}{P(\text{embarazada})} \\ &= \frac{80/99}{85/99} = 0.941 \end{aligned}$$

b. Aquí buscamos $P(\text{embarazada} \mid \text{positivo})$. Si suponemos que la mujer que se seleccionó dio positivo, estamos tratando con las 83 mujeres de la primera columna de la tabla 3-1. De entre estas 83 mujeres, 80 estaban embarazadas; entonces,

$$P(\text{embarazada} \mid \text{positivo}) = \frac{80}{83} = 0.964$$

Otra vez, se calcula el mismo resultado aplicando la fórmula para la probabilidad condicional:

$$\begin{aligned} P(\text{embarazada} \mid \text{positivo}) &= \frac{P(\text{positivo y embarazada})}{P(\text{positivo})} \\ &= \frac{80/99}{83/99} = 0.964 \end{aligned}$$

LA ESTADÍSTICA EN LAS NOTICIAS

¿Aplicación con resultado anticipado?

¿Se afecta la probabilidad de que lo acepten en una universidad si el aspirante opta por un resultado anticipado? Esta pregunta se trata con métodos de estadística; los resultados son algo sorprendentes. Al escribir acerca de las investigaciones del proceso de admisión a las universidades, la reportera Karen Arenson, del *New York Times*, afirma que "esto no sólo documenta que los estudiantes que hacen el examen pidiendo resultado anticipado tienen una clara ventaja sobre aquellos que no lo hacen —el equivalente de añadir 100 puntos en la calificación del SAT de un aspirante instantáneamente—, sino que también sugiere que el proceso es injusto, ya que muchos estudiantes de preparatorias menos prestigiadas no entienden la manera en que el sistema inclina las posibilidades de aceptación". Ella citó evidencia de 10 universidades con aspirantes que tienen calificaciones del SAT entre 1400 y 1490: al 70% de los estudiantes que solicitaron resultados anticipados se les aceptó, en comparación con el 51% de aceptación de quienes no solicitaron resultados anticipados.

Comparando los resultados de los incisos *a* y *b*, veremos que $P(\text{positivo} \mid \text{embarazada})$ no es lo mismo que $P(\text{embarazada} \mid \text{positivo})$.

INTERPRETACIÓN El primer resultado, de $P(\text{positivo} \mid \text{embarazada}) = 0.941$, indica que una mujer embarazada tiene un 0.941 de probabilidad de dar positivo. Esto sugiere que si una mujer no da positivo, no puede confiar en que no está embarazada, así que tiene que continuar con pruebas adicionales. El segundo resultado, de $P(\text{embarazada} \mid \text{positivo}) = 0.964$, indica que para una mujer que dio positivo, hay un 0.964 de probabilidad de que ella esté realmente embarazada. Una mujer que dio positivo sería inteligente si se sometiera a pruebas adicionales.

Note que, en el ejemplo anterior, $P(\text{positivo} \mid \text{embarazada}) \neq P(\text{embarazada} \mid \text{positivo})$. Aunque los dos valores de 0.941 y 0.964 son muy cercanos en este ejemplo, dichos resultados estarían muy apartados en otros casos. El hecho de creer incorrectamente que $P(B \mid A) = P(A \mid B)$, suele llamarse *confusión del inverso*. Hay estudios que muestran que algunas veces los médicos dan información muy enredada cuando padecen de confusión del inverso.

Prueba de independencia

En la sección 3-4 establecimos que los eventos A y B son independientes si la ocurrencia de uno no afecta la probabilidad de la ocurrencia del otro. En la regla de la multiplicación para sucesos dependientes, si $P(B \mid A) = P(B)$; entonces, la ocurrencia del suceso A no tiene efecto en la probabilidad del suceso B y los dos sucesos A y B son independientes. Esto nos sugiere una prueba de independencia: si $P(B \mid A) = P(B)$, entonces A y B son sucesos independientes; sin embargo, si $P(B \mid A) \neq P(B)$, entonces A y B son sucesos dependientes. Otra prueba de independencia implica revisar si $P(A \text{ y } B)$ y $P(A) \cdot P(B)$ son iguales. Si las expresiones son iguales, los sucesos A y B son independientes. Si $P(A \text{ y } B) \neq P(A) \cdot P(B)$; entonces, A y B son sucesos dependientes. Estos resultados se resumen como sigue:

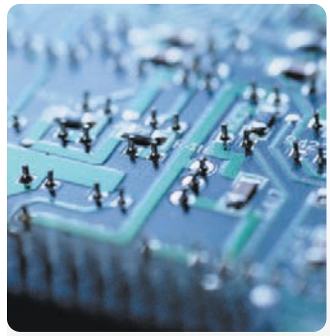
Dos sucesos A y B son *independientes* si	Dos sucesos A y B son *dependientes* si
$P(B \mid A) = P(B)$	$P(B \mid A) \neq P(B)$
o	o
$P(A \text{ y } B) = P(A) \cdot P(B)$	$P(A \text{ y } B) \neq P(A) \cdot P(B)$

3-5 Destrezas y conceptos básicos

Descripción de complementos. *En los ejercicios 1 a 4 haga una descripción escrita del complemento del suceso dado.*

1. Prueba sanguínea Cuando se prueba a 10 estudiantes para determinar su grupo sanguíneo, uno al menos tiene sangre del grupo A.

2. Control de calidad Cuando se envían 50 unidades de HDTV, todas están libres de defectos.



Teorema de Bayes

Thomas Bayes (1702-1761) dijo que las probabilidades deben revisarse cuando averiguamos más sobre un suceso. He aquí una forma del teorema de Bayes:

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(A) \cdot P(B \mid A) + P(\overline{A}) \cdot P(B \mid \overline{A})}$$

Suponga que el 60% de los circuitos integrados para computadora de una compañía se producen en una de sus fábricas (denotada por A) y el 40% en su otra fábrica (denotada por $\overline{A}$). Para un circuito integrado que se selecciona al azar, la probabilidad de que provenga de la fábrica A es de 0.60. Suponga además que se entera de que un circuito integrado está defectuoso y que las tasas de defectos para las dos fábricas son del 35% (para A) y del 25% (para $\overline{A}$). Se utiliza la fórmula anterior para determinar que hay una probabilidad de 0.677 de que el circuito integrado defectuoso provenga de la fábrica A.

3. Auditorías del fisco Cuando un oficial del Tesoro selecciona 12 devoluciones de impuestos de ingresos y les hace auditoría, encuentra que ninguna de las devoluciones es correcta.

4. Éxito con las mujeres Cuando Mike pidió una cita a cinco mujeres diferentes, al menos una aceptó.

5. Probabilidad subjetiva Utilice la probabilidad subjetiva para estimar la probabilidad de seleccionar un adulto al azar y obtener una mujer, puesto que la persona que se seleccionó tiene el pelo más largo por 10 pulgadas. ¿Es la probabilidad suficientemente alta como para presumir que alguien con pelo largo casi con seguridad es mujer?

6. Probabilidad subjetiva Utilice la probabilidad subjetiva para estimar la probabilidad de seleccionar un adulto al azar y obtener un hombre, puesto que la persona que se seleccionó es dueña de una motocicleta. Si un investigador de crímenes encuentra que una motocicleta se registró a nombre de Pat Ryan, ¿será razonable creer que Pat es hombre?

7. Probabilidad de al menos una niña Si una pareja planea tener cinco hijos, ¿cuál es la probabilidad de que tenga al menos una niña? ¿Es dicha probabilidad lo suficientemente alta como para que la pareja sienta mucha confianza de que nacerá al menos una niña en cinco hijos?

8. Probabilidad de al menos una niña Si una pareja planea tener 12 hijos (llega a suceder), ¿cuál es la probabilidad de que nazca al menos una niña? Si la pareja eventualmente tuvo 12 hijos y todos fueron niños, ¿qué concluiría la pareja?

9. Al menos una multa de tránsito Si se pasa en un cruce que se equipó con una cámara de vigilancia, con la luz del semáforo en rojo, hay un 0.1 de probabilidad de recibir una multa de tránsito. Si usted se pasa este cruce cinco veces diferentes con la luz del semáforo en rojo, ¿cuál es la probabilidad de recibir al menos una multa de tránsito?

10. Al menos una respuesta correcta Si usted adivina al azar las respuestas a cuatro preguntas de opción múltiple (cada una con cinco respuestas posibles), ¿cuál es la probabilidad de tener al menos una correcta? Si un maestro poco exigente dice que para aprobar el examen es necesario al menos tener una respuesta correcta, ¿puede usted esperar razonablemente aprobar adivinando?

11. Probabilidad de una niña Calcule la probabilidad de que una pareja tenga una niña cuando nace su tercer hijo, puesto que los primeros dos hijos fueron niños. ¿Es el resultado igual a la probabilidad de que nazcan tres niñas entre tres hijos?

12. Genética de Mendel Remítase a la figura 3-3 en la sección 3-3, que incluye los chícharos que se usaron en un experimento genético. Si se selecciona aleatoriamente uno de los chícharos y se encuentra que tiene vaina verde, ¿cuál es la probabilidad de que tenga flor morada?

13. Pruebas clínicas de embarazo Remítase a la tabla 3-1 y suponga que una de las mujeres se selecciona al azar. Calcule la probabilidad de un resultado de prueba negativo, puesto que la mujer que se seleccionó no está embarazada. ¿Qué debe hacer una mujer si le aplican esta prueba de embarazo y obtiene un resultado negativo?

14. Pruebas clínicas de embarazo Remítase a la tabla 3-1 y suponga que una de las mujeres se selecciona al azar. Calcule la probabilidad de que la mujer que se seleccionó no esté embarazada, ya que la prueba indicó negativo. ¿El resultado es igual a la probabilidad de un resultado de prueba negativo, ya que la mujer que se seleccionó no está embarazada?

15. Redundancia en relojes despertadores Un estudiante pierde muchas clases por el mal funcionamiento de los relojes despertadores. En lugar 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 uno, por separado, tiene un 99% de probabilidad de funcionar correctamente? ¿En realidad gana mucho el estudiante usando tres relojes despertadores en lugar de uno solo?

16. Muestreo de aceptación Con un método del procedimiento que llaman *muestreo de aceptación*, se selecciona aleatoriamente y sin reemplazo una muestra de artículos. Tome en cuenta que el lote completo se rechazará si se encuentra al menos un defecto. La Niko Electronics Company acaba de fabricar 5,000 CD, de los cuales el 3% están defectuosos. Si se seleccionan 10 de los CD y se prueban, ¿cuál es la probabilidad de que se rechace el lote completo?

17. 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 especímenes de sangre. Así, si las muestras de tres personas se combinan y la mezcla da un resultado negativo, 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 positivo es de 0.1 (la probabilidad de la población "en riesgo" de acuerdo con datos del Departamento de Salud del estado de Nueva York).

18. 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*. Con la finalidad de reducir costos de laboratorio, se mezclan las muestras de agua de seis áreas de natación públicas para efectuar una sola prueba, la cual sólo se hará más amplia si la muestra que se mezcla falla. Con base en resultados previos, hay un 2% de probabilidad de encontrar la bacteria *E. coli* en un área de natación pública. Calcule la probabilidad de que una muestra que se combinara de seis áreas de natación públicas revele la presencia de la bacteria *E. coli*.

Probabilidades condicionales. *En los ejercicios 19 a 22 use los datos de mortalidad que hubo en el* Titanic *de la tabla adjunta.*

	Hombres	Mujeres	Niños	Niñas
Sobrevivientes	332	318	29	27
Muertos	1360	104	35	18

19. Si seleccionamos aleatoriamente a una persona que abordó el *Titanic*, ¿cuál es la probabilidad de elegir un hombre, puesto que la persona que se seleccionó murió?

20. Si seleccionamos aleatoriamente a una persona que murió, ¿cuál es la probabilidad de elegir a un hombre?

21. ¿Cuál es la probabilidad de elegir a un niño o una niña, puesto que la persona que se seleccionó al azar es alguien que sobrevivió?

22. ¿Cuál es la probabilidad de elegir a un hombre o una mujer, ya que la persona seleccionada aleatoriamente es alguien que murió?

3-5 Más allá de lo básico

23. 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 se sube a esta montaña rusa una vez, ¿cuál es la probabilidad de obtener el tan codiciado lugar de hasta adelante? ¿Cuántas veces habrá que subirse para tener un mínimo del 95% de probabilidad de que le toque el asiento delantero al menos una vez?

24. ¿Quién fue? La planta en Atlanta de la Medassist Pharmaceutical Company fabricó 400 marcapasos, de los cuales tres están defectuosos. La planta en Baltimore de la misma compañía fabricó 800 marcapasos, dos de los cuales salieron defectuosos. Si se selecciona al azar uno de los 1,200 marcapasos y se encuentra que está defectuoso, ¿cuál es la probabilidad de que lo hayan fabricado en Atlanta?

25. Uso de una tabla de dos entradas El Departamento de Salud del estado de Nueva York reporta una tasa de incidencia de VIH del 0.3% para la población general; en ciertas condiciones, las pruebas de investigación preliminar para el VIH son correctas un 95% de las veces (para verdaderos positivos y verdaderos negativos). Suponga que la población general conste de 100,000 personas.
  a. Construya una tabla con un formato similar al de la tabla 3-1.
  b. Usando la tabla del inciso *a* calcule $P(\text{VIH} \mid \text{positivo})$ para una persona seleccionada aleatoriamente de la población general. Esto es, calcule la probabilidad de seleccionar al azar a una persona con VIH, ya que esa persona dio positivo.

26. Fecha de cumpleaños compartida Calcule la probabilidad de que, de 25 personas que se seleccionaron al azar,
  a. No haya dos que compartan la misma fecha de cumpleaños.
  b. Al menos dos compartan la misma fecha de cumpleaños.

27. Monedas ocultas Un profesor de estadística lanza dos monedas que ningún estudiante logra 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.

3-6 Probabilidades por medio de simulaciones

Los estudiantes que toman un curso introductorio de estadística suelen encontrar que el tema de la probabilidad es el más difícil. Algunos problemas de probabilidad pueden parecer simples, pero sus soluciones son increíblemente complejas. En este capítulo ya identificamos varias reglas básicas e importantes, las cuales suelen usarse para calcular probabilidades, pero en esta sección introducimos un enfoque muy diferente, que logra vencer gran parte de la dificultad que se encuentra en la aplicación de reglas formales. Este enfoque alternativo consiste en desarrollar una simulación.

Definición

Simulación de un procedimiento: proceso que se comporta de la misma forma que el procedimiento, de manera que se producen resultados similares.

Considere los ejemplos siguientes para entender mejor el uso de la simulación.

EJEMPLO Selección del género Cuando se prueban técnicas de selección del género, los investigadores médicos 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 hombre o de una mujer es igualmente probable, describa una simulación que dé como resultado los géneros de 100 bebés recién nacidos.

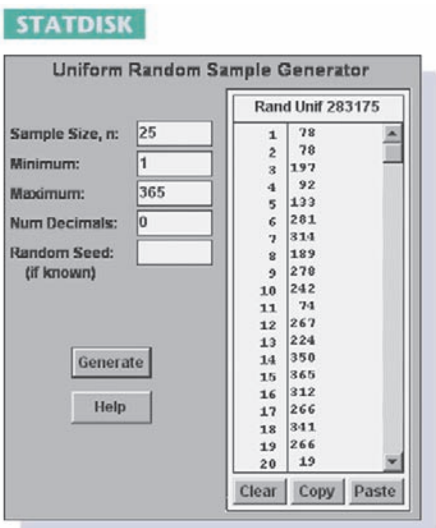
SOLUCIÓN Una opción es simplemente lanzar una moneda al aire 100 veces, con la cara representando a las mujeres, y la cruz, a los hombres. Otra opción es usar una calculadora o computadora para generar aleatoriamente ceros y unos (el 0 representa a un niño y el 1 representa a una niña). Los números deben generarse de forma que sean igualmente probables.

EJEMPLO Las mismas fechas de cumpleaños El ejercicio 26 en la sección 3-5 se refiere al clásico problema de fecha de cumpleaños, en el que encontramos la probabilidad de que, en un grupo 25 personas que se seleccionaron al azar, al menos dos compartan 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.

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, todos entre 1 y 365. Estos números pueden ordenarse, ya que así es fácil estudiar la lista para determinar si dos de las fechas de cumpleaños que se simularon son iguales. Es posible repetir el proceso tantas veces como queramos, hasta quedar 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 de 1 a 365 generados aleatoriamente, incluyendo la siguiente:

- **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 de 1 a 365 en este tipo de tablas. Una consiste en tomar los dígitos en las primeras columnas, ignorando 000 y cualquier número mayor a 365).
- **STATDISK:** Seleccione **Data** de la barra del menú principal, luego seleccione **Uniform Generator**; después, 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 reproduce abajo. Use **copy/paste**, para copiar el conjunto de datos al **Sample Editor**, donde los valores pueden acomodarse en orden creciente. En la pantalla del STATDISK que se presenta aquí, vemos que las primeras dos personas tienen la misma fecha de cumpleaños, que es el día 78 del año.



Monos mecanógrafos

Una aseveración clásica dice que un mono que golpea al azar un teclado, tarde que temprano produciría las obras completas de Shakespeare, suponiendo que continúe tecleando siglo tras siglo. Se utilizó ya la regla de la multiplicación para probabilidades con la finalidad de obtener estimados de esta clase. Algunos consideran muy pequeño un resultado de 1,000,000,000,000,000,000,000,000,000,000,000,000 años. Con algo similar en mente, sir Arthur Eddington escribió este poema: "Había una vez un sesudo babuino, que soplaba y soplaba un fagot; 'pues estoy convencido —decía—, de que en miles de millones de años, si sigo, al soplar me saldrá una canción'".

- **Minitab:** Seleccione **Calc** en la barra del menú principal; después, seleccione **Random Data** e **Integer**. En el cuadro de diálogo, introduzca 25 para el número de filas, guarde los resultados en la columna C1; tras esto, ingrese un mínimo de 1 y un máximo de 365. Entonces ya está listo para usar **Manip** y **Sort** para acomodar los datos en orden creciente. El resultado se verá como se muestra abajo, pero los números no serán los mismos. Este resultado del Minitab de 25 números indica que el 9o y el 10o son iguales.

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:** Haga clic en la celda que se encuentra en la esquina superior izquierda, después haga clicen el icono de función **fx.** Seleccione **Matemáticas y Trigonometría**, y después seleccione **RANDBETWEEN**. En el cuadro de diálogo, escriba 1 para el límite inferior (*bottom*) y ponga 365 como límite

Excel

	A	
1	15	
2	3	
3	15	
4	362	
5	164	
6	184	
7	158	
8	59	
9	143	
10	85	
11	134	

LA ESTADÍSTICA EN LAS NOTICIAS

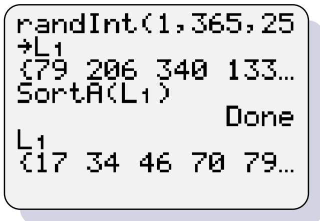
Para ganar, apueste con audacia

El diario *New York Times* publicó un artículo de Andrew Pollack en el cual se reportó que el casino Mirage en Las Vegas tenía menores ganancias que las que se esperaban. Él escribió que "las ganancias del Mirage pueden ser particularmente volátiles, ya que se favorece a los grandes apostadores, jugadores que llegan a 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, ¡ponga una apuesta grande en lugar de muchas apuestas chicas! Con el juego adecuado, por ejemplo el de dados, usted tiene poco menos del 50% de posibilidades de duplicar su dinero si se anima a una apuesta grande. Al hacer muchas apuestas pequeñas, la probabilidad de duplicar su dinero disminuye sustancialmente.

superior (*top*). Después de obtener el número aleatorio en la primera celda, haga clic y mantenga presionado el botón para arrastrar la esquina inferior derecha de la primera celda, luego arrástrela hacia abajo hasta resaltar 25 celdas. Cuando suelte el botón, deben aparecer los 25 números aleatorios. La pantalla que se reproduce aquí, indica que los números 1o y 3o son iguales.

- **Calculadora TI-83 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. Esto es, introduzca randInt(1,365,25). Observe la pantalla de la TI-83 Plus, la cual indica el uso de **randInt** para generar los números; luego se guardaron en la lista L1, donde se sortearon y se mostraron. La imagen de la pantalla que se reproduce aquí indica que no hay números iguales entre los pocos que se alcanzan a ver. Para observar la lista completa de números generados, oprima **STAT** (Estadística) y seleccione **Edit.**

TI-83 Plus



Es en extremo importante construir una simulación que se comporte precisamente como el procedimiento real. En el siguiente ejemplo demostramos una forma correcta y una forma incorrecta de construir una simulación.

EJEMPLO Simulación de dados Describa un procedimiento para simular el acto de tirar un par de dados.

SOLUCIÓN En el procedimiento de tirar un par de dados, cada uno de los dos dados nos dará un número entre 1 y 6 (inclusive); estos dos números se suman. Cualquier simulación debe efectuar lo mismo. Hay una manera correcta y una incorrecta de simular un tiro de dos dados.

La manera correcta: Generar aleatoriamente un número entre 1 y 6, generar aleatoriamente otro número entre 1 y 6; 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 dan 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 facilita mucho el uso de métodos de simulación, tanto que ahora las simulaciones se emplean con frecuencia para determinar valores de probabilidad.

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 aleatoriamente 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 mucho más allá del alcance de este texto.

3-6 Destrezas y conceptos básicos

En los ejercicios 1 a 8 utilice la lista de números que se generaron aleatoriamente y se encuentra al margen. (Se obtiene una lista similar usando calculadoras, computadoras, resultados de la lotería o tablas especiales de números aleatorios).

1. Simulación de respuestas por adivinación Suponga que usted quiere usar los dígitos de la lista adjunta para simular adivinaciones en un examen de verdadero/falso. Si un dígito impar representa "verdadero" y un dígito par representa "falso", haga una lista de cinco respuestas correspondientes con la primera fila de dígitos.

2. Simulación de dados Suponga que usted quiere usar los dígitos de la lista adjunta para simular el hecho de tirar un solo dado. Si se usan los dígitos 1, 2, 3, 4, 5, 6, mientras se ignoran todos los demás, haga una lista de los resultados que se obtuvieron con las primeras dos filas.

46196
99438
72113
44044
86763
00151
64703
78907
19155
67640
98746
29910
82855
25259
14752
85446
75260
92532
87333
55848

3. Simulación de fabricación La compañía Telektronic está experimentando con un nuevo proceso de fabricación de fusibles, en la cual la tasa de defectos es del 20%. Es posible simular fusibles defectuosos usando 0 y 1, mientras que 2, 3, 4, 5, 6, 7, 8, 9 representan fusibles en buen estado (de manera que se consideran el 20% de defectos). Identifique los fusibles aceptables y defectuosos que corresponden con la primera fila de dígitos.

4. Simulación de fechas de cumpleaños En un ejemplo de esta sección se señaló que las fechas de cumpleaños pueden simularse generando enteros entre 1 y 365. Si usamos entradas en una lista de dígitos aleatorios, se representa el 1 de enero como 001, el 2 de enero como 002, . . . , y el 31 de diciembre como 365. Todos las demás ternas de dígitos deben ignorarse. Utilizando este método, la primera fila nos da la fecha de cumpleaños válida de 196. Haga una lista de las siguientes cinco fechas de cumpleaños que se logran obtener de esta forma.

5. Simulación de familias con cinco hijos Use los dígitos aleatorios al margen para desarrollar una simulación con la finalidad de calcular la probabilidad de tener al menos dos niñas en una familia de cinco hijos. Describa la simulación y luego estime la probabilidad con base en sus resultados. ¿De qué forma es posible comparar el resultado con la cifra correcta de 0.813? (*Sugerencia:* Haga que los dígitos impares representen niñas).

6. Simulación de tres dados Use los dígitos aleatorios al margen y desarrolle una simulación para el tiro de tres dados. Describa la simulación, luego proceda a estimar la probabilidad de obtener un total de 10 cuando se tiran tres dados. ¿El resultado es comparable con el resultado correcto de 0.125? (*Sugerencia:* Use sólo los dígitos 1, 2, 3, 4, 5, 6, e ignore todos los demás dígitos).

7. Simulación de zurdos El 10% de las personas son zurdas. En un estudio de destreza, se seleccionan al azar grupos de cinco. Utilice los dígitos aleatorios al margen y desarrolle una simulación para calcular la probabilidad de obtener al menos una persona zurda en un grupo de cinco. ¿El resultado es comparable con el resultado correcto de 0.410, que se puede calcular usando las reglas de probabilidad de este capítulo? (*Sugerencia:* Puesto que el 10% de las personas son zurdas, deje que el dígito 0 represente a alguien que es zurdo y que los otros dígitos representen a alguien que no es zurdo).

8. Simulación de hibridación Cuando Mendel realizó sus famosos experimentos de hibridación, utilizó chícharos con vainas verdes y vainas amarillas. Un experimento incluyó la mezcla de chícharos de tal forma que se esperaba que el 25% de los chícharos vástagos tuvieran vainas amarillas. Use los dígitos aleatorios al margen y desarrolle una simulación para calcular la probabilidad de que cuando se produzcan dos chícharos vástagos, al menos uno de ellos contenga vainas amarillas. ¿El resultado es comparable con la probabilidad correcta de 7/16, que puede calcularse usando las reglas de probabilidad de este capítulo? (*Sugerencia:* Puesto que se espera que el 25% de los vástagos tengan vainas amarillas y el 75% tengan vainas verdes, haga que el dígito 1 represente vainas amarillas y que los dígitos 2, 3, 4 representen vainas verdes; ignore cualquier otro dígito).

T *En los ejercicios 9 a 12 desarrolle una simulación utilizando la calculadora TI-83 Plus, el STATDISK, el Minitab, Excel, o cualquier otro programa o calculadora adecuados.*

9. Simulación de familias de cinco hijos En el ejercicio 5 utilizamos los dígitos al margen para estimar la probabilidad de tener al menos dos niñas en una familia de cinco hijos. En lugar de usar los mismos dígitos, desarrolle su propia simulación para calcular la probabilidad de que haya al menos dos niñas en una familia de cinco hijos. Simule 100 familias. Describa la simulación y después estime la probabilidad con base en los resultados.

10. Simulación de tres dados En el ejercicio 6 utilizamos los dígitos al margen para simular el tiro de dados. En lugar de usar los mismos dígitos, desarrolle su propia simulación para tirar tres dados. Simule tirar tres dados 100 veces. Describa la simulación y maneje el resultado para estimar la probabilidad de tener un total de 10 cuando se tiran tres dados.

11. Simulación de zurdos En el ejercicio 7 utilizamos los dígitos al margen para simular personas que son zurdas o diestras. (El 10% de las personas son zurdas). Desarrolle una simulación para calcular la probabilidad de obtener al menos una persona zurda en un grupo de cinco. Simule 100 grupos de cinco.

12. Simulación de hibridación En el ejercicio 8 utilizamos los dígitos al margen como base para simular la hibridación de chícharos. Suponga otra vez que se espera que el 25% de los chícharos vástagos tengan vainas amarillas, pero desarrolle su simulación y genere 100 pares de vástagos. Con base en sus resultados, estime la probabilidad de tener al menos un chícharo con vainas verdes cuando se obtienen dos chícharos vástagos.

3-6 Más allá de lo básico

13. Simulación del problema de Monty Hall Un problema que ha atraído gran atención es el problema de Monty Hall, que se inspiró en el antiguo programa de concurso de televisión “Let’s Make a Deal”, que presenta Monty Hall. Suponga que usted es un concursante que eligió una de tres puertas, después de que se le dijo que detrás de dos de ellas no hay nada, pero que detrás de una de las tres está un Corvette rojo último modelo. Entonces, se le da la opción de quedarse con su primera selección o cambiarla. ¿Debe quedarse con su primera elección o le conviene cambiar? (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 relacionarse con la toma de decisiones).

14. Simulación de fechas de cumpleaños

a. Elabore una simulación para calcular la probabilidad de que, cuando 50 personas se seleccionan al azar, al menos dos de ellas tengan la misma fecha de cumpleaños. Describa la simulación y estime la probabilidad.

a. Elabore una simulación para calcular la probabilidad de que, cuando 50 personas se seleccionan al azar, al menos tres de ellas tengan la misma fecha de cumpleaños. Describa la simulación y estime la probabilidad.

15. Genética: simulación de control poblacional Un clásico problema de probabilidad se refiere a un rey que quería incrementar la proporción de mujeres decretando que después de que una madre diera a luz a un hijo hombre, se le prohibiera tener más hijos. El rey razona que algunas familias tendrán sólo un varón, mientras que en otras habrá unas cuantas mujeres y un hombre, luego de lo cual la proporción de niñas se incrementará. ¿Es correcto su razonamiento? ¿Se incrementará la proporción de niñas?

3-7 Conteo

¿Cuál es la probabilidad de que usted gane en la lotería? En la lotería de Maine, una lotería típica, usted debe escoger seis números entre 1 y 42, inclusive. Si elige la misma combinación de seis números que los oficiales de la lotería sacan al azar, ganará el premio mayor, que a veces es de millones de dólares. Hay algunos premios menores, pero son relativamente insignificantes. Utilizando el enfoque clásico para la probabilidad (puesto que los resultados son igual de probables), la probabilidad de ganar la lotería se encuentra usando $P(\text{ganar}) = s/n$, donde s es el número de formas en que usted puede ganar y n es el número total de resultados posibles. Con la lotería de Maine $s = 1$, puesto que sólo existe una manera de ganar el premio mayor: escoger la misma combinación de seis números que se saca en la lotería. Sabiendo que sólo hay una manera de ganar, ahora necesitamos calcular n, el número total de resultados, es decir, ¿cuántas combinaciones de seis números son posibles cuando selecciona números de 1 a 42? Escribir una lista de las posibilidades tomaría alrededor de un año de trabajo sin parar; además, ese método no le dejaría tiempo para estudiar estadística. Necesitamos una manera más práctica de calcular el número total de posibilidades. Esta sección introduce métodos eficientes para calcular números de ese tipo, sin hacer listas directamente y contar las posibilidades. Regresaremos a dicho problema de la lotería después de presentar algunos principios básicos. Comencemos por la *regla fundamental de conteo*.

Regla fundamental de conteo

Para una secuencia de dos sucesos en la que el primer suceso puede ocurrir de m formas y el segundo suceso 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 impliquen más de dos eventos, como se ilustra en los ejemplos siguientes.

EJEMPLO Las bases del robo Los sistemas comunes de alarma para casas tienen un código que consta de cuatro dígitos. Los dígitos (0 hasta 9) pueden estar repetidos, aunque deben ingresarse en el orden correcto. Suponga que usted planea tener acceso intentando códigos hasta encontrar el correcto. ¿Cuántos códigos diferentes son posibles?

SOLUCIÓN Hay 10 valores posibles para cada uno de los cuatro dígitos; entonces, el número de códigos posibles distintos es de $10 \cdot 10 \cdot 10 \cdot 10 = 10{,}000$. Aunque los 10,000 códigos pueden intentarse en alrededor de 11 horas, los sistemas de alarma normalmente se diseñaron para que el sistema rechace intentos subsecuentes después de unas cuantas entradas incorrectas. Además de los problemas morales y legales de ser ladrón profesional, parece que hay un asunto matemático que también sugiere hacer otra carrera.

EJEMPLO **Cotinina en fumadores** El conjunto de datos 6 del Apéndice B lista niveles de cotinina que se midieron en una muestra de personas de cada uno de los 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 el promedio del nivel de cotinina de cada uno de los tres grupos y luego acomodamos dichos promedios en orden creciente, obtendremos 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 se incrementa a medida que la exposición y el uso del tabaco 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 tener la secuencia NEF? ¿Es la probabilidad lo suficientemente baja como para concluir que la secuencia NEF indica que la presencia de cotinina se incrementa a medida que la exposición y el uso del tabaco 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 aleatoriamente una de las seis secuencias posibles, la probabilidad de obtener la secuencia NEF es de $1/6$. Puesto que la probabilidad de $1/6$ es relativamente alta, sabemos que la secuencia NEF puede ocurrir con facilidad por posibilidad. La probabilidad no es suficientemente baja como para concluir que la secuencia NEF indique que la presencia de cotinina se incrementa a medida que la exposición y uso del tabaco también lo hacen. 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 específica se generaliza utilizando la siguiente notación para el símbolo ! y la siguiente *regla factorial*.

¿Cuántas veces hay que barajar?

Después de 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 acomodos posibles de los naipes son igualmente probables. Barajar más de siete veces no surtirá un efecto significativo, en tanto que 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 se mezclan totalmente. Algunos jugadores expertos aprovechan las mezclas incompletas que resultan de barajar menos de siete veces.

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, $0! = 1$. (Muchas calculadoras traen una tecla factorial. En la calculadora TI-83 Plus, ingrese primero el número, luego presione **MATH**; luego, seleccione **PRB** Y el elemento 4 del menú; por último, presione la tecla **ENTER**).

Elección de códigos de seguridad

Utilizamos códigos de seguridad personales para tener acceso a máquinas contestadoras telefónicas, cuentas de Internet de computadora y sistemas de seguridad para casas, entre otros sistemas. La seguridad de tales códigos depende del gran número de posibilidades diferentes, aunque ahora los piratas informáticos cuentan con sofisticadas herramientas capaces de superar este obstáculo con creces. Los investigadores encontraron ya que usando variaciones del nombre y apellido del usuario, además de otros 1,800 nombres, identificarían 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 palabra más corta que siete caracteres, ni números telefónicos, ni números del Seguro Social. Incluya caracteres no alfabéticos, como números o símbolos de puntuación.

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 de $n - 1$ maneras y así sucesivamente).

Los problemas de ruta con frecuencia implican la aplicación de la regla factorial. AT&T quiere hacer sus llamadas telefónicas a través de las redes más cortas. Federal Express desea encontrar las rutas más cortas para sus entregas. American Airlines busca la ruta más corta para regresar a los miembros de la tripulación a sus casas. Véase el siguiente ejemplo.

EJEMPLO Rutas a 50 capitales Por su éxito en el curso de estadística, a usted lo contrató la organización Gallup. Su primer trabajo consiste en realizar una encuesta en cada una de las 50 capitales de los estados de Estados Unidos. Como usted se encuentra planeando su ruta de viaje, quiere determinar el número de rutas diferentes posibles. ¿Cuántas rutas diferentes son posibles?

SOLUCIÓN Aplicando la regla factorial sabemos que 50 elementos pueden acomodarse en 50! formas diferentes. Es decir, a las 50 capitales de estado es posible acomodarlas en 50! formas, o sea

$$30{,}414{,}093{,}201{,}713{,}378{,}043{,}612{,}608{,}166{,}064{,}768,$$
$$844{,}377{,}641{,}568{,}960{,}512{,}000{,}000{,}000{,}000$$

Ahora sí tenemos un número grande.

El ejemplo anterior es una variación del clásico problema que se conoce como *problema del vendedor viajero*, que es especialmente interesante, pues el gran número de posibilidades existentes significa que no estamos en condiciones de utilizar una computadora para calcular la distancia de cada ruta. El tiempo que tomaría calcular la ruta más corta, aun con la computadora más rápida, es de alrededor de

$$1{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000{,}000 \text{ siglos.}$$

Constantemente se dedican esfuerzos considerables para tratar de encontrar métodos eficientes que resuelvan problemas de este tipo.

De acuerdo con la regla factorial, n diferentes elementos pueden ser acomodados de $n!$ diferentes maneras. Algunas veces tenemos n elementos diferentes, pero necesitamos seleccionar sólo algunos de ellos en lugar de todos. Si hay que realizar encuestas en capitales estatales, como en el ejemplo anterior, pero sólo tenemos tiempo de visitar cuatro capitales, el número de posibles rutas diferentes es de $50 \cdot 49 \cdot 48 \cdot 47 = 5{,}527{,}200$. Otra forma de obtener este mismo resultado es evaluar

$$\frac{50!}{46!} = 50 \cdot 49 \cdot 48 \cdot 47 = 5{,}527{,}200$$

En este cálculo, advierta que al dividir el número factorial del numerador entre el número factorial del denominador, sólo permanecen los factores de 50, 49, 48 y 47.

Se generaliza este resultado observando que si tenemos n elementos disponibles diferentes y queremos seleccionar un número r de ellos, el número de combinaciones es de $n!/(n-r)!$ como en $50!/46!$ Dicha generalización se conoce comúnmente como *regla de las permutaciones*.

Regla de las permutaciones (cuando todos los elementos son diferentes)

El número de **permutaciones** (o secuencias) de r elementos que se seleccionan entre n elementos disponibles (sin reemplazo) es

$$_nP_r = \frac{n!}{(n-r)!}$$

Muchas calculadoras logran evaluar expresiones de $_nP_r$.

Es muy importante reconocer que la regla de las permutaciones requiere las siguientes condiciones:

- Debemos tener un total de *n diferentes* elementos disponibles. (Esta regla no se aplica si algunos de los elementos son idénticos a otros).
- Debemos seleccionar r entre los n elementos (sin reemplazo).
- Debemos considerar que los reacomodamientos de los mismos elementos son secuencias diferentes.

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 cuentan como secuencias distintas. Las letras ABC se pueden acomodar de seis formas distintas: ABC, ACB, BAC, BCA, CAB y CBA. (Más adelante nos referiremos a las combinaciones, en las cuales tales acomodos no se consideran distintos). En el ejemplo siguiente se nos pide calcular el número total de secuencias distintas posibles. Eso sugiere el uso de la regla de las permutaciones.

EJEMPLO Programación de televisión Usted acaba de ser contratado para conformar la programación de la cadena de televisión Fox. Cuando está seleccionando los programas a transmitir el lunes por la noche, encuentra que tiene 27 programas disponibles y que debe seleccionar cuatro de ellos. El orden de los programas es importante, por los efectos de liderazgo. ¿Cuántas secuencias diferentes de cuatro programas son posibles cuando hay 27 programas disponibles?

SOLUCIÓN Necesitamos seleccionar $r = 4$ programas de $n = 27$ que están disponibles.

$$_nP_r = \frac{n!}{(n-r)!} = \frac{27!}{(27-4)!} = 421{,}200$$

Hay 421,200 arreglos posibles diferentes de cuatro programas que se seleccionaron de entre 27 disponibles.

Ganar centavos en la lotería

Muchas personas gastan grandes cantidades de dinero comprando billetes de lotería, a pesar de 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 correctamente 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 que alcanzan esos jets de altos vuelos; además, 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.

Escasez de números telefónicos

Las compañías telefónicas con frecuencia separan regiones con un código de área por regiones 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 un número telefónico de siete dígitos no es posible que 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, los equipos de fax y la Internet, todos los números gratuitos tenían un prefijo de 800. Estos números 800 duraron 29 años antes de que todos se asignaran. Se introdujo el prefijo 888 con el objetivo de contribuir a satisfacer la demanda de números gratuitos, pero se estimó que pasarían sólo 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 satisfacer las demandas del futuro.

En ocasiones necesitamos calcular el número de permutaciones cuando algunos de los elementos son idénticos a otros. La siguiente variación de la regla de las permutaciones se aplica a tales casos.

Regla de las permutaciones (cuando algunos elementos son idénticos a otros)

Si hay n elementos con n_1 iguales, n_2 iguales, . . . , n_k iguales, el número de permutaciones de los n elementos es

$$\frac{n!}{n_1!n_2! \cdots n_k!}$$

EJEMPLO Invertir en acciones Los ejemplos clásicos de la regla de permutaciones son aquellos que muestran, por ejemplo, que las letras de la palabra Mississippi es posible acomodarlas de 34,650 formas diferentes o que las letras de la palabra estadística lo serían de 2,494,800 formas. En lugar de ello, consideraremos las letras *DDDDDEEEE*, que representan una secuencia de años recientes en los que el promedio industrial Dow Jones estuvo por debajo (*D*) de la media o por encima (*E*) de la media. ¿De cuántas maneras se acomodarían las letras *DDDDDEEEE*? ¿Parece que la secuencia es aleatoria? ¿Existe un patrón que sugiera que sería prudente invertir en acciones?

SOLUCIÓN En la secuencia *DDDDDEEEE* tenemos $n = 9$ elementos, con $n_1 = 5$ iguales y otros $n_2 = 4$ iguales. El número de permutaciones se calcula como sigue:

$$\frac{n!}{n_1!n_2!} = \frac{9!}{5!4!} = \frac{362{,}880}{2880} = 126$$

Hay 126 diferentes formas en que las letras *DDDDDEEEE* pueden acomodarse. Ya que hay 126 diferentes posibles arreglos y sólo dos de ellos (*DDDDDEEEE* y *EEEEDDDDD*) resultan en todas las letras agrupadas juntas, parece que la secuencia no es aleatoria. Puesto que todos los valores por *debajo* suceden al principio y todos los valores por *encima* al final, parece haber un patrón de incremento del valor de las acciones. Esto sugiere que sería prudente invertir en acciones. (Véase en la sección 13-7 la *prueba de rachas para detectar aleatoriedad*, un procedimiento formal que se usa con frecuencia para identificar tendencias económicas).

El ejemplo anterior comprende n elementos, cada uno perteneciente a una de dos categorías. Cuando sólo hay dos categorías, es posible estipular que x de los elementos son iguales y que los otros $n-x$ elementos también son iguales; entonces, 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 4-3.

Cuando tenemos la intención de seleccionar r elementos de entre n elementos diferentes, *sin tomar en cuenta el orden*, nos preocupan en realidad las combinaciones posibles, más que las permutaciones. Es decir, **cuando diferentes ordenamientos de los mismos elementos cuentan por separado, tenemos un problema de permutaciones, pero cuando los diferentes ordenamientos de los mismos elementos no cuentan por separado, nos enfrentamos a un problema de combinaciones;** en tal caso se aplica la regla siguiente:

Regla de las combinaciones

El número de **combinaciones** de r elementos que se seleccionaron de entre n elementos diferentes es

$$_nC_r = \frac{n!}{(n-r)!r!}$$

Muchas calculadoras se diseñaron para evaluar $_nC_r$.

Es muy importante reconocer que, en la aplicación de la regla de las combinaciones, se aplican las siguientes condiciones:

- Debemos tener un total de n elementos diferentes disponibles.
- Debemos seleccionar r de los n elementos (sin reemplazo).
- Debemos considerar que los reacomodos de tales elementos son los mismos. (La combinación *ABC* es la misma que *CBA*).

Puesto que quizá resulte confuso escoger entre la regla de las permutaciones y la regla de las combinaciones, damos el ejemplo siguiente, que tiene la intención de poner énfasis en la diferencia entre ellas.

EJEMPLO Oficinas electas El consejo de fondos de inversión de la universidad a la que asistió el autor se integra con nueve miembros. Cada año, ellos eligen un comité de tres personas para supervisar los edificios y los terrenos. También cada año eligen un presidente, un vicepresidente y un secretario.

a. Cuando el consejo elige el comité de edificios y terrenos, ¿cuántos distintos comités de tres personas son posibles?

b. Cuando el consejo elige a los tres funcionarios (presidente, vicepresidente y secretario), ¿cuántas diferentes planillas de candidatos son posibles?

SOLUCIÓN Note que el orden es irrelevante cuando se elige el comité de edificios y terrenos. Sin embargo, cuando se elige a los funcionarios, los diferentes acomodos cuentan por separado.

a. Puesto que el orden no cuenta para los comités, queremos el número de combinaciones de $r = 3$ personas que se seleccionarán de entre las $n = 9$ disponibles. Tenemos

$$_nC_r = \frac{n!}{(n-r)!r!} = \frac{9!}{(9-3)!3!} = \frac{362{,}880}{4320} = 84$$

continúa

b. Ya que el orden sí cuenta en las planillas de candidatos, queremos el número de secuencias (o permutaciones) de $r = 3$ personas que se seleccionarán de entre las $r = 9$ disponibles. Tenemos

$$_nP_r = \frac{n!}{(n - r)!} = \frac{9!}{(9 - 3)!} = \frac{362{,}880}{720} = 504$$

Hay 84 diferentes comités de tres miembros del consejo posibles, pero 504 diferentes planillas de candidatos posibles.

Las técnicas de conteo que se presentan en esta sección se usan algunas veces en problemas de probabilidad. Los ejemplos siguientes ilustran dichas aplicaciones.

EJEMPLO **Lotería de Maine** En la lotería de Maine se extraen seis diferentes números del 1 al 42. Un jugador gana o comparte el premio mayor cuando escoge la combinación correcta de seis números. Si un jugador elige una combinación particular de seis números, calcule la probabilidad de ganar el premio mayor. (No se requiere que el jugador seleccione los seis números en el mismo orden en que se sacaron, por lo que el orden es irrelevante).

SOLUCIÓN Puesto que se seleccionan seis números diferentes de 42 posibilidades diferentes, el número total de combinaciones es

$$_{42}C_6 = \frac{42!}{(42 - 6)!6!} = \frac{42!}{36!6!} = 5{,}245{,}786$$

Si el jugador sólo selecciona una combinación, la probabilidad de ganar es de $1/5{,}245{,}786$.

EJEMPLO **La lotería Powerball** La lotería Powerball se juega en 21 estados de Estados Unidos. Usted debe seleccionar cinco números entre 1 y 49, además de otro número especial Powerball entre 1 y 42. (Se extraen cinco bolas de una tómbola con 49 bolas blancas y una bola roja de una tómbola con 42 bolas rojas). El número especial Powerball puede ser el mismo que uno de los otros cinco números. Para ganar o compartir el premio mayor, tendrá que seleccionar la combinación correcta de cinco números *y* también debe seleccionar el número Powerball correcto. Calcule la probabilidad de ganar o compartir el premio mayor.

SOLUCIÓN Dividamos este problema en tres partes: **1.** tener la combinación correcta de cinco números; **2.** obtener el número Powerball correcto, y **3.** combinar los resultados para calcular la probabilidad de ganar o compartir el premio mayor. Comencemos con el número de combinaciones que son posibles cuando usted selecciona cinco números entre 1 y 49, que es

$$_{49}C_5 = \frac{49!}{(49 - 5)!5!} = \frac{49!}{44!5!} = 1{,}906{,}884$$

La probabilidad de obtener la combinación ganadora de cinco números es, por lo tanto, de $1/1{,}906{,}884$.

continúa

En segundo lugar, seleccionemos también el número Powerball correcto entre 1 y 42. La probabilidad de seleccionar el número Powerball ganador es de $1/42$.

En tercer lugar, puesto que debemos obtener la combinación correcta de cinco números y el número Powerball correcto, la probabilidad de la ocurrencia de ambos sucesos es de $1/1{,}906{,}884 \times 1/42 = 1/80{,}089{,}128$. Este último resultado es una aplicación de la regla de la multiplicación que se explica en la sección 3-4. Para una persona que compre un boleto, la probabilidad de ganar la lotería Powerball es de $1/80{,}089{,}128$.

3-7 Destrezas y conceptos básicos

Cálculo de factoriales, combinaciones y permutaciones. *En los ejercicios 1 a 8 evalúe las expresiones dadas y exprese todos los resultados utilizando el formato normal para escribir números (en lugar de notación científica).*

1. 6! **2.** 15! **3.** $_{25}P_2$ **4.** $_{100}P_3$

5. $_{25}C_2$ **6.** $_{100}C_3$ **7.** $_{52}C_5$ **8 .** $_{52}P_5$

Probabilidad de ganar en la lotería. *Esta sección incluyó un ejemplo que mostró que la probabilidad de ganar la lotería de Maine es de $1/5{,}245{,}786$. En los ejercicios 9 a 12 calcule la probabilidad de ganar la lotería que se indica.*

9. Massachusetts Mass Millions: Seleccionar los seis números ganadores entre 1, 2, . . . , 49.

10. Pennsylvania Super 6 Lotto: Seleccionar los seis números ganadores entre 1, 2, . . . ,69.

11. New York Lotto: Seleccionar los seis números ganadores entre 1, 2, . . . , 59.

12. New York Take Five: Seleccionar los cinco números ganadores entre 1, 2, . . . , 39.

13. Discriminación por edad La empresa Pitt Software Company redujo su personal de ventas de 32 a 28 empleados. La compañía afirmó que seleccionó a cuatro empleados al azar para despedirlos. Sin embargo, los cuatro empleados que eligió son los más viejos de la fuerza de ventas original de 32. Calcule la probabilidad de que cuando se seleccionan cuatro empleados al azar de un grupo de 32, éstos sean los cuatro más viejos. ¿Es la probabilidad lo suficientemente baja como para acusar a la Pitt Software Company de que en lugar de usar selección aleatoria, en realidad sólo despidió a los empleados más viejos?

14. Diseño de computadoras En el diseño de una computadora, un *byte* se define como una secuencia de 8 bits y cada bit debe ser un 0 o un 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 otros?

15. Lotería de Maine La probabilidad de ganar la lotería de Maine es de $1/5{,}245{,}786$. Si las reglas se modificaran para que, además de seleccionar los seis números correctos del 1 a 42, ahora tuviera que elegirlos en el mismo orden en que son extraídos, ¿cuál es la probabilidad de ganar?

16. Prueba de una afirmación Mike afirma que desarrolló la habilidad de obtener un 6 casi siempre que tira un dado. Usted prueba su afirmación haciendo que Mike tire un dado cinco veces, en tanto él obtiene el 6 cada vez. Si Mike no tiene posibilidad de afectar los resultados, calcule la probabilidad de que él tire cinco veces consecutivas y obtenga 6 en todas. ¿Es la probabilidad suficientemente baja como para apoyar la afirmación de Mike?

17. Selección de grupo de tratamiento La empresa Walton Pharmaceuticals quiere probar la eficacia de un nuevo fármaco que se desarrolló para aliviar síntomas de alergia. La prueba inicial se realizará tratando a seis personas que se escogieron de un grupo de 15 voluntarios. Si el grupo de tratamiento se selecciona aleatoriamente, ¿cuál es la probabilidad de que se conforme con las seis personas más jóvenes del grupo? Si se selecciona a los seis más jóvenes, ¿hay evidencia suficiente para concluir que la selección se basó en la edad, en lugar de ser aleatoria?

18. Lo hizo a su manera El legendario cantante Frank Sinatra grabó 381 canciones. Usted debe seleccionar, de una lista de sus 10 más grandes éxitos, tres que serán cantados en un popurrí como un tributo en la próxima ceremonia de los premios MTV Music Awards. El orden de las canciones es importante, ya que tienen que sonar bien juntas. Si usted selecciona tres de las canciones de Sinatra de entre sus 10 mayores éxitos, ¿cuántas secuencias diferentes son posibles?

19. Rutas aéreas Usted acaba de inaugurar su propia empresa de líneas aéreas llamada Air América, cuyo lema es: "Donde su probabilidad de un vuelo seguro es mayor que cero". Trazó un plan para una ruta que conecta Austin, Boise y Chicago. Una ruta es Austin-Boise-Chicago; una segunda ruta es Chicago-Boise-Austin. ¿Cuántas otras rutas son posibles? ¿Cuántas rutas diferentes son posibles si el servicio se expandiera para incluir un total de ocho ciudades?

20. Números del Seguro Social Cada número del Seguro Social es una secuencia de nueve dígitos. ¿Cuál es la probabilidad de generar aleatoriamente nueve dígitos y obtener su número de Seguro Social?

21. Electrizante Para probar la corriente eléctrica en un conductor con cables codificados en cinco colores, el autor utilizó un medidor para probar dos cables a la vez. ¿Cuántas pruebas se requieren para verificar cada posible par de cables?

22. Consejo de administración electo En un consejo de administración del hospital general de Newport hay 12 miembros.
 a. Si ellos deben elegir un presidente, un primer vicepresidente, un segundo vicepresidente y un secretario, ¿cuántas planillas de candidatos diferentes son posibles?
 b. Si tienen que formar un subcomité de ética de cuatro miembros, ¿cuántos subcomités diferentes son posibles?

23. Sopa de letras Muchos periódicos incluyen una "sopa de letras", un crucigrama donde 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 que se codificó y luego determine la probabilidad de obtener este resultado seleccionando al azar un arreglo de las letras dadas.

24. Calcular el 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 con un asterisco (*), en tanto que las notas sucesivas lo hacen con una *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 mediante las primeras 16 notas. Con este esquema, ¿cuántas melodías clásicas diferentes son posibles?

25. Candados de combinación Un candado "de combinación" común se abre con la secuencia correcta de tres números entre 0 y 49, inclusive. (Es posible utilizar un número más de una vez). ¿Cuál es la probabilidad de adivinar los tres números y de abrir el candado en el primer intento?

26. Flor de cinco naipes Un mazo de naipes normal contiene 13 tréboles, 13 diamantes, 13 corazones y 13 espadas. Si se eligen cinco naipes aleatoriamente, calcule la probabilidad de obtener una flor. (Se tiene una flor cuando las cinco cartas son del mismo palo. Es decir, cuando todas son tréboles, diamantes, corazones o espadas).

27. Probabilidades de secuencias de género
a. Si una pareja planea tener ocho hijos, ¿cuántas secuencias de género diferentes son posibles?
b. Si una pareja tiene cuatro niños y cuatro niñas, ¿cuántas secuencias de género diferentes son posibles?
c. Con base en los resultados de los incisos *a* y *b*, cuando una pareja tiene ocho hijos ¿cuál es la probabilidad de que sean cuatro niños y cuatro niñas?

28. ¿El investigador está haciendo trampa? Cuando un investigador de genética selecciona al azar grupos de 20 bebés recién nacidos y aparentemente obtiene 10 niñas y 10 niños con consistencia, usted se vuelve suspicaz. El investigador explica que es común obtener 10 niños y 10 niñas en estos casos.
a. Si se seleccionan 20 bebés recién nacidos, ¿cuántas secuencias de género distintas son posibles?
b. ¿De cuántas formas diferentes pueden acomodarse, en secuencia, 10 niños y 10 niñas?
c. ¿Cuál es la probabilidad de obtener 10 niños y 10 niñas cuando nacen 20 bebés?
d. Con base en los resultados de lo anterior, ¿está de acuerdo con la explicación del investigador de que es común obtener 10 niños y 10 niñas cuando se seleccionan 20 bebés al azar?

29. Calcular el número de códigos de área El reportero Paul Wiseman, del diario *USA Today*, describió las viejas reglas para los códigos de área telefónicos al escribir acerca de “códigos de área posibles con 1 o 0 en el segundo dígito”. (Excluidos: códigos que terminen en 00 y 11, para llamadas con cargo gratuito, servicios de emergencia y otros usos especiales). Los códigos que empiezan con 0 o 1 también deben ser excluidos. ¿Cuántos códigos de área distintos era posible obtener bajo tales viejas reglas?

30. Huevos rotos Una caja contiene 12 huevos, tres de los cuales están rotos. Si seleccionamos al azar cinco de los huevos para cocerlos, ¿cuál es la probabilidad de los sucesos siguientes?:
a. Todos los huevos que se seleccionaron están rotos.
b. Ninguno de los huevos que se seleccionaron está roto.
c. Dos de los huevos que se seleccionaron están rotos.

31. Lotería de California En el juego de lotería Super Lotto Plus de California, ganar el premio mayor requiere que se seleccionen los cinco números correctos del 1 al 47 y, por separado, también elegir un solo número correcto entre 1 y 27. Calcule la probabilidad de ganar el premio mayor.

32. Torneo de basquetbol NCAA Cada año, 64 equipos universitarios de basquetbol compiten en el torneo de la NCAA. Recientemente, Sandbox.com ofreció un premio de 10 millones de dólares a cualquiera que eligiera al ganador en todos y cada uno de los juegos del torneo. El presidente de esa compañía también prometió que, además de entregar el premio en efectivo, se comería los gusanos que cupieran en una cubeta. ¡Qué asco!
a. ¿Cuántos juegos se requieren para obtener un equipo campeón en un campo de 64 equipos?
b. Si alguna persona 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 de dólares, el diario *New York Times* escribió que “aun un experto en basquetbol 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).

3-7 Más allá de lo básico

33. 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 serían cualesquiera de las 26 letras o de los 10 dígitos. Por ejemplo, A, BBB y M3477K son nombres permitidos de variables. ¿Cuántos nombres de variables diferentes son posibles?

34. Saludos y mesas redondas
- **a.** Cinco gerentes se reúnen para una junta. Si cada uno saluda estrechando la mano a los otros gerentes exactamente una vez, ¿cuál es el número total de saludos?
- **b.** Si n gerentes se saludan con cada uno de los otros exactamente una vez, ¿cuál es el número total de saludos?
- **c.** ¿De cuántas maneras diferentes se pueden sentar cinco gerentes en una mesa redonda? (Suponga que si cada uno se mueve a la derecha el acomodo es el mismo).
- **d.** ¿De cuántas formas diferentes se pueden sentar n gerentes en una mesa redonda?

35. 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 $K = (n + 0.5)\log n + 0.39908993 - 0.43429448n$.
- **a.** Evalúe 50! usando la tecla factorial de una calculadora y también la aproximación que damos aquí.
- **a.** El Departamento de Pesca una vez pidió ayuda a los Laboratorios Bell con la finalidad 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?

36. Inteligencia artificial ¿Las computadoras "piensan"? De acuerdo con la *prueba Turing*, se considera que una computadora piensa, cuando alguien se comunica con ella, si la persona que la utiliza cree que se está comunicando con otro y no con una máquina. En un experimento en el Computer Museum de Boston, cada uno de 10 jueces se comunicó con cuatro de estas máquinas y otras cuatro personas; se les pidió que distinguieran entre ellos.
- **a.** 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 identificar correctamente las cuatro computadoras y las cuatro personas?
- **b.** 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 es así).

Repaso

Iniciamos este capítulo con el concepto básico de probabilidad, el cual es de suma importancia para los métodos de estadística inferencial que se introducen más adelante. Aprendimos que un valor de probabilidad, que se expresa como un número entre 0 y 1, refleja la posibilidad de ocurrencia de algún suceso. También que un valor como 0.01 representa un suceso que tiene muy pocas posibilidades de ocurrir. En la sección 3-1, introdujimos la regla del suceso infrecuente para estadística inferencial: si, bajo un supuesto dado, la probabilidad de un suceso particular es extremadamente pequeña, concluimos que quizás el supuesto es incorrecto. Como ejemplo del enfoque básico que se utilizó, considere la prueba de la aseveración de alguien de que una moneda que se usa en un volado está balanceada. Si lanzamos la moneda 10 veces y obtenemos 10 caras consecutivas, de estos resultados de muestra es posible hacer una de dos inferencias:

1. La moneda está realmente balanceada y la cadena de 10 caras consecutivas es una *chiripa*.

2. La moneda no está balanceada.

Los estadísticos usan la regla del suceso infrecuente cuando deciden cuál inferencia es correcta: en este caso, la probabilidad de obtener 10 caras consecutivas es tan pequeña (1/1024) que la inferencia de que la moneda no está balanceada es la mejor opción. Aquí observamos el importante papel de la probabilidad en los métodos estándar de inferencia estadística.

En la sección 3-2 presentamos definiciones y notaciones básicas, incluyendo la representación de sucesos por letras como A. Definimos, asimismo, las probabilidades de un suceso simple como

$$P(A) = \frac{\text{número de veces que ocurre } A}{\text{número de veces que se repite el experimento}} \qquad \text{(frecuencia relativa)}$$

$$P(A) = \frac{\text{número de formas en que puede ocurrir } A}{\text{número de sucesos simples diferentes}} = \frac{s}{n} \qquad \text{(para resultados igualmente probables)}$$

Además, señalamos que la probabilidad de cualquier suceso imposible es 0 y la probabilidad de cualquier suceso inevitable es 1, así como que para cualquier suceso A, $0 \leq P(A) \leq 1$. También, que $\overline{A}$ denota el complemento del suceso A, es decir, indica que el suceso A no ocurre.

En las secciones 3-3, 3-4 y 3-5 consideramos sucesos compuestos, los cuales combinan dos o más sucesos simples. Asociamos "o" con la suma y asociamos "y" con la multiplicación. Siempre tome en cuenta las consideraciones clave siguientes:

- Cuando se realiza un ensayo, ¿queremos la probabilidad del suceso A o B? Si es así, use la regla de la suma, pero sea cuidadoso a fin de evitar contar cualquier resultado más de una 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 ocurrió realmente.

En algunos problemas de probabilidad, el mayor obstáculo es encontrar el número total de resultados posibles. La sección 3-7 se dedicó a las siguientes técnicas de conteo:

- Regla de conteo fundamental
- 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

Ejercicios de repaso

Detectores de mentiras. *En los ejercicios 1 a 8 use los datos de la tabla adjunta (que se basa en datos de la Office of Technology Assessment). Los datos reflejan las respuestas a una pregunta clave que se hizo a 100 sujetos diferentes.*

	El polígrafo indicó verdad	El polígrafo indicó mentira
El sujeto realmente dijo la verdad	65	15
El sujeto realmente mintió	3	17

1. Si se selecciona al azar a 1 de los 100 sujetos, calcule la probabilidad de obtener a alguien que mintió.

2. Si se selecciona al azar a 1 de los 100 sujetos, calcule la probabilidad de obtener a alguien para quien la prueba del polígrafo indicó que dijo una mentira.

3. Si se selecciona al azar a 1 de los 100 sujetos, calcule la probabilidad de obtener a alguien que mintió o tuvo la indicación de la prueba del polígrafo de que lo hizo.

4. Si se selecciona al azar a 1 de los 100 sujetos, calcule la probabilidad de obtener a alguien que dijo la verdad o la prueba del polígrafo indicó que respondió con la verdad.

5. Si dos diferentes sujetos se seleccionan al azar, calcule la probabilidad de que ambos dijeron la verdad.

6. Si dos diferentes sujetos se seleccionan al azar, calcule la probabilidad de que, al realizar la prueba del polígrafo, éste indicó que ambos dijeron una mentira.

7. Si un sujeto se selecciona al azar, calcule la probabilidad de que él o ella dijeron la verdad, puesto que la prueba del polígrafo indicó que era una mentira.

8. Si un sujeto se selecciona al azar, calcule la probabilidad de que él o ella obtuvieron la indicación de la prueba del polígrafo de que dijeron una mentira, puesto que el sujeto realmente dijo la verdad.

9. Probabilidad de fallas de computadora Una encuesta de la revista *PC World* entre 4,000 personas propietarias de computadoras personales, mostró que 992 de ellas se averiaron durante los primeros dos años (se averiaron las computadoras, no las personas). Al seleccionar entre varios distribuidores de computadoras, un agente de compras quiere saber la probabilidad de que una computadora personal se descomponga durante los primeros dos años. Utilice los resultados de la encuesta para estimar esa probabilidad.

a. Si se selecciona una computadora personal al azar, ¿cuál es la probabilidad de que se descomponga durante los primeros dos años?

b. Si se seleccionan dos computadoras personales al azar, ¿cuál es la probabilidad de que ambas se descompongan durante los dos primeros años?

c. Si se seleccionan tres computadoras personales al azar, ¿cuál es la probabilidad de que al menos una de ellas se descomponga durante los dos primeros años?

10. Muestreo de aceptación Con un método de muestreo de aceptación, una muestra de artículos se selecciona aleatoriamente sin reemplazo y el lote completo se rechaza si se encuentra al menos un defecto. La Niko Electronics Company acaba de fabricar 2,500 CD, de los cuales el 2% salieron defectuosos. Si se seleccionan y prueban cuatro de los CD, ¿cuál es la probabilidad de que se rechace el lote completo?

11. Prueba de una aseveración La Biogene Research Company afirma que desarrolló una técnica para asegurar que un bebé será una niña. En una prueba de esa técnica, 12 parejas tuvieron niñas. Calcule la probabilidad de obtener dos niñas por casualidad, suponiendo que los niños y las niñas son igualmente probables, así como que el género de cualquier hijo es independiente del género de los otros. ¿Apoya el resultado la aseveración de la compañía?

12. Selección de miembros El Consejo de Administración del Hartford Investment Fund cuenta con 10 miembros.

a. Si se selecciona al azar a tres miembros para supervisar a los auditores, calcule la probabilidad de que sean seleccionados los tres miembros más acaudalados.

b. Si se eligen miembros para los puestos de presidente, vicepresidente y tesorero, ¿cuántas planillas diferentes son posibles?

13. Ruleta Cuando se apuesta a *pares* en la ruleta, hay 38 resultados igualmente probables, pero sólo 2, 4, 6, . . . , 36 son resultados ganadores.
a. Calcule la probabilidad de ganar cuando se apuesta a pares.
b. Calcule las posibilidades reales en contra de ganar cuando se apuesta a pares.
c. Los casinos pagan las apuestas ganadoras de acuerdo con las posibilidades descritas como 1:1. ¿Cuál sería su ganancia neta si apuesta $5 a pares y gana?

14. ¿Está mintiendo el encuestador? Un encuestador afirma que 12 votantes se seleccionaron aleatoriamente de una población de 200,000 (el 30% de ellos son republicanos) y que los 12 fueron republicanos. El encuestador añade que esto podría suceder fácilmente por casualidad. Calcule la probabilidad de obtener 12 republicanos cuando se seleccionan 12 votantes de dicha población. Con base en el resultado, ¿parece ser correcta la afirmación del encuestador?

15. Seguros de vida La compañía de seguros New England Life expide pólizas anuales a 12 hombres de 27 años de edad. Con base en datos del Department of Health and Human Services, cada uno de estos individuos tiene una posibilidad de un 99.82% de vivir todo el año. ¿Cuál es la probabilidad de que todos ellos sobrevivan el año?

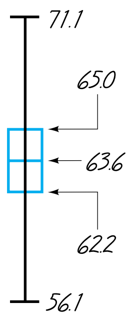
16. Loterías de Illinois Illinois tiene diferentes juegos de lotería. Calcule la probabilidad de ganar el premio mayor en cada juego.
a. Lotto: Seleccionar los seis números ganadores de 1, 2, . . . , 52.
b. Little Lotto: Seleccionar los cinco números ganadores entre 1, 2, . . . , 30.
c. The Big Game: Seleccionar los cinco números ganadores entre 1, 2, . . . , 50 *y*, además, seleccionar también un solo número ganador entre 1, 2, . . . , 36.

Ejercicios de repaso acumulativos

1. Tratamiento del síndrome de fatiga crónica A una muestra de pacientes que padecen el síndrome de fatiga crónica se le trató con medicamentos, después se midió el cambio en su fatiga en una escala de -7 a $+7$, con los valores positivos representando mejoría y con 0 representando ningún cambio. Los resultados se listan abajo (con base en 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

a. Calcule la media.
b. Calcule la mediana.
c. Calcule la desviación estándar.
d. Calcule la varianza.
e. Con base en los resultados, ¿fue efectivo el tratamiento?
f. Si se selecciona aleatoriamente un valor de esta muestra, calcule la probabilidad de que sea positivo.
g. Si se seleccionan aleatoriamente dos diferentes valores de esta muestra, calcule la probabilidad de que ambos sean positivos.
h. Ignore los tres valores de 0 y suponga que sólo son posibles valores positivos o negativos. Suponiendo que el tratamiento no es efectivo y que los valores positivos y negativos son igualmente probables, calcule la probabilidad de que 18 sujetos tengan todos valores positivos (como en este grupo muestral). ¿Es esta probabilidad bastante baja como para justificar el rechazo de la suposición de que el tratamiento no es efectivo?



2. Estaturas de mujeres La gráfica de cuadro anexa describe estaturas (en pulgadas) de una gran colección de mujeres adultas que se seleccionaron aleatoriamente.
   a. ¿Cuál es el promedio de estaturas de las mujeres adultas?
   b. Si se selecciona al azar a una de estas mujeres, calcule la probabilidad de que su estatura se ubique entre 56.1 pulgadas y 62.2 pulgadas.
   c. Si se selecciona al azar a una de estas mujeres, calcule la probabilidad de que su estatura sea más baja que 62.2 pulgadas o que sobrepase 63.6 pulgadas.
   d. Si se seleccionan dos mujeres al azar, calcule la probabilidad de que ambas tengan estaturas entre 62.2 pulgadas y 63.6 pulgadas.
   e. Si se seleccionan cinco mujeres al azar, calcule la probabilidad de que tres de ellas sean más altas que la media y las otras dos sean más bajas que la media.

Actividades de cooperación en equipo

1. Actividad en clase Véase el ejercicio 15 de la sección 3-6. Formen equipos de tres o cuatro y realicen lanzamientos de monedas para desarrollar una simulación que imite al reino que se atiene a tal decreto: después de que una madre dé a luz a un hombre, ella no tendrá ningún otro hijo. Si este decreto se obedece, ¿se incrementará la proporción de mujeres?

2. Actividad en clase Haga equipos de tres o cuatro personas y use tachuelas para estimar la probabilidad de que, cuando se dejan caer, una tachuela quede con la punta hacia arriba. ¿Cuántos intentos son necesarios para obtener un resultado que parezca razonablemente preciso, cuando se redondea al primer espacio decimal?

3. Actividad fuera de clase Los biólogos marinos con frecuencia usan el método captura-recaptura como procedimiento para estimar el tamaño de una población como, por ejemplo, del número de peces en un lago. Este método consiste en capturar una muestra de la población, etiquetar a cada uno de los miembros de la muestra y luego regresarlos a la población.

   En lugar de capturar peces reales, simule el procedimiento utilizando un conjunto uniforme de artículos como, por ejemplo, botones, cuentas de colores, dulces M&M, piezas de cereal de aros de frutas o tarjetas de archivo. Comience con una colección grande de dichos artículos. Obtenga una muestra de 50 y use un marcador para "etiquetar" a cada uno. Reemplace los artículos que se marcaron, revuelva la población completa; luego, 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.

4. Actividad en clase Formen equipos de dos. Remítanse al ejercicio 13 en la sección 3-6 para tener una descripción del "problema Monty Hall". Simulen el concurso, y registren los resultados de quedarse y de cambiar; después, determinen cuál de estas dos estrategias es mejor.

5. Actividad fuera de clase Formen equipos de dos, con el propósito de hacer un experimento que se diseñó con la finalidad de mostrar un enfoque para el manejo de preguntas de encuesta sensibles, que se relacionan con el uso de drogas, la actividad sexual (o la inactividad), el robo o la estafa. En vez de utilizar realmente una pregunta polémica que podría ocasionar ira contra el autor, manejemos 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 es muy sensible y que esos sujetos de la encuesta son reticentes a contestar con honestidad. Encueste pidiéndoles a las personas que lancen una moneda al aire y respondan como sigue:

   - "Sí", si la moneda cae en cruz *o* usted nació en un mes que tiene la letra *r*.
   - "No", si la moneda cae en cara *y* usted nació en un mes que no contiene la letra *r*.

   Supuestamente, quienes responden tienden a ser más honestas, porque sienten que lanzar la moneda al aire protege su privacidad. Encueste personas y analice los resultados para determinar la proporción de aquellos que nacieron en un mes que contiene la letra *r*. La precisión de los resultados se puede cotejar con las fechas de nacimiento reales, que se obtendrán de una segunda pregunta. Es posible repetir el experimento con una pregunta que sea más sensible, pero aquí no se plantean preguntas de este tipo, porque el autor ya recibe suficientes mensajes por correo.

Proyecto tecnológico

Este proyecto ilustra la ley de los números grandes descrita en la sección 3-2. Use una computadora o la calculadora TI-83 Plus para simular 100 nacimientos. Realice esto generando aleatoriamente 100 números, ya sea 0 o 1 (donde 0 = niño y 1 = niña). Utilice los resultados para completar la tabla siguiente. ¿Qué sucede a la proporción de niñas conforme el tamaño de la muestra se va incrementando? ¿Cómo ilustra esto la ley de los números grandes?

Número de nacimientos	10	20	30	40	50	60	70	80	90	100
Proporción de niñas										

Aquí tenemos algunos detalles para las diferentes aplicaciones tecnológicas.

STATDISK Seleccione **Data** en la barra de menú principal. Elija **Uniform Generator**. Realice las entradas para un tamaño de muestra de 100, un mínimo de 0, un máximo de 1 y 0 espacios decimales (puesto que queremos números enteros).

Minitab Seleccione **Calc** de la barra de menú principal en la parte superior. Seleccione **Random Data**; después, **Integer**. Proceda a cargar 100 para el número de filas de datos, C1 para la columna en la que se guardarán los resultados, 0 para el valor mínimo y 1 para el valor máximo. Cuando termine, haga clic en **OK**.

Excel Posicione el cursor en la celda A1. Haga clic en el elemento del menú f_x; luego, seleccione **Math & Trig** y **RANDBETWEEN**. Haga clic en **OK**. En el cuadro de diálogo, ponga 0 para el valor inferior y 1 para el valor superior; luego, haga clic en **OK**. La celda A1 debe contener ahora un 0 o un 1. Haga clic en la esquina inferior derecha de esa celda y, mientras sostiene el botón del ratón, arrastre el cursor hacia abajo hasta la celda A100, ahora suéltelo.

TI-83 Plus **Plus** Oprima la tecla **MATH**. Seleccione **PRB**. Luego, el 5o elemento del menú: **randInt**. Ingrese 0, 1, 100, y presione la tecla **ENTER**. Oprima **STO** y **L1** para guardar los datos en la lista L1. Para ver los nacimientos que se generaron, oprima **STAT** y seleccione **Edit**.

de los DATOS *a la* DECISIÓN

Pensamiento crítico: Cuando usted solicita un empleo, ¿debe preocuparse por la prueba de consumo de drogas?

De acuerdo con la American Management Association, alrededor del 70% de las empresas de Estados Unidos ya realizan la prueba del consumo de drogas al menos a algunos empleados y solicitantes de empleo. El National Insitute on Drug Abuse afirma que alrededor del 15% de las personas entre 18 y 25 años consumen drogas ilegales. Allyn Clark, un universitario graduado de 21 años de edad, cuando solicitó empleo en la compañía Acton Paper, se sometió a la prueba de drogas; en consecuencia, no se le dio el empleo. Él sospechó que tal vez obtuvo un resultado adverso en la prueba de drogas, aunque no las consume. Revisando con el departamento de personal de la empresa, encontró que sólo el 1% de los usuarios de drogas arrojan erróneamente un resultado de prueba negativo y que sólo el 2% de quienes no consumen drogas se identifican incorrectamente como usuarios de drogas. Allyn se sintió aliviado por estas cifras, porque creyó que reflejaban una prueba muy confiable que generalmente daba buenos resultados, pero, ¿es esto realmente cierto?

Análisis de los resultados

La tabla adjunta muestra datos de Allyn y de otros 1999 solicitantes de empleo. Con base en estos resultados, calcule *P*(falso negativo); esto es, la probabilidad de seleccionar al azar a alguna persona que tuvo un resultado de prueba positivo y elegir a alguien que no consume drogas. ¿Son las probabilidades de dichos resultados erróneos suficientemente bajas como para que los solicitantes de empleo y la compañía Acton Paper no tengan de qué preocuparse?

	Usuarios de drogas	No usuarios
Resultado de prueba positivo	297	34
Resultado de prueba negativo	3	1666

PROYECTO DE INTERNET

Proceso de probabilidades por computadora

Calcular las probabilidades cuando se tiran dados es fácil. Con un dado, hay seis posibles resultados, de los cuales cada uno, como por ejemplo tirar un 2, tiene una probabilidad de $1/6$. Para un juego de naipes se necesitan más cálculos, aunque siguen siendo manejables. Pero, ¿qué pasa con un juego más complicado, como por ejemplo 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 y 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 muy grande como para siquiera considerarlos; no obstante, las compañías de seguros, por ejemplo, contemplan probabilidades de este tipo.

El proyecto de Internet para este capítulo considera métodos para calcular probabilidades en situaciones que se complican. Vaya al proyecto de Internet, que encontrará en este sitio:

http://www.pearsoneducacion.net/triola

Se le guiará en la investigación de probabilidades para un juego de mesa. Después, calcule usted mismo probabilidades de este tipo. Finalmente, efectuará un estimado de una probabilidad que se relaciona con la salud utilizando datos empíricos.

La estadística @ en el trabajo

"Con base en estas estadísticas acerca del uso preferido del parque, evaluamos cómo dar mejor servicio a la población más diversa".

Judy Shafer

Segunda superintendente del Parque Nacional Virgin Islands

Judy trabaja para el National Park Service desde hace 17 años y para la administración del parque desde los últimos cuatro.

Como segunda superintendente del Parque Nacional Virgin Islands, ¿utiliza la estadística en su trabajo?

Usamos probabilidad y métodos de estadística en aplicaciones tales como el análisis de la sustentabilidad de una especie particular de coral, bajo las presiones ambientales del uso del parque por el visitante, la contaminación, la sedimentación, etcétera. También, métodos de estadística para determinar cuáles poblaciones étnicas y raciales utilizan el parque, y la manera en la que el parque está logrando sus metas de satisfacción al usuario.

¡Hay tantos usos de la estadística que es imposible describirlos todos! Nuestros guardabosques recolectan datos y usan la estadística para actividades de visita y de protección de los recursos. Nuestros científicos de investigación marina utilizan la estadística a diario para determinar la salud y viabilidad de varias especies marinas, así como para conocer el nivel de amenaza que representan los agentes ambientales (como la contaminación y la sedimentación) y los visitantes del parque. Por ejemplo, algunas de las especies de peces más grandes, como el *Nassau Grouper*, pueden mermarse tanto por la pesca excesiva que tal vez estén cerca de la extinción. Los datos estadísticos de las investigaciones deben guiar las decisiones del parque acerca de cómo proteger dichas especies, ya que se encuentran inextricablemente encadenadas al ecosistema de los grandes arrecifes de coral e incluso al calentamiento global.

¿Podría dar un ejemplo simple y específico de cómo se usa la estadística?

En lo que respecta al uso del parque por el visitante, el National Park Service ha aprendido que distintos grupos étnicos y raciales utilizan los parques nacionales en diferentes formas. Por ejemplo, descubrimos que a algunos grupos les gusta la tradicional caminata campirana mientras que otros prefieren la actividad social de los días de campo en compañía. Estamos considerando la incorporación de más áreas para días de campo en compañía con la finalidad de acomodar mejor a los visitantes de diferentes contextos étnicos y raciales.

¿Es necesario que quienes solicitan empleo en su área tengan un curso de estadística en su historial?

Nuestros empleados deben tener al menos un curso de estadística de nivel universitario. Sin esto, serás un dinosaurio en un siglo XXI que se basa en la estadística. Sin la estadística, no serás tan competitivo como la persona que sí recibió esa preparación.

¿Qué otras habilidades es importante que los empleados posean?

Visión. Liderazgo. Innovación. Creatividad. Y la disposición de tomar algunos riesgos para lograr sus metas y proteger las cosas por las que sienten pasión.

4

Distribuciones de probabilidad

Determinar si un método de selección del género es efectivo

Los avances recientes en medicina, genética y tecnología son sorprendentes. Ya parece posible la clonación de seres humanos. Las operaciones de puenteo arterial cardiaco, hasta hace poco consideradas riesgosas y peligrosas, ahora son rutina. La cirugía con láser es capaz de corregir la visión, de modo que muchas personas pueden deshacerse de sus anteojos o lentes de contacto. En lugar de confiar sólo en la probabilidad aleatoria, las parejas que planean tener bebés disponen de técnicas para determinar el género de sus hijos. En ocasiones, dichos avances van acompañados de una gran polémica. Algunos consideran que las técnicas de clonación o de selección del género conllevan graves implicaciones morales y que deben prohibirse estrictamente, sin importar cuál sea su justificación. Lisa Belkin escribió lo siguiente en el artículo “Getting the girl” (del *New York Times Magazine):* “Si permitimos que los padres elijan el sexo de sus hijos, ¿cuánto tiempo pasará antes de que ordenen el color de ojos y de cabello, rasgos de personalidad y CI? Hay algunos argumentos convincentes que están a favor del uso, aunque sea limitado, de la selección del género. Uno de esos argumentos señala que ciertas parejas son portadoras de genes recesivos que se relacionan con el cromosoma X, lo cual implica que cualquier hijo hombre cuenta con un 50% de probabilidades de heredar una enfermedad grave, pero que ninguna niña heredará el trastorno. Dichas parejas podrían utilizar la selección del género como una forma para asegurarse de que tendrán niñas y de que ninguna heredará un padecimiento grave.

El Genetics and IVF Institute, de Fairfax, Virginia, creó una técnica llamada MicroSort que, se supone, aumenta las posibilidades de que una pareja tenga una niña. En una prueba preliminar, se reunieron 14 parejas que deseaban niñas. Con el uso de la técnica MicroSort, 13 parejas procrearon niñas y una tuvo un niño. Tales resultados nos conducen a plantearnos una interesante pregunta: puesto que 13 de las 14 parejas tuvieron niñas, ¿concluiríamos que la técnica MicroSort es efectiva o sólo explicaríamos el resultado como la consecuencia de una muestra aleatoria? Para responder lo anterior, usaremos principios de probabilidad que determinen si los nacimientos que se observaron difieren de manera significativa de los resultados que se esperarían al azar. Este ejemplo hace surgir un tema que es esencial para la estadística inferencial: ¿De qué forma determinamos si los resultados deben atribuirse al azar o a un factor como el uso de la técnica MicroSort de selección del género?

4-1 Panorama general

En este capítulo combinamos los métodos de *estadística descriptiva* que se presentan en el capítulo 2 y los de *probabilidad* que se estudiaron en el capítulo 3. La figura 4-1 presenta un resumen esquemático de los objetivos. Como se observa en la figura, con el uso de los métodos del capítulo 2, tiraríamos en repetidas ocasiones un dado para reunir datos muestrales, que luego pueden describirse 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). Con los métodos del capítulo 3, calcularíamos la probabilidad de cada resultado. Ahora combinaremos dichos conceptos mientras creamos distribuciones de probabilidad que describan lo que *probablemente* sucederá, en lugar de lo que en realidad *sucedió*. En el capítulo 2 elaboramos tablas de frecuencias e histogramas con los valores muestrales *observados* que se reunieron en realidad, aquí construiremos distribuciones de probabilidad presentando los resultados posibles, junto con las frecuencias relativas que *esperamos*.

El "empleado" de un casino conoce la forma en que un dado debe comportarse. La tabla en el extremo derecho de la figura 4-1 representa una distribución de probabilidad que sirve como modelo para una distribución de frecuencias poblacional teóricamente perfecta. En esencia, es posible describir la tabla de frecuencias relativas para un dado que se tiró un número infinito de veces. Con tal conocimiento de los resultados de la población, seremos capaces de calcular sus características importantes, tales como la media y la desviación estándar. El resto del libro y la esencia de la estadística inferencial se basan en el conocimiento de las distribuciones de probabilidad. Iniciamos examinando el concepto de una variable aleatoria, después estudiaremos distribuciones importantes con muchas aplicaciones reales.

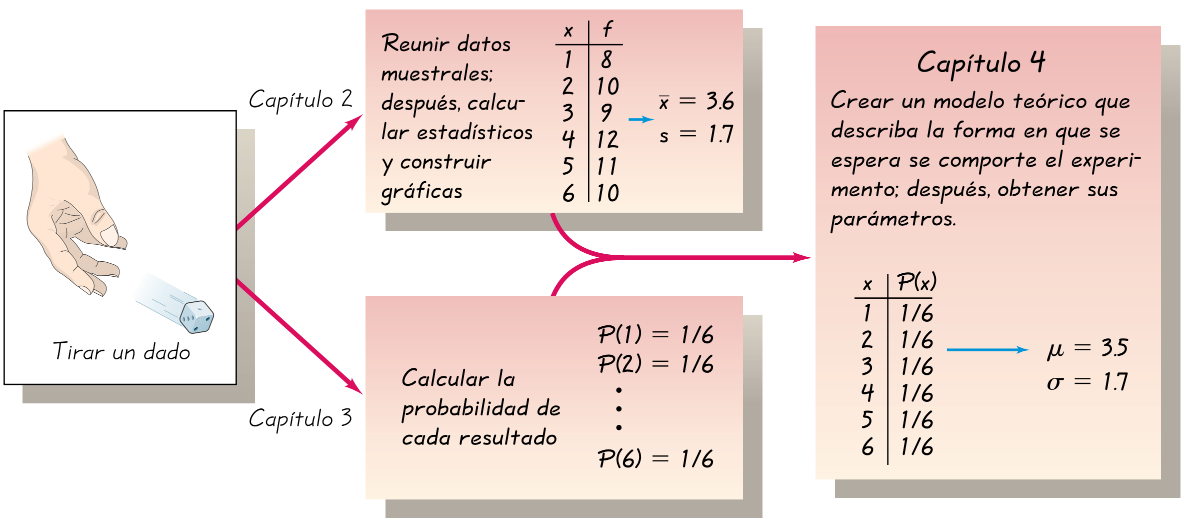


FIGURA 4-1 **Combinación de métodos descriptivos y probabilidades para formar un modelo teórico de comportamiento**

4-2 Variables aleatorias

En esta sección estudiaremos las variables aleatorias, las distribuciones de probabilidad, los procedimientos para calcular la media y la desviación estándar de una distribución de probabilidad, así como los métodos para distinguir entre resultados que pueden ocurrir por azar y aquellos que son “poco comunes”. Iniciaremos con los conceptos que se relacionan de *variable aleatoria* y *distribución de probabilidad*.

Definiciones

Variable aleatoria: variable (casi siempre representada por x) que tiene un solo valor numérico, determinado por el azar, para cada resultado de un procedimiento.

Distribución de probabilidad: gráfica, tabla o fórmula que da la probabilidad de cada valor de la variable aleatoria.

PROBLEMA 4 DEL CAPÍTULO

EJEMPLO **Género de niños** Un estudio consiste en la selección aleatoria de 14 bebés recién nacidos y el conteo del número de niñas (como en el problema del capítulo). Si consideramos que la probabilidad de niños y niñas es la misma, y

$$x = \text{número de niñas de entre 14 bebés}$$

entonces x es una variable aleatoria, porque su valor depende del azar. Los valores posibles de x son 0, 1, 2, . . . , 14. La tabla 4-1 incluye los valores de x, junto con las probabilidades correspondientes. (En la sección 4-3 aprenderemos a calcular los valores de probabilidad, como los que se listan en la tabla 4-1). Ya que la tabla 4-1 incluye la probabilidad para cada valor de la variable aleatoria x, dicha tabla describe una distribución de probabilidad.

Tabla 4-1
Probabilidades de niñas

x (niñas)	$P(x)$
0	0.000
1	0.001
2	0.006
3	0.022
4	0.061
5	0.122
6	0.183
7	0.209
8	0.183
9	0.122
10	0.061
11	0.022
12	0.006
13	0.001
14	0.000

En la sección 1-2 hicimos una distinción entre los datos discretos y los continuos. Las variables aleatorias también pueden ser discretas o continuas, en tanto que las siguientes dos definiciones son consistentes con las que se presentan en la sección 1-2.

Definiciones

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.

Variable aleatoria continua: tiene un número infinito de valores; dichos valores pueden asociarse a mediciones en una escala continua, de manera que no haya huecos o interrupciones.

Este capítulo analiza exclusivamente variables aleatorias discretas, pero los siguientes incluyen variables aleatorias continuas.

a) Variable aleatoria discreta: contador del número de asistentes a un cine.

b) Variable aleatoria continua: voltaje medido de una batería de un detector de humo.

FIGURA 4-2 Aparatos que se utilizan para contar y medir variables aleatorias discretas y continuas

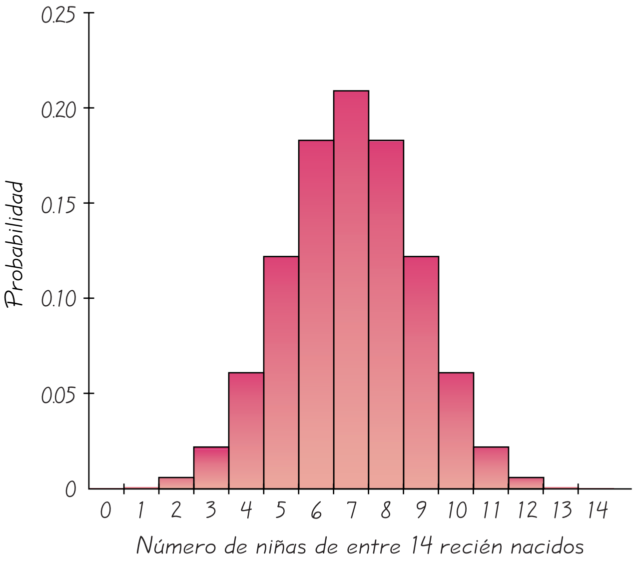
EJEMPLOS Los siguientes son ejemplos de variables aleatorias discretas y continuas:

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 fanáticos que asiste a un concierto de, N Sync es un número entero y, por lo tanto, una variable aleatoria discreta. El aparato de conteo que se muestra en la figura 4-2a 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 llega a producir una cantidad de leche cuyo valor sería cualquiera entre 0 y 5 galones. Es posible obtener 4.123456 galones, debido a que la vaca no se restringe 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 voltios. Por lo tanto, se trata de una variable aleatoria continua. El voltímetro que se muestra en la figura 4-2b es capaz de indicar valores en una escala continua; por lo tanto, se utiliza para obtener valores de una variable aleatoria *continua*.

Gráficas

Hay varias formas para graficar una distribución de probabilidad, aquí consideraremos solamente al **histograma de probabilidad**. La figura 4-3 es un histograma de probabilidad muy similar al histograma de frecuencias relativas que se estudió

FIGURA 4-3 Histograma de probabilidad del número de niñas de entre 14 bebés recién nacidos



en el capítulo 2, pero la escala vertical muestra *probabilidades* en lugar de frecuencias relativas que se basan en resultados muestrales reales.

En la figura 4-3, observe que, a lo largo del eje horizontal, los valores de 0, 1, 2, . . . , 14 se localizan en el centro de los rectángulos, lo cual implica que cada uno de los rectángulos mide una unidad, de modo que las áreas de los rectángulos son 0.000, 0.001, 0.006, etcétera. Las áreas de estos rectángulos son iguales a las *probabilidades* en la tabla 4-1. En el capítulo 5 y en capítulos posteriores veremos que tal correspondencia entre el área y la probabilidad es muy útil en la estadística.

Toda distribución de probabilidad debe satisfacer cada uno de los dos requisitos siguientes:

Requisitos de una distribución de probabilidad

1. $\Sigma P(x) = 1$ donde x toma todos los valores posibles
2. $0 \leq P(x) \leq 1$ para cada valor individual de x

El primer requisito establece que la suma de las probabilidades de todos los valores posibles de la variable aleatoria debe ser igual a 1. Lo anterior tiene sentido cuando nos damos cuenta de que los valores de la variable aleatoria x representan todos los sucesos posibles en el espacio muestral completo, de modo que tenemos la certeza (con probabilidad 1) de que uno de los sucesos ocurrirá. En la tabla 4-1, la suma de todas las probabilidades es de 0.999; sería igual a 1 si elimináramos el pequeño error por redondeo ocupando más decimales. También, la regla de probabilidad que establece que $0 \leq P(x) \leq 1$ para cualquier suceso A (dado en la sección 3-2) implica que $P(x)$ debe estar entre 0 y 1 para cualquier valor de x. Nuevamente, remítase a la tabla 4-1 y observe que cada valor individual de $P(x)$ cae entre 0 y 1 para cualquier valor de x. Puesto que la tabla 4-1 satisface ambos requisitos, es un ejemplo de una distribución de probabilidad. Una distribución de probabilidad llega a describirse como una tabla, tal como la tabla 4-1; una gráfica, como la figura 4-3, o una fórmula.

EJEMPLO ¿Describirá la tabla 4-2 una distribución de probabilidad?

SOLUCIÓN Para ser una distribución de probabilidad, $P(x)$ debe satisfacer los dos requisitos anteriores. Pero

$$\begin{aligned}\Sigma P(x) &= P(0) + P(1) + P(2) + P(3) \\ &= 0.2 + 0.5 + 0.4 + 0.3 \\ &= 1.4 \left[\text{que demuestra que } \Sigma P(x) \neq 1\right]\end{aligned}$$

Como no se satisface el primer requisito, concluimos que la tabla 4-2 no describe una distribución de probabilidad.

Tabla 4-2
Probabilidades de una variable aleatoria

x	$P(x)$
0	0.2
1	0.5
2	0.4
3	0.3

EJEMPLO ¿Determina si $P(x) = x/3$ (donde x puede ser 0, 1 o 2) es 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. $\Sigma P(x) = \frac{0}{3} + \frac{1}{3} + \frac{2}{3} = \frac{3}{3} = 1$

2. Cada uno de los valores $P(x)$ se encuentran entre 0 y 1.

Como ambos requisitos se satisfacen, la función $P(x)$ de este ejemplo es una distribución de probabilidad.

Media, varianza y desviación estándar

Recuerde que en el capítulo 2 describimos las siguientes características importantes de los datos (que pueden recordarse por medio de las siglas CVDDT "**C**uidado con los **V**irus que **D**estruyen **D**atos y **T**rabajo"):

1. *Centro:* valor representativo o promedio 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:* naturaleza o forma de la distribución de los datos (tales como normales, uniformes o sesgadas).
4. *Datos distantes:* valores muestrales que se alejan mucho de la vasta mayoría de los otros valores de la muestra.
5. *Tiempo:* características cambiantes de los datos a través del tiempo.

El histograma de probabilidad nos ofrece información acerca de la naturaleza o forma de la distribución. Además, podemos calcular la media, la varianza y la desviación estándar de los datos, los cuales proporcionan información acerca de otras características. La media, la varianza y la desviación estándar de una distribución de probabilidad se calculan aplicando las fórmulas 4-1, 4-2, 4-3 y 4-4.

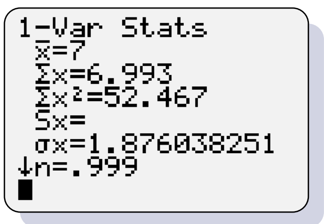
Fórmula 4-1 $\mu = \Sigma[x \cdot P(x)]$ media de una distribución de probabilidad

Fórmula 4-2 $\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)]$ varianza de una distribución de probabilidad

Fórmula 4-3 $\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$ varianza de una distribución de probabilidad

Fórmula 4-4 $\sigma = \sqrt{\Sigma[x^2 \cdot P(x)] - \mu^2}$ desviación estándar de una distribución de probabilidad

TI-83 Plus



Precaución: Evalúe $\Sigma[x^2 \cdot P(x)]$ elevando al cuadrado cada valor de x, multiplicando cada cuadrado por el $P(x)$ correspondiente y sumando.

La calculadora TI-83 Plus resulta útil para calcular la media y la desviación estándar. Aquí se muestra la representación visual de la calculadora TI-83 Plus, de la distribución de probabilidad descrita en la tabla 4-1. En la imagen de la TI-83 Plus, el valor que se muestra como $\bar{x}$ es en realidad el valor de la media μ, en tanto que el valor que se presenta como σx es el valor de la desviación estándar σ. Es decir, $\mu = 7$ y $\sigma = 1.876038251$.

Fundamentos de las fórmulas 4-1 a la 4-4

¿Por qué sirven las fórmulas 4-1 a la 4-4? La fórmula 4-1 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 rescribir la fórmula de la media de una tabla de frecuencias, de modo que se aplique a una población, y luego cambiando su forma, obtendremos

$$\mu = \frac{\Sigma(f \cdot x)}{N} = \Sigma\frac{f \cdot x}{N} = \Sigma x \cdot \frac{f}{N} = \Sigma x \cdot P(x)$$

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, de modo que f/N es la probabilidad del valor de x.

Un razonamiento similar nos permite considerar la fórmula de la varianza del capítulo 2 y aplicarla a una variable aleatoria para una distribución de probabilidad; el resultado es la fórmula 4-2. La fórmula 4-3 es una versión abreviada que siempre producirá el mismo resultado que la fórmula 4-2. Aunque la fórmula 4-3 es más fácil de usar, la fórmula 4-2 es más sencilla de comprender directamente. Con base en la fórmula 4-2, expresamos la desviación estándar como

$$\sigma = \sqrt{\Sigma(x - \mu)^2 \cdot P(x)}$$

o como la forma equivalente dada en la fórmula 4-4.

Cuando utilice las fórmulas 4-1 a la 4-4, 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 debido a circunstancias especiales, tales como resultados que requieren más decimales para tener un significado. Por ejemplo, para aviones de propulsión a chorro de cuatro motores, el número medio de motores que funcionan exitosamente 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 de aviones 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 2-5), 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 dos desviaciones es-

¿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, 7,000 ahogamientos, 900 muertes en bicicletas, 800 muertes por relámpagos y 1,150 muertes por picaduras de abeja. Desde luego, tales cifras no significan necesariamente que el paracaidismo sea más seguro que andar en bicicleta o la natación. En una comparación justa, deben incluirse 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. También voló en una ala delta, un globo aerostático y en el dirigible Goodyear.

tándar de la media; no es frecuente que un valor difiera de la media en más de dos desviaciones estándar. (Por lo general, el uso de dos desviaciones estándar no es un valor absolutamente rígido; en su lugar, se utilizan otros valores, como el 3). De esta manera, identificamos valores "poco comunes" si se determinara que caen fuera de tales límites:

$$\text{valor máximo común} = \mu + 2\sigma$$
$$\text{valor mínimo común} = \mu - 2\sigma$$

PROBLEMA DEL CAPÍTULO 4

EJEMPLO La tabla 4-1 describe la distribución de probabilidad del número de niñas entre 14 bebés recién nacidos que se seleccionaron aleatoriamente. Suponiendo que repetimos el estudio, seleccionamos aleatoriamente 14 bebés recién nacidos y contamos el número de niñas en cada ocasión, calcule el número medio de niñas (de entre 14), la varianza y la desviación estándar. Utilice dichos resultados y la regla práctica del intervalo para obtener los valores máximo y mínimo comunes.

SOLUCIÓN En la tabla 4-3, las dos columnas a la izquierda describen la distribución de probabilidad que se presentó en la tabla 4-1; elaboramos las tres columnas a la derecha con el propósito de lograr los cálculos requeridos.

Tabla 4-3 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	0.000	0.000	0	0.000
1	0.001	0.001	1	0.001
2	0.006	0.012	4	0.024
3	0.022	0.066	9	0.198
4	0.061	0.244	16	0.976
5	0.122	0.610	25	3.050
6	0.183	1.098	36	6.588
7	0.209	1.463	49	10.241
8	0.183	1.464	64	11.712
9	0.122	1.098	81	9.882
10	0.061	0.610	100	6.100
11	0.022	0.242	121	2.662
12	0.006	0.072	144	0.864
13	0.001	0.013	169	0.169
14	0.000	0.000	196	0.000
Total		6.993		52.467
		↑		↑
		$\sum[x \cdot P(x)]$		$\sum[x^2 \cdot P(x)]$

Con el uso de las fórmulas 4-1 y 4-3, así como con los resultados de la tabla, obtendremos

$$\mu = \Sigma[x \cdot P(x)] = 6.993 = 7.0 \quad \text{(redondeado)}$$

$$\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$$

$$= 52.467 - 6.993^2 = 3.564951 = 3.6 \quad \text{(redondeado)}$$

La desviación estándar es la raíz cuadrada de la varianza, entonces

$$\sigma = \sqrt{3.564951} = 1.9 \quad \text{(redondeado)}$$

Sabemos que entre los grupos de 14 bebés recién nacidos, el número medio de niñas es de 7.0, la varianza es de 3.6 "niñas cuadradas" y la desviación estándar es de 1.9 niñas.

Con la aplicación de la regla práctica del intervalo, ahora calculamos los valores máximo y mínimo comunes de la siguiente manera:

$$\text{valor máximo común: } \mu + 2\sigma = 7.0 + 2(1.9) = 10.8$$

$$\text{valor mínimo común: } \mu - 2\sigma = 7.0 - 2(1.9) = 3.2$$

INTERPRETACIÓN Con base en estos resultados, concluimos que para grupos de 14 bebés que se seleccionaron aleatoriamente, el número de niñas debe caer comúnmente entre 3.2 y 10.8.

Identificación de resultados poco comunes con probabilidades

Recomendación importante: La siguiente explicación incluye algunos conceptos difíciles, pero también un método sumamente importante, que se utiliza con frecuencia en la estadística. Debe hacer su mejor esfuerzo para comprender dicha explicación y leerla varias veces si es necesario. Tenga en mente que tal explicación se basa en la *regla del suceso poco común* que se estudió en la sección 3-2:

Si, bajo un supuesto dado (como el de que niñas y niños son igualmente probables), la probabilidad de un suceso particular que se observa (como 13 niñas en 14 nacimientos) es extremadamente pequeña, concluimos que el supuesto probablemente no sea correcto.

En el problema del capítulo señalamos que, con la técnica MicroSort, resultaron 13 niñas entre 14 bebés. ¿Será el resultado poco común? ¿Sugiere que en realidad la técnica es efectiva, o podrían resultar 13 niñas entre 14 bebés sólo por azar? Para resolver esto, utilizamos la regla práctica del intervalo con la finalidad de calcular los resultados probables máximo y mínimo. Pero aquí consideraremos otro método: calcularemos la probabilidad de tener 13 *o más* niñas (no la de tener *exactamente* 13 niñas). Es difícil entender por qué la probabilidad de 13 *o más niñas* es la probabilidad relevante; por lo tanto, trataremos de aclararlo con un ejemplo más sencillo.

Suponga que lanza una moneda para determinar si se ven favorecidas las caras y que 1000 lanzamientos dan como resultado 501 caras, lo cual no es evidencia de que la moneda favorezca las caras, ya que es muy fácil obtener un resultado de 501 caras en 1000 lanzamientos al azar. Sin embargo, la probabilidad de obtener *exactamente* 500 caras en 1000 lanzamientos es bastante baja: 0.0252. Esa baja probabilidad refleja el hecho de que, con 1000 lanzamientos, cada número *es-*

pecífico de caras tendrá una probabilidad sumamente baja. Sin embargo, no consideramos que 500 caras en 1000 lanzamientos sea *poco común*, puesto que la probabilidad de obtener *al menos* 501 caras es alta: 0.487. Este principio se generaliza de la siguiente manera:

Uso de las probabilidades para determinar resultados poco comunes

- **Inusualmente alto:** x éxitos en n ensayos es un número *inusualmente alto* de éxitos, si $P(x \text{ o más})$ es muy pequeño (como 0.05 o menos).
- **Inusualmente bajo:** x éxitos en n ensayos es un número *inusualmente bajo* de éxitos, si $P(x \text{ o menos})$ es muy pequeño (como 0.05 o menos).

PROBLEMA 4 DEL CAPÍTULO

EJEMPLO Selección del género Con el uso de los dos criterios anteriores que se basan en probabilidades, ¿será poco común que resulten 13 niñas en 14 nacimientos? ¿Parecería que la técnica de MicroSort de la selección del género es efectiva?

SOLUCIÓN Trece niñas entre 14 nacimientos es inusualmente alto si P (13 o más niñas) es muy pequeño. Si nos remitimos a la tabla A-1, obtendremos el siguiente resultado:

$$\begin{aligned} P\,(13 \text{ o más niñas}) &= P(13) + P(14) \\ &= 0.001 + 0.000 \\ &= 0.001 \end{aligned}$$

INTERPRETACIÓN Puesto que la probabilidad de 0.001 es demasiado baja, concluimos que es poco común que resulten 13 niñas en 14 nacimientos. Lo anterior sugiere que la técnica MicroSort de selección del género parece ser efectiva, ya que es altamente improbable que el resultado de 13 niñas en 14 nacimientos suceda por azar.

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 pudiesen continuar de manera indefinida. Los usos del valor esperado (también se le llama *esperanza* o *esperanza matemática*), que son extensos y variados, juegan un papel muy importante en una área de aplicación que se denomina *teoría de decisión*.

Definición

Valor esperado (de una variable aleatoria discreta): se denota con E y representa el valor promedio de los resultados. Se obtiene calculando el valor de $\Sigma[x \cdot P(x)]$.

$$E = \Sigma[x \cdot P(x)]$$

Con la fórmula 4-1 vemos que $E = \mu$; es decir, la media de una variable aleatoria discreta es la misma que su valor esperado. Repita el procedimiento de lan-

Elección de números de la 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. Puesto que los números ganadores se seleccionan aleatoriamente, cualquier elección de seis números tendrá la misma posibilidad, 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 de dólares, 52,000 boletos incluían 1, 2, 3, 4, 5, 6; si tal combinación hubiera ganado, el premio hubiese sido de tan sólo 1,000 dólares. Es más sensato elegir combinaciones que muchas otras personas no seleccionan. Evite combinaciones que forman un patrón en el boleto.

zar una moneda cinco veces; el número *medio* de caras es de 2.5; cuando se lanza una moneda cinco veces, el *valor esperado* del número de caras es también de 2.5.

EJEMPLO El juego Pick 3 de Nueva Jersey Hace años, miembros de grupos del crimen organizado realizaban juegos de números. En la actualidad, muchos miembros de gobiernos organizados, así como algunos gobiernos no tan bien organizados, realizan juegos similares. El juego Pick 3 de Nueva Jersey es muy similar a los antiguos juegos ilegales. Una apuesta "legal" funciona de la siguiente manera: apueste 50 centavos y seleccione un número de tres dígitos entre 000 y 999. Si sus tres dígitos coinciden con los que se seleccionan aleatoriamente, recibiría 275 dólares, así que obtendría una ganancia neta de 274.50 dólares (puesto que no se le devuelven sus 50 centavos de apuesta). Suponga que apuesta 50 centavos al número 007, ¿cuál es el valor esperado de ganar o perder?

SOLUCIÓN Para esta apuesta hay dos resultados sencillos: gana o pierde. Como usted eligió el número 007 y hay 1000 posibilidades (desde 000 hasta 999), su probabilidad de ganar es de 1/1000 (o 0.001) y la de perder es de 999/1000 (o 0.999). La tabla 4-4 resume dicha situación.

En la tabla 4-4 observamos que cuando apostamos 50 centavos en el juego Pick 3 de Nueva Jersey, nuestro valor esperado es

$$E = \Sigma[x \cdot P(x)] = -22.5 \text{ centavos}$$

El resultado puede interpretarse de la siguiente manera: a la larga, por cada apuesta de 50 centavos podemos esperar perder un promedio de 22.5 centavos. En cualquier juego individual perderá 50 centavos o ganará 274.50 dólares. Aunque no es posible que usted pierda 22.5 centavos en un juego individual, el valor esperado de −22.5 centavos muestra que, después de muchos juegos, la pérdida promedio por juego es de 22.5 centavos.

Tabla 4-4 Juego Pick 3 de Nueva Jersey

Suceso	x	$P(x)$	$x \cdot P(x)$
Ganancia	\$274.50	0.001	\$0.2745
Pérdida	−\$0.50	0.999	−\$0.4995
Total			−\$0.225 (o −22.5¢)

En esta sección aprendimos que una variable aleatoria tiene un valor numérico asociado a cada resultado de algún procedimiento aleatorio, así como 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. También aprendimos que no debemos esperar enriquecernos con el juego de lotería Pick 3 de Nueva Jersey.

4-2 Destrezas y conceptos básicos

Identificación de variables aleatorias discretas continuas. *En los ejercicios 1 y 2, identifique si la variable aleatoria dada es discreta o continua.*

1. **a.** La estatura de un jugador de basquetbol de la NBA, que se selecciona aleatoriamente.
b. El número de puntos que anota en una temporada un jugador de basquetbol de la NBA, seleccionado aleatoriamente.
c. El tiempo exacto de juego de un jugador de basquetbol de la NBA, que se selecciona aleatoriamente.
d. El número de atletas que participaron en cualquier juego de la NBA en una temporada.
e. El salario de un jugador de basquetbol de la NBA, seleccionado aleatoriamente.

2. **a.** El costo de la realización de una película que se selecciona aleatoriamente.
b. El número de películas que actualmente se exhiben en los cines de Estados Unidos.
c. La duración exacta de una película seleccionada aleatoriamente.
d. El número de actores que aparecen en una película que se selecciona aleatoriamente.
e. El peso del actor principal de una película seleccionada aleatoriamente.

Identificación de distribuciones de probabilidad. *En los ejercicios 3 a 10, determine si se trata de una distribución de probabilidad. En los casos en que no se describa una distribución de probabilidad, identifique los requisitos que no se satisfacen. En los casos en que se describa una distribución de probabilidad, calcule su media y desviación estándar.*

3. Selección de género En un estudio con el método MicroSort de selección del género, las parejas de un grupo control no reciben tratamiento y cada una de ellas tiene tres hijos. La distribución de probabilidad del número de niñas se presenta en la tabla anexa.

x	$P(x)$
0	0.125
1	0.375
2	0.375
3	0.125

4. Control de calidad de DVD Durante la fabricación del DVD de Sony, se seleccionan aleatoriamente grupos de DVD y se calcula el número de defectos x en cada grupo.

x	$P(x)$
0	0.502
1	0.365
2	0.098
3	0.011
4	0.001

5. Renta de videocintas La tabla adjunta se construye de datos que se obtienen en un estudio del número de videocintas rentadas en Blockbuster.

x	$P(x)$
0	0.04
1	0.26
2	0.36
3	0.20
4	0.08

6. Seguro de vida La compañía Telektronic brinda pólizas de seguro de vida a sus cuatro ejecutivos principales; la variable aleatoria x es el número de estos empleados que sobreviven durante el año siguiente.

x	$P(x)$
0	0.0000
1	0.0001
2	0.0006
3	0.0387
4	0.9606

7. Sentencias previas Se selecciona aleatoriamente a un preso, convicto por manejar intoxicado; la distribución de probabilidad del número x de sentencias previas por este delito se describe en la tabla adjunta (según datos del Departamento de Justicia de Estados Unidos).

x	$P(x)$
0	0.512
1	0.301
2	0.132
3	0.055

8. Vuelos sobresaturados La línea Air America tiene, por rutina, la política de vender un número de boletos que rebasa el número de asientos por cada vuelo, ya que la experiencia demuestra que algunos pasajeros no se presentan. La variable aleatoria x representa el número de pasajeros que no pueden abordar porque hay más pasajeros que asientos.

x	$P(x)$
0	0.805
1	0.113
2	0.057
3	0.009
4	0.002

9. Número de juegos en una Serie Mundial de Beisbol Con base en resultados que se encontraron en el *Information Please Almanac*, hay una probabilidad del 0.1809 de que la Serie Mundial de Beisbol dure cuatro juegos, una probabilidad del 0.2234 de que dure cinco juegos, una probabilidad de 0.2234 de que dure seis juegos y una probabilidad del 0.3723 de que dure siete juegos. ¿Será poco común que un equipo se corone ganando cuatro juegos seguidos?

10. 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, en tanto que las probabilidades correspondientes son 0.0016, 0.0250, 0.1432, 0.3892 y 0.4096. ¿Será poco común seleccionar aleatoriamente cuatro consumidores y descubrir que ninguno de ellos reconoce la marca Sony?

11. Cálculo del valor esperado en los dados Cuando usted apuesta 5 dólares en un casino a “pasar la línea” en el juego de dados, hay una probabilidad de 244/495 de que gane \$5 y una probabilidad de 251/495 de que pierda \$5. ¿Cuál es su valor esperado? A la larga, ¿cuánto pierde por cada dólar que apueste?

12. Cálculo del valor esperado en la ruleta Cuando usted apuesta \$5 en un casino al número 7 en la ruleta, tiene una probabilidad de 1/38 de ganar \$175, y una probabilidad de 37/38 de perder \$5. Si apuesta \$5 a que el resultado es un número impar, la probabilidad de ganar \$5 es de 18/38 y la probabilidad de perder \$5 es de 20/38.
   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 7, apostar a número impar o no apostar? ¿Por qué?

13. Cálculo del valor esperado para una póliza de seguro de vida La compañía de seguros CNA le cobra a Mike \$250 por un año de una póliza de seguro de vida de \$100,000. Puesto que Mike es un hombre de 21 años, hay una probabilidad del 0.9985 de que sobreviva durante un año (según datos del National Center for Health Statistics de Estados Unidos).
   a. Desde la perspectiva de Mike, ¿cuáles son los valores de los dos resultados diferentes?
   b. Si Mike compra la póliza, ¿cuál es su valor esperado?
   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 lugar de obtener una ganancia?
   d. Puesto que el valor esperado de Mike es negativo (de modo que la compañía obtiene una ganancia), ¿por qué es aconsejable que Mike o cualquiera otra persona adquiera un seguro de vida?

14. Cálculo del valor esperado de la rifa organizada por una revista Recientemente, 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. 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?

15. Cálculo del valor esperado del juego Pick 4 de Nueva Jersey En el juego de lotería Pick 4 de Nueva Jersey, usted paga 5 centavos para seleccionar una secuencia de cuatro dígitos, como 7273. Si usted gana al seleccionar la misma secuencia de los cuatro dígitos resultantes, obtiene $2788.

a. ¿Cuántas selecciones distintas son posibles?
b. ¿Cuál es la probabilidad de ganar?
c. Si gana, ¿cuál es su ganancia neta?
d. Calcule el valor esperado.
e. Uno de los ejemplos de esta sección demostró que si apuesta 50 centavos en el juego Pick 3 de Nueva Jersey, su valor esperado es -22.5 centavos. ¿Cuál juego es mejor: Pick 3 o Pick 4 de Nueva Jersey? ¿Por qué?

16. Cálculo del valor esperado del juego Pick 3 de Illinois En el juego Pick 3 de Illinois, usted paga 50 centavos para seleccionar una secuencia de tres dígitos, como 911. Si usted gana al seleccionar la misma secuencia de los tres dígitos resultantes, obtendría $250.

a. ¿Cuántas selecciones distintas son posibles?
b. ¿Cuál es la probabilidad de ganar?
c. Si gana, ¿cuál es su ganancia neta?
d. Calcule el valor esperado.
e. Uno de los ejemplos de esta sección demostró que si apuesta 50 centavos en el juego Pick 3 de Nueva Jersey, su valor esperado es -22.5 centavos. ¿Cuál juego es mejor: Pick 3 de Illinois o Pick 3 de Nueva Jersey? ¿Por qué?

17. Determinación de la eficacia de la técnica de selección del género Suponga que, en una prueba de una técnica de selección del género, un ensayo clínico da como resultado nueve niñas en 14 nacimientos. Remítase a la tabla 4-1 y obtenga las probabilidades que se indican.

a. Calcule la probabilidad de que resulten exactamente nueve niñas en 14 nacimientos.
b. Calcule la probabilidad de tener nueve o más niñas en 14 nacimientos.
c. ¿Qué probabilidad es relevante para determinar si nueve niñas en 14 nacimientos es un suceso inusualmente alto: el resultado del inciso *a* o el del inciso *b*?
d. ¿Sugerirá el resultado de nueve niñas en 14 nacimientos que la técnica de selección del género es efectiva? ¿Por qué?

18. Determinación de la eficacia de la técnica de selección del género Suponga que, en una prueba de una técnica de selección del género, un ensayo clínico da como resultado 12 niñas en 14 nacimientos. Remítase a la tabla 4-1 y obtenga las probabilidades que se indican.

a. Calcule la probabilidad de tener exactamente 12 niñas en 14 nacimientos.
b. Calcule la probabilidad de tener 12 o más niñas en 14 nacimientos.
c. ¿Qué probabilidad es relevante para determinar si 12 niñas en 14 nacimientos es un suceso inusualmente alto: el resultado del inciso *a* o del inciso *b*?
d. ¿Sugerirá el resultado de 12 niñas en 14 nacimientos que la técnica de selección del género es efectiva? ¿Por qué?

19. Determinación de la eficacia de la técnica de selección del género Suponga que, en una prueba de una técnica de selección del género, un ensayo clínico tiene como resultado 11 niñas en 14 nacimientos. Remítase a la tabla 4-1 y obtenga las probabilidades que se indican.

continúa

a. ¿Cuál es el valor de probabilidad que debe utilizarse para determinar si el resultado de 11 niñas en 14 nacimientos es inusualmente alto?
b. ¿Sugerirá el resultado de 11 niñas en 14 nacimientos que la técnica de selección del género es efectiva? ¿Por qué?

20. Determinación de la eficacia de la técnica de selección del género Suponga que, en una prueba de una técnica de selección del género, un ensayo clínico tiene como resultado 10 niñas en 14 nacimientos. Remítase a la tabla 4-1 y obtenga las probabilidades que se indican.
a. ¿Cuál es el valor de probabilidad que debe utilizarse para determinar si el resultado de 10 niñas en 14 nacimientos es inusualmente alto?
b. ¿Sugerirá el resultado de 10 niñas en 14 nacimientos que la técnica de selección del género es efectiva? ¿Por qué?

21. Poderes psíquicos Bob, quien se considera un "mentalista", afirma que puede pasar una prueba de verdadero/falso adivinando. Para comprobar su afirmación, se le plantean 14 preguntas con respuestas de verdadero/falso, de las cuales, ocho responde correctamente. Bob asevera que obtener ocho respuestas correctas, en 14 preguntas, es una prueba de sus poderes especiales, ya que obtuvo más respuestas correctas que las siete que se esperarían por el azar. ¿Es válida la aseveración de Bob? ¿Por qué? (*Sugerencia:* Para calcular la probabilidad que se requiere, remítase a la tabla 4-1 y cambie "niñas" por "correcto". Las probabilidades de niñas serán las mismas probabilidades de conjeturas correctas).

22. Poderes psíquicos Bob, quien se considera un "mentalista", afirma que es capaz de pasar una prueba de verdadero/falso adivinando. Para comprobar su afirmación, se le plantean 14 preguntas con respuestas de verdadero/falso, de las cuales dos responde correctamente. ¿Será poco común el número de respuestas correctas? ¿Por qué? ¿Será válida la afirmación de Bob? (*Sugerencia*: Para calcular la probabilidad que se requiere, remítase a la tabla 4-1 y cambie "niñas" por "correcto". Las probabilidades de niñas serán las mismas probabilidades de conjeturas correctas).

4-2 Más allá de lo básico

23. Bonos especulativos Kim Hunter tiene $1000 para invertir, por lo que 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? ¿Debe ella elegir cualquiera de los dos bonos? ¿Por qué?

24. Cálculo de la media y la desviación estándar Sea x la variable aleatoria, la cual representa 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).

25. Partes defectuosas: cálculo de la media y la desviación estándar Sky Ranch es un proveedor de partes para aeronaves. Sus existencias incluyen ocho altímetros que están correctamente calibrados y dos que no lo están. Se seleccionan tres altímetros aleatoriamente sin reemplazo. Sea x la variable aleatoria, la cual representa el número de aparatos que no están calibrados correctamente. Calcule la media y la desviación estándar de la variable aleatoria x.

26. Números generados por computadora, transformados a puntuación z Con frecuencia se utilizan computadoras para generar aleatoriamente los últimos dígitos de números telefónicos de sujetos potenciales a encuestar. Los dígitos se seleccionan de manera que todos sean igualmente probables. La variable aleatoria x es el número elegido.

continúa

a. Calcule la media en la desviación estándar de la variable aleatoria x.
b. Calcule la puntuación z para cada valor de x; después, la media y la desviación estándar de las puntuaciones z.
c. ¿Resultarán la misma media y la desviación estándar del inciso b de cada distribución de probabilidad?

27. Enteros con la misma probabilidad: media y desviación estándar Suponga que una distribución de probabilidad está descrita por la variable aleatoria discreta x, que puede tomar los valores 1, 2, . . . , n, y que dichos valores son igualmente probables. La distribución de probabilidad tiene una media y una desviación estándar como se describe a continuación:

$$\mu = \frac{n+1}{2} \quad \text{y} \quad \sigma = \sqrt{\frac{n^2-1}{12}}$$

a. Demuestre que $\mu = (n+1)/2$ para el caso de $n = 5$.
b. Demuestre que $\sigma = \sqrt{(n^2-1)/12}$ para el caso de $n = 5$.
c. Con la finalidad de probar a un individuo que afirma tener poderes extrasensoriales, usted selecciona aleatoriamente números enteros entre 1 y 20, en tanto que la variable aleatoria x es el número que se selecciona. Calcule la media y la desviación estándar de x.

28. Dados con marcas que permitan una distribución uniforme Suponga que tiene dos dados en blanco, de modo que puede marcar las 12 caras con los números que desee. 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).

4-3 Distribuciones de probabilidad binomial

En la sección 4-2 estudiamos diversas distribuciones discretas de probabilidad, pero en esta sección nos enfocaremos en un tipo específico: la distribución de probabilidad binomial. Las distribuciones de probabilidad binomial son importantes porque nos permiten enfrentar circunstancias en las que los resultados pertenecen a *dos* categorías relevantes, tales como productos aceptables/defectuosos o respuestas sí/no a una pregunta de encuesta. El problema del capítulo implica el conteo del número de niñas en 14 nacimientos; incluye las dos categorías niño/niña, por lo que posee el elemento fundamental requerido de “dualidad”. En la siguiente definición se plantean otros requisitos.

Definición

Distribución de probabilidad binomial: resulta de un procedimiento que cumple con todos los requisitos siguientes:

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 otros ensayos).
3. Todos los resultados de cada ensayo deben estar clasificados en *dos categorías*.
4. Las probabilidades tienen que mantenerse *constantes* para cada ensayo.

Si un procedimiento satisface los cuatro requisitos, la distribución de la variable aleatoria x se denomina *distribución de probabilidad binomial* (o *distribución binomial*). Suele usarse la notación siguiente:

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 modo que

$$P(\text{E}) = p \qquad (p = \text{probabilidad de un éxito})$$

$$P(\text{F}) = 1 - p = q \qquad (q = \text{probabilidad de un fracaso})$$

- n denota el número fijo de ensayos.
- x denota un número específico de éxitos en n ensayos, de modo que x puede ser cualquier número entero entre 0 y n, inclusive.
- p denota la probabilidad de *éxito* en *uno* de n ensayos.
- q denota la probabilidad de *fracaso* en *uno* 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 representa algo bueno. 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 refieran a la misma categoría que se designa 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 sugerencia importante para trabajar con problemas de probabilidad binomial:

Asegúrese de que x y p se refieran a la *misma* categoría que se denomina como un éxito.

Cuando seleccionamos una muestra para algún análisis estadístico, por lo general realizamos el muestreo sin reemplazo. Por ejemplo, al probar artículos manufacturados o realizar encuestas, solemos diseñar el proceso de muestreo de modo que los artículos elegidos no se seleccionan una segunda vez. Estrictamente hablando, el muestreo sin reemplazo implica sucesos dependientes, que violan el segundo requisito de la definición anterior. Sin embargo, la siguiente regla práctica se basa en el hecho de que si la muestra es muy pequeña, en relación con el tamaño de la población, podemos tratar a los ensayos como independientes (aun cuando en realidad sean dependientes), ya que la diferencia en los resultados será insignificante.

Cuando se realiza 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. (Es decir, $n \leq 0.05N$).

Profetas de las ganancias

Muchos libros y programas de computadoras aseguran ser útiles para predecir números ganadores de la lotería. Algunos utilizan la teoría de que ciertos números están "rezagados" (y hay que seleccionarlos), ya que no han salido con frecuencia; otros creen en la teoría de que algunos números son "fríos" (y deben evitarse), puesto que no han salido con frecuencia; incluso existen más que utilizan la astrología, la numerología o los sueños. Ya que la selección de las combinaciones de números ganadores de la lotería son sucesos independientes, dichas teorías 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 forma que si usted gana, no se le obligará a compartir sus ganancias con otros. Por tal razón, la combinación de 1, 2, 3, 4, 5 y 6 es inadecuada, puesto 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.

EJEMPLO Análisis de respuestas de opción múltiple Por su facilidad de corrección, las preguntas de opción múltiple se utilizan con frecuencia para realizar exámenes en el salón de clases, para la prueba SAT, para la prueba MCAT en las escuelas de medicina, la prueba LSAT en las escuelas de leyes y en muchas otras circunstancias. Un profesor que imparte un curso de psicología del comportamiento anormal planea aplicar un examen sorpresa que consta de cuatro preguntas de opción múltiple, cada una con cinco respuestas posibles (a, b, c, d, e), pero sólo una de ellas correcta. Supongamos que un estudiante sin preparación adecuada hace adivinanzas al azar y que queremos calcular la probabilidad de que tenga exactamente tres respuestas correctas en las cuatro preguntas.

a. ¿Resultará este proceso en una distribución binomial?

b. Si el proceso resulta en una distribución binomial, identifique los valores de n, x, p y q.

SOLUCIÓN

a. El procedimiento sí satisface los requisitos de una distribución binomial, como se muestra a continuación:

1. El número de ensayos (4) es fijo.
2. Los cuatro ensayos son independientes, ya que una respuesta correcta o incorrecta para cada pregunta individual no afecta la probabilidad de que otra pregunta sea correcta o incorrecta.
3. Cada uno de los cuatro ensayos tiene dos categorías de resultados posibles: correcta o incorrecta.
4. Para cada pregunta hay cinco respuestas posibles (a, b, c, d, e), pero sólo una de ellas es correcta, por lo que la probabilidad de una respuesta correcta es de $1/5$ o 0.2. La probabilidad de cada uno de los cuatro ensayos permanece constante.

b. Habiendo concluido que el procedimiento dado sí resulta en una distribución binomial, procedamos a identificar los valores de n, x, p y q.

1. Con cuatro preguntas de un examen, tenemos que $n = 4$.
2. Buscamos la probabilidad de exactamente tres respuestas correctas; entonces, $x = 3$.
3. La probabilidad de éxito (respuesta correcta) para una pregunta es de 0.2; por lo tanto, $p = 0.2$.
4. La probabilidad de fracaso (respuesta incorrecta) es de 0.8; por lo tanto, $q = 0.8$.

Nuevamente, es muy importante asegurarse de que tanto x como p se refieren al mismo concepto de "éxito". En este ejemplo, usamos x para contar las respuestas *correctas*, de modo que p debe ser la probabilidad de una respuesta *correcta*. Por consiguiente, x y p sí utilizan aquí el mismo concepto de éxito (respuesta correcta).

Ahora presentaremos tres métodos para calcular las probabilidades correspondientes a la variable aleatoria x en una distribución binomial. El primero, que implica cálculos empleando la *fórmula de probabilidad binomial*, es la base de los otros dos. El segundo implica el uso de la tabla A-1; y el tercero, el uso de un programa de cómputo o una calculadora. Si está utilizando alguna de estas dos herramientas que producen, 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 tales 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, las probabilidades pueden calcularse con el uso de la fórmula de la probabilidad binomial.

Fórmula 4-5

$$P(x) = \frac{n!}{(n - x)!x!} \cdot p^x \cdot q^{n-x} \qquad \text{para } x = 0, 1, 2, \ldots, n$$

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 introdujo en la sección 3-7, denota el producto de factores decrecientes. Dos ejemplos de factoriales son $3! = 3 \cdot 2 \cdot 1 = 6$ y $0! = 1$ (por definición). Muchas calculadoras incluyen una tecla para el factorial, al igual que una tecla con ${}_nC_r$ que permite simplificar los cálculos. Para las calculadoras con esa tecla, utilice esta versión de la fórmula de probabilidad binomial (donde n, x, p y q son iguales en la fórmula 4-5):

$$P(x) = {}_nC_r \cdot p^x \cdot q^{n-x}$$

La calculadora TI-83 Plus se diseñó para calcular de manera automática las probabilidades binomiales por medio de tal fórmula. Los detalles para el manejo de la calculadora TI-83 Plus se explicarán más adelante en esta sección.

EJEMPLO Análisis de respuestas de opción múltiple Aplique la fórmula de la probabilidad binomial para calcular la probabilidad de tener exactamente tres respuestas correctas, cuando se adivina al azar en las cuatro preguntas de opción múltiple. Esto es, calcule $P(3)$, siendo que $n = 4$, $x = 3$, $p = 0.2$ y $q = 0.8$.

SOLUCIÓN Con los valores dados de n, x, p y q en la fórmula de probabilidad binomial (fórmula 4-5), obtenemos

$$\begin{aligned} P(x) &= \frac{4!}{(4 - 3)!3!} \cdot 0.2^3 \cdot 0.8^{4-3} \\ &= \frac{4!}{1!3!} \cdot 0.008 \cdot 0.8 \\ &= (4)(0.008)(0.8) = 0.0256 \end{aligned}$$

La probabilidad de tener exactamente tres respuestas correctas de cuatro, es de 0.0256.

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}, así como luego simplemente multiplicar los tres factores, como se muestra al final de los cálculos del ejemplo anterior. No redondee demasiado al calcular esos tres factores; hágalo al final.

Método 2: Uso de la tabla A-1 del Apéndice A En algunos casos calculamos fácilmente las probabilidades binomiales con sólo remitirnos a la tabla A-1 del Apéndice A. Primero localice n y el valor correspondiente de x deseado. En esta etapa, debe aislarse un renglón de números. Ahora, alinee dicho renglón con la probabilidad correspondiente de p, usando la columna que cruza. El número representa la probabilidad deseada. El 0+ indica una probabilidad tan pequeña como 0.000000345.

Al margen se muestra parte de la tabla A-1. Con $n = 4$ y $p = 0.2$ en una distribución binomial, las probabilidades de 0, 1, 2, 3 y 4 éxitos son de 0.410, 0.410, 0.154, 0.026 y 0.002, respectivamente.

De la tabla A-1:

n	x	p 0.20
4	0	0.410
	1	0.410
	2	0.154
	3	0.026
	4	0.002

↓

Distribución de probabilidad binomial para $n = 4$ y $p = 0.2$

x	$P(x)$
0	0.410
1	0.410
2	0.154
3	0.026
4	0.002

EJEMPLO Use la parte de la tabla A-1 (para $n = 4$ y $p = 0.2$), que está al margen, para calcular lo siguiente:

a. La probabilidad de exactamente tres éxitos.

b. La probabilidad de *al menos* tres éxitos.

SOLUCIÓN

a. En la tabla A-1 se observa que cuando $n = 4$ y $p = 0.2$, la probabilidad de $x = 3$ está dada por $P(3) = 0.026$, que es el mismo valor (excepto por el redondeo) que se calcula con la fórmula de probabilidad binomial en el ejemplo anterior.

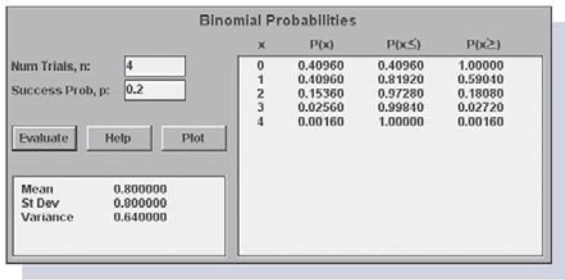
b. "Al menos" tres éxitos, significa que el número de éxitos es 3 o 4.

$$\begin{aligned} P(\text{al menos } 3) &= P(3 \text{ o } 4) \\ &= P(3) + P(4) \\ &= 0.026 + 0.002 \\ &= 0.028 \end{aligned}$$

En el inciso b de la solución anterior, si calculásemos P (al menos 3) por medio de la fórmula de probabilidad binomial, necesitaríamos aplicar la fórmula en dos ocasiones para calcular dos probabilidades diferentes, que después debían sumarse. Al elegir entre la fórmula y la tabla, es más lógico el uso de esta última. Desafortunadamente, la tabla A-1 incluye sólo un número limitado de valores de n y de p, por lo que no siempre sirve; entonces, habrá que calcular las probabilidades con la fórmula de probabilidad binomial, con un programa de cómputo o una calculadora, como en el siguiente método.

Método 3: Uso de las herramientas tecnológicas El STATDISK, Minitab, Excel y la calculadora TI-83 PLus resultan útiles para calcular probabilidades binomiales. Al final de esta sección, presentaremos los detalles para el uso de estas herramientas tecnológicas. Por ahora, observe las representaciones visuales comunes que listan las probabilidades binomiales para $n = 4$ y $p = 0.2$.

STATDISK



Binomial Probabilities

Num Trials, n: 4

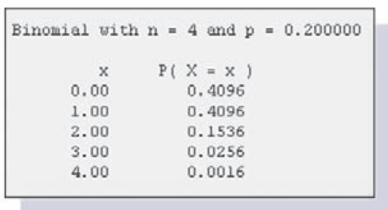
Success Prob, p: 0.2

Evaluate | Help | Plot

x	P(x)	P(x≤)	P(x≥)
0	0.40960	0.40960	1.00000
1	0.40960	0.81920	0.59040
2	0.15360	0.97280	0.18080
3	0.02560	0.99840	0.02720
4	0.00160	1.00000	0.00160

Mean	0.800000
St Dev	0.800000
Variance	0.640000

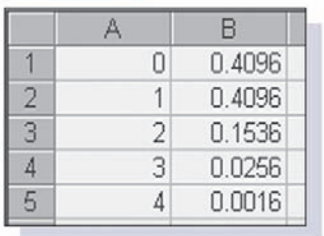
Minitab



Binomial with n = 4 and p = 0.200000

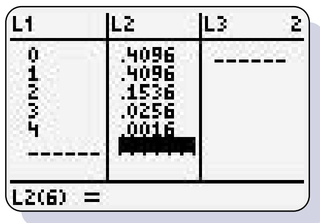
x	P(X = x)
0.00	0.4096
1.00	0.4096
2.00	0.1536
3.00	0.0256
4.00	0.0016

Excel



	A	B
1	0	0.4096
2	1	0.4096
3	2	0.1536
4	3	0.0256
5	4	0.0016

TI-83 Plus



L1	L2	L3 2
0	.4096	------
1	.4096	
2	.1536	
3	.0256	
4	.0016	

L2(6) =

Ya que ahora conocemos tres métodos diferentes para calcular las probabilidades binomiales, he aquí una estrategia efectiva y eficiente:

1. Utilice un programa de cómputo o una calculadora TI-83 Plus, si están disponibles.
2. Si no dispone de un programa de cómputo ni de la calculadora TI-83 Plus, utilice la tabla A-1.
3. Si no dispone de un programa de cómputo ni de una calculadora, y no le es posible hallar 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 que se presentan en esta sección. En lugar de aceptarse y usar la fórmula ciegamente, veamos cómo funciona.

En esta sección ya utilizamos la fórmula de probabilidad binomial para calcular la probabilidad de obtener exactamente tres respuestas correctas, cuando se hacen conjeturas al azar en las preguntas con cuatro opciones. Para cada pregunta

Voltaire gana la lotería

En 1729, el filósofo *Voltaire* se hizo rico al diseñar un esquema para ganar en 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 fuese 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. 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 cubriría todas las combinaciones, pero el estado se rehusó y afirmó que esto cambiaría la naturaleza de la lotería.

Encuestas sensibles

En ocasiones, las personas que contestan las encuestas se rehúsan a responder con honestidad preguntas sobre un tema sensible como el robo al empleador o el sexo. Stanley Warner (York University, Ontario) diseñó un sistema que produce resultados más precisos en casos como éstos. Como ejemplo, pregunte a algunos empleados si cometieron un robo durante el año anterior, luego pídales que lancen una moneda. Instrúyalos para que respondan que no en caso de que no hayan robado y la moneda caiga en cara. Si no es así, deben responder que sí. Los empleados suelen ser más honestos porque el lanzamiento de la moneda los ayuda a proteger su privacidad. Después, es posible utilizar la teoría de la probabilidad para analizar las respuestas, de modo que se obtengan resultados más precisos.

hay cinco respuestas posibles, de modo que la probabilidad de una respuesta correcta es $1/5$ o 0.2. Si utilizamos la regla de la multiplicación, de la sección 3-4, obtenemos el siguiente resultado:

$$\begin{aligned} &P(\text{3 respuestas correctas, seguidas por 1 respuesta incorrecta})\\ &= 0.2 \cdot 0.2 \cdot 0.2 \cdot 0.8\\ &= 0.2^3 \cdot 0.8\\ &= 0.0064 \end{aligned}$$

Este resultado es incorrecto porque supone que las *primeras* respuestas son correctas y que la *última* es incorrecta, pero hay otros acomodos posibles para tres respuestas correctas y una respuesta incorrecta.

En la sección 3-7, vimos que con tres elementos idénticos (como respuestas *correctas*) y otro elemento (como una respuesta *incorrecta*) el número total de acomodos (permutaciones) es $4!/(4-3)!3!$ o 4. Cada uno de estos distintos acomodos tiene una probabilidad de $0.2^3 \cdot 0.8$, de modo que la probabilidad total es la siguiente:

$$P(\text{3 correctas de 4}) = \frac{4!}{(4-3)!3!} \cdot 0.2^3 \cdot 0.8^1$$

Generalice los resultados como sigue: reemplace 4 por n, reemplace x por 3, reemplace 0.2 por p, reemplace 0.8 por q, y exprese el componente de 1 como $4 - 3$, que puede ser reemplazado por $n - x$. El resultado es la fórmula de probabilidad binomial, es decir, la 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. (Véanse los ejercicios 9 y 10).

$$P(x) = \overbrace{\frac{n!}{(n-x)!x!}}^{\text{Número de resultados con exactamente } x \text{ éxitos en } n \text{ ensayos}} \cdot \overbrace{p^x \cdot q^{n-x}}^{\text{Probabilidad de } x \text{ éxitos en } n \text{ ensayos, para cualquier orden}}$$

Utilizando la tecnología

El método 3 de esta sección incluyó el uso del STATDISK, Minitab, Excel o una calculadora TI-83 Plus. Al aplicar los siguientes procedimientos para el cálculo de probabilidades binomiales, aparecen las representaciones visuales del uso del método 3.

STATDISK Seleccione **Analysis** del menú principal; después, seleccione la opción **Binomial Probabilities**. Introduzca los valores requeridos de *n* y *p*, 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 conocer las probabilidades (tales como 0, 1, 2, 3, 4); después, seleccione **Calc** del menú principal y proceda a elegir los siguientes elementos: **Distribuciones de probabilidad** 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, 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, el número de ensayos, la probabilidad y 0 para la distribución binomial (en lugar de 1 para la distribución binomial acumulativa). Debe aparecer un valor en la celda B1. Haga clic y lleve la esquina derecha inferior de la celda B1 hacia abajo de la columna para emparejarla con los datos de la columna A; después, libere el botón del ratón.

TI-83 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*; luego, presione **ENTER;** el resultado será la probabilidad de obtener *x* éxitos en *n* ensayos.

También es posible elegir **binompdf (n, p)** para obtener una lista de *todas* las probabilidades correspondientes a $x = 1, 2, \ldots, n$. 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**).

4-3 Destrezas y conceptos básicos

Identificación de distribuciones normales. *En los ejercicios 1 a 8 determine si el procedimiento dado resulta en una distribución binomial. En aquellos que no sean binomiales, identifique al menos uno de los requisitos que no se satisfacen.*

1. Aplicar una encuesta en la que se le pregunta a las personas lo que piensan del presidente actual.

2. Aplicar una encuesta a 1012 sujetos y registrar si hay una respuesta "no debe" a la siguiente pregunta: "¿La clonación de seres humanos debe o no permitirse?"

3. Tirar un dado balanceado 50 veces.

4. Tirar un dado cargado 50 veces y calcular el número de veces que resulta 5.

5. Registrar el género de 250 bebés recién nacidos.

6. Determinar si cada uno de 3000 marcapasos cardiacos es aceptable o defectuoso.

7. Girar una ruleta 12 veces.

8. Girar una ruleta 12 veces y calcular el número de ocasiones en que el resultado es un número impar.

9. Cálculo de probabilidades con respuestas de adivinación Cada pregunta de opción múltiple tiene cinco posibles respuestas, una de las cuales es la correcta. Suponga que adivina las respuestas de tres de estas preguntas.

continúa

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 adivinaciones?

10. Cálculo de probabilidades con respuestas de adivinación Un examen consta de preguntas de opción múltiple con cuatro respuestas posibles, 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 11 a 16 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 solo ensayo.*

11. $n = 2, x = 0, p = 0.01$

12. $n = 7, x = 2, p = 0.01$

13. $n = 4, x = 3, p = 0.95$

14. $n = 6, x = 5, p = 0.99$

15. $n = 10, x = 4, p = 0.95$

16. $n = 11, x = 7, p = 0.05$

Uso de la fórmula de probabilidad binomial. *En los ejercicios 17 a 20 suponga que un procedimiento produce una distribución binomial, con un ensayo que se repite* 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.*

17. $n = 6, x = 4, p = 0.55$

18. $n = 6, x = 2, p = 0.45$

19. $n = 8, x = 3, p = 1/4$

20. $n = 10, x = 8, p = 1/3$

Uso de resultados de computadora. *En los ejercicios 21 a 14 remítase a la representación visual de Minitab al margen. Las probabilidades se obtuvieron al introducir los valores de n = 6 y p = 0.723. Hay una probabilidad de 0.723 de que un vuelo de American Airlines, que se selecciona aleatoriamente, llegue a tiempo (de acuerdo con datos del Departamento del Transporte de Estados Unidos). En cada caso, suponga que se seleccionan aleatoriamente seis vuelos de American Airlines y calcule la probabilidad indicada.*

Binomial con $n = 6$ y $p = 0.723000$

x	$P(X = x)$
0.00	0.0005
1.00	0.0071
2.00	0.0462
3.00	0.1607
4.00	0.3145
5.00	0.3283
6.00	0.1428

21. Calcule la probabilidad de que al menos cinco vuelos de American Airlines lleguen a tiempo. ¿Es poco común que al menos cinco de seis vuelos de American Airlines lleguen a tiempo?

22. Calcule la probabilidad de que a lo sumo dos vuelos de American Airlines lleguen a tiempo. ¿Es poco común que a lo sumo dos de seis vuelos de American Airlines lleguen a tiempo?

23. Calcule la probabilidad de que más de un vuelo de American Airlines llegue a tiempo. ¿Es poco común que *no* más de uno de seis vuelos de American Airlines llegue a tiempo?

24. Calcule la probabilidad de que al menos un vuelo de American Airlines llegue a tiempo. ¿Es poco común que *no* haya al menos uno de seis vuelos de American Airlines que llegue a tiempo?

25. Ceguera al color El 9% de los hombres y el 0.25% de las mujeres no pueden distinguir entre los colores rojo y verde. Este tipo de problema visual causa dificultades con las señales de tránsito. Si se seleccionan seis hombres aleatoriamente para un estudio de la percepción de las señales de tránsito, calcule la probabilidad de que exactamente dos de ellos no distingan entre el rojo y el verde.

26. Muestreo de aceptación La compañía Telektronic compra grandes embarques de focos fluorescentes y usa el siguiente plan de muestreo de aceptación: seleccionar aleatoriamente y probar 24 focos; después, aceptar el grupo completo sólo si hay uno o ninguno que no funcione. Si un embarque particular de miles de focos tiene en realidad una tasa de defectos del 4%, ¿cuál es la probabilidad de que el embarque completo se acepte?

27. Auditorías de la IRS La Hemingway Financial Company prepara devoluciones de impuestos para individuos. (Su lema: "También escribimos grandiosas novelas de ficción"). Según el Internal Revenue Service, los individuos que ganan entre 425,000 y 50,000 dólares se auditan en una proporción del 1%. La Hemingway Company prepara cinco devoluciones de impuestos para individuos que están en esa categoría de impuestos, en tanto se audita a tres de ellos.

a. Calcule la probabilidad de que, cuando se seleccione aleatoriamente a cinco personas que ganan entre $25,000 y $50,000, se audite exactamente a tres de ellos.

b. Calcule la probabilidad de que se audite al menos a tres.

c. Con base en los resultados anteriores, ¿qué se concluye acerca de los clientes de Hemingway? ¿Sólo son desafortunados o están siendo blanco de las auditorías?

28. Asistencia del directorio telefónico Un artículo de *USA Today* afirma que "encuestas internas que son pagadas por proveedores de asistencia del directorio telefónico, muestran que incluso las compañías más precisas usan los números incorrectos el 15% de las veces". Suponga que prueba a un proveedor de éstos haciendo 10 solicitudes y también que el proveedor le da números telefónicos incorrectos el 15% de las veces.

a. Calcule la probabilidad de obtener un número incorrecto.

b. Calcule la probabilidad de obtener a lo sumo un número incorrecto.

c. Si usted obtiene a lo sumo un número incorrecto, ¿parecería que la tasa de números incorrectos no es del 15%, como se afirma?

29. Vuelos sobresaturados Air America tiene la política de registrar a 15 personas en un avión donde sólo caben 14. (Estudios anteriores revelaron que sólo el 85% de los pasajeros que se registran usan el vuelo). Calcule la probabilidad de que, si Air America registra a 15 personas, no haya suficientes asientos disponibles. ¿Será la probabilidad suficientemente baja, de modo que la sobreventa no sea un problema real para los pasajeros?

30. Reacción al fármaco En una prueba clínica del fármaco Viagra, se encontró que el 4% de los individuos en el grupo placebo sufrieron dolores de cabeza.

a. Suponiendo que la misma tasa del 4% se aplica a quienes toman Viagra, calcule la probabilidad de que, entre ocho usuarios del Viagra, tres experimenten dolores de cabeza.

b. Suponiendo que la misma tasa del 4% se aplica a quienes toman Viagra, calcule la probabilidad de que, entre ocho usuarios de Viagra que se seleccionan aleatoriamente, todos ellos experimenten dolores de cabeza.

c. Si los ocho usuarios de Viagra experimentaran dolores de cabeza, ¿parecería que la tasa de dolores de cabeza de los usuarios de Viagra es diferente de la tasa del 4% de los sujetos del grupo de placebo? Explique.

31. Encuestas a televidentes El programa de televisión *60 minutos*, de la CBS, ha logrado éxito por muchos años. Recientemente registró una audiencia de 20, lo que significa que de todos los televisores en uso el 20% se sintonizan en *60 minutos* (según datos de Nielsen Media Research). Suponga que un anunciante desea verificar dicho valor del 20%, realizando su propia encuesta, y que inicia una encuesta piloto con 10 hogares que tienen la televisión encendida 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 una audiencia del 20%? ¿Por qué?

32. Programas de acción afirmativa Se realizó un estudio para determinar si había diferencias significativas entre estudiantes de medicina que se aceptaron por medio de programas especiales (como el de acción afirmativa) y estudiantes de medicina que se aceptaron a través de los criterios regulares de admisión. Se encontró que el 94% de los estudiantes de medicina que se aceptaron a través de programas especiales se graduaron (según datos del *Journal of the American Medical Association).*

a. Si se seleccionan aleatoriamente 10 de los estudiantes de los programas especiales, calcule la probabilidad de que al menos nueve se gradúen.
b. ¿Sería poco común que de 10 estudiantes de los programas especiales, que se seleccionaron aleatoriamente, sólo se graduaran siete? ¿Por qué?

33. Identificación de la discriminación por género Después de que la rechazaran para un empleo, Kim Kelly se entera de que la Bellevue Advertising Company sólo contrató a dos mujeres entre los últimos 20 empleados nuevos. También, de que el grupo de solicitantes es muy grande, y que incluye un número aproximadamente igual de hombres y mujeres calificados. Ayúdele a presentar cargos por discriminación por género, calculando la probabilidad de que dos o menos mujeres se incluyan en una contratación de 20 personas, suponiendo que no hay discriminación que se basa en el género. ¿Apoya la probabilidad resultante esos cargos?

34. Máquina tragamonedas del autor El autor compró una máquina tragamonedas que se configuró de tal forma que hay 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 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 válida la afirmación del invitado de dos triunfos en cinco juegos? Explique.

35. Prueba de la eficacia de la técnica de selección del género El problema del capítulo describe la distribución de probabilidad del número de niñas x resultantes cuando se seleccionan aleatoriamente 14 bebés recién nacidos. Suponga que otro experimento clínico incluye 12 bebés recién nacidos. Utilice el mismo formato de la tabla 4-1 y construya una tabla para la distribución de probabilidad que resulta de los 12 nacimientos; después, determine si una técnica de selección del género sería efectiva si nacen nueve niñas y tres niños.

36. Cursos de posgrado El Market Research Institute encontró que de los individuos que se graduaron de la universidad desde hace al menos 10 años, con empleo y edades entre 30 y 55 años, el 57% tomaron cursos universitarios tras haberse graduado (según el *USA Today*). Si usted selecciona aleatoriamente a cinco individuos que se graduaron de la universidad desde hace al menos 10 años, con edades entre 30 y 55 años, y descubre que sólo uno de ellos tomó cursos, ¿debe pensar que la tasa del 57% es incorrecta? Explique.

4-3 Más allá de lo básico

37. 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 que se probó.

38. Si realizamos un muestreo sin reemplazo de una población finita pequeña, no debe usarse la distribución binomial porque los sucesos 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 aleatoriamente una combinación de seis números ganadores. Calcule la probabilidad de obtener:

a. Los seis números ganadores.
b. Exactamente cinco de los números ganadores.
c. Exactamente tres de los números ganadores.
d. Ningún número ganador.

39. La distribución binomial se aplica sólo a casos que impliquen dos tipos de resultados, mientras que la **distribución multinomial** supone más de dos categorías. Suponga que tenemos tres tipos de resultados mutuamente excluyentes, que se denotan 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 igual de probables. Si se prueba a 20 descendientes, calcule la probabilidad de obtener con exactitud cinco A, cuatro B, tres C, dos D, tres E y tres F, al expandir la expresión anterior, de modo que se aplique a seis tipos de resultados y no sólo a tres.

4-4 Media, varianza y desviación estándar para la distribución binomial

En el capítulo 2, vimos que cuando se investigan conjuntos de datos reales, las siguientes características suelen ser muy importantes: **1.** las medidas de tendencia central, **2.** las medidas de variación, **3.** la naturaleza de la distribución, **4.** la presencia de datos distantes y **5.** un patrón a lo largo del tiempo. (Recuerde que utilizamos “CVDDT” como herramienta para evocar dichas características). Un aspecto central de este capítulo es que las distribuciones de probabilidad describen lo que *probablemente* sucederá, y no lo que en realidad sucedió. En la sección 4-2 estudiamos métodos para analizar las distribuciones de probabilidad a través del cálculo de la media, la desviación estándar y el histograma de probabilidad. Como una distribución binomial es un tipo especial de distribución de probabilidad, utilizamos las fórmulas 4-1, 4-3 y 4-4 (de la sección 4-2) para calcular la media, la varianza y la desviación estándar. Por fortuna, es posible simplificar mucho tales fórmulas para las distribuciones binomiales, como se muestra a continuación.

Para cualquier distribución de probabilidad discreta		**Para distribuciones binomiales**	
Fórmula 4-1	$\mu = \Sigma[x \cdot P(x)]$	**Fórmula 4-6**	$\mu = np$
Fórmula 4-3	$\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$	**Fórmula 4-7**	$\sigma^2 = npq$
Fórmula 4-4	$\sigma = \sqrt{\Sigma[x^2 \cdot P(x)] - \mu^2}$	**Fórmula 4-8**	$\sigma = \sqrt{npq}$

EJEMPLO Género de los hijos En la sección 4-2 incluimos un ejemplo para ilustrar los cálculos de μ y σ. Utilizamos el ejemplo de la variable aleatoria x que representa el número de niñas entre 14 nacimientos. (Véase la tabla 4-2 en la página 185, con los cálculos que ejemplifican las fórmulas 4-1 y 4-4.) Aplique las fórmulas 4-6 y 4-8 para calcular la media y la desviación estándar del número de niñas en grupos de 14 nacimientos.

SOLUCIÓN Con los valores $n = 14$, $p = 0.5$ y $q = 0.5$, las fórmulas 4-6 y 4-8 se aplican como sigue:

$$\mu = np = (14)(0.5) = 7.0$$
$$\sigma = \sqrt{npq}$$
$$= \sqrt{(14)(0.5)(0.5)} = 1.9 \qquad \text{(redondeado)}$$

Si compara estos cálculos con los que se requieren en la tabla 4-3, resultará evidente que las fórmulas 4-6 y 4-8 son mucho más fáciles de usar.

La fórmula 4-6 para la media es intuitivamente lógica. Si le preguntáramos a cualquier estudiante de estadística cuántas niñas se esperaría en 100 nacimientos, la respuesta común sería 50, que puede generalizarse fácilmente como $\mu = np$. La varianza y la desviación estándar no se justifican tan fácilmente, por lo que omitiremos los complicados manejos algebraicos que conducen a las fórmulas 4-7 y 4-8. En su lugar, remítase de nuevo al ejemplo anterior y a la tabla 4-3 para verificar que, en una distribución binomial, las fórmulas 4-6, 4-7 y 4-8 producirán los mismos resultados que las fórmulas 4-1, 4-3 y 4-4.

EJEMPLO Selección del género El problema del capítulo incluyó un ensayo preliminar con 14 parejas que deseaban tener niñas. Aun cuando el resultado de 13 niñas en 14 nacimientos parece indicar que el método MicroSort es eficaz para la selección del género, tendríamos mucho mayor confianza en dicha conclusión si el tamaño de la muestra fuese considerablemente mayor que 14. Suponga que el método MicroSort se utiliza con 100 parejas que tendrán un bebé. También, que el resultado es de 68 niñas entre 100 bebés.

¿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? El tema se estudió 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 que se estudia en esta sección. Ellos concluyeron que "no encontramos evidencias contundentes de que un sexo prevalezca más en la familia".

a. Suponiendo que el método de selección del género MicroSort no produce efecto alguno, calcule la media y la desviación estándar del número de niñas en grupos de 100 bebés que se seleccionaron aleatoriamente.

b. *Interprete* los valores del inciso *a* para determinar si estos resultados (68 niñas entre 100 bebés) apoyan la afirmación de que el método MicroSort de selección del género es efectivo.

SOLUCIÓN

a. Suponiendo que el método MicroSort no produce efecto alguno, así como que las niñas y los niños son igualmente probables, tenemos $n = 100$, $p = 0.5$ y $q = 0.5$. Podemos calcular la media y la desviación estándar con el uso de las fórmulas 4-6 y 4-8 de la siguiente manera:

$$\mu = np = (100)(0.5) = 50$$
$$\sigma = \sqrt{npq} = \sqrt{(100)(0.5)(0.5)} = 5$$

Para grupos de 100 parejas con un bebé, el número medio de niñas es de 50 y la desviación estándar es de 5.

b. Ahora debemos interpretar los resultados para determinar si 68 niñas, entre 100 bebés, implica algo que podría ocurrir fácilmente por azar, o si es tan improbable que el método MicroSort de selección del género parece ser efectivo. Utilizaremos la regla práctica del intervalo de la siguiente manera:

$$\text{valor máximo común: } \mu + 2\sigma = 50 + 2(5) = 60$$
$$\text{valor mínimo común: } \mu - 2\sigma = 50 - 2(5) = 40$$

INTERPRETACIÓN De acuerdo con la regla práctica del intervalo, los valores se consideran comunes si están entre 40 y 60, de modo que 68 niñas es un resultado poco común, ya que no se encuentra entre 40 y 60. Es muy poco probable que resulten 68 niñas en 100 nacimientos sólo por el azar. Si en realidad obtuviésemos 68 niñas entre 100 nacimientos, deberíamos buscar una explicación alternativa a la del azar. Si las 100 parejas utilizaron el método MicroSort de selección del género, parecería que es efectivo para incrementar la posibilidad de que un hijo sea niña.

Usted debe desarrollar la habilidad para calcular medias y desviaciones estándar con el uso de las fórmulas 4-6 y 4-8, pero es especialmente importante aprender a *interpretar* los resultados con el empleo de esos valores. La regla práctica del intervalo, como se ilustra en el inciso *b* del ejemplo anterior, sugiere que los valores son poco comunes si caen fuera de los siguientes límites:

$$\textbf{valor máximo común} = \boldsymbol{\mu + 2\sigma}$$
$$\textbf{valor mínimo común} = \boldsymbol{\mu - 2\sigma}$$

4-4 Destrezas y conceptos básicos

Cálculo de μ, σ y valores poco comunes. *En los ejercicios 1 a 4 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 des-*

Los estados controlan las selecciones de lotería

Muchos estados permiten una lotería en la que los jugadores seleccionan cuatro dígitos, como 1127 (el cumpleaños del autor). Si un jugador paga un dólar 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 jugadores hacer más apuestas con tal secuencia. Las máquinas de lotería se controlan de tal manera que, una vez que una secuencia popular alcanza cierto nivel de ventas, no aceptarán más dicha secuencia. Lo anterior evita que los estados paguen más de lo que reciben. Los críticos afirman que la práctica es injusta. Según William Thompson, un experto en apuestas de la Universidad de Nevada en Las Vegas, “ellos (los estados) afirman que quieren estar en el negocio del juego, pero no desean ser jugadores. Esto hace una farsa del juego de números”.

viación estándar σ. Además, use la regla práctica del intervalo para calcular el valor mínimo común $\mu - 2\sigma$ y el valor máximo común $\mu + 2\sigma$.

1. $n = 400$, $p = 0.2$

2. $n = 250$, $p = 0.45$

3. $n = 1984$, $p = 3/4$

4. $n = 767$, $p = 1/6$

5. Respuestas de adivinación Varios estudiantes no se prepararon para enfrentar un examen sorpresa de verdadero/falso de 10 preguntas, por lo cual todas sus respuestas son conjeturas.

a. Calcule la media y la desviación estándar del número de respuestas correctas de esos estudiantes.

b. ¿Sería poco común que un estudiante pasara el examen adivinando y que obtuviera al menos siete respuestas correctas? ¿Por qué?

6. Respuestas de adivinación Varios estudiantes no se prepararon para presentar un examen de opción múltiple de 10 preguntas, por lo cual todas sus respuestas son conjeturas. Cada pregunta tiene cinco respuestas posibles, pero sólo una de ellas es correcta.

a. Calcule la media y la desviación estándar del número de respuestas correctas de esos estudiantes.

b. ¿Sería poco común que un estudiante pasara el examen adivinando y que obtuviera al menos siete respuestas correctas? ¿Por qué?

7. Juego de ruleta Si apuesta en cualquier número de la ruleta, su probabilidad de ganar es 1/38. Suponga que apuesta a un solo número en cada uno de 100 giros consecutivos.

a. Calcule la media y la desviación estándar del número de triunfos.

b. ¿Sería poco común no ganar ni una ocasión en los 100 ensayos? ¿Por qué?

8. Personas zurdas El 10% de los adultos estadounidenses son zurdos. Una clase de estadística tiene 25 estudiantes.

a. Calcule la media y la desviación estándar del número de estudiantes zurdos en la clase con 25 estudiantes.

b. ¿Sería poco común hacer una encuesta a una clase con 25 estudiantes y encontrar que 5 de ellos son zurdos? ¿Por qué?

9. Análisis de los resultados de un experimento de selección del género Un experimento, en relación con un método de selección del género, incluye un grupo control de 15 parejas que no reciben ningún tratamiento creado para influir en el género de sus hijos. Cada una de las 15 parejas tiene un hijo.

a. Elabore una tabla que liste los posibles valores de la variable aleatoria x (que representa el número de niñas entre 15 nacimientos) y las probabilidades correspondientes.

b. Calcule la media y la desviación estándar del número de niñas en grupos de 15 nacimientos como éstos.

c. Si las parejas tienen 10 niñas y cinco niños, ¿será poco común este resultado? ¿Por qué?

10. Mensajes descifrados La Central Intelligence Agency tiene especialistas que analizan la secuencia de letras del alfabeto, en un intento por descifrar mensajes que se interceptan. En un texto estándar en inglés, la letra r se utiliza en una proporción del 7.7%.

a. Calcule la media y la desviación estándar del número de veces que la letra r aparecerá en una página común de 2,600 caracteres.

b. En un mensaje que se interceptó que iba hacia Irak, se encontró que en una página con 2,600 caracteres la letra r aparecía 175 veces. ¿Es esto poco común?

11. Determinar si disminuyen las quejas después de un programa de entrenamiento La Newtower Department Store recibió una tasa de quejas de los clientes del 3.2%, e intenta disminuir esta tasa con un programa de entrenamiento para los empleados. Una vez que se completó el programa, se localizó a 850 clientes y se descubrió que sólo siete de ellos se quejaron.

a. Suponiendo que el programa de entrenamiento no tenga efecto alguno, calcule la media y la desviación estándar del número de quejas en grupos de 850 clientes como éste.

b. Con base en los resultados del inciso *a*), ¿es poco común el resultado de siete quejas? ¿Fue efectivo el programa de entrenamiento para disminuir la tasa de quejas?

12. ¿Es azul el 10% de los dulces M&M? Mars, Inc., afirma que el 10% de sus dulces M&M son azules, por lo cual se selecciona aleatoriamente una muestra con 100 de estos dulces.

a. Calcule la media y la desviación estándar del número de dulces azules en grupos de 100 como éste.

b. El conjunto de datos 19, del Apéndice B, consiste en una muestra aleatoria de 100 M&M, de los cuales sólo cinco son azules. ¿Es este resultado poco común? ¿Será incorrecta la tasa del 10%?

13. Teléfonos celulares y cáncer cerebral En un estudio que incluyó 420,000 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 afecta por los teléfonos celulares, la probabilidad de que una persona adquiera tal enfermedad es de 0.000340.

a. Suponiendo que los teléfonos celulares no se relacionan con el cáncer, calcule la media y la desviación estándar del número de personas, en grupos de 420,000, que pueden tener cáncer cerebral o del sistema nervioso.

b. Con base en los resultados del inciso *a*, ¿será poco común que, entre 420,000 personas, se presenten 135 casos de cáncer cerebral o del sistema nervioso? ¿Por qué?

c. ¿Qué sugieren tales 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?

14. Fármaco que reduce el colesterol En un ensayo clínico del Lipitor, un fármaco común que se utiliza para disminuir el colesterol, 863 pacientes recibieron un tratamiento de 10 miligramos de tabletas de Atorvastatin. Este grupo incluyó a 19 pacientes que experimentan síntomas de influenza (según datos de Pfizer, Inc.). La probabilidad de que una persona que no recibe tratamiento alguno presente síntomas de influenza es de 0.019.

a. Suponiendo que el Lipitor no produce efectos sobre los síntomas de la influenza, calcule la media y la desviación estándar del número de personas en grupos de 863 individuos que se esperaría presentaran dichos síntomas.

b. Con base en los resultados del inciso *a*, ¿será poco común encontrar que, de 863 personas, 19 experimentan síntomas de influenza? ¿Por qué?

c. Con base en los resultados anteriores, ¿los síntomas de la influenza parecen ser una reacción adversa que debe preocupar a los usuarios de Lipitor?

15. Opiniones sobre la clonación En una encuesta reciente de Gallup se preguntó a 1012 adultos que se seleccionaron aleatoriamente, si "la clonación humana debe o no permitirse". Los resultados mostraron que el 89% de los encuestados opinaron que la clonación no tiene que permitirse.

a. De los 1012 adultos que se encuestaron, ¿cuántos opinaron que no debe permitirse la clonación.

b. Si suponemos que las personas se muestran indiferentes, de manera que el 50% considera que la clonación humana no tiene que permitirse, calcule la media y la desviación estándar del número de personas en grupos de 1012 que se esperaría que opinaron que la clonación no debe permitirse.

c. Con base en los resultados anteriores, ¿parece inusualmente más alto el resultado del 89% de la encuesta de Gallup que la supuesta tasa del 50%? ¿Parece que una inmensa mayoría de adultos opina que la clonación humana no debe permitirse?

16. Choques de automóviles Los conductores de edades entre 20 y 24 años presentan una proporción del 34% de accidentes automovilísticos durante un año (según datos del National Safety Council). Un investigador de una aseguradora descubre que en un grupo de 500 conductores, en el rango de 20 a 24 años, que viven en la ciudad de Nueva York, que se seleccionaron aleatoriamente, el 42% tuvo accidentes el año anterior.

a. ¿Cuántos conductores, en el grupo de 500 individuos de la ciudad de Nueva York, tuvo accidentes el año anterior?

b. Suponiendo que la misma proporción del 34% se aplica a la ciudad de Nueva York, calcule la media y la desviación estándar del número de personas, en grupos de 500 individuos, que se esperaría tuvieran accidentes.

c. Con base en resultados anteriores, ¿parece inusualmente más alto el resultado del 42% de los conductores de la ciudad de Nueva York, comparado con la proporción del 34% de la población general? ¿Se justifican las tasas de seguro más altas para los conductores de la ciudad de Nueva York?

4-4 Más allá de lo básico

17. Uso de la regla empírica y el teorema de Chebyshev Se diseña un experimento para probar la efectividad del método MicroSort de selección del género y se aplica a 100 parejas que desean tener niñas. En un ejemplo que se incluye en esta sección, se utilizó la regla práctica del intervalo para concluir que, de 100 nacimientos, el número de niñas por lo general debe ubicarse entre 40 y 60.

a. La regla empírica (véase sección 2-5) se aplica a distribuciones normales. ¿Es (aproximadamente) normal la distribución de probabilidad binomial para este experimento? ¿Cómo lo sabe?

b. Suponiendo que la distribución sea normal, ¿qué tan probable es que el número de niñas esté entre 40 y 60 (según la regla empírica)?

c. Suponiendo que la distribución es normal, ¿qué tan probable es que el número de niñas esté entre 35 y 65 (según la regla empírica)?

d. De acuerdo con el teorema de Chebyshev, ¿qué concluimos acerca de la probabilidad de que el número de niñas esté entre 40 y 60?

18. Productos aceptables/defectuosos Mario's Pizza Parlor acaba de inaugurarse. Por 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?

4-5 La distribución de Poisson

Las dos secciones anteriores se ocuparon de la distribución binomial, que es una de las muchas distribuciones de probabilidad que pueden utilizarse en distintas situaciones. Esta sección introduce la *distribución de Poisson*. Dicha distribución es particularmente importante, ya que con frecuencia se utiliza como modelo matemático para describir comportamientos como la disminución radioactiva, la llegada de pasajeros en una línea y la de aviones a un aeropuerto, los automóviles que van a una gasolinera, los comensales que visitan un restaurante, los estudiantes que asisten a una librería y los usuarios de Internet que se conectan a un sitio. Por ejemplo, suponga que su profesor tiene una hora de asesoría cada lunes a las 11:00, y descubre que durante esa hora el número medio de estudiantes que recibe es de 2.3. Calculamos la probabilidad de que exactamente cuatro estudiantes lleguen en una hora de asesoría, que se selecciona aleatoriamente, un lunes a las 11:00. Utilizamos la distribución de Poisson, que se define de la siguiente manera.

Definición

Distribución de Poisson: 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 ocurrencias de 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 4-9.

Fórmula 4-9 $$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} \qquad \text{donde } e \approx 2.71828$$

La distribución de Poisson tiene los siguientes requisitos:

- La variable aleatoria x es el número de ocurrencias de un suceso *durante un intervalo*.
- Las ocurrencias deben ser *aleatorias*.
- Las ocurrencias tienen que ser *independientes* entre sí.
- Las ocurrencias deben estar *uniformemente distribuidas* dentro del intervalo que se emplea.

La distribución de Poisson tiene los siguientes parámetros:

- 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 se afecta por el tamaño de la muestra n y la probabilidad p, mientras que la distribución de Poisson sólo se afecta por la media μ.
2. En una distribución binomial, los valores posibles de la variable aleatoria x son 0, 1, . . . , n, pero los valores posibles x de una distribución de Poisson son 0, 1, 2, . . . , sin límite superior.

EJEMPLO Bombas de la Segunda Guerra Mundial Al analizar los impactos de las bombas V1 en la Segunda Guerra Mundial, el sur de Londres se subdividió en 576 regiones, cada una con un área de 0.25 km^2. En total, 535 bombas estallaron en el área combinada de 576 regiones.

a. Si se selecciona una región aleatoriamente, calcule la probabilidad de que fuese blanco de impactos exactamente en dos ocasiones.

b. Con base en la probabilidad que se calculó en el inciso *a*, ¿cuántas de las 576 regiones se espera que reciban impactos 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 un área de 0.25 km^2). El número medio de impactos por región es

$$\mu = \frac{\text{número de impactos de bomba}}{\text{número de regiones}} = \frac{535}{576} = 0.929$$

continúa

Probabilidad de un suceso que nunca ha ocurrido

Algunos sucesos son posibles, pero tan poco probables que nunca ocurren. He aquí un problema de este tipo, de gran interés para los científicos políticos: estime la probabilidad de que su voto determine al 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 presidentes en prospecto 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.

Ya que buscamos la probabilidad de exactamente dos impactos en una región, sea $x = 2$. Utilizaremos la fórmula 4-9 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 blanco de impactos exactamente dos veces es de $P(2) = 0.170$.

b. Debido a que hay una probabilidad de 0.170 de que una región reciba impactos de bombas exactamente dos veces, esperamos que entre las 576 regiones el número de las que reciben impactos exactamente dos veces sea de $576 \cdot 0.170 = 97.9$.

En el ejemplo anterior también podemos calcular las probabilidades y los valores que se esperan para 0, 1, 2, 3, 4 y 5 impactos. (Nos detenemos en $x = 5$, porque ninguna región recibió impactos de bombas más de cinco ocasiones, y las probabilidades de $x > 5$ son 0.000, cuando se redondea a tres decimales). Esas probabilidades y tales valores que se esperan se muestran en la tabla 4-5. La cuarta columna de la tabla 4-5 describe los resultados reales de la Segunda Guerra Mundial. Hubo 229 regiones sin impactos, 211 que recibieron impactos una vez, y así sucesivamente. Ahora comparemos las frecuencias *predichas* por medio de la distribución de Poisson (tercera columna) con las frecuencias *reales* (cuarta columna), para concluir que hay un acuerdo muy alto. En este caso, la distribución de Poisson hace un buen trabajo al predecir los resultados que ocurrieron en realidad. (La sección 10-2 describe un procedimiento estadístico para determinar si tales frecuencias esperadas constituyen un buen "ajuste" de las frecuencias reales. Ese procedimiento sugiere que, en tal caso, hay un buen ajuste).

Tabla 4-5 Impactos de bombas V1 en 576 regiones del sur de Londres.

Número de impactos de bombas	Probabilidad	Número que se espera 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

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 es utilizar una aproximación como éstas cuando se satisfacen las siguientes dos condiciones:

1. $n \geq 100$
2. $np \leq 10$

Si se cumplen dichas condiciones y deseamos utilizar la distribución de Poisson, como aproximación de la distribución binomial, necesitamos un valor de μ; ese valor se calcula utilizando la fórmula 4-6 (que se presenta originalmente en la sección 4-4):

Fórmula 4-6 $$\mu = np$$

EJEMPLO **Juego Pick 3 de Illinois** En el juego Pick 3 de Illinois, usted paga 50 centavos para seleccionar una secuencia de tres dígitos, como 911. 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 hay un conjunto ganador de números entre los 1000 posibles (del 000 al 999), $p = 1/1000$. Se satisfacen las condiciones $n \geq 100$ y $n\,p \leq 10$, de manera que utilizaríamos la distribución de Poisson como aproximación de la distribución binomial. Primero, necesitamos el valor de μ, que se calcula de la siguiente manera:

$$\mu = np = 365 \cdot \frac{1}{1000} = 0.365$$

Al calcular el valor de μ, ahora hacemos lo mismo con $P(1)$:

$$P(1) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{0.365^1 \cdot 2.71828^{-0.365}}{1!} = 0.253$$

Si aplicamos la distribución de Poisson como aproximación de la distribución binomial, encontraremos que hay una probabilidad de 0.253 de ganar exactamente una vez en 365 días. Si utilizamos la distribución binomial, obtendremos una probabilidad más precisa de 0.254, de modo que observamos que la aproximación de Poisson es bastante buena aquí.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal; después, **Distribuciones de probabilidad** y luego **Poisson**. Haga clic en el botón OPTIONS y proceda a introducir el valor de la media μ. Use el mouse para desplazar a la derecha o a la izquierda los distintos valores de x.

Minitab Primero introduzca el valor deseado de x en la columna C1. Ahora seleccione **Calc** de la barra del menú principal, luego **Probability Distribution** y finalmente **Poisson**. Introduzca el valor de la media μ y C1 en la columna de entrada.

Excel Haga clic en f_x en la barra del menú principal, después seleccione la categoría **Statistical**, luego **POISSON** y haga clic en **OK**. En el cuadro del diálogo introduzca los valores de x y la media, luego 0 para "Cumulative". (Introducir 1 en "Cumulative" da como resultado la probabilidad de los valores hasta el valor que se introdujo de x, inclusive.)

TI-83 Plus Presione **2nd VARS** (para obtener **DISTR**), después seleccione la opción B: **poissonpdf(**. Ahora presione **ENTER** y después proceda a introducir μ, x (incluyendo la coma). Para μ, introduzca el valor de la media; para x introduzca el número deseado de ocurrencias.

4-5 Destrezas y conceptos básicos

Uso de una distribución de Poisson para calcular la probabilidad. *En los ejercicios 1 a 4 suponga que la distribución de Poisson se aplica; después, proceda a emplear la media dada para calcular la probabilidad indicada.*

1. Si $\mu = 2$, calcule $P(3)$.

2. Si $\mu = 0.5$, calcule $P(2)$.

3. Si $\mu = 100$, calcule $P(99)$.

4. Si $\mu = 500$, calcule $P(512)$.

En los ejercicios 5 a 12 utilice la distribución de Poisson para calcular las probabilidades que se indican.

5. Disminución radioactiva Los átomos radioactivos son inestables porque tienen demasiada energía. Cuando liberan su energía sobrante, se dice que disminuyen. Al estudiar el cesio 137, se descubrió que durante el curso de la disminución, durante 365 días, 1,000,000 de átomos radioactivos se reducen a 997,287 átomos radioactivos.
a. Calcule el número medio de átomos radioactivos que se perdieron durante la disminución de un día.
b. Calcule la probabilidad de que en un día dado disminuyan 50 átomos radioactivos.

6. Nacimientos Actualmente nacen 11 bebés cada año en la villa de Westport (con una población de 760) (según datos del National Center for Health Statistics de Estados Unidos).
a. Calcule el número medio de nacimientos por día.
b. Calcule la probabilidad de que en un día dado no haya nacimientos.
c. Calcule la probabilidad de que en un día dado haya al menos un nacimiento.
d. Con base en los resultados anteriores, ¿debe estar en guardia permanente el personal médico o hay que llamarlos cuando sea necesario? ¿Significa esto que las madres de Westport podrían no recibir la atención médica inmediata que probablemente sí recibirían en una área más poblada?

7. Muertes por coces de caballos Un ejemplo clásico de la distribución de Poisson implica el número de muertes de hombres del ejército prusiano causadas por coces de caballo, entre 1875 y 1894. Se combinaron datos de 14 cadáveres durante un periodo de 20 años; los 280 años-cadáveres incluyeron un total de 196 muertes. Después de calcular el número medio de muertes por año-cadáver, determine la probabilidad de que un año-cadáver, que se selecciona aleatoriamente, 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 que se esperan de las probabilidades de Poisson. ¿Sirve la distribución de Poisson como una buena herramienta para predecir los resultados reales?

8. Muertes por homicidio En un año hubo 116 muertes por homicidio en Richmond, Virginia (de acuerdo con “A Classroom Note on the Poisson Distribution: A Model for Homicidal deaths in Richmond, VA for 1991” en *Mathematics and Computer Education*, de Wiston A. Richards). Para un día que se seleccionó aleatoriamente, 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.

9. Ruleta Scott apuesta el número 7 para cada uno de 200 giros de una ruleta. Como $P(7) = 1/38$, él espera ganar aproximadamente cinco veces.

a. Calcule la probabilidad de ningún triunfo en los 200 giros.

b. Calcule la probabilidad de al menos un triunfo en los 200 giros.

c. Scott perderá dinero si el número de triunfos es 0, 1, 2, 3, 4 o 5. Calcule la probabilidad de que Scott pierda dinero después de 200 giros.

d. ¿Cuál es la probabilidad de que Scott obtenga alguna ganancia después de 200 giros?

10. Terremotos Durante un periodo que comprende los últimos 100 años hubo 93 terremotos importantes en el mundo (al menos 6.0 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, la probabilidad de que el número de terremotos en un año que se selecciona 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 (0 terremotos importantes); 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 que se calcularon con los resultados reales, ¿es la distribución de Poisson un buen modelo?

4-5 Más allá de lo básico

11. Aproximación de Poisson a una binomial La distribución de Poisson puede 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, aunque *podemos* calcular la probabilidad de $x = 101$ con la aproximación de Poisson. Calcule dicho valor. ¿Qué tanto coincide el resultado con la imposibilidad de que $x = 101$ en una distribución binomial?

12. Aproximación de Poisson a una 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?

Resumen

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 nociones de probabilidad continua. Se estudiaron los siguientes conceptos básicos:

- Una *variable aleatoria* posee valores que se determinaron al 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:

$$\Sigma P(x) = 1 \text{ y, para cada valor de } x,\ 0 \leq P(x) \leq 1.$$

- Se pueden explorar características importantes de una *distribución de probabilidad* construyendo un histograma de probabilidad, así como calculando su media y desviación estándar por medio de las siguientes fórmulas:

$$\mu = \Sigma[x \cdot P(x)]$$

$$\sigma = \sqrt{\Sigma[x^2 \cdot P(x)] - \mu^2}$$

- En una *distribución binomial*, hay dos categorías de resultados y un número fijo de ensayos independientes con una probabilidad constante. La probabilidad de x éxitos en n ensayos se calcula empleando la fórmula de probabilidad binomial, la tabla A-1, un programa de cómputo (como STATDISK, Minitab o Excel) o una calculadora TI-83 Plus.
- En una distribución binomial, la media y la desviación estándar se obtienen 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 calculan con la fórmula 4-9.
- *Resultados poco comunes:* Este capítulo puso énfasis en la importancia de interpretar los resultados a través de la distinción entre los que son comunes y los que son poco comunes. Utilizamos dos criterios diferentes. Con la regla práctica del intervalo tenemos

$$\textbf{valor máximo común} = \boldsymbol{\mu + 2\sigma}$$

$$\textbf{valor mínimo común} = \boldsymbol{\mu - 2\sigma}$$

También podemos determinar si los resultados son poco comunes por medio de los valores de probabilidad.

Inusualmente alto: x éxitos en n ensayos es un número *inusualmente alto* de éxitos si P (x o más) es muy pequeño (tanto como 0.05 o menos).

Inusualmente bajo: x éxitos en n ensayos es un número *inusualmente bajo* de éxitos si P (x o menos) es muy pequeño (tanto como 0.05 o menos).

Ejercicios de repaso

x	$P(x)$
0	0.08
1	0.05
2	0.10
3	0.13
4	0.15
5	0.21
6	0.09
7	0.19

1. a. ¿Qué es una variable aleatoria?

b. ¿Qué es una distribución de probabilidad?

c. Una gráfica del *USA Today* listó los porcentajes del número de días en una semana que los adultos estadounidenses cocinan en su casa, durante una semana promedio. Por ejemplo, el 13% de los adultos estadounidenses cocinan en su casa tres días, en una semana promedio. La tabla al margen se basa en la gráfica. ¿Describe dicha tabla una distribución de probabilidad? ¿Por qué?

d. Suponiendo que la tabla describe una distribución de probabilidad, calcule su media.

e. Suponiendo que la tabla describe una distribución de probabilidad, calcule su desviación estándar.

f. ¿Será poco común seleccionar aleatoriamente a un adulto estadounidense para averiguar que no cocina en su casa durante una semana promedio? ¿Por qué?

2. Audiencia de televidentes El programa de televisión *West Wing* tiene una audiencia de 15, es decir, mientras se está transmitiendo, el 15% de los televisores sintonizan ese programa (según datos de Nielsen Media Research). Un grupo especial consta de 20 hogares que se seleccionaron al azar (cada uno de ellos con el televisor funcionado durante la transmisión del programa *West Wing*).

continúa

a. ¿Cuál es el número que se espera de televisores que se sintonizan en el programa *West Wing*?
b. En tales grupos de 20, ¿cuál es el número medio de televisores que sintonizan el programa *West Wing*?
c. En tales grupos de 20, ¿cuál es la desviación estándar del número de televisores que sintonizan el programa *West Wing*?
d. Para tal grupo de 20, calcule la probabilidad de que exactamente cinco televisores sintonicen el programa *West Wing*?
e. Para tal grupo de 20, ¿sería poco común descubrir que ningún televisor sintoniza el programa *West Wing*? ¿Por qué?

3. Prueba de drogas a empleados De las compañías que construyen carreteras y puentes, el 80% prueba a sus empleados con respecto al abuso de sustancias (según datos de la Construction Financial Management Association). Un estudio implica la selección aleatoria de 10 compañías de este tipo.
a. Calcule la probabilidad de que 5 de las 10 compañías efectúen pruebas de abuso de sustancias.
b. Calcule la probabilidad de que al menos la mitad de las compañías hagan pruebas de abuso de sustancias.
c. Para tales grupos de 10 compañías, calcule la media y la desviación estándar del número (entre 10) que efectúa pruebas de abuso de sustancias.
d. ¿Sería poco común descubrir que seis de 10 compañías hacen pruebas de abuso de sustancias? ¿Por qué?

4. Razones de despido La incapacidad para llevarse bien con otras personas es la razón que se cita en el 17% de los despidos de trabajadores (de acuerdo con datos de Robert Half International, Inc.). Con preocupación por las condiciones de trabajo de su compañía, el gerente de personal de la Boston Finance Company planea investigar los cinco despidos que ocurrieron durante el año anterior.
a. 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.
b. 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 comunes? ¿Por qué?

5. Muertes Actualmente, un promedio de siete residentes del pueblo de Westport (población 760) mueren cada año (según datos del National Center for Health Statistics de Estados Unidos).
a. Calcule el número medio de muertes por día.
b. Calcule la probabilidad de que en un día dado no haya muertes.
c. Calcule la probabilidad de que en un día dado haya una muerte.
d. Calcule la probabilidad de que en un día dado haya más de una muerte.
e. Con base en los resultados anteriores, ¿debería Westport tener un plan de contingencia para manejar más de una muerte diaria? ¿Por qué?

Ejercicios de repaso acumulativo

1. Distancias de jonrones: análisis de los últimos dígitos La tabla al margen incluye los últimos dígitos de las 73 distancias que se publicaron (en pies) de los jonrones que logró Barry Bonds en el 2001, cuando estableció el récord del mayor número de jonrones en una temporada (según datos de *USA Today*). En ocasiones es posible emplear los últimos dígitos de un conjunto de datos para determinar si éstos se midieron o simplemente se reportaron. La presencia desproporcionada de los dígitos 0 y 5, suele ser un indicador seguro de que los datos se reportaron en lugar de medirse.

x	f
0	47
1	3
2	1
3	0
4	3
5	11
6	3
7	3
8	1
9	1

continúa

a. Calcule la media y la desviación estándar de los últimos dígitos.
b. Construya la tabla de frecuencias relativas que corresponde a la tabla de frecuencias dada.
c. Construya una tabla para la distribución de probabilidad de dígitos que se seleccionaron al azar, con posibilidades iguales. Liste los valores de la variable aleatoria *x* (0, 1, 2, . . . , 9), junto con sus probabilidades correspondientes (0.1, 0.1, 0.1, . . . , 0.1); después, calcule la media y la desviación estándar de tal distribución de probabilidad.
d. Reconociendo que los datos muestrales se desvían naturalmente de los resultados que se esperan teóricamente, ¿habrá un acuerdo burdo de los últimos dígitos dados con la distribución que esperamos con una selección aleatoria? ¿Habrá algo en los datos muestrales (como una desproporción debida a una mayor cantidad de dígitos 0 y 5) que sugiera que los últimos dígitos dados no son aleatorios? (En el capítulo 10 presentaremos un método para responder preguntas como éstas de forma más objetiva).

2. Prueba de contaminación de automóviles La Environmental Protection Agency realizó una prueba en el tubo de escape de 116,667 automóviles para determinar cuáles generaban una gran cantidad de contaminación. Se estima que el 1% de los automóviles no pasan esa prueba.
a. Si seleccionamos al azar 20 automóviles del grupo de 116,667, ¿cuántos se esperaría que no pasaran la prueba del tubo de escape?
b. Calcule la media y la desviación estándar del número de automóviles, en grupos de 20, que no pasan la prueba del tubo de escape.
c. Calcule la probabilidad de que, en un grupo de 20 automóviles que se seleccionaron aleatoriamente, haya al menos uno que no pase la prueba del tubo de escape.
d. ¿Es poco común encontrar que, en un grupo de 20 automóviles que se seleccionaron al azar, haya tres que no pasen la prueba del tubo de escape?
e. Si se seleccionan al azar dos automóviles diferentes, calcule la probabilidad de que ambos no pasen la prueba del tubo de escape.

Actividades de cooperación en equipo

1. Actividad en clase En el capítulo 1 presentamos varios ejemplos de conjuntos de datos confusos. Suponga que deseamos identificar la distribución de probabilidad del número de hijos que nacieron de parejas que se seleccionaron aleatoriamente. Pregunte a cada estudiante de la clase el número de hermanos y hermanas que tiene; ahora, registre el número total de hijos (incluyendo al estudiante) de cada familia. Construya una tabla de frecuencias relativas con el resultado que se obtuvo (los valores de la variable aleatoria *x* serán 1, 2, 3, . . .). ¿Cuál sería el problema si se utilizara esta tabla de frecuencias relativas como un estimado de la distribución de probabilidad del número de hijos que nacieron de parejas que se seleccionaron al azar?

2. Actividad en clase Divídanse en equipos de tres. Seleccionen a una persona a quien probarán su percepción extrasensorial (PES), tratando de identificar correctamente un dígito (0, 1, 2, . . . , 9) que se seleccionará al azar por otro miembro del equipo. Otro participante del equipo debe registrar el dígito que se seleccionará al azar, el dígito que adivinará el sujeto, así como si la adivinación fue correcta o incorrecta. Construyan la tabla de la distribución de probabilidad para dígitos que se generan aleatoriamente, la tabla de frecuencias relativas para dígitos aleatorios seleccionados realmente y una tabla de frecuencias relativas para las adivinaciones. Después de comparar las tres tablas, ¿qué concluyen? ¿Qué proporción de las adivinaciones fue correcta? ¿Parecería que el sujeto tiene la habilidad de seleccionar el dígito correcto, de manera significativa, con mayor frecuencia de lo que se esperaría por el azar?

Proyecto tecnológico

El vuelo 2705 de American Air, que va de Nueva York a San Francisco, incluye asientos para 340 pasajeros. En promedio, el 5% de las personas con reservaciones no se presenta, por lo que American Air vende boletos por encima del cupo y acepta 350 reservaciones para los 340 asientos. Analizamos este sistema a través de una distribución binomial con $n = 350$ y $p = 0.95$ (la probabilidad de que alguien con una reservación sí se presente).

Calcule la probabilidad de que, al aceptar 350 reservaciones en un vuelo particular, haya un mayor número de pasajeros que de asientos. Es decir, calcule la probabilidad de que al menos 341 personas con reservación se presenten, suponiendo que se aceptaron 350 reservaciones. Por el valor de n, no es posible utilizar la tabla A-1; además, los cálculos con la fórmula de probabilidad binomial serían extremadamente largos y tediosos. La mejor opción es utilizar un programa de cómputo de estadística o una calculadora TI-83 Plus. Consulte la sección 4-3 para encontrar instrucciones que describen el uso de STATDISK, Minitab, Excel o una calculadora TI-83 Plus. ¿La probabilidad de sobreventa del vuelo será suficientemente pequeña de modo que no suceda con frecuencia o parece ser demasiado alta, de modo que deben hacerse cambios para disminuirla?

de los DATOS *a la* DECISIÓN

Pensamiento crítico: ¿es mejor jugar en una máquina tragamonedas o en una ruleta?

El autor compró una máquina tragamonedas Las Vegas Mills Golden Nugget con el propósito de determinar su funcionamiento. A pesar de que esta máquina se basa en un diseño de 1940, trabaja según los mismos principios que utilizan las máquina tragamonedas que hay en los casinos de Las Vegas. Dicha máquina tiene tres ruedas que giran de forma independiente y, para cada carrete, se selecciona al azar una de 20 posiciones distintas cada vez que se jala la palanca. Las tablas siguientes resumen los posibles resultados y las posiciones ganadoras. Cada juego cuesta 25 centavos.

a. Con la información de las dos tablas, llene las probabilidades en la segunda tabla. Por ejemplo, hay cuatro resultados que ganan un premio mayor, y 8000 posibles resultados diferentes, de modo que P(premio mayor) = 4/8000.

b. ¿Describe la segunda tabla ya completa una distribución de probabilidad? ¿Por qué?

c. Sea x la variable que represente la ganancia o las pérdidas netas de un solo juego de la máquina tragamonedas y calcule la media de dicha variable aleatoria. Con base en tal resultado, ¿cuál es la cantidad promedio que se gana o pierde cuando un jugador introduce 25 centavos para un juego? ¿Cuál es la recuperación promedio para cada dólar apostado?

d. Cuando se apuesta un dólar al número 7 en la ruleta, hay una probabilidad de 1/38 de ganar, en tanto un triunfo genera una ganancia neta de \$35. ¿Cuál es la recuperación promedio por cada dólar apostado?

e. Compare los resultados de los incisos c y d para determinar si es mejor jugar en una máquina tragamonedas o apostar al 7 en la ruleta. Explique.

Frecuencia de imágenes en los tres carretes

	Carrete 1	Carrete 2	Carrete 3
Pepita de oro	2	2	1
Limón	0	0	4
Campana	1	7	7
Naranja	7	2	5
Ciruela	7	2	3
Cereza	3	7	0

Resultados posibles de los tres carretes

	Ganancia neta	Probabilidad
Premio mayor (3 pepitas de oro)	36.50	
Campana-campana-campana	4.25	
Campana-campana-pepita de oro	4.25	
Ciruela-ciruela-ciruela	3.25	
Ciruela-ciruela-pepita de oro	3.25	
Naranja-naranja-naranja	2.25	
Naranja-naranja-pepita de oro	2.25	
Cereza-cereza-cualquiera	1.00	
Cereza-no cereza-cualquiera	0.25	
Pérdida: cualquier resultado que no se incluya en los nueve renglones anteriores	−0.25	

PROYECTO DE INTERNET

Distribuciones de probabilidad y simulaciones

Las distribuciones de probabilidad se utilizan para predecir el resultado de 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 cinco caras y cinco cruces. Quizá no tengamos el resultado exacto, pero a la larga, después de cientos o miles de lanzamientos, esperamos que la división entre 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 4, 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*, 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 es realmente un programa de educación, pero es ampliamente reconocido debido a que los resultados se han hecho públicos".

Bárbara Carvalho

Directora del 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 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 y el estado de Nueva York, así como a lo largo de toda la nación. No somos partidarios de realizar encuestas para partidos políticos, candidatos políticos o grupos de poder. Recibimos fondos de manera independiente del Marist College y no recibimos ingresos externos que pudiesen 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, por lo que, si seleccionáramos personas al azar a lo largo de todo el estado, obtendríamos un modelo desigual de éste. Hacemos estratos por condado y usamos marcación de dígitos aleatoria, de modo que obtenemos números que se incluyen y no se incluyen en el directorio telefónico.

Acaba 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 las encuestas. Este fenómeno se incrementa con el tiempo y recibe mucha atención por parte de la comunidad de investigación por encuesta. 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%, hay una verdadera preocupación por descubrir quién se rehúsa, por qué no quiere responder y su impacto en la representatividad de los estudios que realizamos.

¿Recomendarían ustedes un curso de estadística para estudiantes?

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 presenten, utilizar datos para mejorar servicios o interpretar resultados para diseñar estrategias. Las encuestas, en particular, están por todas partes. Es vital que como trabajadores, gerentes y ciudadanos seamos 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.

¿Tiene 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.

5

Distribuciones de probabilidad normal

¿Cómo nos adaptamos?

Una disciplina relativamente nueva es la *ergonomía*, que estudia la adaptación de las personas a su ambiente. El buen diseño ergonómico resulta en un ambiente seguro, funcional, eficiente y cómodo. Las aplicaciones de la ergonomía incluyen el diseño de tableros de automóvil, cascos para ciclistas, tapas de botellas, perillas de puertas, cubiertas para registros, teclados, centros de control de tráfico aéreo y líneas de ensamble de computadoras. Por ejemplo, en Vail, Colorado, el teleférico que lleva a los esquiadores a la cima de la montaña tiene un letrero que especifica que la capacidad máxima es de 12 personas o 2004 libras. La lectura del anuncio hace que muchos pasajeros miren a su alrededor y se pregunten si están en peligro porque hay demasiadas personas o porque hay 12 individuos (o incluso menos) que, por variación aleatoria, son exageradamente pesados. ¿Qué tan probable es que 12 personas que se eligen al azar tengan un peso total mayor de 2004 libras?

La habilidad para tolerar un largo vuelo transcontinental se afecta por el ancho del asiento que ocupemos. La mayoría de los aviones comerciales de Estados Unidos contienen asientos que miden entre 17 y 18 pulgadas de ancho, que apenas rebasan las 16 pulgadas requeridas por un pasajero promedio. Los asientos de primera clase y de la clase de negocios suelen tener anchuras entre 19 y 21 pulgadas, de modo que un espacio más grande permite un mayor grado de comodidad. Si American Airlines quiere ganar más incrementando la comodidad de sus pasajeros, ¿qué anchura deben tener sus asientos que rediseña?

Cuando se visitan construcciones que datan de hace cientos de años, muchas personas se sorprenden por el hecho de que las entradas tienen aberturas demasiado bajas para la mayoría de los adultos actuales. Cuando caminamos a través de una entrada moderna, la mayoría de nosotros cabe cómodamente por debajo del umbral, que suele medir 80 pulgadas de alto. Sin embargo, algunas personas son excepcionalmente altas y deben agacharse para evitar golpearse la cabeza. ¿Qué porcentaje de personas son demasiado altas para los estándares de diseño de las entradas actuales?

En años recientes, la Fuerza Aérea de Estados Unidos reconoció que las mujeres son muy buenos pilotos de aviones de guerra. Las cabinas de los aviones de guerra se diseñaron originalmente para hombres, de modo que se requirieron varios cambios para acomodar mejor a las mujeres pilotos. Uno de dichos cambios implicó el rediseño de los asientos de expulsión ACES-II. Puesto que se diseñaron originalmente para hombres que pesaran entre 140 y 211 libras, los asientos de expulsión implicaban un mayor riesgo de daño para cualquier mujer piloto que pesara menos de 140 libras o más de 211 libras. ¿Qué pesos deben utilizarse para el nuevo diseño de la cabina?

En este capítulo resolvemos preguntas como las anteriores.

5-1 Panorama general

En el capítulo 2, consideramos medidas importantes de conjuntos de datos, incluyendo medidas de tendencia central y de variación, así como la distribución de los datos. En el capítulo 3, estudiamos los principios básicos de probabilidad; en el capítulo 4, presentamos los siguientes conceptos:

- Una *variable aleatoria* es una variable con un valor numérico único, que se determina al 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 tomar es 0 o 1, o 2, etcétera.
- Una variable aleatoria *continua* tiene un número infinito de valores, los cuales suelen asociarse con mediciones en una escala continua, sin huecos ni interrupciones.

En el capítulo 4, consideramos únicamente las 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 porque ocurren con gran frecuencia en las aplicaciones reales y porque juegan un papel fundamental en los métodos de estadística inferencial. Las distribuciones normales se utilizarán frecuentemente a lo largo del libro.

Definición

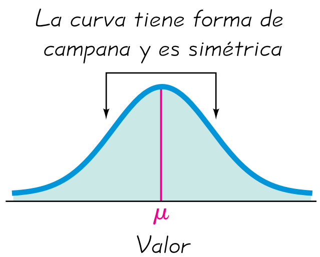
Si una variable aleatoria continua tiene una distribución con una gráfica simétrica y en forma de campana, como la de la figura 5-1, a la vez que puede ser descrita por medio de la ecuación dada como fórmula 5-1, decimos que tiene una **distribución normal**.

Fórmula 5-1

$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

La complejidad de la fórmula 5-1 provoca que muchas personas eleven las cejas mientras pronuncian la expresión "¡oh, oh!", o algo peor. Pero en realidad, te-

FIGURA 5-1 La distribución normal



nemos buenas noticias: no es necesario que utilicemos la fórmula 5-1. Sin embargo, la fórmula muestra que a cualquier distribución normal en particular la determinan dos parámetros: la media, μ, y la desviación estándar, σ. Una vez que se seleccionan valores específicos para μ y σ, se grafica la fórmula 5-1 como graficaríamos cualquier ecuación que relaciona a x con y; el resultado es una distribución de probabilidad continua con forma de campana.

5-2 Distribución normal estándar

El objetivo de este capítulo es el concepto de distribución de probabilidad normal, pero iniciamos con una *distribución uniforme*. La distribución uniforme nos permite ver algunas propiedades muy importantes que también se utilizarán con las distribuciones normales.

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 presenta una forma rectangular.

EJEMPLO Duración de la clase Un profesor de estadística planea sus clases con tanto cuidado que sus duraciones se distribuyen uniformemente entre 50.0 y 52.0 minutos. (Porque las clases de estadística son tan interesantes, que generalmente dan la impresión de ser más cortas). Esto es, cualquier tiempo entre 50.0 y 52.0 minutos es posible, en tanto que todos los valores posibles son igualmente probables. 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 5-2.

Cuando estudiamos las distribuciones de probabilidad *discretas* en la sección 4-2, identificamos dos requisitos: **1.** $\Sigma P(x) = 1$ y 2) $0 \leq P(x) \leq 1$ para todos

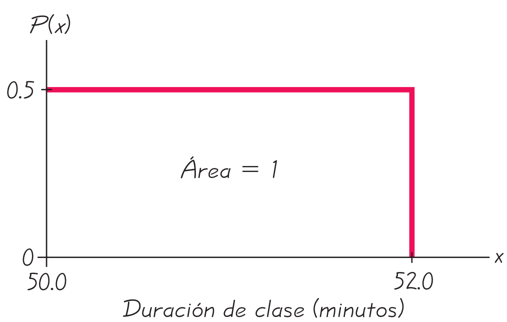


FIGURA 5-2 Distribución uniforme de la duración de las clases

Poblaciones cambiantes

Una de las cinco características más importantes de un conjunto de datos, que se listan en el capítulo 2, es el patrón de cambio de los datos a través del tiempo. Algunas poblaciones cambian y sus estadísticos importantes también. Los estándares de los cinturones de seguridad de los automóviles no han cambiado en 40 años, aun cuando el peso de los estadounidenses se incrementó de manera considerable desde entonces. En 1960, se consideraba que el 12.8% de los adultos estadounidenses tenían sobrepeso, en comparación con el 22.6% de 1994.

Según la National Highway Traffic Safety Administration, los cinturones de seguridad deben ajustarse a un maniquí estándar para choque (diseñado de acuerdo con los datos de 1960), que se colocó en la posición más adelante posible con 4 pulgadas de sobra. En teoría, el cinturón de seguridad tiene que ajustarse al 95% de los hombres y al 99% de las mujeres, pero tales porcentajes son ahora más bajos debido al incremento en el peso que tuvo lugar durante la última mitad del siglo. Algunas compañías proporcionan extensiones para cinturones de seguridad, pero otras no.

Muestreo rechazado para el censo

Se estima que, en el censo de 2000 en Estados Unidos, 7 millones de individuos no se contaron, mientras que otros 4 millones se consideraron dos veces. Dichos errores pueden corregirse al aplicar métodos de estadística que se conocen, aunque ello implica un asunto político. Los conteos poblacionales afectan el número de asientos en la Cámara de Representantes, de modo que los republicanos se oponen al muestreo, ya que las regiones que no se contaron completamente tienden a ser principalmente demócratas, mientras que las regiones que se cuentan en exceso tienen mayorías republicanas. Los demócratas están a favor del uso de métodos de muestreo. Algunas personas argumentan que la Constitución de Estados Unidos especifica que el censo debe ser un "conteo real" (un conteo por cabeza), que no permite métodos de muestreo; la Suprema Corte apoya esta posición. Los métodos estadísticos resultarían útiles para mejorar de manera sustancial los resultados del censo, de forma que los ciudadanos disfruten de una distribución más equitativa de la ayuda federal, junto con una representación más equitativa en el Congreso.

los valores de x. También en la sección 4-2 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 5-2, se llama *curva de densidad*; debe satisfacer dos propiedades similares a los requisitos de las distribuciones de probabilidad discretas, tal como se plantea en la siguiente definición.

Definición

Curva de densidad (o **función de densidad de probabilidad**): 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).

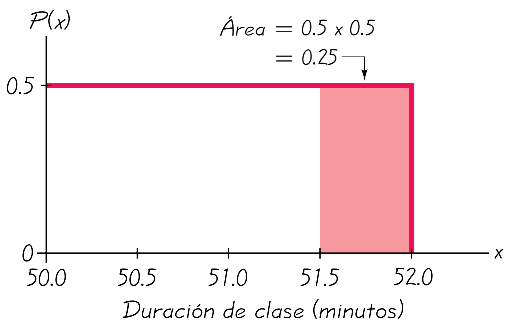
Si establecemos que la altura del rectángulo de la figura 5-2 es 0.5, obligamos a que el área circunscrita sea $2 \times 0.5 = 1$, como se requiere. (En general, el área del rectángulo se convierte en 1, cuando igualamos su altura al valor de $1/\text{rango}$). Esta propiedad (área $= 1$) facilita la solución de problemas de probabilidad, de modo que la siguiente afirmación es importante:

Ya que el área total debajo de la curva de densidad es igual a 1, hay una correspondencia entre *área* y *probabilidad*.

EJEMPLO Duración de la clase Kim, que 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. Dada la distribución uniforme de la figura 5-2, calcule la probabilidad de que una clase que se selecciona aleatoriamente dure más de 51.5 minutos.

SOLUCIÓN Observe la figura 5-3, donde la región sombreada representa duraciones mayores de 51.5 minutos. Puesto que el área total bajo la curva de

FIGURA 5-3 Uso del área para el cálculo de la probabilidad



densidad es igual a 1, hay una correspondencia entre área y probabilidad. Por lo tanto, es posible calcular la probabilidad que se desea utilizando áreas de la siguiente manera:

$$
\begin{aligned}
P\,(\text{clase mayor de 51.5 minutos}) &= \text{área de región sombreada de la figura 5-3}\\
&= 0.5 \times 0.5\\
&= 0.25
\end{aligned}
$$

INTERPRETACIÓN La probabilidad de seleccionar aleatoriamente una clase que dure más de 51.5 minutos es de 0.25. Ya que esa probabilidad es demasiado alta, Kim tiene que considerar hacer 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 forma que es sencillo 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 ve en la figura 5-1, por lo que es más difícil calcular áreas, pero el principio básico es el mismo: *existe una correspondencia entre área y probabilidad*.

Así como hay 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 5-4 incluye curvas de densidad de estaturas de hombres y mujeres adultos. Como los hombres tienen una estatura media mayor, la cima de la curva de densidad de los hombres se ubica 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 5-4 presenta dos posibles distribuciones normales diferentes. Hay una infinidad de posibilidades, pero una es de especial interés.

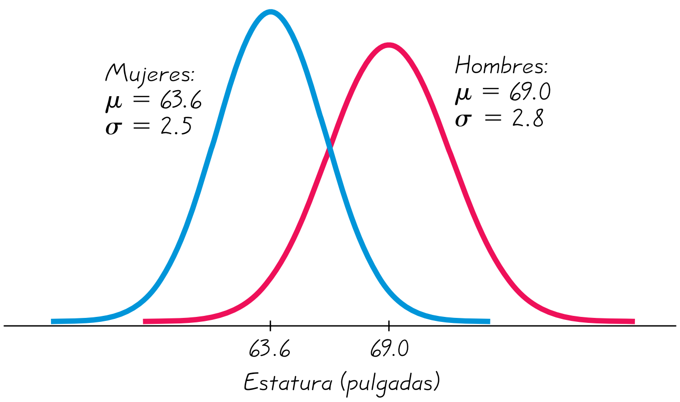


FIGURA 5-4 Estaturas de hombres y mujeres adultos

Confiabilidad y validez

La confiabilidad de los datos se refiere a la consistencia con que se presentan los resultados, mientras que la validez de los datos se refiere a lo bien que los datos miden lo que se supone deben medir. La confiabilidad de una prueba de CI puede juzgarse comparando las puntuaciones de la prueba en una aplicación, con las puntuaciones de la misma prueba que se aplica en otro momento. Para probar la validez de una prueba de CI, habrá que comparar las puntuaciones de la prueba con algún otro indicador de inteligencia, como el desempeño académico. Muchos críticos afirman que las pruebas de CI son confiables, pero no válidas; ofrecen resultados consistentes, aunque no miden realmente la inteligencia.

Definición

Distribución normal estándar: distribución normal de probabilidad con una media de 0 y una desviación estándar de 1, en tanto el área total debajo de su curva de densidad es igual a 1. (Véase la figura 5-5).

Suponga que nos contrataron para realizar cálculos con el uso de la fórmula 5-1. 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 calculan muchas áreas diferentes bajo la curva. Como se aprecia en la figura 5-5, el área bajo la curva es 1; lo anterior 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*

Si empleamos la tabla A-2 (en el Apéndice A y en la tarjeta con *fórmulas y tablas*) TI-83 Plus o programas de cómputo como el STATDISK, Minitab o Excel. Las características más importantes de los distintos métodos se resumen en la tabla 5-1. No es necesario conocer los cinco métodos, sólo necesita aprender el método que utilizará para la clase y los exámenes.

Puesto que los siguientes ejemplos y ejercicios se basan en la tabla A-2, es esencial comprender los siguientes puntos:

1. La tabla A-2 se diseñó 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 otra para las puntuaciones *z positivas*.
3. Cada valor en la tabla es una *área acumulativa desde la izquierda* hasta una frontera vertical, por sobre 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 que se asocia con un valor menor que 1.58. Comience con la puntuación *z* de 1.58, localizando 1.5 en la columna izquierda; después, calcule el valor en el renglón adjunto de probabili-

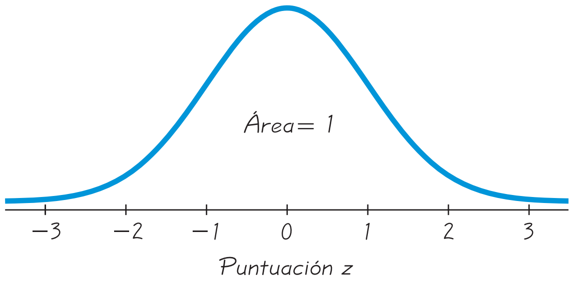
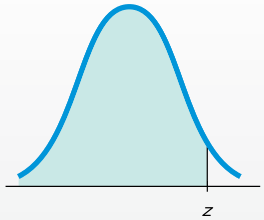
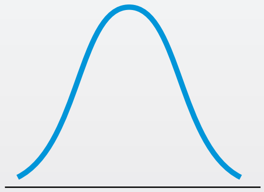
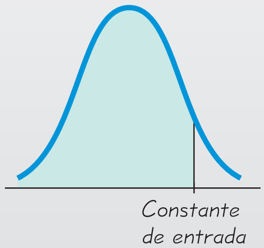
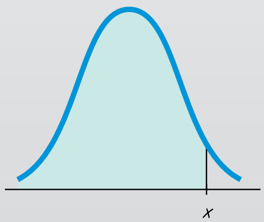
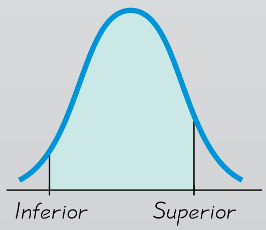


FIGURA 5-5 Distribución normal estándar: $\mu = 0$ y $\sigma = 1$

Tabla 5-1 Métodos para el cálculo de las áreas de la distribución normal

Tabla A-2 Da el área acumulativa de la izquierda hasta una línea vertical por encima de un valor específico de *z*.		El procedimiento para el uso de la tabla A-2 se describe en el texto.
STATDISK Da pocas áreas, incluyendo el área acumulativa de la izquierda y el área acumulativa de la derecha.		Seleccione **Analysis, Probability Distributions, Normal Distribution.** Deslice el mouse hacia la derecha y la izquierda.
Minitab Da el área acumulativa de la izquierda hasta una línea vertical por arriba de un valor específico.		Seleccione **Calc, Probability Distributions, Normal.** En el cuadro de diálogo, seleccione **Cumulative Probability, Input Constant.**
Excel Da el área acumulativa de la izquierda hasta una línea vertical por arriba de un valor específico.		Seleccione **fx, Statistical, NORMDIST.** En el cuadro de diálogo, introduzca el valor y la media, la desviación estándar y "true".
TI-83 Plus Da el área con límites izquierdo y derecho, por medio de líneas verticales, sobre cualesquiera valores específicos.		Presione `2nd` `VARS` **[2: normal cdf(]**; después, introduzca *z* separadas por una coma, como en (puntuación *z* izquierda, puntuación *z* derecha).

z	....	0.08
.		.
.		.
.		.
1.5	....	0.9429

dad que está directamente debajo de 0.08, como se observa en esta porción de la tabla A-2.

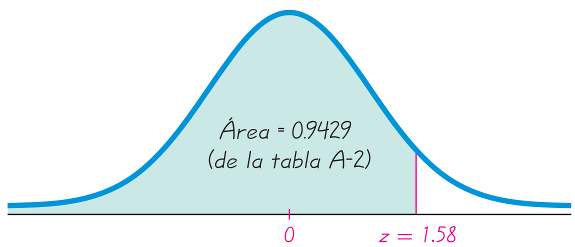
El valor del área (o probabilidad) de 0.9429 indica que hay una probabilidad de 0.9429 de seleccionar aleatoriamente una puntuación z menor que 1.58. (En las siguientes secciones, consideraremos casos en los cuales la media no es 0 ni la desviación estándar es 1.)

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 tales instrumentos reveló que, en el punto de congelación del agua, algunos termómetros daban lecturas por debajo de 0° (que se denotan con números negativos) y otros por encima de 0° (que se denotan con números positivos). Suponga que la lectura media es 0°C y la desviación estándar de las lecturas es 1.00°C. También, 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 $z = 1.58$, en la figura 5-6. 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 dicha área es de 0.9429.

INTERPRETACIÓN La *probabilidad* de seleccionar aleatoriamente 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 que se sombreó en la figura 5-6. Otra forma de interpretar el resultado es concluyendo que el 94.29% de los termómetros tendrán lecturas por debajo de 1.58°.

FIGURA 5-6 Cálculo del área por debajo de $z = 1.58$



EJEMPLO Termómetros científicos Utilice los termómetros del ejemplo anterior y calcule la probabilidad de seleccionar aleatoriamente un termómetro con una lectura, en el punto de congelación del agua, por arriba de −1.23°.

SOLUCIÓN Nuevamente, calculamos la *probabilidad* que se desea encontrando un *área* correspondiente. Buscamos el área de la región que se sombreó en la figura 5-7, pero la tabla A-2 se diseñó para aplicarse sólo 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, como se observa. Sabiendo que el área total bajo la

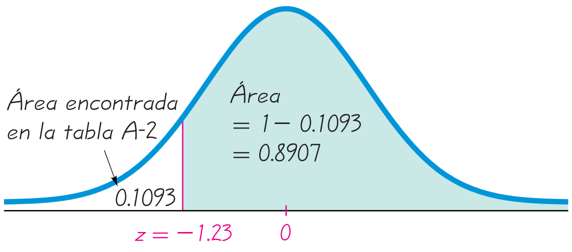


FIGURA 5-7 **cálculo del área por encima de $z = -1.23$**

curva es 1, calcularemos el área que se sombreó si restamos 0.1093 de 1. El resultado es 0.8907. Aun cuando la tabla A-2 se diseñó únicamente para áreas acumulativas de la izquierda, podemos utilizarla para calcular áreas acumulativas de la derecha, tal como se muestra en la figura 5-7.

INTERPRETACIÓN Por la correspondencia entre probabilidad y área, concluimos 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 (correspondiente al *área* que está por arriba de $z = -1.23$). En otras palabras, el 89.07% de los termómetros tienen lecturas por encima de −1.23°.

El ejemplo anterior ilustra una de las formas en que es posible utilizar la tabla A-2 para calcular, de manera indirecta, una área acumulativa de la derecha. El siguiente ejemplo ilustra otra manera para calcular el área con el uso de 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 que se eligió tenga lecturas, en el punto de congelación del agua, entre −2.00° y 1.50°.

SOLUCIÓN Nuevamente tratamos con valores que se distribuyen de manera normal, con una media de 0° y una desviación estándar de 1°. La probabilidad de seleccionar un termómetro con lecturas entre −2.00° y 1.50° corresponde al área que se sombreó en la figura 5-8. La tabla A-2 no puede utilizarse para calcular el área de forma directa, pero sí 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 indica en la figura. Remítase a la figura 5-8 y observe que el área que se sombreó corresponde a la diferencia entre 0.9332 y 0.0228. El área que se sombreó es, por lo tanto, $0.9332 - 0.0228 = 0.9104$.

continúa

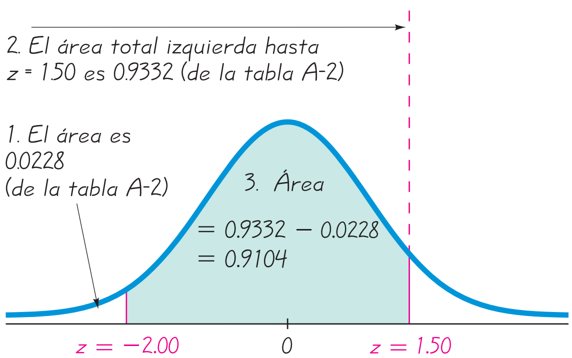


FIGURA 5-8 **Cálculo del área entre dos valores**

INTERPRETACIÓN Con el uso de la correspondencia entre probabilidad y área, concluimos que hay una probabilidad de 0.9104 de seleccionar aleatoriamente uno de los termómetros con una lectura de 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, y se prueban 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 que se localiza entre dos puntuaciones z específicas se obtiene calculando la diferencia entre las dos áreas que se localizan en la tabla A-2. Observe la figura 5-9, que muestra que la región B que se sombreó se obtiene 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 A (que en la tabla A-2 aparece como el área correspondiente a $z_{\text{Izquierda}}$). *Sugerencia*: No trate de memorizar una regla o una fórmula para este caso, ya que es infinitamente mejor *comprender* el procedimiento. Entienda, mejor, cómo funciona la tabla A-2; después, dibuje una gráfica, sombree el área que se desea y piense en una forma para calcular el área, considerando que la tabla A-2 proporciona sólo áreas acumulativas desde la izquierda.

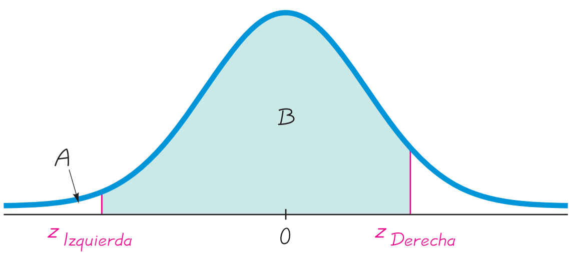
El ejemplo anterior concluyó con la afirmación de que la probabilidad de una lectura entre $-2.00°$ y $1.50°$ es de 0.9104. Probabilidades como ésta, también pueden expresarse con la siguiente notación:

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 el uso de esa notación expresaremos el resultado del último ejemplo de la siguiente manera: $P(-2.00 < z < 1.50) = 0.9104$ que, en símbolos, establece 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 de 0. Es decir, $P(z = a) = 0$.

FIGURA 5-9 Cálculo del área entre dos puntuaciones z



Área sombreada B = (áreas A y B combinadas) − (área A) = (área de la tabla A-2, usando z_{Derecha}) − (área de la tabla A-2, usando $z_{\text{Izquierda}}$)

Por ejemplo, hay una probabilidad 0 de seleccionar aleatoriamente a alguien y obtener una persona cuya estatura sea con exactitud de 68.12345678 pulgadas. En la distribución normal, cualquier punto único sobre la escala horizontal se representa, no por una región bajo la curva, sino por una línea vertical por encima del punto. Para $P(z = 1.50)$, tenemos una línea vertical que está por arriba de $z = 1.50$, pero la línea vertical, por sí misma, no contiene un área, de modo que $P(z = 1.50) = 0$. Para cualquier variable aleatoria continua la probabilidad de un valor exacto es 0; además, 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 siguen el mismo formato: dadas puntuaciones z, calculamos áreas bajo la curva; dichas áreas corresponden a probabilidades. En muchos otros casos, realizamos el proceso contrario, porque ya conocemos el área (o probabilidad), pero necesitamos calcular la puntuación z correspondiente. En estos 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, que se representa con 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), regiones bajo la curva, se representan con los valores en el cuerpo de la tabla A-2. Asimismo, las puntuaciones z que se ubican 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 un área conocida

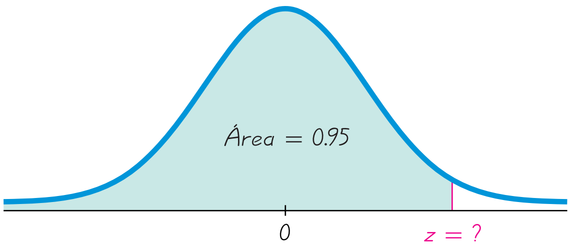
1. Dibuje una curva en forma de campana e identifique la región bajo la curva correspondiente a la probabilidad dada. Si no se trata de una región acumulativa de la izquierda, en su lugar trabaje con una región acumulativa que se conoce de la izquierda.
2. Usando el área acumulativa de la izquierda, 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 que se distribuyen 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, determine la temperatura que separa el 95% inferior del 5% superior. Observe la figura 5.10.

SOLUCIÓN La figura 5-10 incluye la puntuación z correspondiente 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 de la izquierda*. Al remitirnos a la tabla A-2, buscamos el área de

continúa

FIGURA 5-10 Cálculo del percentil 95



0.95 *en el cuerpo* de la tabla y después calculamos la puntuación z correspondiente. En la tabla encontramos las áreas de 0.9495 y 0.9505, donde hay un asterisco con una nota especial que indica que 0.9500 corresponde a una puntuación z de 1.645. Entonces concluimos que la puntuación z, en la figura 5-10, es 1.645, por lo que el percentil 95 es la lectura de la temperatura de 1.645°C.

INTERPRETACIÓN Al realizar pruebas a la temperatura de congelación, el 95% de las lecturas serán menores o iguales que 1.645°C, en tanto que el 5% de ellas será mayor o igual que 1.645°C.

Puntuación z	Área acumulativa de la izquierda
1.645	0.9500
−1.645	0.0500
2.575	0.9950
−2.575	0.0050

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 evitaremos la interpolación si seleccionamos sencillamente el valor más cercano. Hay casos especiales, que se listan en la tabla adjunta, los cuales son importantes porque se utilizan con frecuencia en una amplia variedad de aplicaciones. (El valor de $z = 2.576$ da un 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, es posible seleccionar el valor más cercano en la tabla. (Si un valor que se desea se encuentra entre dos valores de la tabla, seleccione el valor más grande). Además, para las puntuaciones z por arriba de 3.49, utilizaremos 0.9999 como aproximación del área acumulativa de la izquierda; para puntuaciones z por debajo de −3.49, usaremos 0.0001 como aproximación del área acumulativa 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 5-11, que presenta las puntuaciones z que se requieren. Para encontrar la puntuación z que se localiza a la izquierda, remítase a la tabla A-2 y busque un área de 0.025, en el *cuerpo de la tabla*. El resultado es $z = -1.96$. Para encontrar la puntuación z que se localiza a la derecha, remítase al *cuerpo de la tabla* y busque un área de 0.975. (Recuerde que la tabla A-2 siempre da áreas acumulativas 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 muestra la figura 5-11.

INTERPRETACIÓN Al realizar pruebas a la temperatura de congelación, el 2.5% de las lecturas de los termómetros serán iguales o menores que −1.96°, en tanto que 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 se ubicarán entre −1.96° y 1.96°.

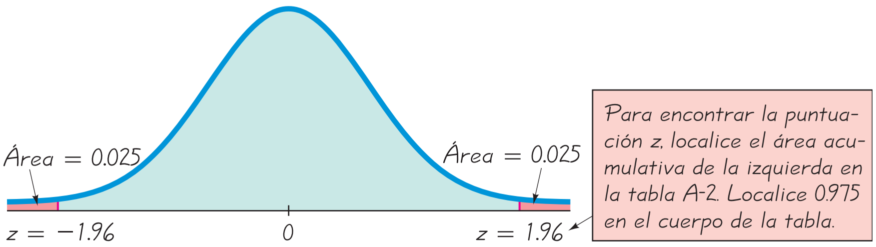


FIGURA 5-11 Cálculo de puntuaciones *z*

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 clásicas incluyen medias distintas de 0 y desviaciones estándar distintas de 1. En la siguiente sección, introducimos métodos para trabajar con este tipo de distribuciones normales, que son más realistas.

Utilizando la tecnología

STATDISK Seleccione **Analysis, Probability Distributions, Normal Distribution**; luego, proceda a deslizar el mouse hacia la derecha o la izquierda, hasta encontrar el valor que se desea. Quizá logre mayor precisión si emplea el mouse para arrastrar parte de la curva, de manera que sea posible amplificarla.

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;** después, introduzca la media de 0 y la desviación estándar de 1, haga clic en el botón de **Input Constant** e inserte la puntuación z.
- Para encontrar la puntuación z correspondiente a una probabilidad que se conoce, seleccione **Calc, Probability Distributions, Normal**; después, elija **Inverse cumulative probabilities** y la opción **Input constant.** Para la constante de entrada, introduzca 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, seleccione **Statistical, NORMSDIST** e introduzca la puntuación z.
- Para encontrar la puntuación z correspondiente a una probabilidad que se conoce, seleccione ***fx*****, Statistical, NORMSINV** e introduzca el área total que se encuentra a la izquierda del valor dado.

TI-83 Plus

- Para calcular el área entre dos puntuaciones z, presione **2nd VARS, 2**(para normalcdf); después, proceda a introducir las dos puntuaciones z, que se separaron con una coma, como en (puntuación z izquierda, puntuación z derecha).
- Para encontrar una puntuación z correspondiente a una probabilidad que se conoce, presione **2nd VARS, 3** (para invNorm), y proceda a introducir 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.

5-2 Destrezas y conceptos básicos

Uso de una distribución uniforme continua*. En los ejercicios 1 a 4, remítase a la distribución uniforme de la figura 5-2; suponga que se selecciona una clase con duración entre 50.0 y 52.0 minutos, calcule la probabilidad de seleccionar el tiempo dado.*

1. Menor que 50.3 min.

2. Mayor que 51.0 min.

3. Entre 50.5 minutos y 50.8 minutos.

4. Entre 50.5 minutos y 51.8 minutos.

***Uso de una distribución uniforme continua**. En los ejercicios 5 a 8, suponga que los voltajes en un circuito varían entre 6 y 12 volts, así como que los voltajes se distribuyen de forma equitativa en el rango de posibilidades, de modo que existe una distribución uniforme. Calcule la probabilidad del rango dado de niveles de voltaje.*

5. Mayor que 10 volts.

6. Menor que 11 volts.

7. Entre siete y 10 volts.

8. Entre 6.5 y ocho volts.

***Uso de la distribución normal estándar.** En los ejercicios 9 a 28, suponga que las lecturas de los 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 en grados.*

9. Menor que -0.25.

10. Menor que -2.75.

11. Menor que 0.25.

12. Menor que 2.75.

13. Mayor que 2.33.

14. Mayor que 1.96.

15. Mayor que -2.33.

16. Mayor que -1.96.

17. Entre 0.50 y 1.50.

18. Entre 1.50 y 2.50.

19. Entre -2.00 y -1.00.

20. Entre 2.00 y 2.34.

21. Entre -2.67 y 1.28.

22. Entre -1.18 y 2.15

23. Entre -0.52 y 3.75.

24. Entre -3.88 y 1.07.

25. Mayor que 3.57.

26. Menor que -3.61.

27. Mayor que 0.

28. Menor que 0.

***Bases de la regla empírica.** En los ejercicios 29 a 32, calcule el área bajo la curva que se indica de la distribución normal estándar; después, conviértala en porcentaje y llene el espacio en blanco. Los resultados conforman la base de la regla empírica que se explicó en la sección 2-5.*

29. Aproximadamente el ____ % del área está entre $z = -1$ y $z = 1$ (o dentro de una desviación estándar a partir de la media).

30. Aproximadamente el ____ % del área está entre $z = -2$ y $z = 2$ (o dentro de dos desviaciones estándar a partir de la media).

31. Aproximadamente el ____ % del área está entre $z = -3$ y $z = 3$ (o dentro de tres desviaciones estándar a partir de la media).

32. Aproximadamente el ____ % 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 los termómetros se distribuyen normalmente, con una media de 0° y una desviación estándar de 1.00°. Calcule la probabilidad que se indica, donde z es la lectura en grados.*

33. $P(-1.96 < z < 1.96)$

34. $P(z < 1.645)$

35. $P(z > -2.575)$

36. $P(1.96 < z < 2.33)$

Cálculo de valores de temperatura. *En los ejercicios 37 a 40, suponga que las lecturas de los 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 lectura de la temperatura correspondiente a la información dada.*

37. Calcule P_{90}, el percentil 90o. Ésta es la lectura de temperatura que separa el 90% inferior del 10% superior.

38. Calcule P_{20}, el percentil 20o.

39. Si se rechaza el 5% de los termómetros porque tienen lecturas muy bajas, pero el resto de los termómetros son aceptables, calcule la lectura que separa a los termómetros que se rechazaron de los otros.

40. Si se rechaza el 3.0% de los termómetros, porque tienen lecturas muy altas, y otro 3.0% se rechaza por registrar lecturas muy bajas, calcule las dos lecturas de los valores que separan a los termómetros que se rechazaron de los otros.

5-2 Más allá de lo básico

41. Para una distribución normal estándar, calcule el porcentaje de datos que están:
a. Dentro de 1 desviación estándar a partir de la media.
b. Dentro de 1.96 desviaciones estándar a partir de la media.
c. Entre $\mu - 3\sigma$ y $\mu + 3\sigma$.
d. Entre 1 desviación estándar por debajo de la media y 2 desviaciones estándar por encima de la media.
e. 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 $\sqrt{3}$.
a. Para esta distribución calcule $P(-1 < x < 1)$.
b. Calcule $P(-1 < x < 1)$ si considera de manera incorrecta que la distribución es normal en lugar de uniforme.
c. Compare los resultados de los incisos *a* y *b*. ¿Afecta mucho la distribución a los resultados?

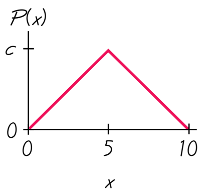
43. Suponga que puntuaciones *z* se distribuyen normalmente, con una media de 0 y una desviación estándar de 1.
a. Si $P(0 < z < a) = 0.3907$, calcule *a*.
b. Si $P(-b < z < b) = 0.8664$, calcule *b*.
c. Si $P(z > c) = 0.0643$, calcule *c*.
d. Si $P(z > d) = 0.9922$, calcule *d*.
e. Si $P(z < e) = 0.4500$, calcule *e*.

44. En una distribución uniforme continua,

$$\mu = \frac{\text{mínimo} + \text{máximo}}{2} \qquad \text{y} \qquad \sigma = \frac{\text{rango}}{\sqrt{12}}$$

Calcule la media y la desviación estándar de la distribución uniforme que se representa en la figura 5-2.

45. Realice el dibujo 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 marginal).
  a. Calcule el valor de la constante *c*.
  a. Calcule la probabilidad de que *x* esté entre 0 y 3.
  a. Calcule la probabilidad de que *x* esté entre 2 y 9.

5-3 Aplicaciones de las distribuciones normales

Todos los ejemplos y ejercicios de la sección 5-2 son poco realistas, ya que incluyeron la distribución normal *estándar* (con una media de 0 y una desviación estándar de 1). En esta sección incluimos distribuciones normales no estándar, de modo que podamos trabajar con aplicaciones reales y prácticas. Sin embargo, es posible transformar valores de una distribución normal no estándar a una distribución normal estándar, para así continuar utilizando los mismos procedimientos de la sección 5-2.

Si convertimos valores en puntuaciones estándar, empleando la fórmula 5-2, entonces los procedimientos para trabajar con todas las distribuciones normales son los mismos que los de la distribución normal estándar.

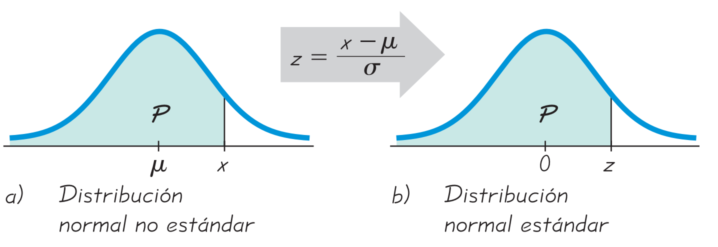
Fórmula 5-2 $z = \frac{x - \mu}{\sigma}$ (redondear puntuaciones *z* hasta dos decimales)

El uso continuo de la tabla A-2 requiere la comprensión y la aplicación del principio anterior. (Si utiliza ciertas calculadoras o programas de cómputo, no será necesaria la transformación a puntuaciones *z*, ya que las probabilidades se calculan de manera directa). Sin importar el método que utilice, debe comprender con claridad el principio básico anterior, puesto que constituye un fundamento importante de los conceptos que se introducen en los siguientes capítulos.

La figura 5-2 ilustra la transformación de una distribución no estándar a una estándar. El área de cualquier distribución normal que se limita por alguna puntuación *x* (como en la figura 5-12a), es igual que el área que se limita por la puntuación *z* equivalente en la distribución normal estándar (como en la figura 5-12b). Lo anterior significa que cuando se trabaja con una distribución normal no estándar, a veces se utiliza la tabla A-2 de la misma forma que se empleó en la sección 5-2, siempre y cuando los valores se conviertan primero a puntuaciones *z*. Cuando calcule áreas en una distribución normal no estándar, utilice este procedimiento:

1. Dibuje una curva normal, etiquete la media y los valores específicos de *x*; después, *sombree* la región que representa la probabilidad que se desea.
2. Para cada valor relevante de *x* que sea un límite de la región que se sombreó, utilice la fórmula 5-2 para transformar el valor a la puntuación *z* equivalente.

FIGURA 5-12 Transformación de una distribución normal no estándar a una distribución normal estándar



3. Remítase a la tabla A-2 para encontrar el área de la región que se sombreó, la cual constituye la probabilidad que se desea.

El siguiente ejemplo aplica estos tres pasos e ilustra la relación entre una distribución no normal típica y la distribución normal estándar.

EJEMPLO Diseño de automóviles La altura, en posición de sentado (del asiento a la cima de la cabeza), de los conductores debe tomarse en cuenta en el diseño de un nuevo modelo de automóvil. Los hombres tienen alturas que se distribuyen normalmente, con una media de 36.0 pulgadas y una desviación estándar de 1.4 pulgadas (según datos de una encuesta antropométrica de Gordon, Clauser *et al*.). Los ingenieros elaboran planes que pueden acomodar a hombres con alturas, estando sentados, de hasta 38.8 pulgadas, pero aquellos con mayor altura no se ajustan. Si se selecciona un hombre aleatoriamente, calcule la probabilidad de que su altura, estando sentado, sea menor que 38.8 pulgadas. Con base en ese resultado, ¿es factible el actual diseño de ingeniería?

SOLUCIÓN

Paso 1: Observe la figura 5-13, donde está marcada la media de 36.0 y la altura máxima de un hombre sentado de 38.8 pulgadas, en la cual el área que representa la probabilidad que se busca se sombreó. (Continuamos utilizando la misma correspondencia entre *probabilidad* y *área*, tal como se introdujo la sección 5-2).

Paso 2: Para usar la tabla A-2, primero hay que aplicar la fórmula 5-2, para convertir la distribución de alturas a una distribución normal estándar. La altura de 38.8 pulgadas se convierte en una puntuación z, de la siguiente manera:

$$z = \frac{x - \mu}{\sigma} = \frac{38.8 - 36.0}{1.4} = 2.00$$

Tal resultado demuestra que la altura de un hombre sentado, de 38.8 pulgadas, se encuentra por arriba de la media de 36.0 pulgadas por 2.00 desviaciones estándar.

Paso 3: Remitiéndonos a la tabla A-2, encontramos que $z = 2.00$ corresponde a un área de 0.9772.

INTERPRETACIÓN Hay una probabilidad de 0.9772 de seleccionar aleatoriamente a un hombre que, sentado, tenga una altura menor que 38.8 pulgadas. En símbolos, esto se expresa como

$$P(x < 38.8 \text{ pulg.}) = P(z < 2.00) = 0.9772$$

continúa

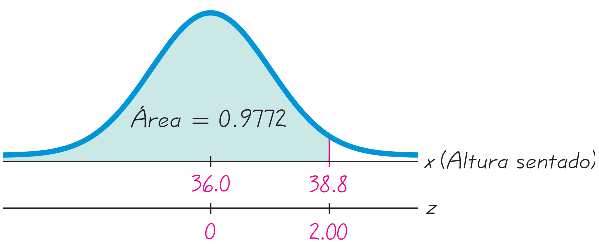


FIGURA 5-13 Distribución normal de la altura de hombres sentados

Atajo para ensayo clínico

¿Qué haría si estuviera probando un tratamiento y, antes de que su estudio termine, se da cuenta de que es claramente efectivo? Debería acortar el estudio e informar a todos los participantes acerca de la efectividad del tratamiento. Lo anterior fue lo que sucedió cuando se probó la hidroxiurea como tratamiento para la anemia falciforme. El estudio se programó para durar cerca de 40 meses, pero la efectividad 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).

Otra forma de interpretar este resultado es concluir que el 97.72% de los hombres tienen alturas menores que 38.8 pulgadas, cuando están sentados en un automóvil. Una consecuencia importante de tal resultado es que el 2.28% de los hombres no se ajustan al automóvil. El fabricante debe decidir ahora si puede costear la pérdida del 2.28% de los conductores de automóviles varones.

PROBLEMA 5 DEL CAPÍTULO **EJEMPLO Asientos de expulsión de aviones de propulsión a chorro** En el problema del capítulo se señaló que la Fuerza Aérea de Estados Unidos estuvo usando los asientos expulsión ACES-II, que se diseñaron para hombres con un peso entre 140 y 211 libras. Siendo que los pesos de mujeres se distribuyen normalmente, con una media de 143 libras y una desviación estándar de 29 libras (según datos del National Health Survey), ¿qué porcentaje de las mujeres tiene pesos que se encuentran dentro de dichos límites?

SOLUCIÓN Observe la figura 5-14, que muestra la región que se sombreó de las mujeres que pesan entre 140 y 211 libras. Es posible encontrar esa área que se sombreó directamente de la tabla A-2, y obtenerla de manera indirecta utilizando los procedimientos básicos presentados en la sección 5-2. Para esto, primero debemos encontrar el área acumulativa de la izquierda, hasta 140 libras, y el área acumulativa de la izquierda, hasta 211 libras; después, se obtiene la diferencia entre ambas áreas.

Obtención del área acumulativa hasta 140 libras:

$$z = \frac{x - \mu}{\sigma} = \frac{140 - 143}{29} = -0.10$$

Si usamos la tabla A-2, encontraremos que $z = -0.10$ corresponde a un área de 0.4602, como se aprecia en la figura 5-14.

Obtención del área acumulativa hasta 211 libras:

$$z = \frac{x - \mu}{\sigma} = \frac{211 - 143}{29} = 2.34$$

Si usamos la tabla A-2, encontraremos que $z = 2.34$ corresponde a un área de 0.9904, como se observa en la figura 5-14.

Obtención del área entre 140 libras y 211 libras:

$$\text{Área sombreada} = 0.9904 - 0.4602 = 0.5302$$

INTERPRETACIÓN Encontramos que el 53.02% de las mujeres tienen pesos que se encuentran entre los límites del asiento de expulsión de 140 y 211 libras. Lo anterior significa que el 46.98% de las mujeres no tienen pesos que estén dentro de los límites actuales; muchas mujeres pilotos correrían el riesgo de dañarse gravemente si tuvieran que utilizar el asiento de expulsión.

Filas

La teoría de las filas es una rama de las matemáticas que se apoya en la probabilidad y la 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 atienden en sus automóviles. Disney realiza extensos estudios de filas en sus parques de diversiones, para mantener contentos a sus visitantes y planear sus expansiones. 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.

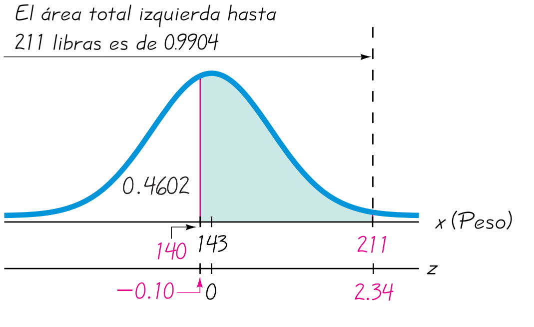


FIGURA 5-14 Pesos de mujeres y límites de los asientos de expulsión

Cálculo de valores de áreas conocidas

Los ejemplos anteriores en esta sección son del mismo tipo: se nos dan valores de límites específicos y debemos encontrar un área (o probabilidad o porcentaje). En muchos casos prácticos y reales, ya se conoce el área (o probabilidad o porcentaje), pero habrá que encontrar el (los) valor(es) relevante(s). Cuando busque valores de áreas conocidas, asegúrese de considerar lo siguiente:

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 localizará 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 del cálculo de valores con el uso de la tabla A-2 y la fórmula 5-2

1. Dibuje una curva de distribución normal, anote la probabilidad o porcentaje dados en la región apropiada de la gráfica e identifique el (los) valor(es) 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 5-2, sustituya los valores de μ, σ y la puntuación z que se obtuvo en el paso 2; después, calcule x. Con base en el formato de la fórmula 5-2, calculamos x de la siguiente manera:

$$x = \mu + (\underset{\uparrow}{z} \cdot \sigma) \qquad \text{(otra forma de la fórmula 5-2)}$$

(Si z se localiza 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 es lógica en el contexto de la gráfica y en el del problema.

El siguiente ejemplo utiliza el procedimiento que se acaba de describir.

PROBLEMA DEL CAPÍTULO 5

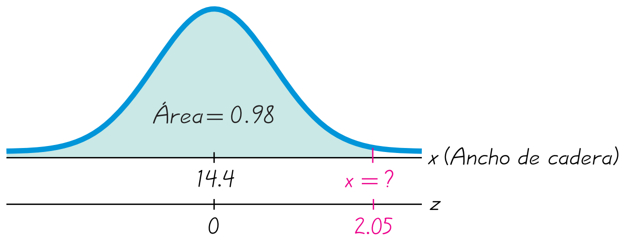
EJEMPLO Ancho de caderas y asientos de aviones Al diseñar los asientos que se habrán de instalar en aviones comerciales, los ingenieros buscan hacerlos con una anchura suficiente para que quepa el 98% de los hombres. (Acomodar al 100% de los hombres requeriría asientos muy anchos, que serían muy caros). El ancho de las caderas de los hombres se distribuye de manera normal, con una media de 14.4 pulgadas y

continúa

El medio de una encuesta puede afectar los resultados

En una encuesta que se realizó a individuos católicos de Boston, se les preguntó si pensaban que debían proporcionarse anticonceptivos a las mujeres solteras. En entrevistas personales, el 44% de los sujetos respondieron afirmativamente. Pero en un grupo similar, que se contactó por correo o por teléfono, el 75% de las personas respondió que sí a la misma pregunta.

FIGURA 5-15 Distribución del ancho de cadera de hombres



una desviación estándar de 1.0 pulgadas (de acuerdo con datos de una encuesta antropométrica de Gordon, Clauser *et al*.). Calcule P_{98}, es decir, obtenga el ancho de cadera de hombres que separa al 98% inferior del 2% superior.

SOLUCIÓN

Paso 1: Iniciamos con la gráfica de la figura 5-15. Ya anotamos la media de 14.4 pulgadas, sombreamos el área que representa al 98% inferior e identificamos el valor que se desea como x.

Paso 2: En el *cuerpo* de la tabla A-2 buscamos un área de 0.9800. (El área de 0.98 que se muestra en la figura 5-15 es un área acumulativa de la izquierda y exactamente el tipo de área que se lista en la tabla A-2). El área más cercana a 0.98, que es 0.9798, corresponde a la puntuación z de 2.05.

Paso 3: Con $z = 2.05$, $\mu = 14.4$ y $\sigma = 1.0$, calculamos x empleando la fórmula 5-2 de manera directa o utilizando la siguiente versión de la fórmula 5-2:

$$x = \mu + (z \cdot \sigma) = 14.4 + (2.05 \cdot 1.0) = 16.45$$

Paso 4: Si permitimos que $x = 16.45$ en la figura 5-15, veremos que esta solución es razonable, ya que el percentil 98 debe ser mayor que la media de 14.4.

INTERPRETACIÓN El ancho de cadera de 16.5 pulgadas (que se redondeó a un decimal, como en μ y σ) separa al 98% inferior del 2% superior. Es decir, los asientos que se diseñan para un ancho de cadera de hasta 16.5 pulgadas se ajustarán al 98% de los hombres. Este tipo de análisis se utiliza para diseñar los asientos que se emplean actualmente en los aviones comerciales.

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 tiene que mover su cuerpo de manera que podría distraerse, lo cual sería peligroso. (No desearíamos que alguien se lastimara tratando 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.

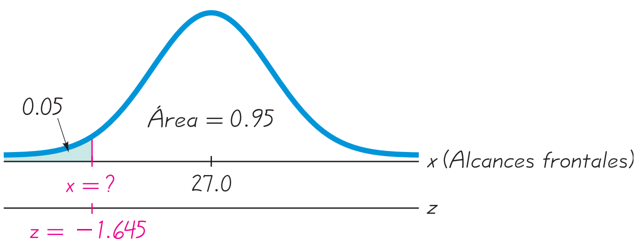


FIGURA 5-16 Cálculo del valor que separa al 95% superior

SOLUCIÓN

Paso 1: Iniciamos con la gráfica de la figura 5-16. 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 un área *izquierda* acumulativa, por lo que restamos 0.95 de 1 para obtener 0.05, que aparece como la región que se sombreó.

Paso 2: En el *cuerpo* de la tabla A-2 buscamos un área de 0.05. Las áreas más cercanas a 0.05 son 0.0505 y 0.0495, pero hay un asterisco que indica que un área de 0.05 corresponde a una puntuación z de -1.645.

Paso 3: Con $z = -1.645$, $\mu = 27.0$ y $\sigma = 1.3$, calculamos x empleando la fórmula 5-2 de manera directa o utilizando la siguiente versión de la fórmula 5-2:

$$x = \mu + (z \cdot \sigma) = 27.0 + (-1.645 \cdot 1.3) = 24.8615$$

Paso 4: Si permitimos que $x = 24.8615$ en la figura 5-16, veremos que tal 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% tienen alcances frontales menores que 24.9 pulgadas.

Utilizando la tecnología

STATDISK Seleccione **Analysis, Probability Distributions, Normal Distribution**; introduzca los valores de la media y de la desviación estándar; después, deslice el mouse a la derecha o a la izquierda hasta obtener el valor que se desea. Puede lograr mayor precisión si emplea el mouse para arrastrar parte de la curva, de manera que pueda amplificarla.

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;** introduzca la media y la desviación estándar; después, haga clic en el botón de **Input Constant,** e introduzca el valor.
- Para encontrar un valor correspondiente a un área que se conoce, seleccione **Calc, Probability Distributions, Normal;** después, **Inverse cumulative probabilities,** e introduzca la media y la desviación estándar. Elija la opción **Input constant** e introduzca 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, introduzca el valor de x, la media y la desviación estándar; finalmente, 1 en el espacio "cumulative".

continúa

- Para encontrar el valor correspondiente a un área que se conoce, seleccione ***fx*, Statistical, NORMINV;** ahora proceda a introducir la información en el cuadro de diálogo. Cuando anote el valor de probabilidad, introduzca el área total a la izquierda del valor dado.

TI-83 Plus

- Para calcular el área entre dos valores, presione **2nd VARS, 2**(para normalcdf); después, proceda a introducir 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 un área que se conoce, presione **2nd VARS, 3** (para invNorm); ahora proceda a introducir 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.

5-3 Destrezas y conceptos básicos

Puntuaciones de CI*. En los ejercicios 1 a 8, suponga que sujetos adultos tienen puntuaciones de CI que se distribuyen normalmente, con una media de 100 y una desviación estándar de 15 (como en la prueba Weschler). (Sugerencia: Dibuje una gráfica en cada caso).*

1. Calcule la probabilidad de que un adulto que se seleccione al azar tenga un CI menor de 115.

2. Calcule la probabilidad de que un adulto que se seleccione al azar tenga un CI mayor de 131.5 (requisito para ser miembro de la organización Mensa).

3. Calcule la probabilidad de que un adulto que se seleccione aleatoriamente tenga un CI entre 90 y 110 (denominado rango *normal*).

4. Calcule la probabilidad de que un adulto seleccionado aleatoriamente tenga un CI entre 110 y 120 (denominado *normal brillante*).

5. Calcule P_{20}, que es la puntuación de CI que separa al 20% inferior del 80% superior.

6. Calcule P_{80}, que es la puntuación de CI que separa al 80% inferior del 20% superior.

7. Calcule la puntuación de CI que separa al 15% superior del resto.

8. Calcule la puntuación de CI que separa al 55% superior del resto.

9. Temperaturas corporales Con base en los resultados muestrales del conjunto de datos 4 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.

a. El hospital Bellevue, en la ciudad de Nueva York, establece que la temperatura más baja que se considera como fiebre es de 100.6°F. ¿Qué porcentaje de personas normales y saludables se consideraría que tiene fiebre? ¿Sugiere este porcentaje que un punto de corte de 100.6°F es apropiado?

b. Los médicos desean seleccionar una temperatura mínima como requisito para aplicar mayor cantidad de exámenes médicos. ¿Qué temperatura debe ser 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).

10. Duración de embarazos La duración de los embarazos se distribuye normalmente, con una media de 268 días y una desviación estándar de 15 días.

a. Un uso clásico de la distribución normal se inspiró en 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, que trabajaba en la Marina. Con esta información, calcule la probabilidad de que un embarazo dure 308 días o más. ¿Qué sugiere el resultado?

b. 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 sería útil para que los administradores de hospitales planeen dichos cuidados.

11. Requisitos de la prueba SAT La combinación de las calificaciones verbales y de matemáticas de mujeres que toman la prueba SAT-I se distribuye de manera normal, con una media de 998 y una desviación estándar de 202 (de acuerdo con datos del College Board). El College of Westport incluye una calificación mínima de 1100 entre sus requisitos.

a. ¿Qué porcentaje de mujeres *no* satisfacen este requisito?

b. Si se cambia el requisito a "una calificación que esté dentro del 40% superior", ¿cuál es la calificación mínima que se requiere? ¿Qué dificultad práctica se crearía si se anunciara que el nuevo requisito es ubicarse en "el 40% superior"?

12. Diseño de cascos Los ingenieros deben tomar en cuenta la anchura de las cabezas de los hombres cuando diseñan cascos para motociclistas. La anchura de las cabezas de los hombres se distribuye normalmente, con una media de 6.0 pulgadas y una desviación estándar de 1.0 pulgadas (según datos de una encuesta antropométrica de Gordon, Churchill *et al*.). Por limitaciones económicas, los cascos se diseñarán para que se ajusten a todos los hombres, excepto al 2.5% con anchuras menores y al 2.5% con anchuras más grandes. Calcule las anchuras de cabeza mínima y máxima que se ajustarán a los cascos.

13. Garantía de televisores El tiempo de reemplazo de los televisores se distribuye normalmente, con una media de 8.2 años y una desviación estándar de 1.1 años (de acuerdo con datos de "Getting Things Fixed", *Consumer Reports*).

a. Calcule la probabilidad de que un televisor que se selecciona aleatoriamente tenga un tiempo de reemplazo menor de 5.0 años.

b. Si usted desea ofrecer una garantía tal que sólo el 1% de los televisores se reemplacen antes de que expire la garantía, ¿cuál debe ser la duración de la garantía?

14. Garantía de reproductores de CD El tiempo de reemplazo de los reproductores de CD se distribuye normalmente, con una media de 7.1 años y una desviación estándar de 1.4 años (con base en datos de "Getting Things Fixed", *Consumer Reports*).

a. Calcule la probabilidad de que un reproductor de CD, que se seleccionó aleatoriamente, tenga un tiempo de reemplazo menor de 8.0 años.

b. Si usted desea ofrecer una garantía tal que sólo el 2% de los reproductores se reemplace antes de que expire la garantía, ¿cuál debe ser la duración de la garantía?

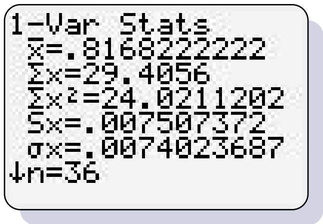
15. M&M A continuación se presentan los resultados de Minitab de los pesos (en gramos) de los 100 dulces M&M listados en el conjunto de datos 19 del Apéndice B. Aun cuando la media y la desviación estándar son estadísticos muestrales, suponga que son los parámetros poblacionales de todos los M&M.

a. Calcule el porcentaje de pesos menores que 8.88925 g. ¿De qué forma coincide el resultado con el valor de 0.88925 presentado como Q_1, el primer cuartil?

b. Calcule el valor de Q_1. ¿De qué forma coincide el resultado con el valor de 0.88925 que se muestra en la representación visual?

```
Variable        N     Media   Mediana Desv. Est.  Desv. Est.  SE Mean
M&M           100   0.91470   0.91050    0.91307     0.03691  0.00369

Variable  Mínimo    Máximo        Q1          Q3
M&M      0.83800   1.03300   0.88925     0.93375
```

TI-83 Plus



16. Pesos de Coca Cola regular Al margen se presenta la representación visual de la calculadora TI-83 Plus con los pesos (en libras) de la Coca Cola regular, tal como se listan en el conjunto de datos 17 del Apéndice B. Aun cuando la media y la desviación estándar son estadísticos muestrales, suponga que son parámetros poblacionales de todas las latas de Coca Cola regular. (Utilice el valor de Sx para la desviación estándar).

a. Si se selecciona aleatoriamente una lata de Coca Cola regular, calcule la probabilidad de que su contenido pese más de 0.8250 libras.

b. Con el propósito de verificar la producción futura de Coca Cola, calcule los pesos que separan al 2.5% inferior y al 2.5% superior.

Estaturas de mujeres. *En los ejercicios 17 a 20, suponga que la estatura de las mujeres se distribuye de forma normal, con una media dada $\mu = 63.6$ pulgadas y una desviación estándar dada $\sigma = 2.5$ pulgadas (de acuerdo con datos del National Health Survey). En cada caso, dibuje una gráfica.*

17. Estatura 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 este requisito?

18. 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 este requisito. ¿Se les negará a muchas mujeres la oportunidad de unirse al ejército porque son muy altas o muy bajas?

19. Estaturas requeridas para las bailarinas *Rockettes* Para estar en una compañía de baile con una apariencia uniforme, las famosas bailarinas *Rockette*, del Radio City Music Hall de Nueva York, deben sujetarse a ciertas restricciones de estaturas. Como las mujeres ahora son más altas, un cambio reciente requiere que una bailarina *Rockette* tenga una estatura entre 66.5 y 71.5 pulgadas. Si se selecciona una mujer al azar, ¿cuál es la probabilidad de que cumpla este nuevo requisito de estatura? ¿Qué porcentaje de mujeres cumplen este nuevo requisito de estatura? ¿Parecería que las *Rockettes* son generalmente más altas que la mujer común?

20. Estaturas requeridas para las bailarinas *Rockettes* El ejercicio 19 identificó requisitos de estatura específicos para las bailarinas *Rockettes*. Suponga que dichos requisitos deben cambiarse porque muy pocas mujeres los cumplen. ¿Cuáles serían las nuevas estaturas mínima y máxima que se permiten si se excluyera al 20% de mujeres más bajas y al 20% de mujeres más altas?

5-3 Más allá de lo básico

21. Unidades de medición Los pesos de las mujeres se distribuyen normalmente, con una media de 143 libras y una desviación estándar de 29 libras.

a. Si los pesos de mujeres individuales se expresan en libras, ¿cuáles serían las unidades que se utilizarían para las puntuaciones *z* correspondientes a los pesos individuales?

b. Si los pesos de todas las mujeres se convierten a puntuaciones *z*, ¿cuál es la media, la desviación estándar y la distribución de estas puntuaciones *z*?

c. ¿Cuál es la distribución, la media y la desviación estándar de los pesos de las mujeres después de convertirlos a kilogramos (1 libra = 0.4536 kg)?

22. Uso de la corrección por continuidad Hay muchas situaciones en las que una distribución normal se utiliza como una buena aproximación de una variable aleatoria con sólo valores *discretos*. En tales casos, podemos emplear esta *corrección por continuidad*: represente cada número entero con el intervalo que va desde 0.5 por debajo del número hasta 0.5 por arriba de él. Suponga que las puntuaciones de CI son todas números enteros con una distribución aproximadamente normal, con una media de 100 y una desviación estándar de 15.

a. Sin utilizar la corrección por continuidad, calcule la probabilidad de seleccionar aleatoriamente a alguien con una puntuación de CI mayor que 105.

b. Utilice la corrección por continuidad y calcule la probabilidad de seleccionar aleatoriamente a alguien con una puntuación de CI mayor que 105.

c. Compare los resultados de los incisos *a* y *b*.

23. Normalización de calificaciones de un examen Una profesora informa a sus alumnos de la clase de psicología que un examen es muy difícil, pero que las calificaciones se normalizarán. Las calificaciones de este examen se distribuyen normalmente, con una media de 25 y una desviación estándar de 5.

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 lugar 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 del inciso *c*? Explique.

24. Calificaciones del SAT Según datos del College Entrance Examination Board, las calificaciones de la prueba SAT-I tienen una media de 1017 y Q_1 es 880, con una distribución aproximadamente normal. Calcule la desviación estándar y después utilice este resultado para obtener P_{99}.

25. 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 el SAT que corresponde al percentil 67, calcule su calificación real en el SAT y su calificación equivalente en el ACT.

b. Si una mujer obtiene una calificación de 1220 en el SAT, calcule su calificación equivalente en el ACT.

5-4 Distribuciones muestrales y estimadores

Comenzamos a embarcarnos en un viaje que nos permitirá conocer las poblaciones al obtener datos de muestras. Las secciones 5-5 y 5-6 proporcionan conceptos importantes que revelan el comportamiento de medias de muestra y proporciones de muestra. Antes de considerar dichos conceptos, observemos el comportamiento de los estadísticos muestrales en general. El principal objetivo de esta sección es aprender lo que conocemos como la *distribución muestral de un estadístico*; otro objetivo importante es aprender un principio básico acerca de la distribución muestral de medias de muestra y la distribución muestral de proporciones de muestra.

Iniciemos con las medias de muestra. En lugar de ser muy abstractos, consideremos la *población* que consiste en los valores 1, 2, 5. El McGwire Electronics Center estuvo abierto sólo durante tres días, por un personal de ventas descortés, a un pobre plan de negocios, publicidad ineficaz y a una ubicación inadecuada. Durante el primer día se vendió un teléfono celular, durante el segundo día dos teléfonos celulares y sólo cinco durante el tercer día. Puesto que 1, 2, 5 constituyen la población completa, es fácil calcular los valores de los parámetros poblacionales: $\mu = 2.7$ y $\sigma = 1.7$. Para calcular el valor de la desviación estándar poblacional σ, empleamos

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} \text{ en lugar de } s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

Es raro que conozcamos todos los valores de una población completa. Lo más común es que exista una población grande desconocida que queremos investigar. Como no es práctico encuestar a cada miembro de la población, obtenemos una muestra; luego, con base en las características de ésta, hacemos estimaciones acerca de las características de la población. Por ejemplo, la Hartford Insurance Company querría conocer la población de las edades de todos los conductores, por medio de la obtención de una muestra de dichas edades.

Ya que los valores 1, 2, 5 constituyen una población completa, consideremos muestras de tamaño 2. Con sólo tres valores poblacionales, hay únicamente nueve posibles muestras diferentes de tamaño 2, suponiendo que el muestreo se realiza con reemplazo. Es decir, se reemplaza cada valor seleccionado antes de realizar una nueva selección.

¿Por qué se hace el muestreo *con* reemplazo? Para muestras pequeñas como la que estudiamos, 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, a lo largo de la presente sección nos interesamos particularmente en el muestreo *con reemplazo,* por las siguientes razones: **1.** Cuando se selecciona una muestra relativamente pequeña, de una población grande, no hay mucha diferencia si realizamos la muestra con reemplazo o sin él. **2.** El muestreo con reemplazo da como resultado sucesos independientes que no se afectan por resultados previos; asimismo, los sucesos independientes son más fáciles de analizar y derivan en fórmulas más simples. Por eso nos enfocamos en el comportamiento de muestras que se seleccionan aleatoriamente *con reemplazo*.

Cuando tomamos una muestra de dos valores con reemplazo, de la población de 1, 2, 5, cada una de las nueve muestras es igualmente posible, con una probabilidad de $1/9$. La tabla 5-2 lista las nueve muestras posibles de tamaño 2, junto con los estadísticos para cada muestra. Esta tabla contiene mucha información, pero consideremos primero la columna de medias de muestra. Puesto que se listan todos los posibles valores de $\bar{x}$,y puesto que se sabe que la probabilidad de cada uno es de $1/9$, tenemos una distribución de probabilidad. (Recuerde, una distribución de probabilidad describe la probabilidad de cada valor de una variable aleatoria, en tanto que la variable aleatoria, en este caso, es el valor de la media muestral). Como varios métodos importantes de estadística inician con una media muestral, que se utiliza subsecuentemente para hacer inferencias acerca de la media poblacional, es importante comprender el comportamiento de tales medias de muestra. Otros métodos de estadística importantes comienzan con una proporción de muestra que se emplea subsecuentemente para hacer inferencias acerca de la proporción

Tabla 5-2 Distribuciones muestrales de diferentes 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	Probabilidad
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	2.7	2.7	1.8	2.9	1.3	0.667	
Parámetro poblacional	2.7	2	4	2.9	1.7	0.667	
¿Coincide el estadístico muestral con el parámetro poblacional?	Sí	No	No	Sí	No	Sí	

poblacional, así que es importante comprender el comportamiento de estas proporciones de muestra. En general, es importante entender el comportamiento de los estadísticos muestrales. El "comportamiento" de un estadístico se puede conocer al comprender su distribución.

Definición

La **distribución muestral de la media** es la distribución de probabilidad de medias muestrales, donde todas las muestras tienen el mismo tamaño *n*. (En general, la distribución de muestreo de cualquier estadístico es la distribución de probabilidad de dicho estadístico).

EJEMPLO Distribución muestral de la media Una *población* se compone de los valores 1, 2, 5, en tanto que la tabla 5-2 incluye todas las distintas muestras posibles del tamaño $n = 2$. La probabilidad de cada muestra se lista en la tabla 5-2 como $1/9$. Identifique la distribución muestral específica de la media de las muestras de tamaño $n = 2$, que se seleccionan aleatoriamente con reemplazo, de la población 1, 2, 5. También, calcule la media de esta distribución muestral. ¿Coinciden las medias de muestra con el valor de la media poblacional?

continúa

Tabla 5-3
Distribución muestral de la media

Media $\bar{x}$	Probabilidad
1.0	1/9
1.5	1/9
3.0	1/9
1.5	1/9
2.0	1/9
3.5	1/9
3.0	1/9
3.5	1/9
5.0	1/9

Esta tabla lista las medias de las muestras en la tabla 5-2, pero podría condensarse listando 1.0, 1.5, 2.0, 3.0, 3.5 y 5.0, junto con sus probabilidades correspondientes de 1/9, 2/9, 1/9, 2/9, 2/9 y 1/9.

SOLUCIÓN La distribución muestral de la media es la distribución de probabilidad que describe la probabilidad para cada valor de la media, y dichos valores se incluyen en la tabla 5-2. Así pues, la distribución muestral de la media se escribe usando la tabla 5-3. La media de la distribución muestral se calcula con dos métodos diferentes: **1.** utilice $\mu = \Sigma[x \cdot P(x)]$, que es la fórmula 4-2, o **2.** puesto que las nueve medias de muestra son igualmente posibles, podríamos simplemente calcular la media de esos nueve valores. Como la media poblacional también es 2.7, parece que las medias de muestra "coinciden" con el valor de la media poblacional, en lugar de subestimar o sobreestimar sistemáticamente la media poblacional.

En el ejemplo anterior, observamos que la media de todas las medias muestrales posibles es igual a la media de la población original, que es $\mu = 2.7$. Podemos generalizar esto como una propiedad de las medias de muestra: para un tamaño de muestras 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, pero antes hagamos una observación evidente, aunque muy importante: *las medias de muestra varían*. Véase la tabla 5-3 y observe que las medias de muestra 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. Tal 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, datos distantes y patrón temporal (resumido con las siglas "CVDDT"). Al examinar las muestras en la tabla 5-2, ya identificamos una propiedad que describe el comportamiento de las medias de muestra: la media de medias de muestra es igual a la media de la población. Esta propiedad pone énfasis en la característica central; investigaremos otras características en la siguiente sección. Veremos que al incrementarse el tamaño de la muestra, la distribución muestral de medias de muestra tiende a convertirse en una *distribución normal*. (Esto no nos sorprende, ya que el título de este capítulo es "Distribuciones de probabilidad normal"). En consecuencia, la distribución normal tiene una importancia que va más allá de las aplicaciones que se ilustran en la sección 5-3. La distribución normal se utilizará en casos en los cuales deseamos emplear una media de muestra con el propósito de hacer alguna inferencia acerca de una media poblacional μ.

Distribución muestral de proporciones

Cuando hacemos inferencias acerca de una proporción poblacional, también es importante comprender el comportamiento de las proporciones muestrales. Definimos la distribución de proporciones muestrales de la siguiente manera.

Definición

Distribución muestral de la proporción: distribución de probabilidad de proporciones muestrales, donde todas las muestras tienen el mismo tamaño muestral *n*.

Uno de los usos clásicos de la estadística inferencial es el cálculo de alguna proporción muestral y su aplicación para hacer una inferencia acerca de la proporción poblacional. Encuestadores de la organización Gallup le preguntaron a 491 adultos que se seleccionaron al azar si estaban a favor de la pena de muerte para una persona que sentenciaron por homicidio. Los resultados mostraron que 319 individuos (o el 65% de ellos) se manifestaron a favor. El resultado muestral conduce a la inferencia de que "el 65% de todos los adultos están a favor de la pena de muerte para una persona sentenciada por homicidio". La proporción muestral de 319/491 se utilizó para estimar una proporción poblacional p, pero aprenderemos mucho más si comprendemos la distribución muestral de dichas proporciones muestrales.

EJEMPLO Distribución muestral de proporciones Una *población* se compone de los valores 1, 2, 5, en tanto que la tabla 5-2 incluye todas las distintas muestras posibles de tamaño $n = 2$, que se seleccionaron con reemplazo. Para cada muestra, considere la proporción de números *impares*. Identifique la distribución muestral para la proporción de números impares y después calcule su media. ¿Coinciden las proporciones muestrales con el valor de la proporción poblacional?

SOLUCIÓN Observe la tabla 5-2, donde las nueve proporciones muestrales que se indican son 1, 0.5, 1, 0.5, 0, 0.5, 1, 0.5, 1. Al combinar tales proporciones muestrales con sus probabilidades de 1/9 en cada caso, obtenemos la distribución muestral de proporciones que se resume en la tabla 5-4. La media de las proporciones *muestrales* es 0.667. Puesto que la población 1, 2, 5 contiene dos números impares, la proporción *poblacional* de números impares es también 2/3 o 0.667. En general, las proporciones muestrales tienden a coincidir con el valor de la proporción poblacional, y no a subestimar o sobreestimar sistemáticamente ese valor.

Tabla 5-4
Distribución muestral de proporciones

Proporción de números impares	Probabilidad
1	1/9
0.5	1/9
1	1/9
0.5	1/9
0	1/9
0.5	1/9
1	1/9
0.5	1/9
1	1/9

La tabla lista las proporciones de las muestras en la tabla 5-2, pero podría condensarse listando las proporciones de 0, 0.5 y 1, junto con sus probabilidades correspondientes de 1/9, 4/9 y 4/9.

El ejemplo anterior incluye una proporción bastante pequeña, de manera que ahora consideraremos los géneros de los senadores en el congreso 107o. Como sólo existen 100 miembros [13 mujeres (M) y 87 hombres (H)], listaremos 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 **M** **M** 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
M 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

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

Tabla 5-5
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 estándar:	0.100

partir de una muestra. Con el propósito de estudiar el comportamiento de las proporciones muestrales, listamos unas cuantas muestras de tamaño $n = 10$:

Muestra 1: H M H H M H H H H H $\rightarrow$ la proporción muestral es 0.2
Muestra 2: H M H H H H H H H H $\rightarrow$ la proporción muestral es 0.1
Muestra 3: H H H H H H M H H H $\rightarrow$ la proporción muestral es 0.1
Muestra 4: H H H H H H H H H H $\rightarrow$ la proporción muestral es 0
Muestra 5: H H H H H H H H M H $\rightarrow$ la proporción muestral es 0.1

Puesto que hay un número muy grande de muestras como éstas, no es posible listarlas todas. El autor seleccionó aleatoriamente 95 muestras adicionales, antes de detener las llantas de su automóvil. Si combinamos las 95 muestras adicionales con las cinco listadas antes, obtendremos las 100 muestras que se incluyen en la tabla 5-5.

Notamos, a partir de la tabla 5-5, que la media de las 100 proporciones muestrales es 0.119, pero si incluyéramos todas las otras posibles muestras de tamaño 10, la media de las proporciones muestrales sería igual a 0.13, que es el valor de la proporción poblacional. La forma de esa distribución se asemeja razonablemente a la que se obtendría con todas las muestras posibles de tamaño 10. Observamos que la distribución que se presenta en la figura 5-17 tiene cierto sesgo hacia la derecha, aunque con un poco de alargamiento se aproximaría a una distribución normal. En la figura 5-18 mostramos los resultados obtenidos a partir de 10,000 muestras, de tamaño 50, que se seleccionaron al azar y con reemplazo, de la lista anterior de 100 géneros. La figura 5-18 sugiere, con énfasis, que la distribución se aproxima a la forma de campana que caracteriza a una distribución normal. Por consiguiente, los resultados de la tabla 5-5 y de la figura 5-18 sugieren lo siguiente.

Propiedades de la distribución de proporciones muestrales

- **Las proporciones muestrales tienden a coincidir con el valor de la proporción poblacional.**
- **En ciertas condiciones, la distribución de proporciones muestrales se aproxima a una distribución normal.**

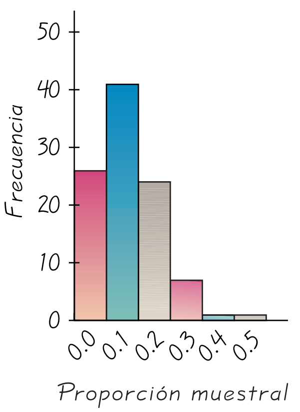


FIGURA 5-17 **100 proporciones muestrales con $n = 10$ en cada muestra**

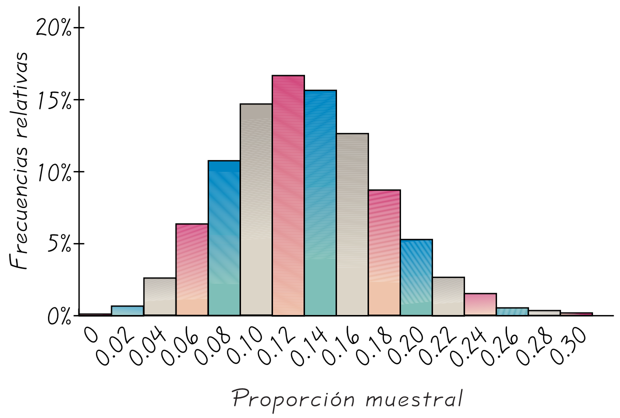


FIGURA 5-18 **10,000 proporciones muestrales con $n = 50$ en cada muestra**

¿Cuáles estadísticos son buenos estimadores de parámetros?

En el capítulo 6 estudiaremos métodos formales para el uso de estadísticos muestrales que nos permitirán hacer estimaciones de los valores de parámetros de población. Algunos estadísticos funcionan mucho mejor que otros, por lo cual es posible juzgar su valor examinando sus distribuciones muestrales, como en el siguiente ejemplo.

EJEMPLO **Distribuciones muestrales** Una *población* se compone de los valores 1, 2, 5. Si seleccionamos aleatoriamente muestras de tamaño 2 con reemplazo, hay nueve distintas muestras posibles, que se listan en la tabla 5-2. Como las nueve muestras distintas son igualmente posibles, cada muestra tiene una probabilidad de $1/9$.

a. Para cada muestra calcule la media, mediana, rango, varianza, 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).

b. Para cada estadístico, calcule la media de los resultados del inciso *a*.

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 de muestra tienden a centrarse alrededor del valor de la media poblacional, que es $8/3 = 2.7$, de manera que la media de muestra coincide con el valor de la media poblacional.

SOLUCIÓN

a. Véase la tabla 5-2, que incluye los estadísticos individuales para cada muestra.

b. Las medias de los estadísticos muestrales aparecen casi al final de la tabla 5-2. La media de las medias de muestra es 2.7, la media de las medianas de muestra es 2.7, y así sucesivamente.

c. El renglón inferior de la tabla 5-2 se basa en una comparación de los parámetros poblacionales y los resultados de los estadísticos muestrales. Por ejemplo, la media poblacional de 1, 2, 5 es $\mu = 2.7$, y las medias de muestra “coinciden” con el valor de 2.7, ya que la media de las medias de muestra también es 2.7.

INTERPRETACIÓN Con base en los resultados de la tabla 5-2, observamos que cuando se utiliza un estadístico muestral para estimar un parámetro de población, algunos estadísticos son buenos porque 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

continúa

Aunque la desviación estándar muestral no coincide con la desviación estándar poblacional, el sesgo es relativamente pequeño en muestras grandes, de manera que con frecuencia utiliza s para estimar σ. Por consiguiente, a medias, proporciones, varianza y desviaciones estándar se les considerará temas importantes en los siguientes capítulos, pero la mediana y el rango se utilizarán en pocas ocasiones.

El objetivo más importante de esta sección es introducir 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 muy grande, no es práctico medir la temperatura de cada adulto. En su lugar, obtenemos una muestra de temperaturas corporales y la utilizamos para estimar la media poblacional. El conjunto de datos 4 del Apéndice B incluye una muestra de 106 temperaturas corporales; la media de esta muestra es $\bar{x} = 98.20°F$. Las conclusiones que hagamos acerca de la temperatura media poblacional de todos los adultos requiere que comprendamos el comportamiento de la distribución muestral de todas las medias de muestra de este tipo. Aunque no es práctico obtener cada muestra posible y nos conformamos con una sola muestra, es posible obtener algunas conclusiones muy importantes y con significado acerca de la población de todas las temperaturas corporales. Uno de los objetivos principales de las secciones y los capítulos que siguen es aprender el uso eficaz de una muestra para sacar conclusiones acerca de una población. En la sección 5-5 tendremos en cuenta más detalles acerca de la distribución muestral de medias de muestra, en tanto en la sección 5-6 estudiaremos más detalles acerca de la distribución muestral de las proporciones de muestra.

5-4 Destrezas y conceptos básicos

1. 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?

2. Distribución muestral de Harry Potter El conjunto de datos 14 del Apéndice B incluye una muestra de los niveles de lectura medidos para 12 páginas que se seleccionaron aleatoriamente del libro *Harry Potter y la piedra filosofal*, de J. K. Rowling. La media de los 12 valores del nivel de Flesch-Kincaid es 5.08. El valor de 5.08 forma parte de una distribución muestral. Describa esta distribución muestral.

3. Distribución muestral de temperaturas corporales El conjunto de datos 4 del Apéndice B incluye una muestra de 106 temperaturas corporales de adultos. Si construyésemos un histograma para describir la forma de la distribución de dicha muestra, ¿mostraría el histograma la forma de una *distribución muestral* de medias de muestra? ¿Por qué?

4. Distribución muestral de resultados de encuesta La organización Gallup realizó una encuesta a 1015 estudiantes que se seleccionaron al azar, desde jardín de niños hasta preparatoria, y encontró que el 10% acudía a escuelas privadas o religiosas.
a. ¿Será el resultado del 10% (o 0.10) un estadístico o un parámetro? Explique.
b. ¿Cuál es la distribución muestral sugerida por los datos?
c. ¿Tendría más confianza en los resultados si el tamaño de muestra hubiese sido de 2000 en lugar de 1025? ¿Por qué?

5. Centro telefónico La Nome Ice Company abrió únicamente durante tres días (adivine por qué). He aquí el número de llamadas que se recibieron durante cada uno de esos días: 10, 6, 5. Suponga que se seleccionan aleatoriamente muestras de tamaño 2, *con reemplazo*, de esta población de tres valores.

a. Liste las nueve muestras diferentes posibles y calcule la media de cada una de ellas.

b. Identifique la probabilidad de cada muestra y describa la distribución muestral de las medias de muestra. (*Sugerencia*: Véase la tabla 5-3).

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 que se listan? ¿Serán estas medias *siempre* iguales?

6. Telemercadeo A continuación se presenta el número de ventas por día de Kim Ryan, un cortés vendedor que trabajó durante cuatro días antes de que lo despidieran: 1, 11, 9, 3. Suponga que se seleccionan al azar muestras de tamaño 2, *con reemplazo*, de esta población de cuatro valores.

a. Liste las 16 diferentes muestras posibles y calcule la media de cada una de ellas.

b. Identifique la probabilidad de cada muestra y después describa la distribución muestral de medias de muestra. (*Sugerencia*: Véase la tabla 5-3).

c. Calcule la media de la distribución muestral.

d. ¿Es la media de la distribución muestral (del inciso *c*) igual a la media de la población de los cuatro valores que se listan? ¿Serán estas medias *siempre* iguales?

7. Estaturas de los Lakers de L.A. A continuación se presentan las estaturas (en pulgadas) de cinco jugadores de basquetbol estelares de los Lakers de Los Ángeles: 85, 79, 82, 73, 78. Suponga que se seleccionan al azar muestras de tamaño 2, *con reemplazo*, de la población de cinco estaturas.

a. Después de identificar las 25 distintas muestras posibles, calcule la media de cada una de ellas.

b. Describa la distribución muestral de medias. (*Sugerencia:* Véase la tabla 5-2).

c. Calcule la media de la distribución muestral.

d. ¿Es la media de la distribución muestral (del inciso *c*) igual a la media de la población de las cinco estaturas que se listan? ¿Serán estas medias *siempre* iguales?

8. 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.

a. Después de identificar las 25 distintas muestras posibles, calcule la media de cada una de ellas.

b. Describa la distribución muestral de medias. (*Sugerencia:* Véase la tabla 5-2).

c. Calcule la media de la distribución muestral.

d. ¿Es la media de la distribución muestral (del inciso *c*) igual a la media de la población de las cinco edades? ¿Serán estas medias *siempre* iguales?

9. Genética Un experimento en genética incluye una población de moscas de la fruta consistente en un macho que llamaron Mike y tres hembras cuyos nombres son Anna, Bárbara y Chris. Suponga que se seleccionan dos moscas de la fruta al azar, *con reemplazo*.

a. Después de identificar las 16 diferentes muestras posibles, calcule la proporción de hembras en cada una de ellas.

b. Describa la distribución muestral de proporciones de hembras. (*Sugerencia*: Véase la tabla 5-2).

c. Calcule la media de la distribución muestral.

d. ¿Es la media de la distribución muestral (del inciso *c*) igual a la proporción poblacional de hembras? ¿Coincide *siempre* la media de la distribución muestral de proporciones con la proporción poblacional?

10. Control de calidad Después de construir una máquina nueva, se producen cinco faros prototipo para automóvil y se descubre que dos están defectuosos (D) y tres son aceptables (A). Suponga que se seleccionan dos faros al azar, *con reemplazo*, de esta población.

a. Después de identificar las 25 distintas muestras posibles, calcule la proporción de defectos en cada una de ellas.

b. Describa la distribución muestral de proporciones de defectos. (*Sugerencia*: Véase la tabla 5-2).

c. Calcule la media de la distribución muestral.

d. ¿Es la media de la distribución muestral (del inciso *c*) igual a la proporción poblacional de defectos? ¿Coincide *siempre* la media de la distribución muestral de proporciones con la proporción poblacional?

11. Senadoras Permitamos que una población se conforme con 10 senadoras demócratas y tres republicanas en el Congreso 107o de Estados Unidos.

a. Desarrolle un procedimiento para seleccionar aleatoriamente (con reemplazo) una muestra de tamaño 5 de la población de 10 demócratas y tres republicanas; después, seleccione una muestra de éstas y liste los resultados.

b. Calcule la proporción de demócratas en la muestra del inciso *a*.

c. ¿Es la proporción del inciso *b* un estadístico o un parámetro?

d. ¿Es la proporción muestral del inciso *b* igual a la proporción poblacional de demócratas? ¿Podrá cualquier muestra aleatoria de tamaño 5 resultar en una proporción muestral que iguale la proporción poblacional?

e. Suponga que se listan todas las distintas muestras posibles de tamaño 5 y que se calcula la proporción muestral de cada una de ellas. ¿Qué se concluye acerca del valor de la media de estas proporciones muestrales?

12. Senadoras Permitamos que una población conste de los siguientes estados de residencia de las tres senadoras republicanas del Congreso 107o de Estados Unidos: Maine, Maine, Texas. Suponiendo que las muestras de tamaño 2 se seleccionan aleatoriamente de esta población, *sin* reemplazo, liste las distintas muestras posibles. Calcule la probabilidad de cada muestra. Además, para cada muestra calcule la proporción de senadoras de Maine. Por ejemplo, la muestra de "Maine y Texas" resulta en una proporción muestral de $1/2$ (porque una de las dos senadoras es de Maine). Calcule la media de la distribución muestral y verifique que sea igual a la proporción poblacional de senadoras de Maine.

5-4 Más allá de lo básico

13. 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 todas las muestras se seleccionan *sin reemplazo*.

a. Después de listar todas las muestras posibles de tamaño $n = 2$, calcule la media y la desviación estándar de las medias de muestra.

b. Después de listar todas las muestras posibles de tamaño $n = 3$, calcule la media y la desviación estándar de las medias de muestra.

c. Después de listar todas las muestras posibles de tamaño $n = 4$, calcule la media y la desviación estándar de las medias de muestra.

d. Cuando se realiza muestreo sin reemplazo, ¿tienden las medias de muestra a coincidir con el valor de la media poblacional?

e. Con base en los resultados anteriores, ¿de qué manera se afecta la *variación* de la distribución muestral de medias de muestras al incrementar el tamaño de la muestra?

14. Desviación media absoluta La población de 1, 2, 5 se utilizó para elaborar la tabla 5-2. Identifique la distribución muestral de la desviación media absoluta (que se definió en la sección 2-5); después, determine si la desviación media absoluta de una muestra es un buen estadístico para estimar la desviación media absoluta de la población.

15. La mediana como estimador En la tabla 5-2 la distribución muestral de las medianas tiene una media de 2.7. Puesto que la media poblacional también es de 2.7, parecería que la mediana es un buen estadístico para estimar el valor de la media poblacional. Utilice los valores poblacionales 1, 2, 5, y calcule las 27 muestras de tamaño $n = 3$ que es posible seleccionar sin reemplazo; después, calcule la mediana y la media de cada una de las 27 muestras. Una vez con estos resultados, calcule la media de la distribución muestral de la mediana y la media de la distribución muestral de la media. Compare los resultados con la media poblacional de 2.7. ¿Qué concluye?

5-5 Teorema del límite central

Esta sección es *sumamente* importante porque presenta el teorema del límite central, que establece los fundamentos para estimar parámetros poblacionales y pruebas de hipótesis (temas que se estudian con profundidad en los siguientes capítulos). Mientras estudia esta sección, trate de evitar la confusión que causa el hecho de que el teorema del límite central implica dos distribuciones diferentes: la distribución de la población original y la distribución de las medias de muestra.

Los siguientes términos y conceptos clave se presentaron en secciones anteriores:

- Una *variable aleatoria* es una variable que tiene un solo valor numérico, determinado por el azar, para cada resultado de un procedimiento (sección 4-2).
- Una *distribución de probabilidad* es una gráfica, tabla o fórmula que da la probabilidad para cada valor de una variable aleatoria (sección 4-2).
- La *distribución muestral de la media* es la distribución de probabilidad de medias de muestra, donde todas las muestras tienen el mismo tamaño n (sección 5-4).

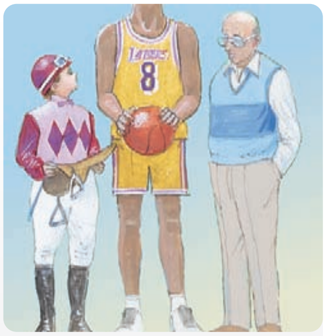
Véase el siguiente ejemplo, a modo de ilustración de estos conceptos abstractos.

EJEMPLO Dígitos aleatorios Considere la población de los dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, los cuales se seleccionan aleatoriamente con reemplazo.

a. *Variable aleatoria:* Si realizamos ensayos que consisten en la selección aleatoria de un solo dígito, y si representamos el valor del dígito que seleccionamos con x, entonces x es una variable aleatoria (porque su valor depende del azar).

b. *Distribución de probabilidad:* Suponiendo que los dígitos se seleccionan aleatoriamente, la probabilidad de cada dígito es $1/10$, que puede expresarse con la fórmula $P(x) = 1/10$. Ésta es una distribución de probabilidad (ya que describe la probabilidad de cada valor de la variable aleatoria x).

c. *Distribución muestral:* Ahora suponga que seleccionamos aleatoriamente todas las distintas muestras posibles de tamaño $n = 4$. (Recuerde que estamos haciendo muestreo con reemplazo, de modo que cualquier muestra particular tendría el mismo dígito más de una vez). En cada muestra calculamos la media de muestra $\bar{x}$ (que en sí misma es una variable aleatoria porque su valor depende del azar). La distribución de probabilidad de las medias muestrales es una distribución muestral.



El teorema del límite central borroso

En *The Cartoon Guide to Statistics*, de Gonick y Smith, los autores 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: fluctuaciones del mercado de las acciones, pesos de estudiantes, promedios anuales de temperatura, 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.

El inciso *c* del ejemplo ilustra una distribución muestral específica de medias de muestra. En la sección 5-4 observamos que la media de medias de muestra es igual a la media de la población y que si, el tamaño de la muestra aumenta, las medias de muestra correspondientes tienden a variar menos. El *teorema del límite central* nos indica que si el tamaño de la muestra es lo suficientemente grande, la distribución de medias de muestra puede aproximarse a una *distribución normal*, aun si la población original no se distribuye normalmente. Aunque hablamos de un "teorema", no incluimos pruebas rigurosas; en su lugar, nos enfocamos en los *conceptos* y la forma de aplicarlos. A continuación se presentan los puntos clave, que conforman un fundamento importante para los siguientes capítulos.

El teorema del límite central y la distribución muestral de $\bar{x}$

Puesto 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 modo que todas las muestras posibles de tamaño n tengan la misma posibilidad de seleccionarse).

Conclusiones:

1. La distribución de las medias de muestra $\bar{x}$ se aproximará a una distribución *normal*, conforme el tamaño de la muestra aumente.
2. La media de todas las medias de muestra 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 de muestra es $\sigma/\sqrt{n}$. (Es decir, la distribución normal de la conclusión 1 tiene una desviación estándar $\sigma/\sqrt{n}$).

Reglas prácticas de uso común

1. Si la población original no se distribuye normalmente, la siguiente es una guía común: para muestras de tamaño n mayores que 30, la distribución de las medias de muestra puede aproximarse razonablemente bien a una distribución normal. (Hay excepciones, como las poblaciones con distribuciones muy diferentes a la normal, que requieren tamaños de muestra mucho más grandes que 30, aunque éstas son relativamente raras). La aproximación mejora conforme el tamaño muestral n se incrementa.
2. Si la población original se distribuye normalmente, entonces las medias de muestra se distribuirán 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 de muestra. 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 medias de muestra.

Notación para la distribución muestral de $\bar{x}$

Si se seleccionan todas las muestras aleatorias de tamaño n de una población con media μ y desviación estándar σ, la media de las medias de muestra se denota con $\mu_{\bar{x}}$, de modo que

$$\mu_{\bar{x}} = \mu$$

También la desviación estándar de las medias de muestra 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 Dígitos aleatorios Nuevamente considere la población de dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, que se seleccionan aleatoriamente con reemplazo. Suponga que seleccionamos al azar muestras de tamaño $n = 4$. En la población original de dígitos, todos los valores son igualmente probables. Con base en las "Reglas prácticas de uso común" (que se incluyen en el recuadro del teorema del límite central), *no podemos* concluir que las medias de muestra están distribuidas normalmente, ya que la población original no tiene una distribución normal y el tamaño de muestra 4 no es mayor que 30. Sin embargo, exploraremos la distribución muestral para ver qué aprendemos de ella.

La tabla 5-6 se elaboró registrando los cuatro últimos dígitos de los números de seguridad social de cada uno de 50 estudiantes. Los últimos cuatro dígitos de los números del seguro social son aleatorios, a diferencia de los dígitos iniciales, que se utilizan para codificar información particular. Si combinamos los cuatro dígitos de cada estudiante en un conjunto grande de 200 números, obtendremos una media $\bar{x} = 4.5$, una desviación estándar $s = 2.8$, y una distribución como la gráfica que se presenta en la tabla 5-19. Ahora note qué ocurre cuando calculamos las 50 medias de las muestras, como se observa en la tabla 5-6. (Por ejemplo, el primer estudiante tiene los dígitos 1, 8, 6 y 4, en tanto que la media de los cuatro dígitos es 4.75). Aun cuando el conjunto original de datos *no* tiene una distribución normal, las medias de muestra presentan una distribución aproximadamente *normal*. Esto tal vez resulte confuso, por lo que conviene detenerse aquí y estudiar el párrafo hasta que su idea central quede clara: el conjunto original de los 200 números individuales *no* tiene una distribución normal (porque los dígitos 0-9 ocurren con frecuencias aproximadamente iguales), pero las 50 medias de muestra *sí* tienen una distribución aproximadamente normal. (Una de las "Reglas prácticas de uso común" establece que las muestras con $n > 30$ pueden aproximarse a una distribución normal, pero las muestras pequeñas, como $n = 4$ de este ejemplo, en ocasiones tienen distribuciones aproximadamente normales). El hecho de que al hacer un muestreo de una distribución lleguemos a crear una distribución de medias de muestra que sea normal, o al menos aproximadamente normal, es un fenómeno verdaderamente fascinante e intrigante de la estadística.

Tabla 5-6

SSN dígitos				$\bar{x}$
1	8	6	4	4.75
5	3	3	6	4.25
9	8	8	8	8.25
5	1	2	5	3.25
9	3	3	5	5.00
4	2	6	2	3.50
7	7	1	6	5.25
9	1	5	4	4.75
5	3	3	9	5.00
7	8	4	1	5.00
0	5	6	1	3.00
9	8	2	2	5.25
6	1	5	7	4.75
8	1	3	0	3.00
5	9	6	9	7.25
6	2	3	4	3.75
7	4	0	7	4.50
5	7	5	6	5.75
4	1	5	7	4.25
1	2	0	6	2.25
4	0	2	8	3.50
3	1	2	5	2.75
0	3	4	0	1.75
1	5	1	0	1.75
9	7	4	0	5.00
7	3	1	1	3.00
9	1	1	3	3.50
8	6	5	9	7.00
5	6	4	1	4.00
9	3	9	5	6.50
6	0	7	3	4.00
8	2	9	6	6.25
0	2	8	6	4.00
2	0	9	7	4.50
5	8	9	0	5.50
6	5	4	9	6.00
4	8	7	6	6.25
7	1	2	0	2.50
2	9	5	0	4.00
8	3	2	2	3.75
2	7	1	6	4.00
6	7	7	1	5.25
2	3	3	9	4.25
2	4	7	5	4.50
5	4	3	7	4.75
0	4	3	8	3.75
2	5	8	6	5.25
7	1	3	4	3.75
8	3	7	0	4.50
5	6	6	7	6.00

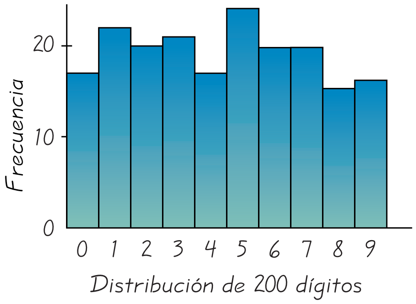


FIGURA 5-19 Distribución de 200 dígitos de los números del seguro social

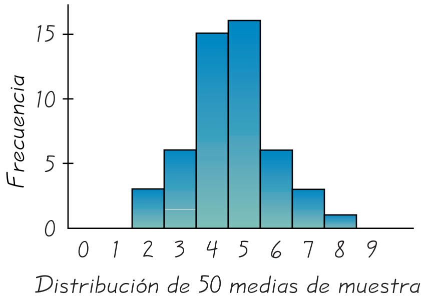


FIGURA 5-20 Distribución de 50 medias de muestra

En la figura 5-20 vemos que la distribución de las medias de muestra del ejemplo anterior es aproximadamente normal, a pesar de que la población original no tiene una distribución normal y de que el tamaño de $n = 4$ de las muestras individuales no excede de 30. Si examina con atención la figura 5-20 verá que no se trata de una distribución normal exacta, pero se acercaría a una distribución normal exacta conforme el tamaño de la muestra se incrementa más allá de 4.

Conforme aumenta el tamaño de la muestra, la distribución muestral de medias de muestra se aproxima a una distribución *normal*.

Aplicación del teorema del límite central

Es posible resolver muchos problemas prácticos importantes con el teorema del límite central. Cuando trabaje en 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 de muestra como si fuera una distribución normal con media μ y desviación estándar $\sigma/\sqrt{n}$.

En el siguiente ejemplo, el inciso *a*) incluye un valor *individual*, pero el inciso *b* incluye la media de una *muestra* de 36 mujeres, por lo cual usaremos 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 que se utilizaron 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 que se distribuye normalmente, utilice los métodos de la sección 5-3. Use $z = \dfrac{x - \mu}{\sigma}$.**
- **Cuando trabaje con una media de alguna *muestra* (o grupo), asegúrese de utilizar el valor de $\sigma/\sqrt{n}$ para la desviación estándar de las medias de muestra. Use $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$.**

PROBLEMA 5 DEL CAPÍTULO

EJEMPLO Seguridad del teleférico En el problema del capítulo señalamos que un teleférico en Vail, Colorado, lleva a los esquiadores a la cima de la montaña. Hay una placa que indica que su capacidad máxima es de 12 personas o 2004 libras. Dicha capacidad se excedería si 12 personas tienen pesos con una media mayor que $2004/12 = 167$ libras.

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 libras y una desviación estándar de 29 libras (según datos del National Health Survey).

a. Calcule la probabilidad de que, al seleccionar aleatoriamente a un hombre, su peso sea mayor de 167 libras.

b. Calcule la probabilidad de que 12 hombres que se seleccionaron al azar tengan una media mayor de 167 libras (de manera que su peso total sea mayor que la máxima capacidad del teleférico de 2004 libras).

SOLUCIÓN

a. *Aproximación: Utilice los métodos presentados en la sección 5-3* (ya que estamos trabajando con un valor *individual* de una población distribuida normalmente). Buscamos el área de la región que se sombreó de la figura 5-21a. Antes de emplear la tabla A-2, convertimos el peso de 167 a su puntuación z correspondiente:

$$z = \frac{x - \mu}{\sigma} = \frac{167 - 172}{29} = -0.17$$

Ahora nos remitimos a la tabla A-2, usando $z = -0.17$; encontramos que el área acumulativa a la izquierda de 167 libras es 0.4325. La región que se sombreó es, por tanto, $1 - 0.4325 = 0.5675$. La probabilidad de que un hombre que se selecciona aleatoriamente pese más de 167 libras es de 0.5675.

b. *Aproximación: Utilice el teorema del límite central* (ya que estamos trabajando con la *media de una muestra* de 12 hombres, no de un solo hombre). Aun cuando el tamaño de la muestra no es mayor de 30, empleamos 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 ahora estamos trabajando con una distribución de medias de muestra, debemos emplear los parámetros $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$, que se evalúan de la siguiente manera:

$$\mu_{\bar{x}} = \mu = 172$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{29}{\sqrt{12}} = 8.37158$$

continúa

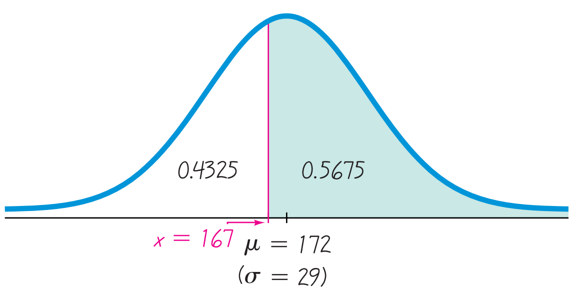


FIGURA 5-21(a) **Distribución de pesos individuales de hombres**

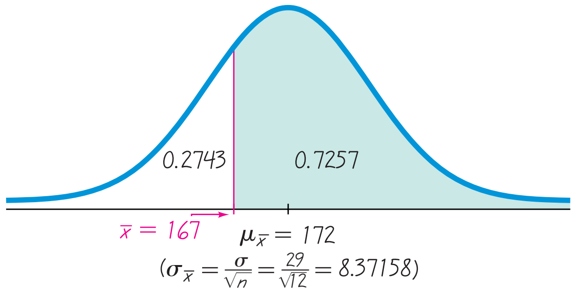


FIGURA 5-21(b) Medias de muestras de los pesos de 12 hombres

El siguiente es un punto importante: debemos usar la desviación estándar que se calculó de 8.37158, no la desviación estándar original de 29 (porque estamos trabajando con la distribución de medias de muestra, con una desviación estándar de 8.37158 y no con la distribución de pesos individuales, cuya desviación estándar es de 29). Deseamos encontrar el área que se sombreó de la figura 5-21b. En la tabla A-2 obtenemos la puntuación z relevante, que se calcula de la siguiente manera:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{167 - 172}{\frac{29}{\sqrt{12}}} = \frac{-5}{8.37158} = -0.60$$

Si nos remitimos a la tabla A-2, encontramos que $z = -0.60$ corresponde a un área izquierda acumulativa de 0.2743, por lo que la región que se sombreó es $1 - 0.2743 = 0.7257$. La probabilidad de que los 12 hombres tengan un peso medio mayor que 167 libras es de 0.7257.

INTERPRETACIÓN Hay una probabilidad de 0.5675 de que un hombre pese más de 167 libras y otra de 0.7257 de que 12 hombres tengan un peso medio mayor que 167 libras. Puesto que la capacidad máxima del teleférico es de 2004 libras, es probable (con una probabilidad de 0.7257) que se sobrecargue si se llena con 12 hombres que se seleccionaron al azar. Sin embargo, la seguridad de los pasajeros no es tan mala, por factores como: **1.** es probable que los hombres esquiadores tengan un peso medio menor que la media de 172 libras de la población general de hombres; **2.** es probable que también suban mujeres esquiadoras, y tienden a pesar menos que los hombres; **3.** a pesar de que la capacidad máxima que se señaló es de 2004 libras, el teleférico se diseñó para operar con seguridad con pesos muy por encima de la carga conservadora de 2004 libras. No obstante, los operadores del teleférico harían bien en evitar una carga de 12 hombres, especialmente si parecen ser muy pesados. Los cálculos utilizados aquí son exactamente los mismos que emplean los ingenieros cuando diseñan teleféricos, elevadores, escaleras eléctricas, aviones y otros aparatos que cargan personas.

Interpretación de resultados

El siguiente ejemplo ilustra otra aplicación del teorema del límite central, pero examine con atención la conclusión a la que se llega. Este ejemplo muestra el tipo de pensamiento que es fundamental para el importante procedimiento de prueba de hipótesis (que se estudia en el capítulo 7). Este ejemplo ilustra la regla del suceso poco común de la estadística inferencial, que se presentó inicialmente en la sección 3-1.

Regla del suceso poco común

Si, bajo cierto supuesto, la probabilidad de un suceso particular que se observa es excepcionalmente pequeña, 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 que la desviación estándar de la población es 0.62°F (de acuerdo con 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; observe las temperaturas de medianoche del día 2 en el conjunto de datos 4 del Apéndice B).

SOLUCIÓN No se nos dio la distribución de la población pero, porque el tamaño de la muestra $n = 106$ excede a 30, utilizamos el teorema del límite central y concluimos que la distribución de medias de muestra es normal, con estos parámetros:

$$\mu_{\bar{x}} = \mu = 98.6 \qquad \text{(por supuesto)}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.62}{\sqrt{106}} = 0.0602197$$

La figura 5-22 muestra el área que se sombreó (observe la pequeña cola izquierda de la gráfica), correspondiente a la probabilidad que buscamos. Con los parámetros que se aplican a la distribución de la figura 5-22, hallaremos el área que se sombreó utilizando los mismos procedimientos desarrollados en la sección 5-3. En la tabla A-2 primero encontramos la puntuación z:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{98.20 - 98.6}{0.0602197} = -6.64$$

continúa

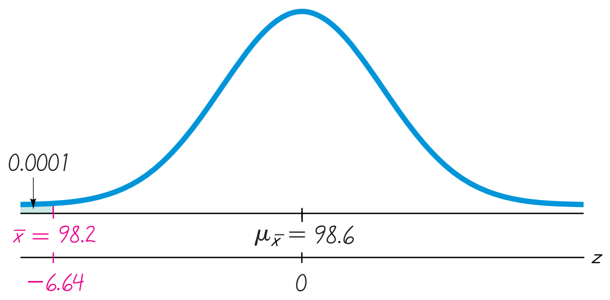


FIGURA 5-22 **Distribución de temperaturas corporales medias, para muestras de tamaño $n = 106$**

Si nos remitimos a la tabla A-2, encontraremos que $z = -6.64$ no aparece, pero para los valores de z que están por debajo de -3.49 utilizamos un área de 0.0001 para el área izquierda acumulativa hasta $z = -3.49$. Por lo tanto, concluimos que la región que se sombreó de la figura 5.22 es 0.0001. (Las tablas más precisas de los resultados de la calculadora TI-83 Plus indican que el área de la región que se sombreó es cercana a 0.00000000002, pero incluso tales resultados son sólo aproximaciones. Con seguridad reportaríamos que la probabilidad es muy baja, menor que 0.001).

INTERPRETACIÓN Los resultados demuestran que si la media de nuestra temperatura corporal es en realidad 98.6°F, entonces hay una probabilidad sumamente baja de obtener una media de muestra de 98.2°F o menor, cuando se seleccionan 106 sujetos aleatoriamente. Los investigadores de la Universidad de Maryland obtuvieron una media muestral como ésta, y existen dos explicaciones posibles: o la media de la población es en realidad de 98.6°F y su muestra representa un suceso aleatorio extremadamente poco común, o 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 la *prueba de hipótesis*, que se estudiará en el capítulo 7. Por ahora, habrá que enfocarnos en el uso del teorema del límite central para calcular la probabilidad de 0.0001, pero debemos tomar en cuenta que dicho teorema se utilizará 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, que se regresa cada elemento que se eligió antes de hacer la siguiente selección), la población es efectivamente infinita. Aunque muchas aplicaciones realistas implican un muestreo sin reemplazo, tales muestras sucesivas dependen de resultados previos. En la fabricación, los inspectores de control de calidad suelen muestrear elementos de una racha finita 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 de muestra $\sigma_{\bar{x}}$, multiplicándola por el *factor de corrección de población finita:*

$$\sqrt{\frac{N-n}{N-1}}$$

Con excepción de los ejercicios 21 y 22, los ejemplos y los ejercicios de esta sección suponen que el factor de corrección de población finita *no* se aplica, porque estamos tomando una muestra con reemplazo, porque la población es infinita o porque el tamaño de 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 6 aplicaremos el teorema cuando utilicemos datos muestrales para estimar medias de poblaciones. En el capítulo 7 lo aplicaremos cuando usemos datos muestrales para probar aseveraciones hechas sobre medias poblacionales. Dichas aplicaciones para estimar parámetros de población y probar aseveraciones, constituyen usos sumamente importantes de la estadística, en tanto que el teorema del límite central los hace posibles.

5-5 Destrezas y conceptos básicos

Uso del teorema del límite central. *En los ejercicios 1 a 6, suponga que los pesos de los hombres se distribuyen de manera normal, con una media dada por μ = 172 libras y una desviación estándar dada por σ = 29 libras (según datos del National Health Survey).*

1. a. Si se selecciona un hombre al azar, calcule la probabilidad de que pese menos de 167 libras.
b. Si se seleccionan 36 hombres al azar, calcule la probabilidad de que tengan un peso medio menor de 167 libras.

2. a. Si se selecciona un hombre al azar, calcule la probabilidad de que su peso sea mayor de 180 libras.
b. Si se seleccionan 100 hombres al azar, calcule la probabilidad de que tengan un peso medio mayor de 180 libras.

3. a. Si se selecciona un hombre al azar, calcule la probabilidad de que su peso se ubique entre 170 y 175 libras.
b. Si se seleccionan 64 hombres al azar, calcule la probabilidad de que tengan un peso medio entre 170 y 175 libras.

4. a. Si se selecciona un hombre al azar, calcule la probabilidad de que pese entre 100 y 165 libras.
b. Si se seleccionan 81 hombres al azar, calcule la probabilidad de que tengan un peso medio entre 100 y 165 libras.

5. a. Si se seleccionan 25 hombres al azar, calcule la probabilidad de que tengan un peso medio mayor de 160 libras.
b. ¿Por qué puede usarse el teorema del límite central en el inciso *a*, a pesar de que el tamaño de muestra no excede a 30?

6. a. Si se seleccionan cuatro hombres al azar, calcule la probabilidad de que tengan un peso medio entre 160 y 180 libras.
b. ¿Por qué puede usarse el teorema del límite central en el inciso *a*, a pesar de que el tamaño de muestra no excede a 30?

7. Rediseño de asientos de expulsión En el problema del capítulo se señaló que los ingenieros estaban rediseñando los asientos de expulsión de aviones de combate para que se ajustaran mejor a las mujeres. Antes de que las mujeres se convirtieran en pilotos de aviones de combate, los asientos expulsores ACES-II se diseñaron para hombres que pesaran entre 140 libras y 211 libras. La población de mujeres tiene pesos distribuidos normalmente, con una media de 143 libras y una desviación estándar de 29 libras (de acuerdo con datos del National Health Survey).
a. Si se selecciona una mujer al azar, calcule la probabilidad de que pese entre 140 libras y 211 libras.
b. Si se seleccionan 36 mujeres diferentes al azar, calcule la probabilidad de que su peso medio se ubique entre 140 y 211 libras.
c. 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?

8. Diseño de cascos para motociclistas Los ingenieros deben tomar en cuenta la anchura de las cabezas de los hombres cuando diseñan cascos para motociclistas. Las anchuras de las cabezas de los hombres se distribuyen normalmente, con una media de 6.0 pulgadas y una desviación estándar de 1.0 pulgada (según datos de una encuesta antropométrica de Gordon, Churchill *et al*.).

a. Si se selecciona un hombre al azar, calcule la probabilidad de que el ancho de su cabeza sea menor que 6.2 pulgadas.

b. La compañía Safeguard Helmet planea una racha de producción inicial de 100 cascos. Calcule la probabilidad de que 100 hombres, que se seleccionaron al azar, tengan una anchura media de cabeza menor que 6.2 pulgadas.

c. 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 pulgadas, porque se ajustarían a casi todos los hombres. ¿Por qué es incorrecto este razonamiento?

9. Diseño de montaña rusa El Rock'n'Roller Coaster de los estudios MGM de Disney, en Orlando, tiene dos asientos en cada fila. Al diseñar esa montaña rusa, debía determinarse la anchura total de los dos asientos de cada fila. En el "peor de los casos", ambos asientos los ocupan hombres. Los hombres tienen anchuras de cadera distribuidas normalmente, con una media de 14.4 pulgadas y una desviación estándar de 1.0 pulgadas (según datos de una encuesta antropométrica de Gordon, Churchill *et al*.). Suponga que se seleccionan dos hombres al azar.

a. Calcule la probabilidad de que su anchura media de cadera sea mayor que 16.0 pulgadas.

b. Si cada fila de dos asientos se diseña para ajustarse a dos hombres, sólo si éstos tienen una anchura media de cadera de 16.0 pulgadas o menos, ¿muchos individuos serían incapaces de ajustarse? ¿Parece aceptable este diseño?

10. Generador uniforme de números aleatorios El generador de números aleatorios de la calculadora TI-83 Plus, así como el de muchas otras calculadoras y computadoras producen números de una distribución uniforme de valores entre 0 y 1, con una media de 0.500 y una desviación estándar de 0.289. Si se generan 100 números aleatorios, calcule la probabilidad de que su media sea mayor que 0.57. ¿Sería poco común generar 100 de estos números y obtener una media mayor que 0.57? ¿Por qué?

11. Cantidades de Coca Cola Suponga que latas de Coca Cola se llenan de tal manera que las cantidades reales tienen una media de 12.00 onzas y una desviación estándar de 0.11 onzas.

a. Calcule la probabilidad de que una muestra de 36 latas tenga una cantidad media de al menos 12.19 onzas, como en el conjunto de datos 17 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 onzas? Si la media no es 12.00 onzas, ¿se está engañando a los consumidores?

12. Puntuaciones de CI Para formar parte de la organización Mensa se requiere una puntuación de CI por arriba de 131.5. Nueve candidatos toman pruebas de CI y el resumen de sus resultados indica que su puntuación media de CI es 133. (Las puntuaciones de CI se distribuyen de manera normal, con una media de 100 y una desviación estándar de 15).

a. Si se selecciona una persona al azar, calcule la probabilidad de elegir a alguien con una puntuación de CI de al menos 133.

b. Si se seleccionan nueve personas al azar, calcule la probabilidad de que su puntuación media de CI sea de al menos 133.

c. Aunque el resumen de los resultados está disponible, se perdieron las puntuaciones individuales de CI. ¿Se puede concluir que los nueve candidatos tienen puntuaciones de CI mayores de 133, de manera que todos son elegibles para formar parte de Mensa?

13. Tiempo medio de reemplazos El gerente de la tienda Portland Electronics se preocupa porque sus distribuidores le están entregando televisores con una calidad menor al promedio. Su investigación revela que los tiempos de reemplazo de televisores tienen una media de 8.2 años y una desviación estándar de 1.1 años (según datos de "Getting Things Fixed", *Consumer Reports*). Entonces, selecciona al azar 50 televisores que se vendieron en el pasado y encuentra que el tiempo de reemplazo es de 7.8 años.

a. Suponiendo que el tiempo de reemplazo de televisores tiene una media de 8.2 años y una desviación estándar de 1.1 años, calcule la probabilidad de que 50 televisores que se seleccionaron aleatoriamente tengan un tiempo medio de reemplazo de 7.8 años o menos.

b. Con base en el resultado del inciso *a*, ¿parecería que la tienda Portland Electronics vende televisores con una calidad menor al promedio?

14. Presión sanguínea La presión sistólica (en mm Hg) de mujeres entre 18 y 24 años se distribuye normalmente, con una media de 114.8 y una desviación estándar de 13.1 (de acuerdo con datos del National Health Survey). La hipertensión suele definirse como una presión sistólica que rebasa 140.

a. Si se selecciona al azar a una mujer de entre 18 y 24 años, calcule la probabilidad de que su presión sistólica sea mayor que 140.

b. Si se seleccionan al azar cuatro mujeres del mismo rango de edad, calcule la probabilidad de que su presión sistólica media sea mayor que 140.

c. Puesto que el inciso *b* incluye un tamaño de muestra no mayor que 30, ¿por qué se puede utilizar el teorema del límite central?

d. Si un médico recibe un reporte que afirma que cuatro mujeres tienen una presión sistólica media menor que 140, ¿concluiría que ninguna de las mujeres es hipertensa (con una presión sanguínea mayor que 140)?

15. Reducción de nicotina en cigarrillos Las cantidades de nicotina en los cigarrillos Dytusoon tienen una media de 0.941 g y una desviación estándar de 0.313 g (con base en el conjunto de datos 5 del Apéndice B). La Huntington Tobacco Company, que produce los cigarrillos Dytusoon, afirma que redujo la cantidad de nicotina. La evidencia consiste en una muestra de 40 cigarrillos con una cantidad media de nicotina de 0.882 g.

a. Suponiendo que la media y la desviación estándar dadas no cambiaron, calcule la probabilidad de seleccionar al azar 40 cigarrillos con una cantidad media de nicotina de 0.882 g o menos.

b. Con base en el resultado del inciso *a*, ¿se valdrá afirmar que la cantidad de nicotina es menor? ¿Por qué?

16. 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 que se seleccionaron aleatoriamente se les da el Columbian Review Course, antes de tomar la prueba SAT. Suponga que el curso no tiene efecto alguno.

a. Si se selecciona a uno de los hombres al azar, calcule la probabilidad de que su calificación sea de al menos 590.

b. Si se selecciona a 16 de los hombres al azar, calcule la probabilidad de que su calificación media sea de al menos 590.

c. En el cálculo de la probabilidad del inciso *b*, ¿por qué puede usarse el teorema del límite central si el tamaño muestral no excede de 30?

d. Si la muestra aleatoria de 16 hombres arroja una calificación media de 590, ¿habría una fuerte evidencia que apoye la afirmación de que el curso es realmente eficaz? ¿Por qué?

17. Sobrecarga de un depósito de desperdicios La ciudad de Newport opera un depósito de basura que se sobrecarga si las descargas de desperdicios de sus 4872 hogares exceden una media de 27.88 libras en una semana. En muchas semanas se observa que las muestras de 4872 hogares tienen pesos que se distribuyen de manera normal, con una media de 27.44 libras y una desviación estándar de 12.46 libras (según datos del Garbage Project de la Universidad de Arizona). ¿Cuál es la proporción de semanas cuando el depósito de basura se sobrecarga? ¿Será un nivel aceptable o se deben tomar acciones para corregir un problema de un sistema que se sobrecargó?

18. Etiquetas de paquetes de M&M Los dulces M&M sencillos tienen un peso medio de 0.9147 g y una desviación estándar de 0.0369 g (con base en el conjunto de datos 19 del Apéndice B). Los dulces M&M que se utilizan en el conjunto de datos 19 provienen de un paquete que contenía 1498 dulces y la etiqueta del paquete establecía que su peso neto era de 1361 g. (Si cada paquete tiene 1498 dulces, el peso medio de los dulces debe exceder $1361/1498 = 0.9085$ g del contenido neto, para pesar al menos 1361 g).

a. Si se selecciona al azar un dulce M&M sencillo, calcule la probabilidad de que pese más de 0.9085 g.

b. Si se seleccionan al azar 1498 dulces M&M sencillos, calcule la probabilidad de que su peso medio sea de al menos de 0.9085 g.

c. Con estos resultados, ¿está Mars Company ofreciendo a los consumidores de M&M la cantidad que anuncia en la etiqueta?

19. Diseño de elevadores Los pesos de las mujeres se distribuyen normalmente, con una media de 143 libras y una desviación estándar de 29 libras, y los pesos de los hombres se distribuyen normalmente, con una media de 172 libras y una desviación estándar de 29 libras (datos del National Health Survey). Usted necesita diseñar un elevador para Wesport Shopping Center, el cual debe llevar a salvo a 16 personas. Suponiendo que "en el peor de los casos" se suban 16 pasajeros hombres, calcule el peso máximo total que se permite si deseamos un probabilidad de 0.975 de que este máximo no se rebase cuando se seleccione aleatoriamente a 16 hombres.

20. Diseño de asiento Usted necesita construir una banca que sentará a 18 jugadores universitarios de futbol 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 pulgadas y una desviación estándar de 1.0 pulgadas.

a. ¿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 de 18 hombres que se seleccionaron al azar?

b. ¿Por qué sería incorrecto utilizar realmente el resultado del inciso *a* como longitud de la banca?

5-5 Más allá de lo básico

21. Corrección para una población finita El club Boston Women necesita un elevador que se limite a ocho pasajeros. El club tiene 120 miembros mujeres con pesos que se aproximan a una distribución normal, con una media de 143 libras y una desviación estándar de 29 libras. (*Sugerencia:* Véase la explicación del factor de corrección para una población finita).

a. Si se seleccionan al azar ocho miembros diferentes, calcule la probabilidad de que su peso total no rebase la capacidad máxima de 1300 libras.

b. Si buscamos una probabilidad de 0.99 de que el elevador no se sobrecargue siempre que se seleccione aleatoriamente a ocho miembros como pasajeros, ¿cuál debe ser el peso máximo permitido?

22. Parámetros de población Una *población* se compone de los valores: 2, 3, 6, 8, 11, 18.

a. Calcule μ y σ.

b. Liste todas las muestras de tamaño $n = 2$ que es posible obtener con reemplazo.

c. Calcule la población de todos los valores de $\bar{x}$ al obtener la media de cada muestra del inciso *b*.

d. Calcule la media $\mu_{\bar{x}}$ y la desviación estándar $\sigma_{\bar{x}}$ para la población de medias de muestra obtenidas en el inciso *c*).

e. Verifique que

$$\mu_{\bar{x}} = \mu \qquad \text{y} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

23. Generador uniforme de números aleatorios En el ejercicio 10 se señaló que muchas calculadoras y computadoras tienen generadores de números aleatorios que producen números de una distribución uniforme de valores entre 0 y 1, con una media de 0.500 y una desviación estándar de 0.289. Si se generan 100 números aleatorios, calcule la probabilidad de que su media caiga entre 0.499 y 0.501. Si generásemos 100 de estos números y encontrásemos que la media cae entre 0.499 y 0.501, ¿concluiríamos que el resultado es " poco común", de tal manera que el generador de números aleatorios está defectuoso? ¿Por qué?

5-6 La distribución normal como aproximación de la distribución binomial

En lugar de "la distribución normal como aproximación de la distribución binomial", el encabezado apropiado para esta sección debería ser "uso de la distribución normal como aproximación de la distribución binomial", pero preferimos el primer título. El segundo refleja mejor el propósito de esta sección. Comencemos revisando las condiciones requeridas para una *distribución de probabilidad binomial*, que se analizan en la sección 4-3:

1. El procedimiento debe tener un *número fijo de ensayos*.
2. Los ensayos deben ser *independientes*.
3. Todos los resultados de cada ensayo deben estar clasificados en *dos categorías*.
4. Las probabilidades deben permanecer *constantes* para cada ensayo.

En la sección 4-3 presentamos tres métodos para calcular probabilidades binomiales: **1.** uso de la fórmula de probabilidad binomial, **2.** uso de la tabla A-1 y **3.** uso de programas de cómputo (tales como STATDISK, Minitab o Excel) o una calculadora TI-83 Plus. Sin embargo, en muchos casos ninguno de estos métodos es práctico, ya que los cálculos requieren demasiado tiempo y esfuerzo. Ahora presentamos un nuevo método que utiliza una distribución normal como aproximación de la distribución binomial. El siguiente recuadro resume el punto más importante de esta sección.

La distribución normal como aproximación de la distribución binomial

Si $np \geq 5$ y $nq \geq 5$, entonces la variable aleatoria binomial tiene una distribución de probabilidad que puede aproximarse con una distribución normal, donde la media y la desviación estándar están dadas por

$$\mu = np$$
$$\sigma = \sqrt{npq}$$

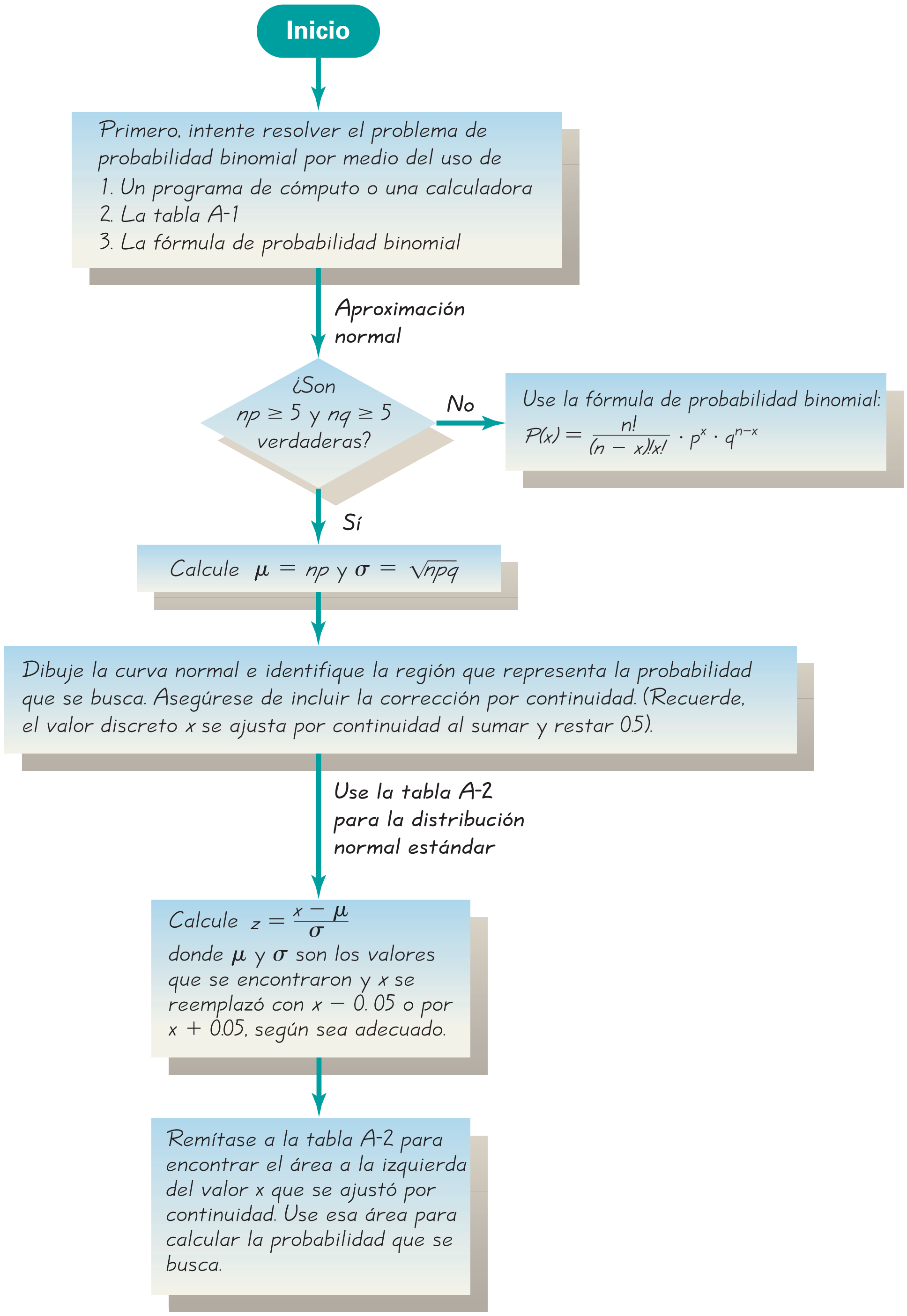


FIGURA 5-23 Uso de la distribución normal como aproximación de la distribución binomial

Para comprender mejor la manera en que la distribución normal, puede usarse para aproximar una distribución binomial, remítase a la figura 5-18 de la sección 5-4. Esta figura presenta un histograma de frecuencias relativas de los valores de 10,000 *proporciones* de muestra, donde cada una de las 10,000 muestras consta de 50 géneros que se seleccionaron al azar con reemplazo, de una población en la que la proporción de mujeres es de 0.13. Dichas proporciones muestrales pueden considerarse probabilidades binomiales, de manera que la figura 5-18 indica que, en condiciones adecuadas, las probabilidades binomiales tienen una distribución muestral aproximadamente normal. La justificación formal que nos permite emplear la distribución normal como aproximación de la distribución binomial resulta de matemáticas más avanzadas; la figura 5-18 es un argumento visual convincente que apoya esa aproximación.

Al resolver problemas de probabilidad binomial, primero intente obtener resultados más exactos por medio de un programa de cómputo, una calculadora, la tabla A-1 o la fórmula de probabilidad binomial. Si la probabilidad binomial no puede calcularse con estos procedimientos más exactos, intente la técnica del uso de la distribución normal como aproximación de la distribución binomial. Este método implica el siguiente procedimiento, que también se presenta en un diagrama de flujo en la figura 5-23.

Procedimiento para el uso de una distribución normal como aproximación de la 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 ambas condiciones no se satisfacen, 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). Reemplace el valor *discreto* x con el *intervalo* desde $x - 0.5$ hasta $x + 0.5$. (Para mayor explicación consulte la sección titulada "Correcciones por continuidad", más adelante en esta sección). Dibuje una curva normal e introduzca 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 $x + 5$, según sea apropiado.
5. Utilice $x - 0.5$ o $x + 0.5$ (según sea apropiado) en lugar de x, calcule el área correspondiente a la probabilidad deseada encontrando primero la puntuación z: $z = (x - \sigma)/\sigma$. Ahora use esa puntuación z para encontrar el área a la izquierda de $x - 0.5$ o $x + 0.5$, según sea apropiado. Ahora el área puede emplearse para identificar el área correspondiente a la probabilidad que se desea.

Ilustraremos este procedimiento de aproximación normal con el siguiente ejemplo.

EJEMPLO Cargas de aviones Cuando un avión se carga con pasajeros, equipaje, carga y combustible, el piloto debe verificar que el peso completo no rebase el límite máximo que se permite, por lo cual el peso tiene que distribuirse de forma conveniente para que el equilibrio del avión esté dentro de

continúa

Los niños y las niñas no son igualmente probables

En muchos cálculos de probabilidad, se obtienen buenos resultados suponiendo que los niños y las niñas tienen la misma posibilidad de nacer. En realidad, un niño tiene mayores posibilidades de nacer (probabilidad de 0.5117) que una niña (probabilidad de 0.4883). Tales resultados se basan en datos recientes del National Center for Health Statistics de Estados Unidos, institución que mostró que los 4,058,814 nacimientos que ocurrieron en un año incluyeron 2,076,969 niños y 1,981,845 niñas. Los investigadores están pendientes de dichas probabilidades por los cambios que podrían sugerir factores; entre otros, en el ambiente y la exposición a agentes químicos.

los límites aceptables de seguridad. Air America estableció un procedimiento según el cual debe reducirse la carga extra siempre que un avión con 200 pasajeros incluya al menos 120 hombres. Calcule la probabilidad de que, de 200 pasajeros seleccionados al azar, haya al menos 120 hombres. Suponga que la población de pasajeros potenciales consiste en un número igual de hombres y mujeres.

SOLUCIÓN Remítase a la figura 5-23 que presenta el procedimiento que se realizó para esta solución. El problema dado implica una distribución binomial con un número fijo de ensayos ($n = 200$), que se presume son independientes, con dos categorías de resultados (hombres, mujeres) para cada ensayo, y con la probabilidad de un hombre ($p = 0.5$) que se supone permanece constante de un ensayo a otro.

Supondremos que no disponemos de un programa de cómputo ni de una calculadora. La tabla A-1 no puede aplicarse, porque termina en $n = 15$. La fórmula de probabilidad binomial no es práctica, ya que tendríamos que utilizarla en 81 ocasiones (una para cada valor de x desde 120 hasta 200, inclusive), y nadie en su sano juicio desearía hacerlo.

Procedamos con el método de los cinco pasos para el uso de una distribución normal como aproximación de la distribución binomial.

Paso 1: Primero debemos verificar que es razonable aproximar la distribución binomial con la distribución normal, porque $np \geq 5$ y $nq \geq 5$. Con $n = 200$, $p = 0.5$ y $q = 1 - p = 0.5$, verificamos las condiciones requeridas como sigue:

$$np = 200 \cdot 0.5 = 100 \qquad \text{(Por lo tanto, } np \geq 5\text{).}$$

$$nq = 200 \cdot 0.5 = 100 \qquad \text{(Por lo tanto, } nq \geq 5\text{).}$$

Paso 2: Ahora procedamos a calcular los valores de μ y σ, necesarios para la distribución normal. Obtendremos lo siguiente:

$$\mu = np = 200 \cdot 0.5 = 100$$

$$\sigma = \sqrt{npq} = \sqrt{200 \cdot 0.5 \cdot 0.5} = 7.0710678$$

Paso 3: El valor discreto de 120 se representa con la franja que se limita con 119.5 y 120.5. (Véase la explicación sobre correcciones por continuidad, que sigue de este ejemplo).

Paso 4: Ya que buscamos la probabilidad de *al menos* 120 hombres, queremos el área que representa el número discreto de 120 (la región limitada por 119.5 y 120.5), así como también el área a la derecha, como se muestra en la figura 5-24.

Paso 5: Ahora es posible proceder a la búsqueda del área que se sombreó de la figura 5-24, utilizando los mismos métodos que se emplearon en la sección 5-3. Para usar la tabla A-2 de la distribución normal estándar, primero habrá que transformar 119.5 a una puntuación z; después, usar la tabla para encontrar el área a la izquierda de 119.5, que posteriormente se resta de 1. La puntuación z se obtiene como sigue:

$$z = \frac{x - \mu}{\sigma} = \frac{119.5 - 100}{7.0710678} = 2.76$$

LA ESTADÍSTICA EN LAS NOTICIAS

Ganadores múltiples de la lotería

Evelyn Marie Adams ganó la lotería de Nueva Jersey en dos ocasiones en cuatro meses. Este feliz suceso fue reportado en los medios 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 hay 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; pero 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 Pennsylvania cuatro veces en 12 años, con premios de \$2.5 millones, \$68,000; \$206,217 y \$71,037, respectivamente.

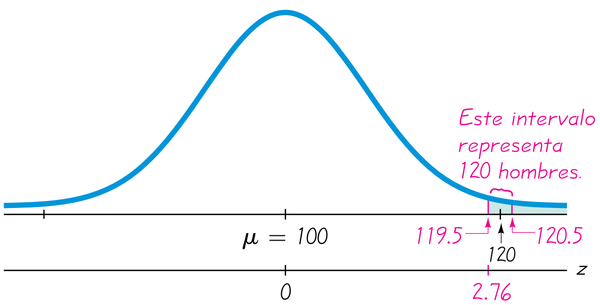


FIGURA 5-24 Búsqueda de la probabilidad de "al menos" 120 hombres entre 200 pasajeros

Al emplear la tabla A-2, encontramos que $z = 2.76$ corresponde a un área de 0.9971, de manera que la región que se sombreó es de $1 - 0.9971 = 0.0029$.

INTERPRETACIÓN Hay una probabilidad de 0.0029 de obtener al menos 120 hombres entre 200 pasajeros. Como esa probabilidad es muy baja, concluimos que, en muy pocas ocasiones, una lista de 200 pasajeros incluirá al menos 120 hombres, por lo que no es necesario preocuparse mucho por reducir la carga extra.

Correcciones por continuidad

El procedimiento que implica el uso de la distribución normal como aproximación de la distribución binomial incluye un paso en el que cambiamos un número discreto por un intervalo que está 0.5 por abajo y 0.5 por arriba del número discreto. Observe la solución anterior, donde cambiamos 120 por el intervalo entre 119.5 y 120.5. Este paso particular, que se denomina *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 un 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* desde $x - 0.5$ hasta $x + 0.5$ (es decir, sumando y restando 0.5).

Las siguientes sugerencias prácticas deben ayudarlo a utilizar las correcciones por continuidad en forma apropiada.

Procedimiento para correcciones por continuidad

1. Cuando use la distribución normal como aproximación de la distribución binomial, *siempre* aplique la corrección por continuidad. (Esto se requiere, 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 120 hombres entre 200 personas que se seleccionaron aleatoriamente, el número entero discreto relevante sería $x = 120$. Primero enfoque su atención en el valor x e ignore temporalmente si busca al menos x, más que x, menos que x, u otro.

3. Dibuje una distribución normal centrada alrededor de μ; después, una *franja vertical* alrededor de x. Marque el lado izquierdo de la franja con el número igual a $x - 0.5$ y el lado derecho con el número igual a $x + 0.5$. Por ejemplo, para $x = 120$, dibuje una franja desde 119.5 hasta 120.5. *Considere el área completa de la franja para representar la probabilidad del número 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). Después, 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 a la izquierda de la franja, según sea apropiado; también el interior de la franja *si y sólo si* x se incluirá. Esta región total que se sombreó corresponde a la probabilidad buscada.

Para ver cómo resulta este procedimiento en las correcciones por continuidad, observe los casos comunes que se ilustran en la figura 5-25. Esos casos corresponden a los enunciados de la siguiente lista.

Enunciado	**Área**
Al menos 120 (incluye 120 y números mayores)	A la *derecha* de 119.5
Más de 120 (no incluye 120)	A la *derecha* de 120.5
A lo sumo 120 (incluye 120 y números menores)	A la *izquierda* de 120.5
Menos de 120 (no incluye 120)	A la *izquierda* de 119.5
Exactamente 120	Entre 119.5 y 120.5

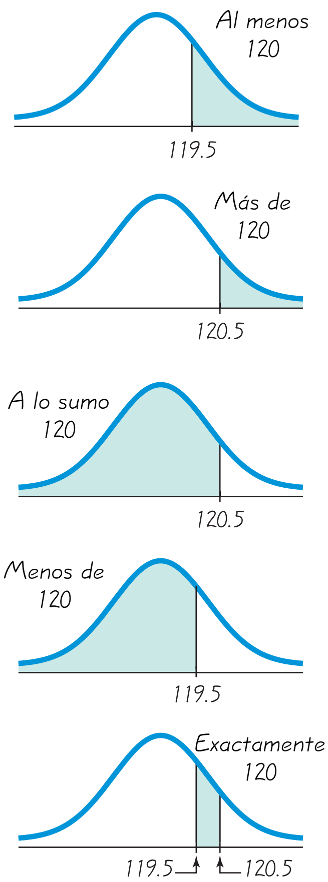


FIGURA 5-25 Uso de las correcciones por continuidad

EJEMPLO Audiencia televisiva El programa de televisión *60 minutos*, de la CBS, tuvo recientemente una audiencia de 20, lo que significa que, de los televisores en uso, el 20% estaba sintonizando *60 minutos* (de acuerdo con datos de Nielsen Media Research). Un anunciante desea verificar este valor del 20% de audiencia realizando su propia encuesta a 200 hogares que tengan su televisión encendida en el momento de la transmisión de *60 minutos*. Los resultados indican que, de los 200 televisores en uso, el 16% (o 32 televisores) están sintonizando *60 minutos*. Suponiendo que el valor de audiencia del 20% sea correcto, calcule la probabilidad de que en una encuesta de 200 hogares, *exactamente* 32 televisores estén sintonizando *60 minutos*. Puesto que el resultado muestral del 16% es menor que el valor de audiencia que se anunció del 20%, ¿hay evidencia fuerte para concluir que el valor de audiencia del 20% es incorrecto?

SOLUCIÓN Tenemos $n = 200$ ensayos independientes, $x = 32$ televisores sintonizando *60 minutos* y una proporción poblacional de $p = 0.20$. Para los propósitos de este ejemplo, suponemos que no se nos permite el acceso a un programa de cómputo ni a una calculadora TI-83 Plus. Tampoco es posible utilizar la tabla A-1, porque $n = 200$ excede el valor más alto de la tabla de

$n = 15$. Si utilizamos la fórmula de probabilidad binomial, deberíamos evaluar una expresión que incluya 200!, pero muchas calculadoras y programas de cómputo no manejan tantos datos. Por consiguiente, procedemos a emplear una distribución normal para aproximar la distribución binomial.

Paso 1: Primero verificamos si es posible la aproximación:

$$np = 200 \cdot 0.20 = 40 \qquad \text{(Por lo tanto, } np \geq 5.)$$

$$nq = 200 \cdot 0.80 = 160 \qquad \text{(Por lo tanto, } nq \geq 5.)$$

Paso 2: Ahora procedemos a calcular los valores de μ y σ, necesarios para la distribución normal. Obtenemos lo siguiente:

$$\mu = np = 200 \cdot 0.20 = 40$$

$$\sigma = \sqrt{npq} = \sqrt{200 \cdot 0.20 \cdot 0.80} = 5.6568542$$

Paso 3: Dibujamos la curva normal de la figura 5-26. La región sombreada de la figura representa la probabilidad que buscamos. La aplicación de la corrección por continuidad da como resultado la representación de 32, ubicada entre 31.5 y 32.5.

Paso 4: He aquí el método que se empleó para calcular la región sombreada de la figura 5-26: primero calcule el área total a la izquierda de 32.5; después, obtenga el área total a la izquierda de 31.5; luego, calcule la *diferencia* entre ambas áreas. Iniciando con el área total a la izquierda de 32.5, debemos obtener la puntuación z que corresponde a 32.5. Si nos remitimos a la tabla A-2, obtendremos

$$z = \frac{32.5 - 40}{5.6568542} = -1.33$$

Usamos la tabla A-2 para encontrar que $z = -1.33$ corresponde a una probabilidad de 0.0918, que es el área total a la izquierda de 32.5. Ahora, procedemos a obtener el área a la izquierda de 31.5, calculando primero la puntuación z correspondiente a 31.5:

$$z = \frac{31.5 - 40}{5.6568542} = -1.50$$

En la tabla A-2 encontramos que $z = -1.50$ corresponde a una probabilidad de 0.0668, que es el área total a la izquierda de 31.5. El área sombreada es $0.0918 - 0.0668 = 0.0250$.

continúa

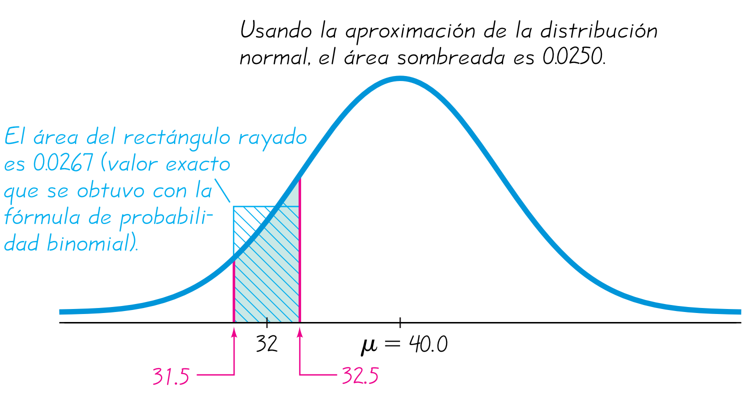


FIGURA 5-26 **Uso de la corrección por continuidad de la audiencia televisiva**

INTERPRETACIÓN La probabilidad de que exactamente 32 televisores sintonicen *60 minutos* (de un total de 200) es de aproximadamente 0.0250. El planteamiento del problema también nos pide determinar si el resultado muestral del 16% constituye una evidencia suficiente para concluir que el valor de audiencia del 20% es incorrecto. Sin embargo, en lugar de considerar la probabilidad de *exactamente* 32 televisores que sintonizan *60 minutos*, debemos considerar la probabilidad de *32 o menos*. [En la sección 4-2, señalamos que x éxitos en n ensayos es un número de éxitos *infrecuentemente bajo*, si P(x o menor) es muy pequeña, como 0.05 o menos]. En la solución anterior vemos que la probabilidad de 32 o menos éxitos es P(menos que 32.5), que es 0.0918. Puesto que 0.0918 no es muy pequeño, *no* tenemos suficiente evidencia para concluir que el valor de audiencia del 20% sea incorrecto.

Si resolvemos el ejemplo anterior por medio de STATDISK, Minitab o una calculadora, obtendremos un resultado de 0.0267, pero el método de aproximación normal arrojó un valor de 0.0250. La discrepancia de 0.0017 sucede porque el uso de la distribución normal da como resultado un valor que se *aproxima* al que corresponde al área de la región sombreada en la figura 5-26, mientras que el área correcta exacta es un rectángulo que se centra por arriba de 32. (La figura 5-26 ilustra tal discrepancia). El área del rectángulo es 0.0267, pero el área aproximada de la región sombreada es 0.0250.

Interpretación de los resultados

En realidad, cuando utilizamos una distribución normal como aproximación de la distribución binomial, nuestra meta no es sencillamente calcular un número de probabilidad. Con frecuencia necesitamos hacer algún *juicio* con base en el valor de probabilidad, como en la conclusión final del ejemplo anterior. 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 posibles y sucesos imposibles. El siguiente criterio (de la sección 4-2) describe la aplicación de las probabilidades para distinguir resultados que pueden ocurrir fácilmente por el azar, de aquellos que son en extremo 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 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 o menos) es muy pequeña (como 0.05 o menos).

5-6 Destrezas y conceptos básicos

Aplicación de la corrección por continuidad. *En los ejercicios 1 a 8, los valores dados son discretos. Utilice la corrección por continuidad y describa la región de la distribución normal que corresponde a la probabilidad que se indica. Por ejemplo, la probabilidad de "más que 20 artículos defectuosos" corresponde al área de la curva normal descrita en esta respuesta: "el área a la derecha de 20.5".*

1. Probabilidad de que más de 15 personas en prisión quiten las etiquetas de advertencia a las almohadas.
2. Probabilidad de que al menos 24 estudiantes comprendan la corrección por continuidad.
3. Probabilidad de que haya menos de 100 pasajeros en su siguiente vuelo comercial.
4. Probabilidad de que el número de distribuidores automáticos en Estados Unidos sea exactamente 27.
5. Probabilidad de no más de cuatro estudiantes ausentes en una clase de estadística.
6. Probabilidad de que el número de CD defectuosos de Wayne Newton sea de 15 a 20, inclusive.
7. Probabilidad de que el número de senadores estadounidenses ausentes sea de ocho a 10, inclusive.
8. Probabilidad de exactamente tres respuestas "sí" en peticiones de citas.

Uso de la aproximación normal. *En los ejercicios 9 a 12, haga lo siguiente: a) Calcule la probabilidad binomial que se indica por medio de la tabla A-1 del Apéndice A. b) Si* $np \geq 5$ *y* $nq \geq 5$*, también estime la probabilidad que se indica con el uso de la distribución normal como aproximación de la distribución normal; si* $np < 5$ *o* $nq < 5$*, entonces establezca que la aproximación normal no es adecuada.*

9. Con $n = 14$ y $p = 0.5$, calcule $P(9)$.
10. Con $n = 12$ y $p = 0.8$, calcule $P(7)$.
11. Con $n = 15$ y $p = 0.9$, calcule P(al menos 14).
12. Con $n = 13$ y $p = 0.4$, calcule P(menor que 3).
13. Probabilidad de más de 55 niñas Estime la probabilidad de que resulten más de 55 niñas en 100 nacimientos. Suponga que los niños y las niñas son igualmente probables. ¿Es poco común que resulten más de 55 niñas en 100 nacimientos?
14. Probabilidad de al menos 65 niñas Estime la probabilidad de que resulten al menos 65 niñas en 100 nacimientos. Suponga que los niños y las niñas son igualmente probables. ¿Es poco común que resulten al menos 65 niñas en 100 nacimientos?
15. Probabilidad de al menos aprobar Estime la probabilidad de aprobar un examen de verdadero/falso de 100 preguntas, si el 60% (o 60 respuestas correctas) es la calificación mínima de aprobación y si todas las respuestas son conjeturas. ¿Es la probabilidad lo suficientemente alta como para arriesgarse a aprobar adivinando en lugar de estudiar?
16. Examen de opción múltiple Un examen de opción múltiple consta de 25 preguntas con las respuestas posibles a, b, c, d y e. Estime la probabilidad de que, al adivinar, el número de respuestas correctas sea de tres a 10, inclusive.
17. Experimento de hibridación de Mendel Cuando Mendel realizó sus famosos experimentos de hibridación, utilizó chícharos con vainas verdes y vainas amarillas. Uno de los experimentos implicó una cruza de chícharos, de manera que se esperaba que el 25% (o 145) de los 580 chícharos vástagos tuvieran vainas amarillas. En lugar de obtener 145 chícharos con vainas verdes, obtuvo 152. Suponiendo que el porcentaje del 25% de Mendel es correcto, estime la probabilidad de obtener al menos 152 chícharos con vainas amarillas, entre los 580 chícharos vástagos. ¿Existirá una fuerte evidencia que sugiera que la probabilidad del 25% de Mendel es incorrecta?
18. Fármaco que reduce el colesterol La probabilidad de que una persona que no recibe ningún tratamiento tenga síntomas de gripe es de 0.019. En un ensayo clínico de Lipitor, un fármaco común que se utilizó 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. ¿Sugieren estos resultados acerca de los síntomas de gripe que hay una reacción adversa al fármaco?

19. Probabilidad de al menos 50 hombres daltónicos El 9% de los hombres y el 0.25% de las mujeres no pueden distinguir entre los colores rojo y verde. Este tipo de daltonismo causa problemas con las señales de tránsito. Los investigadores necesitan al menos 50 hombres con este tipo de ceguera al color, de manera que seleccionan aleatoriamente a 600 hombres para un estudio de percepción de las señales de tránsito. Estime la probabilidad de que al menos 50 de los hombres no distingan entre el rojo y el verde. ¿Es el resultado lo suficientemente alto como para qué los investigadores puedan confiarse de obtener al menos 50 hombres con daltonismo?

20. Teléfonos celulares y cáncer cerebral En un estudio de 420,000 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, hay 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 420,000 personas seleccionadas al azar. Estime la probabilidad de 135 o menos casos de este cáncer en un grupo de 420,000 personas. ¿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?

21. Vuelos sobresaturados Air America está considerando la nueva política de registrar 400 personas en un avión que tiene sólo 350 asientos. (Estudios anteriores han revelado que sólo el 85% de los pasajeros registrados llegan al vuelo). Estime la probabilidad de que, si Air America registra a 400 personas, no haya suficientes asientos disponibles. ¿Es esta probabilidad lo suficientemente baja para ser funcional, o deberá modificarse la política?

22. Vuelos a tiempo Recientemente, el 72.3% de los vuelos de American Airlines llegaron a tiempo (según datos del Departamento del Transporte de Estados Unidos). Al verificar 40 vuelos de American Airlines, seleccionados al azar, 19 llegaron a tiempo. Estime la probabilidad de que 19 vuelos o menos, entre 40, lleguen a tiempo suponiendo que el porcentaje del 72.3% sea correcto. ¿Será poco común que 19 vuelos o menos, entre 40 vuelos de American Airlines seleccionados aleatoriamente lleguen a tiempo?

23. Identificación de discriminación por género Después de que la rechazaron para un empleo, Kim Kelly se entera de que la Bellevue Advertising Company contrató únicamente a 21 mujeres entre sus 62 empleados nuevos. También de que el grupo de solicitantes es muy grande, con igual número de hombres y mujeres calificados. Ayúdela a hacer una acusación por discriminación, estimando la probabilidad de obtener 21 mujeres o menos cuando se contrata a 62 personas, suponiendo que no hay discriminación por género. ¿En realidad apoya la probabilidad resultante una acusación como ésta?

24. Dulces M&M: ¿el 10% son azules? Según un representante de asuntos de consumo de Mars (la compañía de dulces), el 10% de todos los dulces sencillos M&M son azules. El conjunto de datos 19 del Apéndice B indica que de 100 M&M elegidos, cinco son azules. Estime la probabilidad de seleccionar al azar 100 dulces M&M y obtener cinco o menos que sean azules. Suponga que el porcentaje de azules del 10%, establecido por la compañía, es correcto. Con base en el resultado, ¿será poco común obtener cinco o menos M&M azules cuando se seleccionan 100 al azar?

25. Grupo sanguíneo El 45% de nosotros tiene sangre del grupo O, según datos que proporcionó el Great New York Program. El Providence Memorial Hospital está realizando una campaña de donación de sangre, ya que su abastecimiento de sangre del grupo O es bajo y necesita 177 donadores de este tipo de sangre. Si 400 voluntarios donan sangre, estime la probabilidad de que el número de personas con sangre del grupo O sea al menos de 177. ¿Es probable que el grupo de 400 voluntarios sea suficiente?

26. Muestreo de aceptación En la sección 3-4 establecimos que 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 específico de defectos. La Dayton Machine Company compra tornillos de máquina en lotes de 5000 y rechaza un lote si, cuando se saca una muestra de 50, al menos dos son defectuosos. Estime la probabilidad de rechazar un lote si el abastecedor está fabricando los tornillos con una tasa de defectos del 10%. ¿Es posible que el plan de verificación identifique la tasa inaceptable de defectos?

27. Choques de automóviles Entre los conductores de 20 a 24 años de edad hay una tasa del 34% de accidentes automovilísticos en un año (según datos del National Safety Council de Estados Unidos). Un investigador de seguros encuentra que en un grupo de 500 conductores con edades que fluctúan entre 20 y 24 años, que se seleccionó aleatoriamente, y que viven en la ciudad de Nueva York, el 40% tuvo accidentes el año anterior. Si el porcentaje del 34% es correcto, estime la probabilidad de que en un grupo de 500 conductores seleccionados al azar, al menos el 40% tuvieran accidentes el año anterior. Con base en el resultado, ¿existe fuerte evidencia que apoye la aseveración de que la tasa de accidentes en la ciudad de Nueva York es mayor al 34%?

28. Encuesta sobre clonación Una reciente encuesta de Gallup incluyó 1012 adultos que se seleccionaron al azar, a quienes se les preguntó si "la clonación humana debe o no permitirse". Los resultados mostraron que el 89% de los encuestados indicaron que no debe permitirse. 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 dicha clonación. Suponiendo que el 50% de todas las personas se oponga, estime la probabilidad de obtener al menos 89% de oposición en una encuesta de 1012 personas seleccionadas al azar. Con base en el resultado, ¿hay fuerte evidencia que apoye la afirmación de que la mayoría se opone a la clonación de humanos?

5-6 Más allá de lo básico

29. Ganar en la ruleta Marc Taylor planea hacer 200 apuestas, de \$1 cada una, al número 7 en la ruleta. Un triunfo paga con posibilidades 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.

30. 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 (de acuerdo con datos de "Getting Fixed", *Consumer Reports*). Estime la probabilidad de que, para 250 televisores seleccionados al azar, al menos 15 de ellos tengan tiempos de reemplazo mayores de 10.0 años.

81. 'Joltin' Joe Suponga que un jugador de beisbol pega de "hit" .350, de manera que su probabilidad de 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.

a. Calcule la probabilidad de al menos un "hit" en cuatro intentos, en un juego.

b. Suponiendo que este bateador pasa a batear cuatro veces cada juego, estime la probabilidad de obtener un total de al menos 56 "hits" en 56 juegos.

c. Suponiendo que este bateador pasa a batear cuatro veces cada juego, estime la probabilidad de al menos un "hit" en cada uno de 56 juegos consecutivos (que es 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?

32. Vuelos sobresaturados 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 una capacidad de 250, si hay al menos una probabilidad de 0.95 de que a todos los individuos que reservaron y se presenten se les acomode?

5-7 Determinación de la normalidad

Los siguientes capítulos incluyen algunos métodos estadísticos muy importantes que requieren que los datos muestrales se seleccionen aleatoriamente a partir de una población con una distribución *normal*. Por consiguiente, es necesario determinar si los datos muestrales parecen provenir de una población distribuida normalmente. En esta sección introducimos la *gráfica cuantilar normal* como una herramienta que nos ayuda a determinar si aparentemente se satisfacen los requisitos de una distribución normal.

Definición

Una **gráfica cuantilar normal** es una gráfica de puntos (x, y) donde cada valor x proviene del conjunto original de datos muestrales, así como cada valor y es una puntuación z correspondiente a un valor cuantilar 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. *Datos distantes:* Identifique datos distantes. Rechace la normalidad si hay más de un dato distante presente. (La presencia de un solo dato distante podría ser un error o el resultado de la variación por el azar, pero tenga cuidado porque incluso un solo dato distante llega a producir un efecto importante en los resultados).
3. *Gráfica cuantilar normal:* Si el histograma es básicamente simétrico y existe a lo sumo un dato distante, construya una *gráfica cuantilar normal*. Los siguientes pasos describen la construcción de una gráfica cuantilar normal, pero el procedimiento es tan confuso que solemos utilizar un programa de cómputo o una calculadora para generar la gráfica. Al final de esta sección se incluyen instrucciones para el uso de STATDISK, Minitab, Excel y la calculadora TI-83 Plus, para obtener gráficas cuantilares normales.
   - **a.** Primero ordene los datos del más bajo al más alto.
   - **b.** Con una muestra de tamaño n, cada valor representa una proporción de $1/n$ de la muestra. Utilizando el tamaño muestral n que se conoce, 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 (tabla A-2) para calcular las puntuaciones z correspondientes a las áreas izquierdas acumulativas que se obtuvieron en el paso *b*.
   - **d.** Una los valores originales de los datos ordenados con sus puntuaciones z correspondientes, que se calcularon en el paso *c*, después grafique los puntos (x, y), donde cada x es un valor muestral original, en tanto y es la puntuación z correspondiente.
   - **e.** Examine la gráfica cuantilar 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.

Los pasos 1 y 2 son directos, pero ilustramos la construcción de una gráfica cuantilar normal (paso 3) en el siguiente ejemplo.

EJEMPLO **Edades de presidentes** Los ejercicios 8 y 13 de la sección 5-4 incluyen las edades de cinco presidentes de Estados Unidos con profesiones militares en el momento de tomar posesión: 62, 46, 68, 64, 57. Construya una gráfica cuantilar normal para las edades y determine si parecen provenir de una población que se distribuye normalmente.

SOLUCIÓN Los siguientes pasos corresponden a los listados en el procedimiento anterior para la construcción de una gráfica cuantilar normal.

1. Primero hay que ordenar los datos: 46, 57, 62, 64, 68.
2. Con una muestra de tamaño $n = 5$, cada valor representa una proporción de $1/5$ de la muestra, por lo que procedemos e 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 áreas específicas para el presente ejemplo, con $n =$ 5: $1/10$, $3/10$, $5/10$, $7/10$ y $9/10$. Tales áreas izquierdas acumulativas, que se expresan en forma decimal, son 0.1, 0.3, 0.5, 0.7 y 0.9.
3. Ahora buscamos en 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.
4. Ahora unimos las edades ordenadas con sus puntuaciones z correspondientes; obtenemos las siguientes coordenadas (x, y), que están graficadas en la figura 5.27: $(46, -1.28)$, $(57, -0.52)$, $(62, 0)$, $(64, 0.52)$ y $(68, 1.28)$.

INTERPRETACIÓN Examinamos la gráfica cuantilar normal de la figura 5-27. Como los puntos parecen estar razonablemente cerca de una línea recta, concluimos que las edades dadas parecen provenir de una población que se distribuye normalmente.

Puesto que la construcción de una gráfica cuantilar normal requiere que ordenemos los datos muestrales y que luego hagamos un proceso complicado para calcular las puntuaciones z correspondientes, entonces la construcción manual de la gráfica es

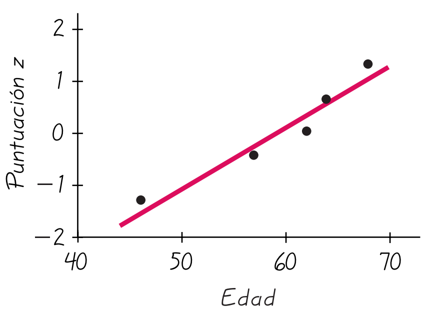


FIGURA 5-27 Gráfica normal cuantilar de las edades de presidentes

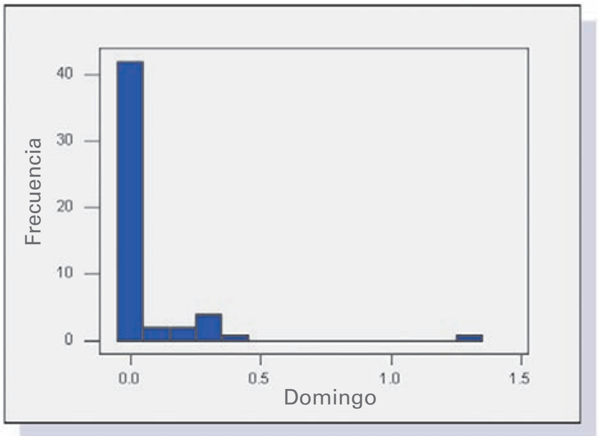
difícil con conjuntos grandes de datos. El siguiente ejemplo ilustra el uso de programas de cómputo.

EJEMPLO Lluvia en Boston En el conjunto de datos 11 del Apéndice B, utilice las 52 cantidades de lluvia de los domingos en Boston y haga una prueba de normalidad.

SOLUCIÓN

Paso 1: Construya un histograma. La siguiente pantalla de Minitab incluye el histograma de las 52 cantidades de lluvia, el cual presenta un sesgo extremo, lo que sugiere que dichas cantidades no se distribuyen de manera normal.

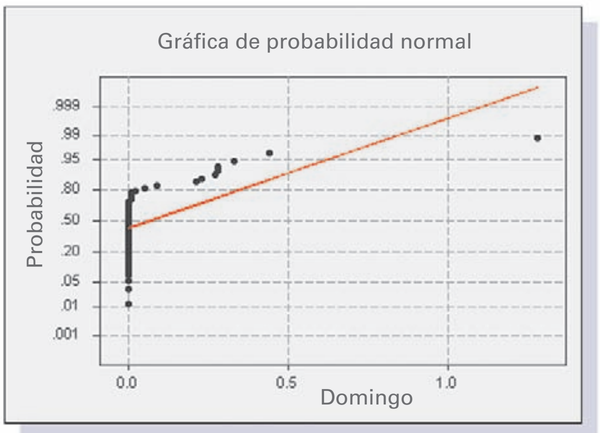
Minitab Histograma de las cantidades de lluvia de los domingos en Boston



Paso 2: Identifique datos distantes. Si examinamos la lista de 52 cantidades de lluvia, encontramos que 1.28 pulgadas parece ser el único dato distante. Debido a que sólo hay un dato distante, no sacamos conclusiones sobre la normalidad de los datos, con base en los datos distantes.

Paso 3: Construya una gráfica cuantilar normal. La siguiente pantalla de Minitab incluye una gráfica *de probabilidad* normal. (Puesto que muchos

Minitab Gráfica de probabilidad normal de las cantidades de lluvia de los domingos en Boston



valores de datos son iguales, la gráfica de probabilidad normal original incluye únicamente 12 puntos distintos en lugar de 52, de manera que la gráfica se modificó para mostrar los 52 puntos). La gráfica de probabilidad normal es igual a la gráfica cuantilar normal, excepto por la escala del eje vertical. Una gráfica de probabilidad normal se interpreta con los mismos criterios que una gráfica cuantilar normal. El examen de una gráfica de probabilidad normal revela un patrón muy diferente al patrón de una línea recta, lo que sugiere que los datos no provienen de una población distribuida normalmente.

INTERPRETACIÓN Puesto que el histograma no parece tener forma de campana, y porque la gráfica de probabilidad normal no produce un patrón de puntos que se aproxime razonablemente a una línea recta, concluimos que las cantidades de lluvia en Boston los domingos no se distribuyen de manera normal. Algunos de los procedimientos estadísticos en los capítulos posteriores requieren que los datos muestrales se distribuyan de manera normal, pero ese requisito no se satisface para las cantidades de lluvia de Boston los domingos, por lo que tales procedimientos no pueden aplicarse.

A continuación presentamos unos comentarios finales acerca de los procedimientos que se emplean para determinar si los datos provienen de una población distribuida de manera normal:

- Si el requisito de una distribución normal no es muy estricto, el examen de un histograma y de los datos distantes podría ser todo lo que necesite para determinar la normalidad.
- Las gráficas cuantilares normales en ocasiones resultan difíciles de construir, pero pueden generarse con una calculadora TI-83 Plus o con un programa de cómputo como STATDISK, Minitab y Excel.
- Además de los procedimientos estudiados en esta sección, hay otros procedimientos más avanzados, como la chi cuadrada, la prueba de bondad de ajuste, la prueba de Kolmogorov-Smirnov y la prueba de Lilliefors. (Véase "Beyond Basic Statistics with the Graphing Calculator, Part I: Assessing Goodness-of-fit", de Calzada y Scariano, *Mathematics and Computer Education*).

Utilizando la tecnología

STATDISK STATDISK puede utilizarse para generar una gráfica cuantilar normal. Primero seleccione **Data** de la parte superior de la barra del menú principal, luego seleccione **Normal Quantile Plot.** Proceda a introducir los datos y haga clic en **Evaluate.**

Minitab Minitab puede emplearse para generar una *gráfica de probabilidad normal*, que se interpreta de la misma forma que la gráfica cuantilar normal. Es decir, los datos que se distribuyen de manera normal deben aproximarse a una línea recta. Primero introduzca 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 de probabilidad normal*, que se interpreta de la misma manera que una gráfica cuantilar normal. Primero

continúa

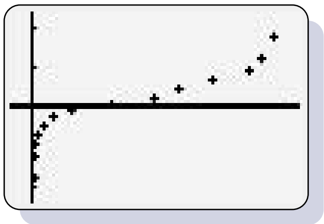
introduzca los valores muestrales en la columna A, después haga clic en **DDXL.** (Si DDXL no aparece en la barra del menú, instale el complemento Data Desk XL). Seleccione **Charts and Plots**, después seleccione la función de **Normal Probability Plot.** Haga clic en el icono del lápiz para "Quantitative Variable", luego introduzca rangos de valores, tales como A1:A36. Presione **OK**.

TI-83 Plus La calculadora TI-83 Plus permite generar una gráfica cuantilar normal de la siguiente manera: primero introduzca los datos muestrales en la lista L1, presione **2nd** y la tecla **Y =** (para **STAT PLOT**), y después, **ENTER**. Seleccione **ON**, seleccione el elemento "type", que es el último del segundo renglón de opciones, luego **L1** para la lista de datos. Luego de hacer todas las selecciones, presione **ZOOM** y luego **9**.

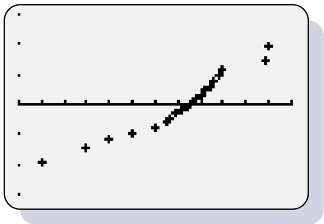
5-7 Destrezas y conceptos básicos

Interpretación de gráficas cuantilares normales. *En los ejercicios 1 a 4, examine la gráfica cuantilar normal y determine si describe datos que tienen una distribución normal.*

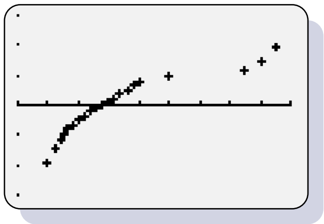
1.



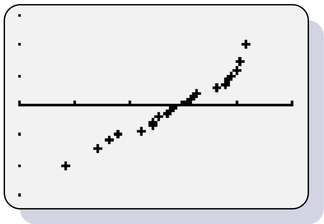
2.



3.



4.



Determinación de normalidad. *En los ejercicios 5 a 8, remítase al conjunto de datos que se indican 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.*

5. Lluvia en Boston Las cantidades de lluvia que caen en Boston los miércoles, como se lista en el conjunto de datos 11 del Apéndice B.

6. Circunferencia de cabezas Las circunferencias de las cabezas de hombres, como se lista en el conjunto de datos 3 del Apéndice B.

7. Pesos de M&M Los pesos de los dulces M&M color café, como se lista en el conjunto de datos 19 del Apéndice B.

8. Conductividad del agua Los niveles de conductividad de los Everglades de Florida, como se lista en el conjunto de datos 12 del Apéndice B.

Generación de gráficas cuantilares normales. *En los ejercicios 9 a 12, utilice los datos del ejercicio que se indica en esta sección. Emplee una calculadora TI-83 Plus o un programa de cómputo (como STATDISK, Minitab o Excel), capaces 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.*

9. Ejercicio 5

10. Ejercicio 6

11. Ejercicio 7

12. Ejercicio 8

13. Comparación de conjuntos de datos Con las estaturas y los niveles de colesterol de mujeres, que se listan en el conjunto de datos 1 del Apéndice B, analice cada uno de 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.

14. Comparación de conjuntos de datos Con los niveles de presión sanguínea histórica y las anchuras del codo de mujeres, que se listan en el conjunto de datos 1 del Apéndice B, analice cada uno de 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 15 y 16, 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 una distribución normal.*

15. Estaturas de los Lakers de L.A. Utilice esta muestra de estaturas (en pulgadas) de los jugadores de la alineación estelar del equipo profesional de basquetbol de los Lakers de Los Ángeles: 85, 79, 82, 73, 78.

16. Monitoreo del plomo en el aire En los días siguientes a la destrucción que causaron los ataques terroristas del 11 de septiembre de 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.

5-7 Más allá de lo básico

17. Uso de puntuaciones estándar Al construir una gráfica cuantilar normal, suponga que en lugar de calcular las puntuaciones z por medio del procedimiento descrito en esta sección, cada valor en una muestra se transforma a su puntuación estándar correspondiente a través de $z = (x - \bar{x})/s$. Si los puntos (x, y) se marcan en una gráfica, ¿es posible usar esta gráfica para determinar si la muestra proviene de una población distribuida normalmente? Explique.

18. Distribución log normal Se considera que la variable aleatoria x tiene una *distribución log normal,* si los valores de ln x se distribuyen normalmente. Pruebe la normalidad de las siguientes duraciones de llamadas telefónicas (en segundos), después pruebe la normalidad de los logaritmos naturales de las duraciones. ¿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

Repaso

En el capítulo 4 estudiamos el concepto de distribuciones de probabilidad, pero sólo incluimos las distribuciones *discretas*. En este capítulo estudiamos las distribuciones de probabilidad *continua*, enfocándonos en su categoría más importante: las distribuciones normales. Las distribuciones normales se utilizarán continuamente en los siguientes capítulos.

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 hay una correspondencia conveniente entre áreas y probabilidades. Las áreas específicas pueden encontrarse por medio de la tabla A-2, de una calculadora TI-83 Plus o de un programa de cómputo. (No utilizamos la fórmula 5-1, que es la ecuación utilizada para definir la distribución normal).

En este capítulo 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:

- Puesto que las puntuaciones de CI se distribuyen normalmente, con $\mu = 100$ y $\sigma = 15$, calcule la probabilidad de seleccionar aleatoriamente a un individuo con un CI por arriba de 90.
- Puesto 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 5-4 presentamos el concepto de distribución muestral. La distribución muestral de la media es la distribución de probabilidad de medias de muestra, donde 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 dicho estadístico.

En la sección 5-5 presentamos los siguientes puntos importantes, que se asocian con el teorema del límite central:

1. La distribución de medias de muestra se aproxima a la distribución normal, conforme el tamaño de muestra n se incrementa.
2. La media de las medias de muestra es la media poblacional μ.
3. La desviación estándar de las medias de muestra es $\sigma/\sqrt{n}$.

En la sección 5-6 señalamos que en ocasiones podemos aproximar una distribución de probabilidad binomial con una distribución normal. Si $np \geq 5$ y $nq \geq 5$, la variable aleatoria binomial x se distribuye de manera aproximadamente normal, con la media y la 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.

Finalmente, en la sección 5-7 presentamos un procedimiento para determinar si los datos muestrales parecen provenir de una población con distribución normal. Algunos de los métodos estadísticos que se estudiarán posteriormente en este libro requieren, de forma flexible, de una población que se distribuya normalmente. En estos casos es probable que lo único que se necesite sea el examen de un histograma y de los datos distantes. En otros casos se necesitarían gráficas cuantilares normales, porque es muy estricto el requisito de que la población tenga una distribución normal.

Ejercicios de repaso

1. 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 se basan en el National Health Survey.
   a. 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 que 260, valor que se considera “moderadamente alto”.
   b. 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.
   c. Si se selecciona al azar a nueve hombres entre 18 y 24 años, calcule la probabilidad de que su nivel medio de colesterol sérico esté entre 170 y 200.
   d. 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?

2. Bebés en riesgo El peso de los bebés recién nacidos en Estados Unidos se distribuye normalmente, con una media de 3420 g y una desviación estándar de 495 g (según datos de “Birth Weight and Prenatal Mortality”, de Wilcox *et al*., *Journal of the American Medical Association,* vol. 273, núm. 9).
   a. Se considera que un recién nacido con un peso menor de 2200 g se encuentra en riesgo, porque la tasa de mortalidad de este grupo es al menos del 1%. ¿Qué porcentaje de recién nacidos se encuentra en la categoría “de riesgo”? Si el Chicago General Hospital tiene 900 nacimientos en un año, ¿cuántos de estos bebés se encuentran en la categoría “de riesgo”?
   b. Si redefinimos que un bebé se encuentra en riesgo si su peso al nacer está en el 2% inferior, calcule el peso que se convierte en el punto de corte que separa a los bebés en riesgo de los que no lo están.
   c. Si se seleccionan al azar 16 bebés recién nacidos, calcule la probabilidad de que su peso medio sea mayor de 3700 g.
   d. Si se seleccionan al azar 49 bebés recién nacidos, calcule la probabilidad de que su peso medio esté entre 3300 g y 3700 g.

3. Genes azules Algunas parejas poseen características genéticas que se configuran de manera que una cuarta parte de sus descendientes tienen ojos azules. Se realiza un estudio con 100 parejas en las que se sospechan dichas características; resulta 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 esta probabilidad, ¿parece que la tasa de un cuarto es incorrecta? ¿Por qué?

4. Estatura requerida para hombres de la Marina La Marina de Estados Unidos requiere que los hombres tengan una estatura de entre 64 y 78 pulgadas. (La National Health Survey indica que la estatura de los hombres se distribuye normalmente, con una media de 69.0 pulgadas y una desviación estándar de 2.8 pulgadas).
   a. Calcule el porcentaje de hombres que cumplen con la estatura requerida. ¿Habrá demasiados hombres a quienes se les negará la oportunidad de unirse a la Marina porque son muy bajos o muy altos?
   b. Si a usted se le designa secretario de Defensa y desea modificar el requisito de modo que sólo se rechace al 2% de los hombres más bajos y al 2% de los hombres más altos, ¿cuáles serían las nuevas estaturas mínima y máxima requeridas?
   c. Si se seleccionan 64 hombres al azar, calcule la probabilidad de que su estatura media sea mayor que 68.0 pulgadas.

5. **Distribución uniforme** La San Francisco Supply Company diseñó una máquina que tiene contenedores de café de modo que los contenidos se distribuyen *uniformemente*, con un mínimo de 11.8 onzas y un máximo de 12.2 onzas. Si se selecciona un contenedor al azar, calcule la probabilidad de que la cantidad de café sea
   a. Menor que 12.0 onzas.
   b. Entre 11.2 y 12.7 onzas.
   c. Mayor que 12.2 onzas.
   d. Entre 11.9 y 12.0 onzas.

6. **Distribuciones muestrales**
   a. Se seleccionan al azar muchas muestras diferentes, de tamaño 100, de los pesos de los automóviles que se registran actualmente en Estados Unidos. ¿Qué se concluye acerca de la forma de la distribución de las medias de las distintas muestras?
   b. Si los pesos de todos los automóviles que se registran en Estados Unidos tienen una desviación estándar de 512 libras, ¿cuál es la desviación estándar de las medias de muestra calculadas de muchas muestras diferentes de tamaño 100?
   c. Se seleccionan al azar muchas muestras diferentes, de tamaño 1200, de la población de todos los adultos de Estados Unidos. En cada muestra se registra la proporción de personas que votaron en las últimas elecciones. ¿Qué concluye acerca de la forma de la distribución de estas proporciones de muestra?

7. **Discriminación por género** Cuando a varias mujeres no las contrató la Telektronics Company, se dieron a la tarea de realizar una investigación y encontraron que, entre la gran cantidad de personas que solicitaron empleo, el 30% eran mujeres. Sin embargo, las 20 personas que sí contrataron incluyen sólo dos mujeres y 18 hombres. Calcule la probabilidad de seleccionar al azar 20 personas de un grupo grande de solicitantes (30% de las cuales son mujeres) y obtener dos o menos mujeres. De acuerdo con el resultado, ¿parece que la compañía está discriminando con base en el género?

8. **Prueba de normalidad** Remítase a los pesos de paquetes de azúcar que se listan en el conjunto de datos 28 del Apéndice B. ¿Provienen dichos pesos de una población con distribución normal? Explique.

Ejercicios de repaso acumulativos

1. **Estadísticas de movimientos oculares** La lista de distancias muestreadas (en milímetros) se obtuvo con el uso de un pupilómetro para medir las distancias entre las pupilas de adultos (según datos que reunió un alumno del autor).

   67 66 59 62 63 66 66 55

   a. Calcule la media de las distancias en esta muestra.
   b. Calcule la mediana de las distancias en esta muestra.
   c. Calcule la moda de las distancias en esta muestra.
   d. Calcule la desviación estándar s de esta muestra.
   e. Transforme la distancia de 59 mm a una puntuación z.
   f. Calcule el porcentaje real de los valores de esta muestra que excede los 59 mm.
   g. Suponiendo una distribución normal, calcule los porcentajes de las distancias *poblacionales* que exceden los 59 mm. Use los valores muestrales de x y s como estimados de μ y σ.
   h. ¿Qué nivel de medición (nominal, ordinal, intervalo, razón) describe este conjunto de datos?
   i. Las mediciones listadas parecen redondeadas al milímetro más cercano, pero ¿las distancias exactas que no se redondearon son datos discretos o continuos?

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 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* a través de la aproximación normal de la distribución binomial?
   d. Si se seleccionan al azar grupos de 50 personas, ¿cuál es el número medio de individuos zurdos en estos grupos?
   e. Si se seleccionan al azar grupos de 50 personas, ¿cuál es la desviación estándar del número de personas zurdas en estos grupos?
   f. ¿Sería infrecuente obtener ocho sujetos zurdos en un grupo seleccionado aleatoriamente de 50 personas? ¿Por qué?

Actividades de cooperación en equipo

1. Actividad fuera de clase Forme grupos de tres o cuatro estudiantes. En cada grupo, diseñe 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 dichas muestras tienden a distribuirse *normalmente*, sin importar la naturaleza de la distribución poblacional. En la sección 5-5, por ejemplo, utilizamos los últimos cuatro dígitos de números del seguro social como fuente de muestras de una población de dígitos igualmente probables; procedimos a demostrar que, aun cuando la población original no tenía una distribución normal, las medias de muestras tendían a distribuirse normalmente.

2. Actividad en clase Forme grupos de tres o cuatro estudiantes. Utilice una moneda para simular nacimientos y pida que cada miembro de un grupo simule 25 nacimientos y registre el número de niñas simuladas. Combine todos los resultados del grupo y registre n = número total de nacimientos y x = número de niñas. Con los lotes de n nacimientos, calcule la media y la desviación estándar del número de niñas. ¿Es común o poco común el resultado simulado? ¿Por qué?

3. Actividad en clase Forme grupos de tres o cuatro estudiantes. Ubique los números de lotería en el conjunto de datos 26 del Apéndice B. Hay seis números que se seleccionaron aleatoriamente para cada uno de los 40 diferentes juegos de lotería. Combine los 240 números en un gran conjunto de datos y realice una prueba de normalidad. Después, calcule las 40 medias correspondientes a los 40 diferentes juegos de lotería y realice una prueba de normalidad. ¿Qué concluye? ¿Qué concepto realmente importante se ilustra en este proyecto?

4. Actividad en clase Forme grupos de tres o cuatro estudiantes. Seleccione un conjunto de datos del Apéndice B (excluya los conjuntos de datos 1, 3, 11, 12, 19 y 28, que se utilizaron como ejemplos o ejercicios en la sección 5-7). Aplique los métodos de la sección 5-7 y construya un histograma y una gráfica cuantilar normal; después, determine si el conjunto de datos parece provenir de una población distribuida normalmente.

Proyecto tecnológico

En la sección 5-3 incluimos un ejemplo sobre el diseño de automóviles. La solución en el ejemplo demostró que el 97.72% de los hombres tienen estaturas, al sentarse, menores que 38.8 pulgadas. Esa solución implicó cálculos teóricos que se basaron en el supuesto de que los hombres tienen estaturas, al sentarse, que se distribuyen normalmente, con una media de 36.0 pulgadas y una desviación estándar de 1.4 pulgadas (de acuerdo con datos de una encuesta antropométrica de Gordon, Clauser *et al*.). Este proyecto describe un método de solución diferente, que se basó en una técnica de *simulación*: usaremos una computadora o una calculadora TI-83 Plus para generar aleatoriamente 500 estaturas de hombres sentados (de una población distribuida normalmente con $\mu = 36.0$ y $\sigma = 1.4$), después calcularemos el porcentaje de los pesos simulados que sean menores que 38.8 pulgadas. A continuación se describen los procedimientos para el STATDISK, Minitab, Excel y la calculadora TI-83 Plus.

STATDISK Seleccione **Data** de la barra del menú principal, después elija la opción de **Normal Generator**. Proceda a generar 500 valores con una media de 36.0 y una desviación estándar de 1.4. (Use la opción **Format** para especificar un decimal). Después, ordene los datos con las opciones **Data**, **Sampler Editor** y luego **Format**. Con esta lista es más fácil contar el número de estaturas menores que 38.8 pulgadas. Se divide ese número entre 500 para obtener el porcentaje de las estaturas que se simulan de hombres sentados, que son menores que 38.8. Compare los resultados con el valor teórico de 97.72%, que se calculó en la sección 5-3.

Minitab Seleccione la opción **Calc**, también **Random Data**, luego **Normal**. Introduzca 500 en el número de renglones, C1 en la columna para almacenar los datos, 36.0 en el valor de la media y 1.4 en el valor de la desviación estándar. Ahora seleccione la opción **Manip**, luego **Sort** y proceda a ordenar la columna C1, con la columna ordenada y almacenada en la columna C1 y con el ordenamiento por hacer en la columna C1. Examine los valores en la columna C1 y determine el número de estaturas que se simularon que son menores que 38.8, después divida ese número entre 500 para obtener el porcentaje menor a 38.8 pulgadas. Compare los resultados con el valor teórico de 97.72% que se obtuvo en la sección 5-3.

Excel Seleccione **Tools** de la barra del menú principal, después **Data Analysis** y **Random Number Generation**. Tras hacer clic en **OK**, utilice el cuadro de diálogo para introducir uno para el número de variables y 500 para la cantidad de números aleatorios; tras esto, seleccione “normal” para el tipo de distribución. Introduzca 36.0 en la media y 1.4 en la desviación estándar. Examine los valores que se desplegaron y determine el número de estaturas que se simularon y que sean menores que 38.8, después divida ese número entre 500 para obtener el porcentaje menor a 38.8 pulgadas. Compare los resultados con el valor teórico de 97.72% que se obtuvo en la sección 5-3.

TI-83 Plus Presione **MATH**, luego **PRB**, también introduzca **randNorm** (36.0, 1.4, 500) para generar 500 valores de una población distribuida normalmente, con $\mu = 36.0$ y $\sigma = 1.4$. Presione **STO→L1** para almacenar los datos en la lista L1. Ahora presione **STAT** y luego **SortA(L1)** para ordenar los datos. Examine los datos de la lista L1 para determinar el número de estaturas simuladas que son menores que 38.8, luego hay que dividir ese número entre 500 para obtener el porcentaje menor que 38.8 pulgadas. Compare los resultados con el valor teórico de 97.72% que se obtuvo en la sección 5.3.

de los DATOS a la DECISIÓN

Pensamiento crítico: diseño de un asiento de automóvil

Si el asiento de un coche es muy bajo o muy alto, éste será incómodo y posiblemente peligroso. La mayoría de los asientos son ajustables, de manera que hombres y mujeres con diferentes estaturas al estar sentados pueden seleccionar una posición cómoda. Al diseñar asientos de automóviles, la altura de las rodillas de hombres y mujeres es muy importante. Los hombres tienen alturas de rodillas que se distribuyen normalmente, con una media de 22.0 pulgadas y una desviación estándar de 1.1 pulgadas; las mujeres tienen alturas de rodillas distribuidas normalmente, con una media de 20.3 pulgadas y una desviación estándar de 1.0 pulgadas (según datos de una encuesta antropométrica de Gordon, Churchill *et al.*). Calcule las alturas de rodillas máxima y mínima que incluyan al menos al 95% de todos los hombres y al menos al 95% de todas las mujeres, pero trate de calcular límites costo-beneficio tan cercanos como sea posible. Obtenga el porcentaje de hombres con alturas de rodillas entre los límites que ha determinado; después calcule el porcentaje de mujeres con alturas de rodillas entre los mismos límites. ¿Favorecen sus límites a un género, a expensas del otro? ¿Por qué no es práctico diseñar sencillamente asientos de automóvil que se ajusten a cualquiera? Si usted fuese un ingeniero de diseño de General Motors, ¿qué porcentaje de la población estaría dispuesto a excluir en su diseño de asientos para automóvil? Además de la altura de las rodillas, ¿qué otro componente de diseño importante debe tomarse en cuenta cuando se determina el rango de ajuste de los asientos de automóviles?

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. No importa qué distribución de probabilidad subyace a un experimento, hay una distribución correspondiente de medias que tendrá una forma aproximadamente normal.

http://www.pearsoneducacion.net/triola

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 elemental*, le permitirá hacerlo. 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 está limitado en lo que se puede hacer".

Joel B. Obermayer

Reportero de The News & Observer

Joel B. Obermayer escribe acerca de temas médicos y asuntos de salud para *The News & Observer*, un periódico que cubre la región este de Carolina del Norte. Realiza reportes sobre la administración de la salud, salud pública e investigaciones en centros médicos académicos, que incluyen a Duke University y University of North Carolina en Chapel Hill.

¿Qué conceptos de estadística utiliza?

Utilizo ideas como la significancia estadística, porcentajes de error y probabilidad. No necesito hacer cosas increíblemente sofisticadas, 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 y 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 se supone que bajan de peso disolviendo las células de grasa. Pues bien, yo dudo que estas cremas funcionen. El estudio no era muy bueno tampoco. Estaban tratando de hacer aseveraciones con base en un estudio de sólo 11 personas. El investigador argumentó que 11 individuos eran suficientes para sacar buenas conclusiones empíricas sobre la salud. Eso no fue muy impresionante. Las personas tratan de manipular 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 en nuevas terapias que pueden estar en la etapa de ensayo clínico se está incrementando, en parte por el énfasis en investigaciones sobre el SIDA y en la obtención de nuevos fármacos que aprobarán y recetarán pronto a los pacientes. Es más importante que nunca que un escritor médico utilice la estadística para asegurarse de que los estudios realmente prueban lo que la gente de relaciones públicas asegura que prueban.

¿Deben tener estudios de estadística los prospectos de empleados?

Es posible ser un periodista y no sentirse cómodo con la estadística, pero definitivamente se está limitado en lo que se puede hacer. Si usted escribe acerca de la eficacia de programas educativos que financia 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 lo que se puede hacer una vez que se obtiene el empleo.

6

Estimados y tamaños de muestra

Resultados de la encuesta "cámara vigilante": ¿Qué nos dicen?

El *Star Tribune*, un periódico de Minneapolis-Saint Paul, patrocinó una encuesta que se diseñó para revelar opiniones acerca de la "cámara vigilante", consistente en cámaras que se colocan para identificar conductores que se pasan la luz roja. Las cámaras fotografían las placas de los automóviles que no respetan las luces rojas y, tiempo después, los propietarios de dichos autos reciben las respectivas multas de tránsito por correo. El periódico patrocinó la encuesta porque la legislación pendiente de Minnesota aprobaría el uso de cámaras para expedir multas de tránsito (agradecemos a Beth Hentges, quien proporcionó la información del periódico).

Los encuestadores preguntaron a 829 adultos de Minnesota y encontraron que el 51% se oponía a legalizar las cámaras vigilantes. Estos resultados de encuesta, como la mayoría de los resultados de este tipo, plantean preguntas interesantes como las siguientes:

- Si sólo se encuestó a 829 adultos, ¿sería posible concluir algo acerca de la población de todos los adultos de Minnesota?
- Puesto que sólo 829 adultos fueron encuestados, ¿qué tan precisos son los resultados?
- ¿Es el tamaño de muestra de 829 suficientemente grande como para arrojar resultados significativos?
- ¿Cómo se seleccionó a las personas que respondieron la encuesta? ¿Se seleccionaron de forma que sean representativas de la población?

Las encuestas son un componente importante del modo actual de vida. Afectan directamente los programas de televisión que vemos, los productos que compramos, los funcionarios que elegimos y la ropa que usamos. Si bien ya son parte integral de nuestra vida, por desgracia la mayoría no somos capaces de interpretar correctamente los resultados de las encuestas. Este capítulo contiene los conceptos de estadística que necesitamos para tales interpretaciones. Plantearemos preguntas del tipo de las que listamos. Analizaremos los resultados de la encuesta del *Star Tribune* y, en el proceso, aprenderemos mucho acerca de las encuestas en general.

6-1 Panorama general

En este capítulo empezamos trabajando con el verdadero núcleo de la estadística inferencial, en tanto que usamos datos muestrales para hacer inferencias acerca de las poblaciones. En específico, usaremos datos muestrales para hacer estimados de parámetros de población. Por ejemplo, el problema del capítulo se refiere a los resultados de una encuesta que se aplicó a 829 adultos de Minnesota, el 51% de los cuales se manifestaron contra el uso de cámaras para expedir multas de tránsito. Con base en el estadístico muestral del 51%, estimaremos el porcentaje de adultos en la población de Minnesota que se oponen a la legislación de la cámara vigilante.

Las dos aplicaciones principales de la estadística inferencial implican el uso de datos muestrales para **1.** estimar el valor de un parámetro de la población, y **2.** probar alguna aseveración (o hipótesis) acerca de una población. En este capítulo introducimos métodos para estimar valores de dichos importantes parámetros de población: proporciones, medias y varianzas. También presentamos métodos para determinar los tamaños de muestra necesarios para estimar tales parámetros. En el capítulo 7 introduciremos los métodos básicos para probar las aseveraciones (o hipótesis) que se hicieron acerca de un parámetro de la población.

Este capítulo, al igual que el 7, incluye importantes métodos inferenciales que implican proporciones de población, medias de población y varianzas de población (o desviaciones estándar). En ambos capítulos comenzamos con proporciones por las siguientes razones:

1. Todos vemos proporciones con frecuencia en los medios de comunicación.
2. Por lo general, las personas tienden a interesarse más en datos que se expresan como proporciones.
3. Por lo general, las proporciones son más fáciles de trabajar que las medias o las varianzas, así que nos enfocaremos mejor en los importantes principios de estimación de parámetros y prueba de hipótesis, tan pronto como nos ocupemos de ellos.

6-2 Estimación de la proporción de una población

Una estrategia de estudio: Esta sección contiene mucha información e introduce muchos conceptos. El tiempo que se dedique a esta sección será muy productivo, ya que introducimos el concepto de un intervalo de confianza, concepto general que se aplicará también en las demás secciones de este capítulo. Sugerimos que utilice esta estrategia de estudio: primero, lea la sección con el objetivo limitado de tratar simplemente de entender qué son los intervalos de confianza, para qué sirven y por qué se necesitan. Segundo, trate de desarrollar la habilidad de construir estimados 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 e intente comprender la teoría que subyace. Siempre tendrá una sensación de mayor éxito si entiende lo que está haciendo, en lugar de aplicar a ciegas pasos mecánicos para lograr una respuesta que puede o no tener sentido.

He aquí el principal objetivo de esta sección: dada una proporción de muestra, estimar el valor de la proporción poblacional p. Por ejemplo, el problema del capítulo incluye resultados que se basan en 829 adultos que se encuestaron, de los cuales el 51% se opone al sistema de cámara vigilante que utiliza cámaras para multar a conductores que se pasan la luz roja. El estadístico muestral de 51% puede representarse como la proporción muestral de 0.51. Mediante el uso del tamaño de muestra $n = 829$

y la proporción muestral de 0.51 procederemos a estimar la proporción p de *todos* los adultos de Minnesota que se oponen a la legislación de las cámaras vigilantes.

Esta sección considerará sólo casos en los que la distribución normal puede usarse como aproximación de la distribución muestral de proporciones muestrales. En la sección 5-6 señalamos que en un procedimiento binomial con n ensayos y probabilidad p, si $np \geq 5$ y $nq \geq 5$, entonces la variable aleatoria binomial tiene una distribución de probabilidad que puede aproximarse por medio de una distribución normal (recuerde que $q = 1 - p$). Tales condiciones se incluyen entre los siguientes supuestos que se aplican a los métodos de esta sección.

Supuestos

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 4-3).
3. La distribución normal resulta útil para aproximar la distribución de proporciones muestrales, ya que $np \geq 5$ y $nq \geq 5$ se satisfacen. (Puesto que p y q no se conocen, usaremos la proporción muestral para estimar sus valores. Además, hay procedimientos para tratar con situaciones en las cuales la distribución normal no es una aproximación adecuada. Véase el ejercicio 48).

Recordemos de la sección 1-4 que una muestra aleatoria simple de n valores se obtiene si cada muestra posible de tamaño n tiene la misma probabilidad de seleccionarse. Este requisito de la selección aleatoria significa que los métodos de esta sección no pueden usarse con ningún otro tipo de muestreo, como los muestreos estratificado, por racimos y de conveniencia. Debemos ser especialmente claros acerca de este importante punto:

Los datos reunidos con descuido pueden ser absolutamente inútiles, aunque la muestra sea bastante grande.

Sabemos que muestras diferentes naturalmente producen resultados diferentes. Los métodos de esta sección suponen que esas diferencias muestrales son consecuencia de la posibilidad de fluctuaciones aleatorias, no de algún método insensato de muestreo. Si usted fuese a realizar una encuesta de opinión acerca de las leyes de conducir en estado de ebriedad, seleccionando una muestra de clientes de un bar, no debe usar los resultados para hacer un estimado de la proporción de todos los adultos estadounidenses. Es muy probable que la muestra de clientes del bar sea una muestra sesgada, en el sentido de que no es representativa de todos los estadounidenses.

Suponiendo que hay una muestra aleatoria simple y se satisfacen los demás supuestos que ya se listaron, procedemos con nuestro objetivo principal: el uso de la muestra como base para estimar el valor de la proporción poblacional p. Introducimos la nueva notación $\hat{p}$ (llamada "p sombrero") para la proporción muestral.

Notación para proporciones

p = proporción de la *población*

$\hat{p} = \dfrac{x}{n}$ 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

LA ESTADÍSTICA EN LAS NOTICIAS

Precisión del conteo de votos

La elección presidencial de 2000 se convirtió en la elección presidencial más cerrada en la historia estadounidense y en la primera elección de este tipo en resolverse por decisiones de la Corte. La elección se llevó a cabo el 7 de noviembre de 2000, pero George G. Bush no se determinó como ganador sino hasta el 12 de diciembre de 2000. El retraso se debió en gran medida a los votos que se impugnaron en el cambiante estado de Florida, donde los recuentos levantaron graves polémicas sobre su exactitud. El 17 de noviembre de 2000, Ford Fessenden y Christopher Drew, reporteros del *New York Times*, escribieron que "las personas que venden los sistemas de votación... dicen que las máquinas pueden ser, en condiciones ideales, 99.99% exactas... El fabricante de un tipo de lector de tarjetas aseguró que la exactitud de su máquina sería de 99.9%...".
Si bien la cuenta final de 50,996,582 votos para Bush y de 50,456,062 votos para Al Gore parece muy precisa, los sistemas de votación causaron que dichos totales fuesen estimados, no conteos precisos.

Muestra pequeña

El Children's Defense Fund se organizó para promover el bienestar de los niños. El grupo publicó *Out of School in America*, donde se informó que, en un área, el 37.5% de los jóvenes entre 16 y 17 años ya no asistían a la escuela. Esta estadística recibió mucha cobertura de los medios de comunicación, pero se basó 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).

Proporción, probabilidad y porcentaje Aunque esta sección se enfoca en la proporción poblacional p, los procedimientos que aquí se analizan pueden aplicarse también a probabilidades o porcentajes, pero los porcentajes deben convertirse a proporciones quitando el signo porcentual y dividiendo entre 100. Por ejemplo, el 51% se expresa en forma decimal como 0.51. El símbolo p puede, por lo tanto, representar una proporción, una probabilidad o el equivalente decimal de un porcentaje. Por ejemplo, si usted entrevista a 200 estudiantes de estadística y encuentra que 80 de ellos compraron calculadoras TI-83 Plus, entonces la proporción muestral es $\hat{p} = x/n = 80/200 = 0.400$ y $\hat{q} = 0.600$ (calculada de $1 - 0.400$). En lugar de calcular el valor de x/n, en ocasiones el valor de $\hat{p}$ ya se conoce, puesto que la proporción muestral o porcentaje se da directamente. Por ejemplo, si se reporta que se encuestaron 829 adultos de Minnesota y el 51% de ellos se oponen a la ley de la cámara vigilante, entonces $\hat{p} = 0.51$ y $\hat{q} = 0.49$.

Si queremos estimar una proporción de una población con un solo valor, el mejor estimado es $\hat{p}$. Puesto que $\hat{p}$ consiste en un solo valor, se llama un 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 porque es el más consistente de los estimadores que puede usarse. No está sesgado en el sentido de que la distribución de las proporciones muestrales tiende al centro para el valor de p; esto es, en las proporciones muestrales $\hat{p}$ no tiende sistemáticamente a subestimar ni a sobreestimar p. (Véase sección 5-4). La proporción muestral $\hat{p}$ es el estimador más consistente en el sentido de que la desviación estándar de la proporción muestral tiende a ser menor que la desviación estándar de cualquier otro estimador sin sesgo.

EJEMPLO Respuestas de la encuesta de la cámara vigilante En el problema del capítulo señalamos que se entevistó a 829 adultos de Minnesota y que el 51% de ellos se opone al uso de la cámara vigilante para expedir multas de tránsito. Utilizando estos resultados de encuesta, encuentre el mejor estimado puntual de la proporción de todos los adultos de Minnesota que se oponen al uso de la cámara vigilante.

SOLUCIÓN Puesto que la proporción muestral es el mejor estimado puntual de la proporción de la población, concluimos que el mejor estimado puntual de p es 0.51. Cuando se usan los resultados de encuesta para estimar el porcentaje de todos los adultos de Minnesota que se oponen al uso de la cámara vigilante, nuestro mejor estimado es 51%.

¿Por qué necesitamos intervalos de confianza?

En el ejemplo anterior vimos que 0.51 era nuestro *mejor* estimado puntual de la proporción poblacional p, pero no tenemos indicación precisa de qué tan *bueno* era nuestro mejor estimado. Si tuviésemos una muestra de sólo 20 adultos de Minnesota y 12 se opusieran al uso de la cámara vigilante, nuestro mejor estimado puntual sería la proporción muestral de $12/20 = 0.6$, pero no esperaríamos que este estimado puntual sea muy bueno, puesto que se basa en una muestra tan pequeña como ésa. Ya que el estimado puntual tiene el grave defecto de no revelar nada acerca de qué tan bueno es, los estadísticos diseñaron ingeniosamente otro tipo de estimado. Este estimado, que se conoce como *intervalo de confianza* o *estimado del intervalo*, consiste en un rango (o un intervalo) de valores en lugar de un solo valor.

Definición

Un **intervalo de confianza** (o **estimado del intervalo**) es una gama (o un intervalo) de valores que se usan 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 sucesos 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 veces se expresa 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 conoce como **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 de 95% es la más común, puesto que provee un buen balance entre precisión (como se refleja en el ancho del intervalo de confianza) y confiabilidad (como se expresa por el nivel de confianza).

A continuación se presenta un ejemplo de un intervalo de confianza que se basa en los datos muestrales de 829 adultos de Minnesota que se encuestaron, 51% de los cuales se oponen al uso de la cámara vigilante:

El intervalo de confianza estimado de 0.95 (o 95%) de la proporción poblacional p es $0.476 < p < 0.544$.

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 predispuestas que se diseñan para desacreditar a dichos 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, sabiendo que el gobernador Romer, al entrar 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 los sondeos de empuje como poco éticos, pero algunos encuestadores profesionales no censuran la práctica en tanto las preguntas no incluyan mentiras directas.

6 Interpretación de un intervalo de confianza

Debemos ser cuidadosos para interpretar los intervalos de confianza correctamente. Hay una interpretación correcta, así como muchas diferentes y creativas interpretaciones erróneas del intervalo de confianza $0.476 < p < 0.544$.

Correcta: "Tenemos una confianza del 95% de que el intervalo de 0.476 a 0.544 realmente contiene el valor verdadero de p". Lo anterior significa que si seleccionamos muchas muestras diferentes de tamaño 829 y construimos los intervalos de confianza correspondientes, el 95% de ellos podrían contener realmente el valor de la proporción poblacional p. (Nótese que en esta interpretación correcta, el nivel de 95% se refiere a la tasa de éxitos del *proceso* que se utiliza para estimar la proporción y no a la proporción de la población en sí).

Errónea: "Existe un 95% de posibilidades de que el valor real de p esté entre 0.476 y 0.544".

Para cualquier punto específico en el tiempo, hay un valor fijo y constante de p, la proporción de todos los adultos de Minnesota que se oponen al uso de la cámara vigilante. Si utilizamos datos muestrales para encontrar límites específicos, como 0.476 y 0.544, tales límites abarcan o no la proporción poblacional p, por lo cual no podemos determinar si lo hacen o no sin conocer el valor real de p. Es erróneo decir que p tiene un 95% de posibilidades de estar entre los límites específicos de 0.476 y 0.544, puesto que p es una constante fija (pero desconocida), no una variable aleatoria, p estará entre estos límites o no; no existe una probabilidad implicada. Éste es un concepto confuso, así que considere el ejemplo más fácil donde queremos encontrar la probabilidad de que un bebé que nace sea niña. Si el bebé ya nació, pero el médico todavía no ha anunciado el género, no podemos asegurar que hay un 0.5 de probabilidad de que el bebé sea una niña, porque el bebé es ya una niña o no. No existe posibilidad implicada, puesto que el género ya se determinó. De manera similar, una proporción poblacional p ya se determinó, y los límites del intervalo de confianza contienen p o no; entonces, es erróneo decir que hay un 95% de posibilidades de que p esté entre 0.476 y 0.544.

Un nivel de confianza de 95% nos dice que el proceso que usaremos, a la larga, derivará en límites del intervalo de confianza que contienen la proporción de la población real el 95% del tiempo. Suponga que la proporción real de todos los adultos de Minnesota que se oponen a la cámara vigilante es $p = 0.520$. Entonces, el intervalo de confianza que se obtuvo de los datos muestrales dados contendría la proporción de la población, puesto que la proporción real de la población 0.520 está entre 0.476 y 0.544. Lo anterior se ilustra en la figura 6-1, que muestra el primer intervalo de confianza para los datos de encuesta reales dados en el problema del capítulo (con el 51% de 829 personas que se entrevistaron quienes se oponen a la cámara vigilante), pero los otros 19 intervalos de confianza representan muestras hipotéticas. Con un 95% de confianza, esperamos que 19 de las 20 muestras resulten en intervalos de confianza que contienen el valor real de p; la figura 6-1 ilustra esto con 19 de los intervalos de confianza que contienen p, mientras un intervalo de confianza no contiene p.

Cuidado: Los intervalos de confianza pueden usarse de manera informal para comparar conjuntos de datos diferentes, pero el traslape de intervalos de confianza no debe manejarse para elaborar conclusiones formales y finales acerca de la igualdad de las proporciones. El análisis de traslape entre dos intervalos de

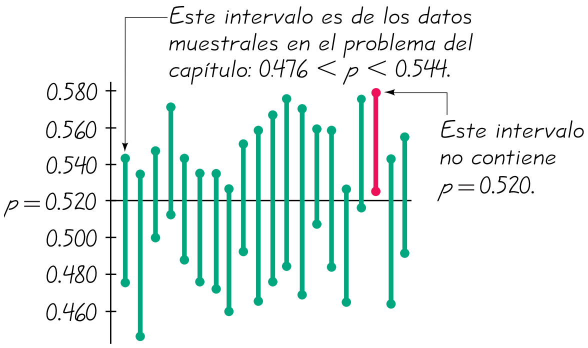


FIGURA 6-1 Intervalos de confianza de 20 muestras diferentes

confianza individuales se asocia con dificultades descritas en capítulos posteriores. (Véase "On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals", de Schenker y Gentleman, *The American Statistician*, vol. 55, núm. 3). En los capítulos siguientes describiremos procedimientos para determinar si las poblaciones tienen proporciones iguales, aunque esos procedimientos no tendrán las dificultades que se asocian con las conclusiones que se basan en el traslape de intervalos de confianza.

No use el *traslape* de intervalos de confianza como base para sacar conclusiones finales acerca de la igualdad de las proporciones.

Valores críticos

Los métodos de esta sección y muchos de los otros métodos estadísticos que se encuentran en los capítulos siguientes incluyen el uso de una puntuación z estándar que puede emplearse para distinguir entre estadísticos muestrales que tienen posibilidades de ocurrir y aquellos que son improbables. Una puntuación z de este tipo se llama *valor crítico* (que se definirá después). Los valores críticos se basan en las siguientes observaciones:

1. Sabemos, desde la sección 5-6, que en ciertas condiciones la distribución muestral de las proporciones muestrales puede ser aproximada por una distribución normal, como en la figura 6-2.
2. Las proporciones muestrales tienen una posibilidad relativamente pequeña (con probabilidad que se denota por α) de caer en una de las colas que se sombreó de la figura 6-2.
3. Denotando el área de cada cola sombreada por $\alpha/2$, vemos que hay una probabilidad total de α de que una proporción muestral caiga en cualquiera de las dos colas que se sombrearon.
4. Por la regla de los complementos (del capítulo 3), concluimos que hay una probabilidad de $1 - \alpha$ de que una proporción muestral caiga dentro de la región interior que se sombreó en la figura 6-2.
5. La puntuación z que separa la región de la cola derecha se denota por lo común por $z_{\alpha/2}$, y se conoce como *valor crítico*, puesto que está en la frontera que separa proporciones muestrales que son probables de ocurrir de aquellas que no son probables.

Dichas observaciones pueden formalizarse con la notación y definición siguientes.

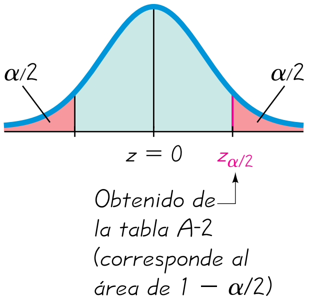


FIGURA 6-2 Valor crítico $z_{\alpha/2}$ en la distribución normal estándar

Notación para el valor crítico

El valor crítico $z_{\alpha/2}$ es el valor z positivo que está en la frontera vertical que separa un 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 un área de $\alpha/2$ en la cola derecha de la distribución normal estándar.

Definición

Un **valor crítico** es el número que está en la frontera que separa las estadísticas de muestra que probablemente ocurrirán, de aquellos que no tienen posibilidades de ocurrir. El número $z_{\alpha/2}$ es un valor crítico que es una puntuación z con la propiedad de que separa un área de $\alpha/2$ en la cola derecha de la distribución normal estándar. (Véase la figura 6-2).

EJEMPLO Calcular un valor crítico Calcule el valor crítico $z_{\alpha/2}$ que corresponde a un nivel de confianza del 95%.

SOLUCIÓN *Cuidado*: Para calcular el valor crítico z para un grado de confianza del 95%, *no* busque 0.95 en el cuerpo de la tabla A-2. Un nivel de confianza del 95% corresponde a $\alpha = 0.05$. Véase la figura 6-3, donde mostramos que el área en cada una de las colas que se sombreó es $\alpha/2 = 0.025$. Calculamos $z_{\alpha/2} = 1.96$, señalando que toda el área a su derecha debe ser $1 - 0.025$, o 0.975. Es posible remitirnos a la tabla A-2 y 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 lo tanto, $z_{\alpha/2} = 1.96$. Línea inferior: para encontrar la puntuación z crítica para un nivel de confianza de 95%, busque 0.9750 en el cuerpo la tabla A-2, no 0.95.

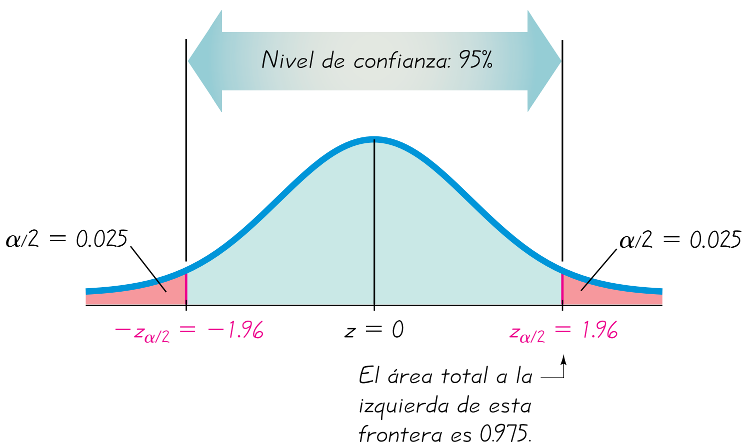


FIGURA 6-3 Cálculo de $z_{\alpha/2}$ para un nivel de confianza del 95%

El ejemplo anterior mostró que un nivel de confianza del 95% resulta en 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 tabla siguiente.

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 de encuesta dados en el problema del capítulo (con el 51% de 829 personas que se oponen a la cámara vigilante), calculamos la proporción muestral $\hat{p}$ y dicha proporción muestral es diferente, por lo regular, de la proporción poblacional p. La diferencia entre la proporción muestral y la proporción de la población puede considerarse 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 **margen de error**, que se denota con $\boldsymbol{E}$, es la diferencia máxima probable (con probabilidad $1 - \alpha$) entre la proporción muestral $\hat{p}$ que se observa y el valor real de la proporción poblacional p. El margen de error E también se conoce como *error máximo del estimado* y se calcula multiplicando el valor crítico por la desviación estándar de las proporciones muestrales, como se muestra en la fórmula 6-1.

Fórmula 6-1 $$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$ margen de error para proporciones

Dada la forma en que se define el margen de error E, hay una probabilidad de $1 - \alpha$ de que una proporción muestral será errónea (diferente de la proporción poblacional p) por no más de E, y existe una probabilidad α de que la proporción muestral será errónea por más de E.

Intervalo de confianza (o estimado de intervalo) para la proporción poblacional *p*

$$\hat{p} - E < p < \hat{p} + E \qquad \text{donde} \qquad E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

El intervalo de confianza suele expresarse en los formatos equivalentes siguientes:

$$\hat{p} \pm E$$

o

$$(\hat{p} - E, \hat{p} + E)$$

En el capítulo 3, cuando las probabilidades se daban en forma decimal, redondeábamos a tres dígitos significativos. Aquí utilizaremos 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, es posible 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 que se requieren se satisfacen. (La muestra es aleatoria simple, las condiciones para la distribución binomial se satisfacen y la distribución normal puede utilizarse para aproximar la distribución de las proporciones muestrales, puesto que $np \geq 5$ y $nq \geq 5$ se satisfacen).
2. Remítase a la tabla A-2 y encuentre el valor crítico $z_{\alpha/2}$ que corresponde al nivel de confianza que se desea. (Por ejemplo, si el nivel de confianza es 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 que ya se conoce 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 Respuestas de encuesta sobre la cámara vigilante En el problema del capítulo señalamos que se encuestaron 829 adultos de Minnesota y que el 51% de ellos se oponen al uso de la cámara vigilante para expedir multas de tránsito. En un ejemplo previo, señalamos que el mejor estimado puntual de la proporción de la población es 0.51. Use estos mismos resultados de encuesta para lo siguiente:

a. Encuentre el margen de error E que corresponde a un nivel de confianza del 95%.
b. Calcule el estimado del intervalo de confianza del 95% de la proporción poblacional p.
c. Con base en los resultados, ¿concluiríamos con seguridad que la mayoría de los adultos de Minnesota se oponen al uso de la cámara vigilante?

SOLUCIÓN Primero debemos verificar que los supuestos que se requieren se satisfagan. Suponiendo que la muestra es aleatoria simple, vemos que las condiciones para una distribución binomial se satisfacen. Con $n = 829$ y $\hat{p} = 0.51$, tenemos $n\hat{p} = 422.79 \geq 5$ y $n\hat{q} = 406.21 \geq 5$; entonces, la distribución normal puede utilizarse como aproximación de la distribución binomial.

a. El margen de error se calcula usando la fórmula 6-1 con $z_{\alpha/2} = 1.96$ (como se calculó en el ejemplo anterior), $\hat{p} = 0.51$, $\hat{q} = 1 - 0.51 = 0.49$, y $n = 829$. Los dígitos de más se utilizan para que el error de redondeo se minimice en los límites del intervalo de confianza que se calcula en el inciso *b*.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{(0.51)(0.49)}{829}} = 0.03403000$$

b. Construir el intervalo de confianza es bastante fácil ahora que tenemos los valores de $\hat{p}$ y E. Simplemente sustituimos dichos valores para obtener este resultado:

$$\hat{p} - E < p < \hat{p} + E$$

$$0.51 - 0.03403000 < p < 0.51 + 0.03403000$$

$$0.476 < p < 0.544 \text{ (redondeado a tres dígitos significativos)}$$

El mismo resultado podría expresarse en el formato de 0.51 ± 0.034 o (0.476, 0.544). Si queremos el intervalo de confianza del 95% para el *porcentaje* de la población real, es posible expresar el resultado como $47.6\% < p < 54.4\%$. Tal intervalo de confianza suele reportarse con una afirmación como ésta: "Se estima que el 51% de los adultos de Minnesota se oponen al uso de la cámara vigilante, con un margen de error de más o menos 3.4 puntos porcentuales". Dicha declaración es una expresión verbal del formato para el intervalo de confianza: $51\% \pm 3.4\%$. El nivel de confianza debe reportarse también, pero rara vez se hace en los medios de comunicación. Los medios de comunicación usan normalmente un nivel de confianza del 95%, pero omiten cualquier referencia a él. Sin embargo, la información provista por el *Star Tribune* acerca de esta encuesta incluyó la aseveración de que "el margen máximo de error muestral por porcentajes que se basan en 829, es de 3.4 puntos porcentuales, más o menos, a un nivel de confianza del 95%, si no se incluye el efecto del diseño muestral". ¡Bien hecho, *Star Tribune*!

c. Con base en los resultados de la encuesta, tenemos una confianza del 95% de que los límites de 47.6% y 54.4% contienen el porcentaje real de adultos de Minnesota que se oponen a la cámara vigilante. Es probable que el porcentaje de adultos de Minnesota que se oponen sea cualquier valor entre 47.6% y 54.4%. Sin embargo, una mayoría requiere un porcentaje mayor que el 50%; entonces, *no podemos* concluir con seguridad que la mayoría se oponen (puesto que el intervalo de confianza completo no es mayor que el 50%).

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), utilizamos los resultados de la sección 5-6 para concluir que μ y α 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:

$$\text{Media de proporciones muestrales: } \mu = \frac{np}{n} = p$$

$$\text{Desviación estándar de proporciones muestrales: } \sigma = \frac{\sqrt{npq}}{n} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}$$

El primer resultado parecería trivial, puesto que ya estipulamos que la porción de la población real es p. El segundo resultado no es trivial y es 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 6-1 para el margen de error refleja el hecho de que $\hat{p}$ tiene una probabilidad de $1 - \alpha$ de estar entre $z_{\alpha/2}\sqrt{pq/n}$ de p. El intervalo de confianza para p, como se dio previamente, refleja el hecho de que hay 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 de la muestra

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 6-1), y luego resolvemos para n, obtendremos la fórmula 6-2. La fórmula 6-2 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 $\hat{q}$ por 0.5, con el resultado de la fórmula 6-3.

Tamaño de muestra para la estimación de la proporción *p*

Cuando se conoce un estimado $\hat{p}$: **Fórmula 6-2** $n = \dfrac{[z_{\alpha/2}]^2\,\hat{p}\hat{q}}{E^2}$

Cuando se desconoce el estimado $\hat{p}$: **Fórmula 6-3** $n = \dfrac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2}$

Regla de redondeo para determinar el tamaño de muestra

Para asegurar que el tamaño de muestra requerido sea al menos tan grande como debe ser, si el tamaño de muestra que se calculó no es un número entero, redondee al siguiente número entero *mayor*.

Utilice la fórmula 6-2 cuando se puedan hacer estimados razonables de $\hat{p}$ usando muestras previas, un estudio piloto o el conocimiento experto de alguna persona. Cuando no se pueden hacer estimados de este tipo, asignamos el valor de 0.5 a $\hat{p}$ y a $\hat{q}$ para que el tamaño de la muestra resultante sea al menos tan grande como debe ser. La razón que sustenta la asignación de 0.5 es ésta: el producto $\hat{p} \cdot \hat{q}$ tiene 0.25 como su mayor valor posible, que ocurre cuando $\hat{p} = 0.5$ y $\hat{q} = 0.5$. (Ensaye experimentando con diferentes valores de $\hat{p}$ para verificar que $\hat{p} \cdot \hat{q}$ tenga 0.25 como el mayor valor posible). Observe que las fórmulas 6-2 y 6-3 no incluyen el tamaño poblacional N, así que el tamaño de la población es irrelevante.

(*Excepción:* Cuando el muestreo es sin reemplazo y de una población finita relativamente pequeña. Véase el ejercicio 46).

EJEMPLO **Tamaño de muestra para una encuesta por correo electrónico** Las formas en las que nos comunicamos se afectaron en forma drástica por el uso de máquinas contestadoras telefónicas, fax, correo de voz y correo electrónico. Suponga que un sociólogo quiere determinar el porcentaje actual de hogares de Estados Unidos que utiliza el correo electrónico. ¿Cuántos hogares deben encuestarse para tener una confianza del 95% de que el porcentaje muestral es erróneo por no más de cuatro puntos porcentuales?

a. 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*).

b. Suponga que no tenemos información previa que sugiera un posible valor de $\hat{p}$.

SOLUCIÓN

a. El estudio previo sugiere que $\hat{p} = 0.169$, entonces $\hat{q} = 0.831$ (que se calculó 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 hay un valor estimado de $\hat{p}$, usamos la fórmula 6-2 como sigue:

$$n = \frac{[z_{\alpha/2}]^2\,\hat{p}\hat{q}}{E^2} = \frac{[1.96]^2(0.169)(0.831)}{0.04^2}$$

$$= 337.194 = 338 \qquad \text{(redondeado)}$$

Debemos encuestar al menos 338 hogares seleccionados aleatoriamente.

b. 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 6-3 como sigue.

$$n = \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2} = \frac{[1.96]^2 \cdot 0.25}{0.04^2}$$

$$= 600.25 = 601 \qquad \text{(redondeado)}$$

INTERPRETACIÓN Para tener un 95% de confianza 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 de muestra de 338 que se calculó en el inciso *a*, veremos que si no tenemos conocimiento de un estudio previo, se requiere una muestra más grande para obtener los mismos resultados que cuando se puede estimar el valor de $\hat{p}$. Pero ahora usemos algo de sentido común: sabemos que el uso del correo electrónico está creciendo tan rápidamente que lo que se estimó en 1997 es muy viejo como para ser de utilidad. El día de hoy, mucho más del 16.9% de los hogares utilizan correo electrónico. Siendo realistas, necesitamos una muestra mayor que 338 hogares. Suponiendo que en verdad no conocemos la tasa actual de uso de correo electrónico, habría que seleccionar al azar a 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 se calcula el tamaño de muestra usando las fórmulas 6-2 o 6-3, asegúrese de sustituir la puntuación z crítica por $z_{\alpha/2}$. Por ejemplo, si usted trabaja con una confianza del 95%, asegúrese de reemplazar $z_{\alpha/2}$ con 1.96. (He aquí la secuencia lógica: $95\% \Rightarrow \alpha = 0.05 \Rightarrow z_{\alpha/2} = 1.96$ encontrado en la tabla A-2). No cometa el error de reemplazar $z_{\alpha/2}$ por 0.95 o 0.05 ni el de usar $E = 4$ en lugar de $E = 0.04$, que causa que el tamaño de la muestra sea $1/10{,}000$ de lo que debe ser, pues se puede terminar con un tamaño de muestra de sólo 1 cuando la respuesta se redondea. Es imposible estimar realmente una proporción de una población encuestando a una sola persona (aunque hay individuos que afirman que lo saben todo).

Tamaño de población El inciso *b* del ejemplo precedente requirió la aplicación de la fórmula 6-3, la misma fórmula que utilizan con frecuencia Nielsen, Gallup y otros encuestadores profesionales. Mucha gente cree, incorrectamente, que el tamaño de la muestra debe ser algún porcentaje de la población, pero la fórmula 6-3 nos enseña 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 46). La mayoría de las encuestas que aparecen en periódicos, revistas y transmisiones de medios de comunicación implican tamaños de la muestra en el rango de 1000 a 2000. Aunque encuestas como éstas quizá se basan en un porcentaje muy pequeño de la población total, a veces ofrecen resultados que son bastante buenos. Cuando Nielsen entrevista a 4000 hogares que tienen televisión de una nación de 104 millones de hogares, sólo se entrevista al 0.004% de los hogares; aun así, podemos tener una confianza del 95% de que el porcentaje de la muestra estará dentro de un punto porcentual del porcentaje verdadero de la población.

Calcular el estimado puntual y E desde un intervalo de confianza Algunas veces queremos entender mejor un intervalo de confianza que se obtuvo de un artículo de una revista o que se generó 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:

$$\hat{p} = \frac{(\text{límite de confianza más alto}) + (\text{límite de confianza más bajo})}{2}$$

margen de error:

$$E = \frac{(\text{límite de confianza más alto}) - (\text{límite de confianza más bajo})}{2}$$

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 abstuvieron de fumar durante ocho semanas [intervalo de confianza (IC) de 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 de 95% es $0.58 < p < 0.81$. El estimado puntual es el valor medio entre los límites del intervalo de confianza más alto y más bajo; entonces, tenemos

$$\hat{p} = \frac{(\text{límite de confianza más alto}) + (\text{límite de confianza más bajo})}{2}$$

$$= \frac{0.81 + 0.58}{2} = 0.695$$

El margen de error se calcula como sigue:

$$E = \frac{(\text{límite de confianza más alto}) - (\text{límite de confianza más bajo})}{2}$$

$$= \frac{0.81 - 0.58}{2} = 0.115$$

Utilizando la tecnología para intervalos de confianza

STATDISK Seleccione **Analysis**, luego **Confidence Intervals**, después **Population Proportion.** 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 que se desea (el predeterminado es 95%), y haga clic en el cuadro con esta frase: “Use test and interval based on normal distribution”.

Excel Utilice el complemento de programa 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 elija **Summ 1 Var Prop Interval** (que es la forma abreviada de “intervalo de confianza para una proporción utilizando datos que se resumen para una variable”). Haga clic en el icono que muestra un lápiz para “Num successes” e ingrese A1. Haga clic en el icono que muestra un lápiz para “Num trials” e ingrese B1. Haga clic en **OK**. En el cuadro de diálogo, seleccione el nivel de confianza, y después haga clic en **Compute Interval**.

TI-83 Plus Oprima **STAT**, seleccione **TESTS**, luego **1-PropZInt** y proceda a ingresar los elementos que se piden.

Utilizando la tecnología para la determinación del tamaño de muestra

STATDISK Seleccione **Analysis**, luego **Sample Size Determination**, y **Estimate Proportion**. Proceda a ingresar los elementos que se le piden en el cuadro de diálogo.

La determinación del tamaño de muestra no está disponible como una función incluida en Minitab, Excel o la calculadora TI-83 Plus.

6-2 Destrezas y conceptos básicos

Calcular valores críticos. *En los ejercicios 1 a 4, calcule el valor crítico $z_{\alpha/2}$ que corresponde al nivel de confianza dado.*

1. 99%

2. 90%

3. 98%

4. 92%

5. Exprese el intervalo de confianza $0.220 < p < 0.280$ en la forma de $\hat{p} \pm E$.

6. Exprese el intervalo de confianza $0.456 < p < 0.496$ en la forma de $\hat{p} \pm E$.

7. Exprese el intervalo de confianza (0.604, 0.704) en la forma de $\hat{p} \pm E$.

8. Exprese el intervalo de confianza 0.742 ± 0.030 en la forma de $\hat{p} - E < p < \hat{p} + E$.

Interpretación de los límites del intervalo de confianza. *En los ejercicios 9 a 12, utilice los límites de intervalo de confianza dados para calcular el estimado puntual $\hat{p}$ y el margen de error E.*

9. (0.444, 0.484)

10. $0.278 < p < 0.338$

11. $0.632 < p < 0.678$

12. $0.887 < p < 0.927$

Calcular el margen de error. *En los ejercicios 13 a 16, suponga que una muestra se utiliza para estimar una proporción poblacional p. Calcule el margen de error E que corresponde al estadístico y el nivel de confianza dados.*

13. $n = 800$, $x = 200$, 95% de confianza

14. $n = 1200$, $x = 400$, 99% de confianza

15. 99% de confianza; el tamaño de la muestra es 1000, del cual el 45% son éxitos.

16. 95% de confianza; el tamaño de la muestra es 500, del cual el 80% son éxitos.

Construcción de intervalos de confianza. *En los ejercicios 17 al 20, use los datos muestrales y el nivel de confianza para construir el intervalo de confianza que se estimó de la proporción poblacional p.*

17. $n = 400$, $x = 300$, 95% de confianza

18. $n = 1200$, $x = 200$, 99% de confianza

19. $n = 1655$, $x = 176$, 98% de confianza

20. $n = 2001$, $x = 1776$, 90% de confianza

Determinación del tamaño *de la muestra**.*** *En los ejercicios 21 a 24, utilice los datos dados para calcular el tamaño de muestra mínimo que se requiere para estimar una proporción o un porcentaje de una población.*

21. Margen de error: 0.060; nivel de confianza: 99%; $\hat{p}$ y $\hat{q}$ *desconocidas*

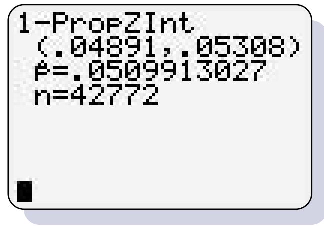
22. Margen de error: 0.038; nivel de confianza: 95%; $\hat{p}$ y $\hat{q}$ *desconocidas*

23. Margen de error: cinco puntos porcentuales; nivel de confianza: 95%; de un estudio previo, $\hat{p}$ se estima por el equivalente decimal del 18.5%.

24. Margen de error: tres puntos porcentuales; nivel de confianza: 90%; de un estudio previo, $\hat{p}$ se estima por el equivalente decimal del 8%.

25. Interpretación de pantalla de calculadora El Insurance Institute of America quiere estimar el porcentaje de conductores de 18 a 20 años de edad que conducen un automóvil en estado de ebriedad. En un estudio grande, se entrevistaron 42,772 hombres de 18 a 20 años de edad; el 5.1% de ellos dijeron haber conducido, en el mes anterior, en estado de ebriedad (según datos de "Prevalence of Alcohol-Impaired Driving", de Liu, Siegel *et al*., *Journal of the American Medical Association*, vol. 277, núm. 2). Utilizando los datos muestrales y un nivel de confianza del 95%, la pantalla de la calculadora TI-83 Plus queda como se muestra.

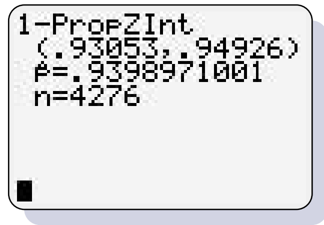
TI-83 Plus



a. Escriba una afirmación que interprete el intervalo de confianza.
b. Con base en el resultado anterior, ¿parece que el conducir en estado de ebriedad es un problema para los hombres de 18 a 20 años de edad? (Todos los estados de Estados Unidos prohíben ahora la venta de alcohol a personas menores de 21 años de edad).
c. Cuando se determinan tasas de seguros para conductores varones de 18 a 24 años de edad, ¿qué porcentaje de conductores en estado de ebriedad usaría usted si trabajara para la compañía de seguros y quisiera ser conservador utilizando el peor de los casos posible?

26. Interpretación de pantalla de calculadora En 1920 sólo el 35% de los hogares de Estados Unidos tenía teléfono, pero esta tasa es ahora mucho más alta. Una encuesta reciente de 4276 hogares seleccionados al azar mostró que 4019 de ellos tienen teléfono (de acuerdo con datos del Census Bureau de Estados Unidos). Utilizando dichos resultados de encuesta y un nivel de confianza del 99%, la pantalla de la calculadora TI-83 Plus aparece como se muestra.

TI-83 Plus



a. Escriba un enunciado que interprete el intervalo de confianza.
b. Con base en el resultado anterior, ¿los encuestadores deberían preocuparse por los resultados de las encuestas por teléfono?

27. Compras en Internet En una encuesta de Gallup se encuestaron 1025 adultos, que se seleccionaron aleatoriamente; el 29% de ellos dijeron que usaban Internet para comprar al menos cinco veces al año.
a. Calcule el estimado puntual del *porcentaje* de adultos que usan Internet para hacer compras.
b. Encuentre un estimado del intervalo de confianza del 99% del *porcentaje* de adultos que usan Internet para hacer compras.
c. Si un almacén tradicional de ventas al menudeo quiere 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 compradores por Internet debe utilizarse?

28. Encuesta de la pena de muerte En una encuesta de Gallup, que se realizó entre 491 adultos seleccionados al azar, se les preguntó si estaban a favor de la pena de muerte para una persona convicta por homicidio; el 65% de ellos dijeron que estaban a favor.
a. Calcule el estimado puntual del porcentaje de adultos que están a favor de la pena de muerte.
b. Calcule un estimado de intervalo de confianza del 95% de adultos que están a favor de la pena de muerte.
c. ¿Podemos concluir con seguridad que la mayoría de los adultos están a favor de la pena de muerte? Explique.

29. Genética mendeliana Cuando Mendel realizó sus famosos experimentos genéticos con chícharos, una muestra de vástagos consistió en 428 chícharos verdes y 152 chícharos amarillos.
a. Calcule un estimado del intervalo de confianza del 95% del porcentaje de chícharos amarillos.
b. Con base en su teoría genética, Mendel esperaba que el 25% de los chícharos vástagos fueran amarillos. Puesto que el porcentaje de chícharos vástagos amarillos no es el 25%, ¿contradicen los resultados la teoría de Mendel? ¿Por qué?

30. Respuestas de encuesta confusas En una encuesta de 1002 personas, 701 dijeron que votaron 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.
a. Calcule un estimado del intervalo de confianza del 99% de la proporción de personas que dijeron que votaron.
b. ¿Son consistentes los resultados de encuesta con los votos reales del 61%? ¿Por qué?

31. Prueba de fármacos El fármaco Ziac se utiliza para tratar la hipertensión. En un ensayo clínico, el 3.2% de 221 usuarios de Ziac experimentaron mareo (según datos de Lederle Laboratories).

a. Construya un estimado de intervalo de confianza del 99% del porcentaje de todos los usuarios de Ziac que experimentaron mareo.

b. En el mismo ensayo clínico, las personas en el grupo placebo no tomaron Ziac, pero el 1.8% de ellos reportaron mareo. Con base en el resultado del inciso *a*, ¿qué podemos concluir acerca del mareo como una reacción adversa al Ziac?

32. Tabaquismo y educación universitaria La industria del tabaco vigila de cerca todas las encuestas relacionadas con el tabaquismo. Una encuesta mostró que entre 785 sujetos que se seleccionaron al azar y completaron cuatro años de universidad, el 18.3% fuma (con base en datos de la American Medical Association).

a. Construya el intervalo de confianza del 98% para el porcentaje real de fumadores, entre todas las personas que completaron cuatro años de universidad.

b. Con base en el resultado del inciso *a*, ¿parece ser sustancialmente diferente la tasa de tabaquismo en individuos con cuatro años de universidad, de la tasa de 27% de la población general?

33. Tamaño de muestra para compras por Internet Muchos estados consideran con cuidado los pasos que les ayudarían a recolectar impuestos por ventas en artículos que se compran a través de Internet. ¿Cuántas transacciones de ventas que se seleccionaran aleatoriamente deben registrarse para determinar el porcentaje que se lleva a cabo por 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.

34. Tamaño de la muestra para jugadores del golf zurdos Como fabricante de equipos de golf, la Spalding Corporation quiere estimar la proporción de golfistas que son zurdos. (La compañía piensa usar tal información en la planeación del número de juegos de palos de golf a fabricar para diestros y zurdos). ¿Cuántos golfistas deben encuestarse si queremos un nivel de confianza del 99% de que la proporción muestral tenga un margen de error de 0.025?

a. Suponga que no hay información disponible que pueda usarse como un estimado de $\hat{p}$.

b. Suponga que tenemos un estimado de $\hat{p}$ que se encontró en un estudio previo, el cual sugiere que el 15% de los golfistas son zurdos (según un reporte de *USA Today*).

c. Suponga que en lugar de usar golfistas que se seleccionaron al azar, los datos muestrales se obtienen pidiendo a los televidentes del canal de golf que llamen a un número telefónico "800" para reportar si ellos son zurdos o diestros. ¿De qué forma se afectan los resultados?

35. Tamaño de muestra para propiedad de vehículos motorizados A usted lo contrató la Ford Motor Company para hacer investigación de mercado, por lo que debe estimar el porcentaje de hogares que poseen un vehículo. ¿Cuántos hogares debe entrevistar si desea tener una confianza del 94% de que su porcentaje muestral tiene un margen de error de tres puntos porcentuales?

a. Suponga que un estudio previo sugiere que el 86% de los hogares poseen vehículos.

b. Suponga que no hay información disponible que pueda usarse para estimar el porcentaje de hogares en los que poseen un vehículo.

c. Suponga que en lugar de utilizar hogares que se seleccionaron al azar, los datos muestrales se obtienen pidiendo a los lectores del periódico *Washington Post* que envíen por correo un formato de encuesta. ¿De qué forma se afectan los resultados?

36. Tamaño de muestra de armas en el campus Interesados por la seguridad en el *campus*, los oficiales universitarios quieren estimar el porcentaje de estudiantes que porta una pistola, un cuchillo u otra arma. ¿Cuántos estudiantes que se seleccionen al azar deben encuestarse para tener una confianza del 95% de que el porcentaje de la muestra tiene un margen de error de tres puntos porcentuales?

a. Suponga que otro estudio indicó que el 7% de los estudiantes universitarios llevan armas (según un estudio de Cornell University).

b. Suponga que no existe información disponible que pueda utilizarse para estimar el porcentaje de estudiantes universitarios que portan armas.

37. Daltonismo En un estudio de percepción, se examinó a 80 hombres y se encontró que siete tenían ceguera al color rojo/verde (según datos de *USA Today*).

a. Construya un estimado del intervalo de confianza de 90%, de la población de todos los hombres con daltonismo.

b. ¿Qué tamaño de muestra se necesitaría para estimar la proporción de varones con daltonismo, si queremos un nivel de confianza del 96% de que la proporción muestral es errónea por no más de 0.03? Utilice la proporción muestral como un estimado que se conoce.

c. Las mujeres tienen una tasa de daltonismo de 0.25%. Con base en el resultado del inciso *a*, ¿concluiríamos con seguridad que las mujeres padecen una tasa más baja de daltonismo que los hombres?

38. Audiencia de televisión El programa televisivo *60 minutos* de la CBS lleva muchos años de ser exitoso. Este programa tuvo recientemente una audiencia de 20, lo que significa que, de los televisores en funcionamiento, el 20% sintonizaron en *60 minutos* (según datos de Nielsen Media Research). Suponga que esto se basa en un tamaño de muestra de 4000 (típico para encuestas de Nielsen).

a. Construya un estimado del intervalo de confianza del 97%, para la proporción de todos los aparatos en operación que sintonizaban *60 minutos* al momento de la transmisión.

b. ¿Qué tamaño de muestra se requeriría para estimar el porcentaje de aparatos que sintonizaban *60 minutos*, si queremos un nivel de confianza del 99% de que el porcentaje muestral es erróneo por no más de medio punto porcentual? (Suponga que no tiene un estimado de la proporción).

c. En el momento de la difusión particular de *60 minutos*, la ABC transmitió “Expuesto: Lucha profesional”, programa que recibió una audiencia de 11. Con base en el resultado del inciso *a*, ¿podemos concluir que *60 minutos* tiene una mayor proporción de televidentes? ¿Necesitan realmente los luchadores profesionales exponerse?

d. ¿De qué forma se afecta el intervalo de confianza en el inciso *a* si, en lugar de seleccionar sujetos al azar, los datos de la encuesta se basan en 4000 televidentes que llaman voluntariamente a un número “800” para registrar sus respuestas?

39. Teléfonos celulares y cáncer Un estudio de 420,000 daneses usuarios de teléfono celular encontró que 135 de ellos desarrollaron cáncer cerebral o del sistema nervioso. Con anterioridad a este estudio del uso de teléfono celular, se encontró que la tasa de este tipo de cáncer es de 0.0340% para aquellos que no usan teléfonos celulares. Los datos son de la *Journal of the National Cancer Institute*.

a. 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.

b. ¿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?

40. Fatalidad en pilotos Investigadores estudian choques de aeroplanos de aviación general (no comercial y no militar) y encuentran que los pilotos murieron en el 5.2% de

8411 aterrizajes con choque (que se basa en datos de "Risk Factors for Pilot Fatalities in General Aviation Airplane Crash Landings", de Rostykus, Cummings y Mueller, *Journal of the American Medical Association*, vol. 280, núm. 11).

a. Construya un estimado del intervalo de confianza del 95%, del porcentaje de pilotos muertos en todos los choques de aviación general.

b. De los choques con una explosión o fuego en tierra, la tasa de fatalidad de los pilotos se estima por el intervalo de confianza del 95% de (15.5%, 26.9%). ¿Es este resultado sustancialmente diferente del resultado del inciso *a*? ¿Qué concluye usted acerca de una explosión o fuego como factor de riesgo?

c. En la planeación para la asignación de fondos federales para ayudar con los exámenes médicos de pilotos difuntos, ¿qué porcentaje único debe usarse? (Queremos estar razonablemente seguros de que tenemos suficientes recursos para la peor situación posible).

41. Uso de ropa naranja de cazador Un estudio sobre las heridas de caza, y el uso de ropa naranja "de cazador", mostró que entre 123 cazadores heridos cuando fueron confundidos con presas, seis usaban ropa naranja. Entre 1115 cazadores seleccionados aleatoriamente, 811 reportaron que habitualmente usan ropa color naranja. Los datos son de los Centers for Disease Control.

a. Construya un estimado del intervalo de confianza del 95%, del porcentaje de cazadores heridos que usaban ropa naranja.

b. Construya un estimado del intervalo de confianza del 95%, del porcentaje de cazadores que usan ropa naranja habitualmente.

c. ¿Indican estos resultados que es menos probable que a un cazador que usa ropa naranja lo hieran por confundirse con una presa? ¿Por qué?

42. La apariencia cuenta Una encuesta de administración de ventas y mercadeo incluyó a 651 gerentes de ventas; el 94% de ellos dijeron que tener una apariencia descuidada puede hacer más difícil el trabajo del representante de ventas.

a. Construya un estimado del intervalo de confianza del 90%, del porcentaje de gerentes de ventas que dicen que tener una apariencia descuidada puede hacer más difícil el trabajo del representante de ventas.

b. Construya un estimado del intervalo de confianza del 90% del porcentaje de gerentes de ventas que dicen que tener una apariencia sin estilo puede hacer más difícil el trabajo del representante de ventas.

c. Puesto que las proporciones muestrales varían naturalmente, ¿es posible concluir que cuando los gerentes de ventas declaran razones por las que el trabajo de los representantes de ventas se vuelve más difícil, el porcentaje es más alto para una apariencia descuidada que para una apariencia sin estilo? ¿Por qué?

43. Dulces M&M rojos Remítase al conjunto de datos 19 en el Apéndice B y calcule la proporción muestral de dulces M&M rojos. Use este resultado para crear un intervalo de confianza del 95%, del porcentaje de la población de dulces M&M rojos. ¿Es consistente el resultado con la tasa del 20% que reporta el fabricante de dulces Mars?

44. Consumo de alcohol y tabaco en películas para niños Remítase al conjunto de datos 7 en el Apéndice B.

a. Construya un estimado del intervalo de confianza del 95%, para el porcentaje de películas infantiles de dibujos animados que muestran consumo de tabaco.

b. Construya un estimado del intervalo de confianza del 95%, para el porcentaje de películas infantiles de dibujos animados que muestran consumo de alcohol.

c. Compare los resultados anteriores. ¿Aparecen en un alto porcentaje el tabaco o el alcohol en las películas de dibujos animados infantiles?

d. 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?

6-2 Más allá de lo básico

45. Prueba de exactitud Un ejemplo de esta sección utilizó los datos de la encuesta sobre la cámara vigilante con $n = 829$ y $\hat{p} = 0.51$ para construir el intervalo de confianza del 95% de $0.476 < p < 0.544$. Sin embargo, $\hat{p}$ no puede ser exactamente 0.51, puesto que el 51% de 829 personas es 422.79, lo cual no es posible. El estadístico muestral de 51% se redondea al número entero más cercano. Calcule los valores mínimo y máximo de x para los que $x/829$ se redondea a 0.51; después, construya los intervalos de confianza que corresponden a esos dos valores de x. ¿Difieren los resultados sustancialmente del intervalo de confianza de $0.476 < p < 0.544$ que se calculó utilizando 0.51?

46. Uso del factor de corrección por población finita Esta sección presentó las fórmulas 6-2 y 6-3, que se usan para determinar el tamaño de la muestra. En ambos casos suponemos que la población es infinita o muy grande y que tomamos muestras con reemplazo. Cuando tenemos una población relativamente pequeña con tamaño N y muestreo sin reemplazo, modificamos E para incluir el *factor de corrección por población finita* que se muestra aquí; podemos, después, resolver para n y así obtener el resultado dado aquí. Utilice este resultado para repetir el inciso *b* del ejercicio 38, suponiendo que limitamos nuestra población a una localidad con 10,000 televisores en operación.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\sqrt{\frac{N-n}{N-1}} \qquad n = \frac{N\hat{p}\hat{q}[z_{\alpha/2}]^2}{\hat{p}\hat{q}[z_{\alpha/2}]^2 + (N-1)E^2}$$

47. Intervalo de confianza de un factor Un *intervalo de confianza de un factor* para p puede expresarse como $p < \hat{p} + E$ o $p > \hat{p} - E$, donde el margen de error E se modifica reemplazando $z_{\alpha/2}$ con 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 de un factor apropiado y luego calcule el porcentaje en cuestión. Suponga que una muestra aleatoria simple de 750 vuelos resulta en 630 que son puntuales.

48. Intervalo de confianza de muestra pequeña Hay tablas especiales disponibles para encontrar intervalos de confianza para proporciones que incluyen números pequeños de casos, cuando no es posible usar la aproximación por distribución normal. Por ejemplo, dado $x = 3$ éxitos entre $n = 8$ ensayos, el intervalo de confianza del 95% que se encuentra en *Standard Probability and Statistics Tables and Formulae* (CRC Press) es $0.085 < p < 0.755$. Encuentre el intervalo de confianza que resultaría si usara la distribución normal erróneamente como una aproximación de la distribución binomial. ¿Son los resultados razonablemente cercanos?

49. Interpretación de límites de intervalo de confianza Suponga que se modifica una moneda para que favorezca las caras, y de 100 lanzamientos 95 son caras. Calcule el estimado del intervalo de confianza del 99%, de la proporción de caras que ocurrirán con esta moneda. ¿Qué es poco común en los resultados que se obtienen usando los métodos de esta sección? ¿Sugiere el sentido común una modificación del intervalo de confianza resultante?

50. Regla de tres Suponga que en n ensayos de un experimento binomial no se registra ningún éxito. De acuerdo con la *Regla de tres*, hay un 95% de confianza de que la proporción real de la población tenga 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 es posible calcular los límites del intervalo de confianza usando los métodos que se describen en esta sección?

b. Si se trata con un fármaco a 20 pacientes y no hay reacciones adversas, ¿cuál es la frontera superior del 95% para p, la proporción de todos los pacientes que experimentaron reacciones adversas a este fármaco?

51. Estaturas de mujeres Las estaturas de las mujeres se distribuyeron normalmente con una media de 63.6 pulgadas y una desviación estándar de 2.5 pulgadas. ¿Cuántas mujeres deben encuestarse si queremos estimar el porcentaje de las que son más altas que cinco pies? Suponga que queremos un nivel de confianza del 98% de que el error no es más de 2.5 puntos porcentuales. (*Advertencia:* La respuesta es sustancialmente menor que 2172).

52. Exactitud de encuesta Un artículo del *New York Times* acerca de resultados de encuesta afirma: "En teoría, en 19 casos de 20, los resultados de una encuesta como ésta deben diferir por más de un punto porcentual en cualquier dirección de lo que podría obtenerse entrevistando a todos los votantes en Estados Unidos". Calcule el tamaño de la muestra sugerido por esta afirmación.

6-3 Estimación de la media poblacional: σ conocida

En la sección 6-2 estudiamos el estimado puntual y el intervalo de confianza como herramientas para el empleo de una proporción muestral para estimar una proporción poblacional. También mostramos cómo determinar el tamaño de muestra mínimo que se requiere para estimar una proporción de población. En esta sección nuevamente analizamos el estimado puntual, el intervalo de confianza y la determinación del tamaño de la muestra, pero ahora consideramos el objetivo de la estimación de una media poblacional μ.

Supuestos

1. La muestra es aleatoria simple. (Todas las muestras del mismo tamaño tienen una posibilidad igual de seleccionarse).
2. Se conoce el valor de la desviación estándar poblacional σ.
3. Cualquiera o ambas de tales condiciones se satisface: la población está normalmente distribuida o $n > 30$.

En los supuestos de arriba, vemos que queremos estimar una media de población μ que no se conoce, pero debemos saber el valor de la desviación estándar poblacional σ. Se trataría de un extraño conjunto de circunstancias que nos permitirían conocer σ desconociendo μ. Después de todo, la única forma de encontrar el valor de σ es calcularlo a partir de todos los valores que se conocen de la población; entonces, el cálculo de μ también sería posible y, si encontráramos el valor real de μ, no hay necesidad de estimarlo. Aunque los métodos del intervalo de confianza de esta sección no son muy realistas, revelan los conceptos básicos del importante razonamiento estadístico y constituyen los cimientos para la determinación del tamaño de la muestra que se analiza más tarde en esta sección.

Supuestos de normalidad En esta sección usaremos los supuestos de que tenemos una muestra aleatoria simple, el valor de σ que se conoce y la población que se distribuye normalmente o $n > 30$. Desde el punto de vista técnico, la población no necesita tener una distribución que sea exactamente normal, aunque sí debe ser aproximadamente normal, es decir, que la distribución sea un tanto simétrica, con una moda y sin datos distantes. Investigue la normalidad utilizando los datos muestrales para construir un histograma; después, determine si tiene aproximadamente forma de campana. Puede construirse una gráfica cuantilar normal (sección 5-7), pero se considera que los métodos de la sección son robustos, lo que significa que no afectaron de forma importante por desviaciones de la normalidad, siempre que tales desviaciones no sean muy extremas. Por lo regular podemos considerar que la población se distribuye normalmente después de utilizar los datos muestrales para confirmar que no hay datos distantes y que el histograma tiene una forma que no es muy distinta de la de una distribución normal.

Supuesto del tamaño de la muestra que se requiere La presente sección utiliza la distribución normal como la distribución de las medias muestrales. Si la población original, en sí misma, normalmente se distribuye, entonces las medias de muestras de cualquier tamaño se distribuirán normalmente. Si la población original no se distribuye normalmente, decimos que las medias de las muestras con tamaño $n > 30$ tienen una distribución que llega a aproximarse a una distribución normal. La condición de que el tamaño de la muestra sea $n > 30$ se usa por lo regular como lineamiento, pero no es posible identificar un tamaño de muestra mínimo específico que sea suficiente para todos los casos. El tamaño de muestra mínimo realmente depende de cuánto se desvía la distribución de la población de una distribución normal. Tamaños de la muestra 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 de la muestra de 50, 100 o más altos. 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 6-2 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.

Aunque utilicemos otro estadístico, como la mediana muestral, la mitad del rango o la moda como un estimado de la media poblacional μ, los estudios muestran que por lo regular la media muestral $\bar{x}$ proporciona el mejor estimado, por las siguientes dos razones:

1. Para muchas poblaciones, la distribución de medias muestrales $\bar{x}$ tiende a ser más consistente (con *menor variación*) que las distribuciones de otros estadísticos muestrales. (Esto es, si utiliza medias muestrales para estimar la media poblacional μ, estas medias muestrales tendrán una desviación estándar menor que la de otros estadísticos muestrales, como son la mediana o la moda. Por lo tanto, las diferencias entre $\bar{x}$ y μ tienden a ser menores que las diferencias que se obtienen con algún otro estadístico, como la mediana).
2. Para muchas poblaciones, la media muestral $\bar{x}$ es un **estimador sin sesgo** de la media poblacional μ, lo que significa que la distribución de las medias muestrales tiende a centrarse alrededor del valor de la media poblacional μ. (Esto es, las medias muestrales no tienden sistemáticamente a sobreestimar el valor de μ ni a subestimar μ. En lugar de ello, tienden a apuntar al propio valor de μ. Véase la sección 5-4, donde ilustramos el principio de que las medias muestrales tienden a apuntar al valor de la media poblacional).

Los números de serie de tanques capturados revelan el tamaño de la población

Durante la Segunda Guerra Mundial, especialistas en espionaje de los aliados querían determinar el número de tanques que Alemania estaba produciendo. Las técnicas de espionaje tradicionales produjeron resultados poco confiables, pero los estadísticos obtuvieron estimados exactos analizando los números de serie de los tanques capturados. Por ejemplo, los registros muestran que Alemania realmente produjo 271 tanques en junio de 1941. El estimado que se basó en los números de serie fue de 244, pero los métodos de espionaje tradicionales dieron como resultado el estimado extremo 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).

Estimación del tamaño de las poblaciones silvestres

La ley que se conoce como National Forest Management Act protege a las especies en peligro, entre las que está el búho manchado septentrional. Como resultado de la aplicación de esta ley, no se permitió a la industria silvícola talar vastas regiones de árboles en el noroeste del Pacífico. Se pidió a biólogos y estadísticos analizar el problema; ellos concluyeron que las tasas de supervivencia y los tamaños de las poblaciones estaban disminuyendo en el caso de los búhos hembra, que se sabe que desempeñan un papel importante en la supervivencia de la especie. Los biólogos y estadísticos 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 con el uso de habilidades de modelaje que son el sello de la buena estadística. Los estadísticos ganan al ser introducidos a la realidad de los problemas por los biólogos, que conocen cuáles son los asuntos cruciales".

EJEMPLO Temperaturas corporales El conjunto de datos 4 del Apéndice B incluye 106 temperaturas corporales que se tomaron a las 12:00 h del día 2. He aquí estadísticos para dicha muestra: $n = 106$, $\bar{x} = 98.20°F$ y $s = 0.62°F$. Utilice esta muestra para calcular el mejor estimado puntual de la media poblacional μ de todas las temperaturas corporales.

SOLUCIÓN Para los datos muestrales, $\bar{x} = 98.20°F$. Puesto que la media muestral $\bar{x}$ es el mejor estimado puntual de la media poblacional μ, concluimos que el mejor estimado puntual de la media poblacional μ de todas las temperaturas corporales es de 98.20°F.

Intervalos de confianza

En la sección 6-2 vimos que 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. Los estadísticos desarrollaron el intervalo de confianza o estimado de intervalo, consistente en un rango (o intervalo) de valores, en lugar de sólo un valor. El intervalo de confianza se asocia con un nivel de confianza, como es 0.95 (o 95%). El nivel de confianza nos ofrece la tasa de éxitos del procedimiento que se usa para construir el intervalo de confianza. Como se describió en la sección 6-2, el nivel de confianza suele expresarse como la probabilidad o área $1 - \alpha$, donde α es el complemento del *nivel de confianza*. Para un nivel de confianza del 0.95 (o 95%), $\alpha = 0.05$. Para un nivel de confianza del 0.99 (o 99%), $\alpha = 0.01$.

Margen de error Cuando reunimos un conjunto de datos muestrales, como el de las 106 temperaturas corporales que se listaron para las 12:00 h del día 2, en el conjunto de datos 4 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 5-5 vimos que $\sigma/\sqrt{n}$ es la desviación estándar de las medias muestrales. Utilizando $\sigma/\sqrt{n}$ y la notación $z_{\alpha/2}$ que se introdujo en la sección 6-2, ahora podemos usar el *margen de error E* que se expresa como sigue.

Fórmula 6-4 $E = z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$ margen de error para media (que se basa en σ conocida)

La fórmula 6-4 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 estándar $\sigma/\sqrt{n}$, siempre y cuando la población tenga una distribución normal con media μ y desviación estándar σ. Si la población no se distribuye normalmente, las muestras mayores producen medias muestrales con una distribución que *se aproxima* a la normal.

Dada la forma en que se define el margen de error E, existe una probabilidad de $1 - \alpha$ de que una media muestral sea errónea (diferente de la media poblacional α) por más de E. El cálculo del margen de error E, como se dio en la fórmula 6-4, requiere que usted conozca la desviación estándar poblacional σ, aunque la sección 6-4 presentará un método para calcular el margen de error E cuando σ no se conoce.

Utilizando el margen de error E ahora identificaremos el intervalo de confianza para la media poblacional μ (si se satisfacen las condiciones supuestas de esta sección). Los tres formatos que suelen usarse para expresar el intervalo de confianza se muestran en el siguiente cuadro.

Estimación del intervalo de confianza de la media poblacional μ (con σ conocida)

$$\bar{x} - E < \mu < \bar{x} + E \qquad \text{donde } E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

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 se conoce y la población parece distribuirse normalmente, o $n > 30$).
2. Remítase a la tabla A-2 y calcule el valor crítico correspondiente al nivel de confianza que se desea (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 $E = z_{\alpha/2}\,\sigma/\sqrt{n}$.
4. Utilizando el valor que se calculó del margen de error E 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 valores resultantes usando la siguiente regla de redondeo.

Regla de redondeo para intervalos de confianza que se utilizan 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 lugar decimal más del que se usa para el conjunto de datos original.
2. Cuando el conjunto de datos original se desconoce y sólo se utilizan los estadísticos resumidos (n, $\bar{x}$, s), redondee los límites del intervalo de confianza al mismo número de espacios decimales que se usan para la media muestral.

Interpretación de un intervalo de confianza Igual que en la sección 6-2, debemos ser cuidadosos para interpretar correctamente los intervalos de confianza. Después de obtener un estimado del intervalo de confianza de la media poblacional μ, tal como un intervalo de confianza del 95% de $98.08 < \mu < 98.32$, hay una interpretación correcta y muchas interpretaciones erróneas.

Correcta: "Tenemos una confianza del 95% de que el intervalo de 98.08 a 98.32 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 6-2, tal 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 posibilidades de que μ caiga entre 98.08 y 98.32". El intervalo de confianza no describe el comportamiento de valores muestrales individuales; entonces, también sería incorrecto afirmar que "el 95% de todos los valores de los datos están entre 98.08 y 98.32". Además, el intervalo de confianza no describe el comportamiento de medias muestrales individuales; asimismo, sería incorrecto decir que "el 95% de las medias muestrales caen entre 98.08 y 98.32".

EJEMPLO Temperaturas corporales Para la muestra de temperaturas corporales del conjunto de datos 4 del Apéndice B (para las 12:00 h del día 2), tenemos $n = 106$ y $\bar{x} = 98.20°F$. Suponga que la muestra es una muestra aleatoria simple y que, por alguna razón, se conoce que σ es 0.62°F. Utilizando un nivel de confianza del 0.95, calcule lo siguiente:

a. El margen de error E.

b. El intervalo de confianza para μ.

SOLUCIÓN Primero verifique que los supuestos requeridos se satisfacen. Se asume que se conoce el valor de σ (0.62°F) y el tamaño de la muestra $n = 106$ es mayor que 30. Además, no hay datos distantes. (Puesto que $n > 30$, no hay necesidad de revisar que la muestra provenga de una población que se distribuye normalmente, pero un histograma de las 106 temperaturas corporales mostraría que los datos muestrales tienen una distribución que toma casi la forma de campana, lo que sugiere que la población de temperaturas corporales se distribuye normalmente). Por lo tanto, los supuestos requeridos se satisfacen y hacen posible proceder 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 6-2). El margen de error E se calcula usando la fórmula 6-4 como sigue. Los lugares decimales de más se utilizan para minimizar los errores de redondeo en el intervalo de confianza que se calculó en el inciso b.

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{0.62}{\sqrt{106}} = 0.118031$$

b. Con $\bar{x} = 98.20$ y $E = 0.118031$, construimos el intervalo de confianza como sigue:

$$\bar{x} - E < \mu < \bar{x} + E$$

$$98.20 - 0.118031 < \mu < 98.20 + 0.118031$$

$$98.08 < \mu < 98.32 \quad \text{(que se redondeó a dos decimales como en } \bar{x}\text{)}$$

INTERPRETACIÓN Este resultado también podría expresarse como 98.20 ± 0.12 o como (98.08, 98.32). Con base en la muestra con $n = 106$, $\bar{x} = 98.20$, y suponiendo que σ es 0.62, el intervalo de confianza para la media de la población μ es $98.08°\text{F} < \mu < 98.32°\text{F}$; este intervalo tiene un nivel de confianza del 0.95. Lo anterio significa que si vamos a seleccionar muchas muestras diferentes de tamaño 106 y a construir los intervalos de confianza, como lo hicimos aquí, el 95% de ellos contendrían realmente el valor de la media poblacional μ. Note que los límites del intervalo de confianza de 98.08°F y 98.32°F no contienen a 98.6°F, el valor que generalmente se cree que es la temperatura corporal media. Con base en estos resultados, parece muy poco probable que 98.6°F sea el valor correcto de μ.

Fundamentos del intervalo de confianza La idea básica que subyace a la construcción de intervalos de confianza se relaciona con el teorema del límite central, que indica que si tenemos una muestra aleatoria simple de una población que se distribuye normalmente, o una muestra aleatoria simple de tamaño $n > 30$ de cualquier población, la distribución de medias muestrales es aproximadamente normal, con media μ y desviación estándar $\sigma/\sqrt{n}$. 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 $\sigma/\sqrt{n}$, sustituya $\mu_{\bar{x}}$ por μ; luego resuelva μ para obtener

$$\mu = \bar{x} - z\frac{\sigma}{\sqrt{n}}$$

Usando los valores positivo y negativo para z se obtienen los límites del intervalo de confianza que estamos empleando.

Consideremos el caso específico de un nivel de confianza del 95%, entonces $\alpha = 0.05$ y $z_{\alpha/2} = 1.96$. Para este caso, hay una probabilidad de 0.05 de que una media muestral esté más allá de 1.96 desviaciones estándar (o $z_{\alpha/2}\sigma/\sqrt{n}$, lo cual denotamos por E) de la media poblacional μ. Por el contrario, hay una probabilidad del 0.95 de que una media muestral esté dentro de 1.96 desviaciones estándar (o $z_{\alpha/2}\sigma/\sqrt{n}$) de μ. (Véase la figura 6-4 en la página siguiente). Si la media muestral $\bar{x}$ está dentro de $z_{\alpha/2}\sigma/\sqrt{n}$ de la media poblacional μ, entonces μ debe estar dentro de $\bar{x} - z_{\alpha/2}\sigma/\sqrt{n}$ y $\bar{x} + z_{\alpha/2}\sigma/\sqrt{n}$; esto se expresa en el formato general de nuestro intervalo de confianza (con $z_{\alpha/2}\sigma/\sqrt{n}$ que se denota como E): $\bar{x} - E < \mu < \bar{x} + E$.

Método alternativo (no se usa en este libro) Cuando se construye un estimado del intervalo de confianza de la media poblacional μ, un método alternativo que *no se usa en este libro*, es el uso de los procedimientos descritos arriba,

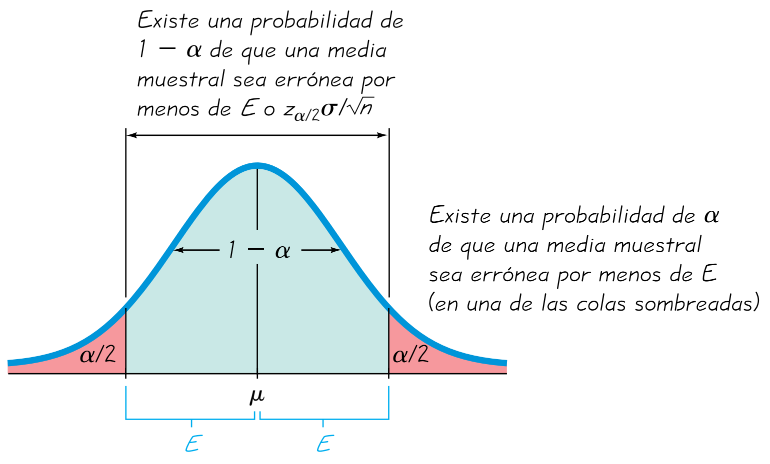


FIGURA 6-4 Distribución de medias muestrales con σ conocida

incluso si no se conoce σ, pero la muestra es aleatoria simple con $n > 30$. Con este método alternativo, usamos s como un estimado de σ, siempre que $n > 30$. En la sección 6-4 listamos las razones por las cuales este método alternativo no se usa en el presente libro; dichas razones incluyen el hecho de que este método alternativo por lo regular no se usa en el mundo real. El mundo real utiliza los métodos descritos en este libro.

Una característica clave de los métodos que estamos usando en esta sección es que queremos estimar una media poblacional μ que se desconoce, en tanto que conocemos la desviación estándar de la población σ. En la siguiente sección presentamos un método para estimar una media poblacional μ desconocida cuando la desviación estándar de la población no se conoce. Las condiciones de la siguiente sección son mucho más probables de ocurrir en circunstancias reales. Aunque los métodos de esta sección no son realistas, puesto que se basan en el conocimiento de la desviación estándar poblacional σ, nos capacitan para ver el método básico para construir un estimado del intervalo de confianza de μ utilizando la misma distribución normal que se usa con frecuencia en el capítulo 5 y en la sección 6-2. Además, los métodos que ya estudiamos en esta sección conducen a un método muy práctico para determinar el tamaño de muestra.

Determinación del tamaño de muestra requerido para estimar μ

Ahora queremos plantear esta pregunta clave: Cuando planeamos reunir una muestra aleatoria simple de datos que se usarán para estimar una media poblacional μ, ¿*cuántos* valores muestrales deben obtenerse? En otras palabras, calcularemos el tamaño de muestra n que se requiere para estimar el valor de una media poblacional. Por ejemplo, suponga que queremos estimar una media del peso 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 punto muy importante, puesto que muestras que son innecesariamente grandes desperdician tiempo y dinero, en tanto que muestras que son muy pequeñas pueden conducir a resultados pobres. En muchos casos podemos encontrar el tamaño mínimo necesario para estimar algún parámetro, como la media poblacional μ.

Si comenzamos con la expresión para el margen de error E (fórmula 6-4) y resolvemos para el tamaño de muestra n, obtenemos lo siguiente.

Tamaño de muestra para estimar la media μ

Fórmula 6-5

$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2$$

donde $z_{\alpha/2}$ = puntuación z crítica que se basa en el nivel de confianza deseado
E = margen de error que se desea
σ = desviación estándar poblacional

La fórmula 6-5 es relevante puesto que indica que el tamaño de muestra no depende del tamaño de la población (N); el tamaño de muestra depende del nivel de confianza deseado, del margen de error deseado y del valor de la desviación estándar σ. (Véase el ejercicio 33 para tratar con casos en los que se selecciona una muestra relativamente grande sin reemplazo de una población finita).

El tamaño de muestra debe ser un número entero, ya que representa el número de valores muestrales que deben encontrarse. Sin embargo, cuando usamos la fórmula 6-5 para calcular el tamaño de muestra n, por lo regular obtenemos un resultado que no es un número entero. Cuando esto sucede, usamos la siguiente regla de redondeo. (Esta regla se basa en el principio de que cuando es necesario redondear, el tamaño de muestra que se requiere debe redondearse *hacia arriba* para que sea al menos adecuadamente grande en oposición a un poco más pequeño).

Regla de redondeo para el tamaño de muestra n

Cuando se calcula el tamaño de muestra n, si el uso de la fórmula 6-5 no resulta en un número entero, siempre *incremente* el valor de n al siguiente número entero *mayor*.

Cálculo del tamaño de muestra con σ desconocida Cuando se aplica la fórmula 6-5, surge un dilema muy práctico: la fórmula requiere que sustituyamos algún valor de la desviación poblacional σ, pero en realidad suele desconocerse. Cuando se determina un tamaño de muestra que se requiere (sin construir un intervalo de confianza), hay algunos procedimientos que pueden funcionar para este problema:

1. Use la regla práctica del intervalo (véase la sección 2-5) para estimar la desviación estándar como sigue: $\sigma \approx \text{intervalo}/4$. (Con una muestra de 87 valores o más que se seleccionó aleatoriamente de una población normalmente distribuida, el intervalo/4 nos da un valor que es mayor que o igual que 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, *The American Statistician*, vol. 55, núm. 4).

2. Realice un estudio piloto empezando por el proceso de muestreo. Con base en la primera colección de al menos 31 valores seleccionados aleatoriamente, calcule la desviación estándar muestral s y úsela en lugar de σ. Entonces el valor estimado de σ mejorará conforme se obtengan más datos muestrales.
3. Estime el valor de σ utilizando los resultados de algún otro estudio hecho con anterioridad.

En suma, algunas veces podemos ser creativos en nuestro uso de otros resultados que se conocen. Por ejemplo, por lo regular las pruebas de CI se diseñaron 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 de muestra mayor del necesario, pero utilizar un valor de σ que sea muy pequeño generaría un tamaño de muestra inadecuado. *Cuando se calcula el tamaño de muestra* n, *cualquier error debe siempre ser conservador, en el sentido de que hace a* n *muy grande en lugar de muy pequeña*.

EJEMPLO Puntuaciones de CI para profesores de estadística Suponga que queremos estimar la media de la puntuación de CI para 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 6-5 se calculan como sigue:

$z_{\alpha/2} = 1.96$ (Esto se resuelve convirtiendo el nivel de confianza del 95% a $\alpha = 0.05$, y luego calculando la puntuación crítica z, como se describe en la sección 6-2).

$E = 2$ (Puesto que queremos que la media muestral esté dentro de dos puntos de CI de μ, el margen de error que se desea 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 6-5 como sigue:

$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2 = \left[\frac{1.96 \cdot 15}{2}\right]^2 = 216.09 = 217 \quad \text{(redondeado hacia \textit{arriba})}$$

INTERPRETACIÓN Entre los miles de profesores de estadística, necesitamos obtener una muestra aleatoria simple de al menos 217 de ellos, y luego 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 resolver para obtener resultados menos precisos utilizando un margen de error más grande, como por ejemplo 4, el tamaño de muestra disminuye a 54.0225, el cual se redondea *hacia arriba* a 55. La duplicación del margen de error provoca que el tamaño de la muestra requerido disminuya a un cuarto de su valor original. Por el contrario, dividir a la mitad el margen de error cuadruplica el tamaño de la muestra. En consecuencia, si usted desea resultados más precisos, el tamaño de la muestra debe incrementarse sustancialmente. Ya que los muestreos grandes por lo regular requieren de más tiempo y dinero, con frecuencia existe la necesidad de ponderar un balance entre el tamaño de la muestra y el margen de error E.

Utilizando la tecnología

Intervalos de confianza Véase al final de la sección 6-4 los procedimientos del intervalo de confianza que se aplican a los métodos de esta sección, así como también los de la sección 6-4. STATDISK, Minitab, Excel y la calculadora TI-83 Plus resultan útiles para calcular intervalos de confianza cuando necesitamos estimar la media de una población y se satisfacen todos los supuestos de esta sección (inclusive el valor conocido de σ).

Determinación del tamaño de la muestra Los cálculos para el tamaño de la muestra no se incluyen en la calculadora TI-83 Plus ni en Minitab ni en Excel. A continuación se describe el procedimiento de STATDISK para determinar el tamaño de muestra requerido para estimar una media poblacional μ.

STATDISK Seleccione **Analysis** de la parte superior de la barra del menú principal, luego elija **Sample Size Determination**, seguido por **Estimate Mean**. Ahora debe ingresar el nivel de confianza (tal como 0.95), el error E, y la desviación estándar poblacional σ. También hay una opción que le permite ingresar el tamaño poblacional N, suponiendo que está haciendo el muestreo sin reemplazo de una población finita. (Véase el ejercicio 34).

6-3 Destrezas y conceptos básicos

Encontrar valores críticos. *En los ejercicios 1 a 4, encuentre el valor crítico $z_{\alpha/2}$ correspondiente al nivel de confianza dado.*

1. 98% **2.** 95%

3. 96% **4.** 99.5%

Verificación de supuestos. *En los ejercicios 5 a 8, determine si las condiciones dadas justifican el uso del margen de error $E = z_{\alpha/2}\sigma/\sqrt{n}$, cuando se calcula un estimado del intervalo de confianza de la media poblacional μ.*

5. El tamaño de la muestra es $n = 200$ y $\sigma = 15$.

6. El tamaño de la muestra es $n = 5$ y σ se desconoce.

5. El tamaño de la muestra es $n = 5$, $\sigma = 12.4$, y la población original se distribuye normalmente.

8. El tamaño de la muestra es $n = 9$, σ no se conoce, y la población original se distribuye normalmente.

Cálculo del margen de error y del intervalo de confianza. *En los ejercicios 9 a 12, use el nivel de confianza y los datos muestrales dados para encontrar a) el margen de error E y b) el intervalo de confianza para estimar la media poblacional μ.*

9. Salarios de profesores de estadística: confianza del 95%; $n = 100$, $\bar{x} = \$95{,}000$ (ya quisiéramos), y se sabe que σ es \$12,345.

10. Las edades de los conductores que ocupan el carril para rebasar mientras conducen a 25 millas/h con la luz intermitente izquierda funcionado: confianza del 99%; $n = 50$, $\bar{x} = 80.5$ años, y se sabe que σ es 4.6 años.

11. El tiempo entre operaciones de un control remoto de televisión por hombres durante los cortes comerciales: confianza del 90%; $n = 25$, $\bar{x} = 5.24$ seg, la población se distribuye normalmente y se conoce que σ es 2.50 seg.

12. Salarios iniciales de graduados universitarios que tomaron un curso de estadística: confianza del 95%; $n = 28$, $\bar{x} = \$45{,}678$, la población se distribuye normalmente y se sabe que σ es \$9900.

Calcular el tamaño de la muestra. *En los ejercicios 13 a 16, use el margen de error, el nivel de confianza y la desviación estándar poblacional σ dados, para calcular el tamaño de muestra mínimo que se requiere para estimar una media poblacional μ desconocida.*

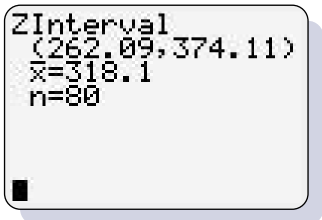
13. Margen de error: \$125, nivel de confianza: 95%, $\sigma = \$500$.

14. Margen de error: 3 lb, nivel de confianza: 99%, $\sigma = 15$ lb.

15. Margen de error: 5 min, nivel de confianza: 90%, $\sigma = 48$ min.

16. Margen de error: \$500, nivel de confianza: 94%, $\sigma = \$9877$

TI-83 Plus



Interpretación de resultados. *En los ejercicios 17 a 20 haga referencia a la representación adjunta de la pantalla de la calculadora TI-83 Plus que muestra un intervalo de confianza del 95%, que se generó con el uso de los métodos de esta sección. La pantalla presenta el resultado de usar una muestra de 80 niveles de colesterol medidos en adultos seleccionados al azar.*

17. Identifique el valor del estimado puntual de la media poblacional μ.

18. Exprese el intervalo de confianza en el formato de $\bar{x} - E < \mu < \bar{x} + E$.

19. Exprese el intervalo de confianza en el formato de $\bar{x} \pm E$.

20. Escriba una afirmación que interprete el intervalo de confianza del 95%.

21. Temperaturas de los Everglades Para controlar la salud ecológica de los Everglades de Florida, se registran varias mediciones en tiempos diferentes. Las temperaturas inferiores se registran en la estación Garfield Bight y se obtiene la media de 30.4°C para 61 temperaturas que se registraron en 61 días diferentes. Suponiendo que $\sigma = 1.7$°C, encuentre un estimado del intervalo de confianza del 95%, de la media poblacional de todas estas temperaturas. ¿Qué aspecto de este problema no es realista?

22. Pesos de osos La salud de la población de osos en el parque nacional de Yellowstone se controla por mediciones periódicas que se toman a osos anestesiados. Una muestra de 54 osos tiene un peso medio de 182.9 lb. Suponga que se sabe que σ es 121.8 lb, encuentre un estimado del intervalo de confianza del 99% de la media poblacional de todos estos pesos de osos. ¿Qué aspecto del problema no es realista?

23. Niveles de cotinina en fumadores Cuando las personas fuman, la nicotina que absorben se convierte en cotinina, que 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 el 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?

24. Circunferencias de la cabeza Para ayudar a identificar patrones de crecimiento anormales en los bebés, necesitamos construir un estimado del intervalo de confianza de la media de la circunferencia de la cabeza de todos los bebés con dos meses de vida. Se obtiene una muestra aleatoria de 100 bebés, y se encuentra que la media de la circunferencia de la cabeza es 40.6 cm. Suponiendo que se sabe que la desviación estándar poblacional es de 1.6 cm, calcule un estimado del intervalo de confianza del 99%

de la media de las circunferencias de la cabeza de todos los bebés de dos meses de edad. ¿Qué aspecto de este problema no es realista?

25. Tamaño de muestra para la media del CI de estudiantes de estadística La prueba Weschler de CI se diseñó para que la media sea 100 y la desviación estándar sea 15, para la población de adultos normales. Calcule el tamaño de la muestra necesario para estimar la media de la puntuación de CI de estudiantes de estadística. Queremos tener un nivel de 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. La desviación estándar para esta población es probablemente menor a 15, porque éste es 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 que se requiere.

26. Tamaño de muestra de pesos de 25 centavos de dólar La Tyco Video Game Corporation encontró que está perdiendo ingresos por las fichas que se usan en sus juegos de video. Las máquinas deben ajustarse para aceptar monedas sólo si caen entre límites que se fijaron desde antes. Para ajustar estos límites, debe estimarse la media del peso de monedas de un cuarto de dólar en circulación. Una muestra de monedas de un cuarto de dólar se pesará para determinar la media. ¿Cuántas monedas de un cuarto de dólar hay que 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 de la población real, para todas las monedas de un cuarto de dólar? Con base en los resultados de la muestra de monedas de un cuarto de dólar del conjunto de datos 29 en el Apéndice B, estimaríamos que la desviación estándar de la población es 0.068 g.

27. Tamaño de muestra para estimar ingresos Un economista quiere estimar la media de los ingresos por el primer año de trabajo de los graduados universitarios que demostraron gran sabiduría al tomar un curso de estadística. ¿Cuántos ingresos de este tipo deben encontrarse si queremos tener un nivel de confianza del 95% de que la media muestral está dentro de \$500 de la media poblacional real? Suponga que un estudio previo reveló que para ingresos de este tipo, $\sigma = \$6250$.

28. Tamaño de muestra para ver televisión Nielsen Media Research quiere estimar la media de la cantidad de tiempo (en minutos) que los estudiantes universitarios que estudian tiempo completo emplean viendo la televisión cada día de la semana. Calcule el tamaño de muestra necesario para estimar esta media con un margen de error de 15 minutos. Suponga que se desea un nivel de confianza del 96%. Suponga también que un estudio piloto mostró que la desviación estándar se estima en 112.2 minutos.

29. Tamaño de muestra utilizando la regla práctica del intervalo A usted lo acaba de contratar la división de mercadeo de General Motors, para estimar la media de la cantidad de dinero que se gasta ahora 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 que se gastan. Es razonable suponer que el rango típico de cantidades va desde \$12,000 hasta \$70,000. Luego use esa desviación estándar estimada para determinar el tamaño de muestra correspondiente a un nivel de confianza del 95% y a un margen de error de \$100. ¿Es práctico el tamaño de muestra? Si no es así, ¿qué se debe cambiar para obtener un tamaño de muestra práctico?

30. Tamaño de muestra con el uso de la regla práctica del intervalo Estime las duraciones mínima y máxima para los libros de texto típicos que se usan en cursos universitarios; después, use la regla práctica del intervalo para estimar la desviación estándar. Luego, encuentre el tamaño de muestra que se requiere para estimar la duración media (en años) de los libros de texto que se usan regularmente en cursos universitarios. Use un nivel de confianza del 90% y suponga que la media muestral tendrá un error no mayor de 0.25 años.

T **31.** Tamaño de muestra utilizando datos muestrales Quiere estimar la media del pulso de adultos hombres. 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 seleccionar al azar y examinar si quiere lograr un nivel de confianza del 95% de que la media muestral del pulso está dentro de dos latidos (por minuto) de la media poblacional μ real? Si en lugar de la regla práctica del intervalo se usa la desviación estándar de los pulsos de hombres del conjunto de datos 1 como un estimado de σ, ¿es muy diferente el tamaño de muestra que se requiere? ¿Qué tamaño de muestra parece estar más cerca del tamaño de muestra correcto?

T **32.** Tamaño de muestra con el uso de datos muestrales Usted quiere estimar la media del nivel de presión sanguínea diastólica de mujeres adultas. Remítase al conjunto de datos 1 en el Apéndice B, y calcule el nivel máximo y mínimo de presión sanguínea diastólica de mujeres; después, use tales valores con la regla práctica del intervalo para estimar σ. ¿Cuántas mujeres adultas debe seleccionar al azar y examinar si quiere tener un nivel de confianza del 95% de que la media muestral del nivel de presión sanguínea diastólica está dentro de 3 mm Hg de la media poblacional μ real? Si en lugar de la regla práctica del intervalo, se usa la desviación estándar de los niveles de presión sanguínea diastólica de mujeres del conjunto de datos 1 del Apéndice B como estimado de σ, ¿es muy diferente el tamaño de muestra que se requiere? ¿Qué tamaño de muestra parece estar más cercano al tamaño de muestra correcto?

6-3 Más allá de lo básico

33. Intervalo de confianza con factor de corrección por población finita El error estándar de la media es $\sigma/\sqrt{n}$, siempre y cuando el tamaño de la población sea infinito. Si el tamaño de la población es finito y se denota por 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 6-4, para que el margen de error sea como se muestra 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 $\sigma = 15$.

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

34. Tamaño de muestra con factor de corrección por población finita En la fórmula 6-4 para el margen de error E, suponemos que la población es infinita, que estamos realizando un muestreo sin 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 33, donde N es el tamaño de la población. Esta expresión del margen de error se resuelve para n y así obtener

$$n = \frac{N\sigma^2(z_{\alpha/2})^2}{(N-1)E^2 + \sigma^2(z_{\alpha/2})^2}$$

Repita el ejercicio 25, suponiendo que los estudiantes de estadística se seleccionan al azar y sin reemplazo, de una población de $N = 200$ estudiantes de estadística.

6-4 Estimación de la media poblacional: σ desconocida

En la sección 6-3 presentamos los métodos para construir un estimado del intervalo de confianza de una media poblacional μ que se desconoce, pero sólo consideramos casos en los que la desviación estándar poblacional σ se conoce. Señalamos que la suposición de una σ conocida no es muy realista, porque el cálculo de σ requiere que conozcamos todos los valores de la población, pero si conociésemos todos los valores de la población calcularíamos fácilmente el valor de la media poblacional μ; por lo tanto, no habría necesidad de estimar μ. En esta sección presentamos un método para construir estimados del intervalo de confianza de σ, sin el requisito de que σ se conozca. El procedimiento habitual consiste en reunir los datos muestrales y calcular el valor de los estadísticos n, $\bar{x}$ y s. Como los métodos de esta sección se basan en dichos estadísticos y no se requiere σ, son muy realistas, prácticos y se usan con frecuencia. Note que los siguientes supuestos para los métodos de esta sección *no* incluyen el requisito de que σ se conozca.

Supuestos

1. La muestra es aleatoria simple.
2. La muestra proviene de una población que se distribuye normalmente o $n > 30$.

Como en la sección 6-3, el requisito de una población que se distribuye 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 hay datos distantes y que el histograma tiene una forma que no es muy lejana a la de una distribución normal. Además, como en la sección 6-3, el requisito de que el tamaño de la muestra sea $n > 30$ suele usarse como un lineamiento, pero el tamaño de muestra mínimo realmente depende de cuánto se aleja la distribución de la población de la distribución normal. Usaremos el criterio simplificado de $n > 30$ como justificación del tratamiento de la distribución de medios muestrales, como si fuese una distribución normal. La distribución muestral de medias muestrales $\bar{x}$ es *exactamente* una distribución normal con media μ y desviación estándar $\sigma/\sqrt{n}$, siempre y cuando la población tenga una distribución normal, con media μ y desviación estándar σ. Si la población no está distribuida normalmente, muestras mayores producen medias muestrales con una distribución que es *aproximadamente* normal, con media μ y desviación estándar $\sigma/\sqrt{n}$.

Igual que en la sección 6-3, la media muestral $\bar{x}$ es el mejor estimado puntual (o estimado de un solo valor) de la media poblacional μ. Así como en la sección 6-3, la distribución de las medias muestrales $\bar{x}$ tiende a ser más consistente (con *menor variación*) que las distribuciones de otros estadísticos muestrales, y la media muestral $\bar{x}$ es un *estimador sin sesgo* coincidente con la media poblacional μ.

La media muestral $\bar{x}$ es el mejor estimado puntual de la media poblacional μ.

En las secciones 6-2 y 6-3 señalamos que hay una grave limitación de la utilidad de un estimado puntual: el valor individual de un estimado puntual no revela qué tan bueno es ese estimado. Los intervalos de confianza nos dan información más significativa porque proporcionan un rango de valores asociado con un grado de posibilidad de que el intervalo verdaderamente contenga el valor real de μ.

He aquí el punto clave de esta sección: si σ no se conoce, pero las condiciones de arriba se satisfacen, en lugar de usar la distribución normal, utilizamos la *distribución t de Student*, que desarrolló William Gosset (1876-1937). Gosset fue un empleado de la cervecería Guinness Brewery, que necesitaba una distribución para utilizar con muestras pequeñas. La cervecería irlandesa donde trabajaba no permitía la publicación de resultados de investigaciones, entonces Gosset publicó bajo el seudónimo de *Student*.

Estimación de azúcar en las naranjas

En Florida, los miembros de la industria de los cítricos usan extensamente métodos estadísticos. Una aplicación específica tiene que ver con la forma en que se les paga a los agricultores por las naranjas que se emplean para elaborar jugo de naranja. Cuando llega un camión con naranjas, primero se pesa la fruta en la planta receptora, luego se elige aleatoriamente una muestra de una docena de naranjas. La muestra se pesa, después 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 que contiene 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.

Extractos de una circular del departamento de transporte

Los siguientes extractos de una circular del Departamento del Transporte de Estados Unidos atañen a algunos de los requisitos de exactitud para el equipo de navegación que se emplea en aviones. Observe el uso del intervalo de confianza. "El total de las contribuciones de error del equipo a bordo, si se combinaran con los errores técnicos de vuelo apropiados incluidos en la lista, no debe exceder del 95% con un nivel de confianza (2-sigma), durante un periodo igual al ciclo de actualización." "El sistema de vías y rutas aéreas de Estados Unidos tiene anchos de ruta de protección que se utilizan en un sistema VOR con una exactitud de $\pm$ 4.5 grados basada en una probabilidad del 95%".

Puesto que no conocemos el valor de σ, lo estimamos con el valor de la desviación estándar muestral s, lo cual introduce otra fuente de falta de confiabilidad, en especial con muestras pequeñas. Para mantener un intervalo de confianza en algún nivel que se desea, como el 95%, compensamos esta falta de confiabilidad adicional haciendo más ancho el intervalo de confianza: utilizamos valores críticos mayores que los valores críticos de $z_{\alpha/2}$ que se utilizaron en la sección 6-3 donde se conocía σ. En lugar de valores críticos de $z_{\alpha/2}$, utilizamos los valores críticos mayores de $t_{\alpha/2}$ que calculamos con la distribución t de Student.

Distribución *t* de Student

Si la distribución de una población es esencialmente normal (con forma aproximada de campana), entonces la distribución de

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

es esencialmente una **distribución *t* de Student** para todas las muestras de tamaño n. La distribución t de Student, que a menudo se refiere como la **distribución *t***, se utiliza para calcular valores críticos denotados por $t_{\alpha/2}$.

Pronto analizaremos algunas de las propiedades importantes de la distribución t, primero 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. Para encontrar un valor crítico $t_{\alpha/2}$ en la tabla A-3, localice el número apropiado de *grados de libertad* en la columna izquierda y avance a través de la fila correspondiente hasta encontrar el número que se encuentra directamente abajo del área adecuada en la parte superior.

Definición

El número de **grados de libertad** para un conjunto de datos muestrales que se recolectaron es el número de valores muestrales que pueden variar tras 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, asignamos con libertad valores a las primeras nueve puntuaciones, pero la 10a se calcula. La suma de las 10 puntuaciones debe ser 800, entonces la 10a será igual a 800 menos la suma de las primeras nueve puntuaciones. Puesto que estas primeras nueve puntuaciones las seleccionamos con libertad para que sean cualquier valor, decimos que hay nueve grados de libertad disponibles. Para las aplicaciones de esta sección, el número de grados de libertad es simplemente el tamaño de la muestra menos 1.

$$\textbf{grados de libertad} = \boldsymbol{n - 1}$$

EJEMPLO Calcular un valor crítico Una muestra de tamaño $n = 15$ es una muestra aleatoria simple seleccionada de una población que se distribuyó normalmente. Calcule el valor crítico $t_{\alpha/2}$ correspondiente a un nivel de confianza del 95%.

SOLUCIÓN Puesto que $n = 15$, el número de grados de libertad está dado por $n - 1 = 14$. Utilizando la tabla A-3, localizamos la fila 14, que se refiere a la columna de la extrema izquierda. Como en la sección 6-2, un nivel de confianza del 95% corresponde a $\sigma = 0.05$, entonces encontramos los valores que se listan en la columna para *un área de 0.05 en dos colas*. El valor correspondiente a la fila para 14 grados de libertad y a la columna para un área de 0.05 en dos colas es 2.145; luego, $t_{\alpha/2} = 2.145$.

Ahora que sabemos cómo encontrar valores críticos denotados por $t_{\alpha/2}$ describimos el margen de error E de este intervalo de confianza.

Margen de error E para la estimación de μ (con σ desconocida)

Fórmula 6-6
$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

donde $t_{\alpha/2}$ tiene $n - 1$ grados de libertad.

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 procedimiento siguiente utiliza el margen de error de arriba 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 se satisfacen los supuestos requeridos. (Tenemos una muestra aleatoria simple; además, la población parece distribuirse 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 que se desea.
3. Evalúe el margen de error $E = t_{\alpha/2} s/\sqrt{n}$.

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 dichos 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 trabaja con estadísticos resumidos (n, $\bar{x}$, s), redondee los límites del intervalo de confianza al mismo número de lugares decimales que se utilizaron para la media muestral.

EJEMPLO Construcción de un intervalo de confianza En la sección 6-3 incluimos un ejemplo que ilustró la construcción de un intervalo de confianza para estimar μ. Utilizamos la muestra de temperaturas corporales del conjunto de datos 4 del Apéndice B (para las 12:00 h del día 2), con $n = 106$ y $\bar{x} = 98.20°F$; también supusimos que la muestra era una muestra aleatoria simple y que σ "de alguna forma se conoce que es 0.62°F". En la realidad, σ no se conoce. Con los estadísticos $n = 106$, $\bar{x} = 98.20°F$ y $s = 0.62°F$ (con σ desconocida) que se obtuvieron de una muestra aleatoria simple, calcule lo siguiente utilizando otra vez un nivel de confianza del 95%:

a. El margen de error E

b. El intervalo de confianza para μ.

SOLUCIÓN

1. Primero debemos verificar que los dos supuestos para esta sección se satisfacen. Tenemos una muestra aleatoria simple y $n > 30$. (Puesto que $n > 30$, no es necesario revisar que la muestra parezca proveniente de una población que se distribuye normalmente). Por lo tanto, procedemos a construir un intervalo de confianza del 95%, utilizando la distribución t.
2. Después calculamos el valor crítico de $t_{\alpha/2} = 1.984$; éste se encuentra en la tabla A-3, como el valor crítico que corresponde a $n - 1 = 105$ grados de libertad (columna izquierda de la tabla A-3) y un área de dos colas de 0.05. (Recuerde, un nivel de confianza del 95% corresponde a $\alpha = 0.05$, que se divide por igual entre las dos colas). La tabla A-3 no incluye 105 grados de libertad, entonces seleccionamos el número más cercano de grados de libertad, que es 100. El valor correcto de $t_{\alpha/2}$ para 105 grados de libertad es 1.983; por lo tanto, el uso del valor más cercano en la tabla A-3 de 1.984 produce aquí un error despreciable.
3. *Calcule el margen de error E*: El margen de error $E = 0.11947593$ se calcula usando la fórmula 6-2 como se muestra abajo, con los espacios decimales de más para minimizar el error de redondeo en el intervalo de confianza que se calculará en el paso 4.

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 1.984 \cdot \frac{0.62}{\sqrt{106}} = 0.11947593$$

4. *Calcule el intervalo de confianza:* Ahora el intervalo de confianza puede calcularse usando $\bar{x} = 98.20$ y $E = 0.11947593$ como se muestra abajo:

$$\bar{x} - E < \mu < \bar{x} + E$$
$$98.20 - 0.11947593 < \mu < 98.20 + 0.11947593$$
$$98.08052407 < \mu < 98.31947593$$

5. Redondee los límites del intervalo de confianza. Como la media muestral de 98.20 utiliza dos espacios decimales, redondee el resultado a dos espacios decimales para obtener: $98.08 < \mu < 98.32$.

INTERPRETACIÓN Este resultado también podría expresarse en la forma de 98.20 ± 0.12 o (98.08, 98.32). Con base en los resultados muestrales que ya se dieron, tenemos un nivel de confianza del 95% de que los límites de 98.08°F y 98.32°F realmente contienen el valor de la media poblacional μ. Note que los límites del intervalo de confianza no contienen a 98.6°F, el valor que suele considerarse la media de la temperatura corporal. Con base en estos resultados, parece que el valor que se considera comúnmente de 98.6°F, es *erróneo*.

El intervalo de confianza que se calculó en el ejemplo anterior parece ser el mismo que el de la sección 6-3, donde usamos la distribución normal y el supuesto de que se sabe que σ es 0.62°F. En realidad, los dos intervalos de confianza son iguales sólo después de redondearlos. Sin redondeo, el intervalo de confianza de la sección 6-3 es (98.08196934, 98.31803066), y el intervalo de confianza que se calculó aquí es (98.08052407, 98.31947593). En algunos otros casos, las diferencias pueden ser mucho mayores.

Ahora listamos las propiedades importantes de la distribución t, que utilizamos en esta sección.

Propiedades importantes de la distribución t de Student

1. La distribución t de Student es diferente para distintos tamaños de muestra. (Véase la figura 6-5 para los casos $n = 3$ y $n = 12$).

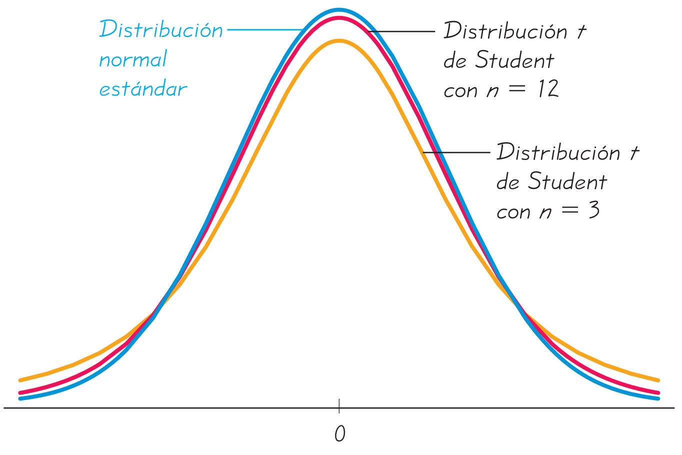


FIGURA 6-5 Distribuciones t de Student para $n = 3$ y $n = 12$

La distribución t de Student tiene la misma forma y simetría general de la distribución normal estándar, pero refleja una mayor variabilidad, de la que se espera con muestras pequeñas.

2. 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 la que se espera con muestras pequeñas.
3. La distribución t de Student tiene una media de $t = 0$ (precisamente 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 de la muestra, pero es mayor que 1 (no como la distribución normal estándar, que tiene $\sigma = 1$).
5. Conforme el tamaño de la muestra 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 6-6 y el de la tabla 6-1 adjunta el resumen los puntos clave a considerarse cuando se construyen intervalos de confianza para estimar μ, la media poblacional. En la figura 6-6 o en la tabla 6-1, note que si tenemos una muestra pequeña ($n \leq 30$) que se seleccionó de una distribución que difiera drásticamente de una distribución normal, no es posible usar los métodos descritos en este capítulo. Una alternativa es utilizar métodos no paramétricos (véase capítulo 12) y otra es usar el método de *bootstrap* por computadora. En ambos métodos no se hacen supuestos acerca de la población original.

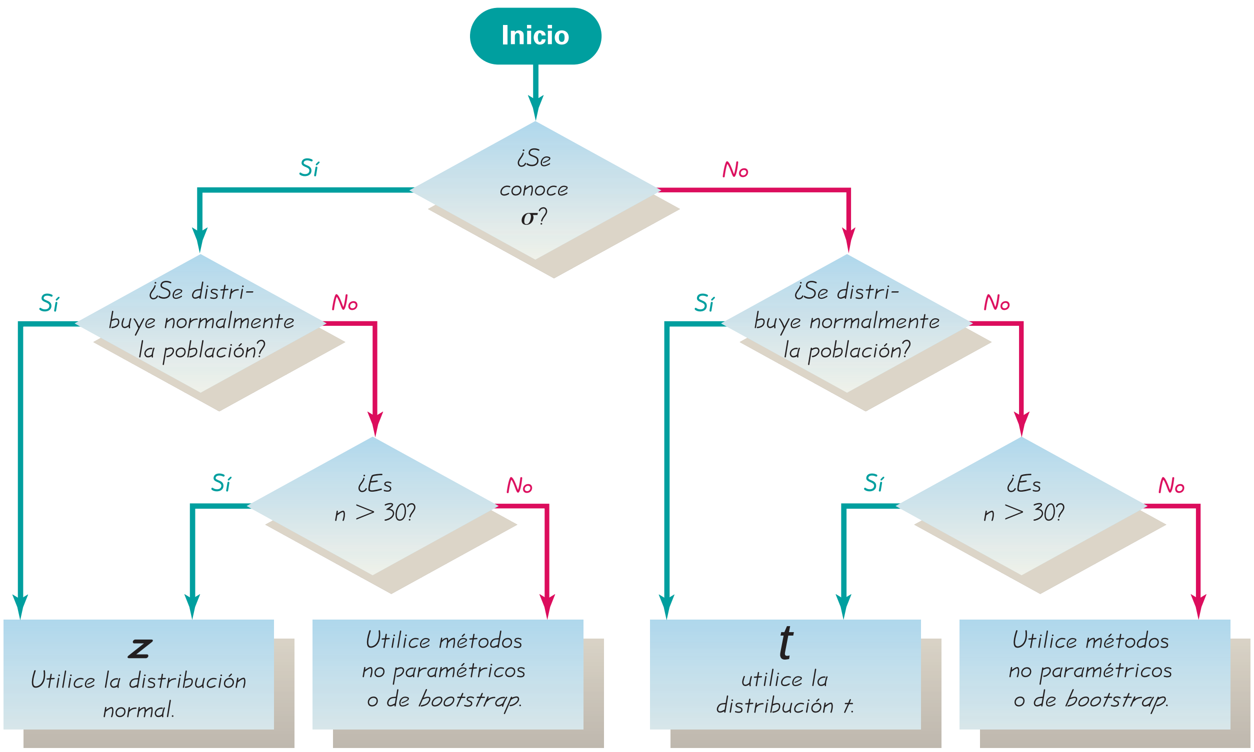


FIGURA 6-6 **Elección entre z y t**

Tabla 6-1 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 *bootstrap*.	La población no se distribuyó normalmente y $n \leq 30$

Notas: 1. **Criterio para decidir si la población se distribuye normalmente:** La población no necesita ser normal exactamente, pero debe parecer un tanto simétrica, con una moda y sin datos distantes.

2. **Tamaño de muestra $n > 30$:** Éste es un lineamiento que se usa regularmente, pero tamaños de la muestra de 15 a 30 son adecuados si la población parece tener una distribución normal y no hay datos distantes. Para algunas distribuciones poblacionales que estén muy alejadas de la normal, puede requerirse que el tamaño de la muestra sea mayor de 50 o aun de 100.

El método de *bootstrap* se describe en el proyecto tecnológico al final de este capítulo.

El siguiente ejemplo se enfoca en escoger la aproximación correcta utilizando los métodos de esta sección y la sección 6-3.

EJEMPLO Elección de distribuciones Suponiendo que usted planea construir un intervalo de confianza para la media poblacional μ, utilice los datos dados 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 éstos (es decir, los métodos de la sección 6-3 y de esta sección no se pueden utilizar).

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).

continúa

SOLUCIÓN Remítase a la figura 6-6 o a la tabla 6-1 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 6-6.

b. Puesto que la desviación estándar poblacional σ no se conoce y la población se distribuye normalmente, el margen de error se calcula usando $t_{\alpha/2}$ en la fórmula 6-6.

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 6-3 ni los de esta sección.

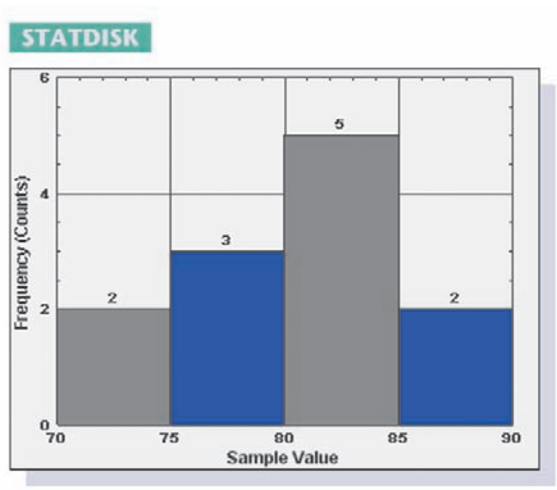
d. Puesto que la desviación estándar poblacional σ se conoce y la muestra es grande ($n > 30$), el margen de error se calcula usando $z_{\alpha/2}$ en la fórmula 6-4.

e. Puesto que la población no se distribuye normalmente y 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 6-3 ni los de esta sección.

EJEMPLO Intervalo de confianza para *Harry Potter* El conjunto de datos 14 en el Apéndice B incluye las puntuaciones de facilidad de lectura de Flesch para 12 páginas diferentes que se seleccionaron aleatoriamente de *Harry Potter y la piedra filosofal*, de J. K. Rowling. Utilice la muestra aleatoria simple de estos dos valores para construir un estimado del intervalo de confianza del 95% de μ, la media de la puntuación de facilidad de lectura de Flesch para todas las páginas del libro.

SOLUCIÓN

1. Primero debemos verificar que se satisfagan los dos supuestos para esta sección. Tenemos una muestra aleatoria simple. Ya que el tamaño de la muestra $n = 12$ no excede de 30, asegurémonos de que la población tenga una distribución que sea aproximadamente normal. La representación adjunta



de la pantalla del STATDISK exhibe que los 12 valores muestrales generan un histograma con una forma de campana aproximada, entonces veremos que la población tiene una distribución que es aproximadamente normal. Mientras se ejecuta STATDISK, también encontramos que $\bar{x} = 80.75$ y $s = 4.68$ para la muestra de 12 puntuaciones de lectura. Con σ desconocida y una población que se distribuye normalmente, ahora procedemos a construir un intervalo de confianza del 95% utilizando la distribución t.

2. Después encontramos el valor crítico de $t_{\alpha/2} = 2.201$. Éste se encuentra en la tabla A-3, como el valor crítico correspondiente a $n - 1 = 11$ grados de libertad (columna izquierda de la tabla A-3) y un área de dos colas de 0.05. (Recuerde, un nivel de confianza del 95% corresponde a $\alpha = 0.05$, que se divide por igual entre las dos colas).

3. *Calcule el margen de error E:* El margen de error $E = 2.97355$ se calcula con la fórmula 6-6, como se muestra abajo, utilizando lugares decimales de más para minimizar el error de redondeo en el intervalo de confianza que se calculará en el paso 4.

$$E = t_{\alpha/2}\frac{s}{\sqrt{n}} = 2.201 \cdot \frac{4.68}{\sqrt{12}} = 2.97355$$

4. *Calcule el intervalo de confianza:* El intervalo de confianza puede ahora calcularse utilizando $\bar{x} = 80.75$ y $E = 2.97355$, como se muestra abajo:

$$\bar{x} - E < \mu < \bar{x} + E$$

$$80.75 - 2.97355 < \mu < 80.75 + 2.97355$$

$$77.77645 < \mu < 83.72355$$

5. Redondee los límites del intervalo de confianza. Puesto que los datos muestrales originales utilizan un decimal, el resultado se redondea a un espacio *adicional* para dar este resultado con dos espacios decimales: $77.78 < \mu < 83.72$.

INTERPRETACIÓN Con base en los datos muestrales, tenemos un nivel de confianza del 95% de que los límites de 77.78 y 83.72 realmente contienen el valor de la media de la puntuación de facilidad de lectura de Flesch para todas las páginas de *Harry Potter y la piedra filosofal*.

Cálculo del estimado puntual y *E* desde un intervalo de confianza

Posteriormente en esta sección describiremos cómo pueden utilizarse los programas de cómputo y las calculadoras para encontrar un intervalo de confianza. Un uso común requiere que usted ingrese un nivel de confianza y estadísticos muestrales; la pantalla mostrará los límites del intervalo de confianza. La media muestral $\bar{x}$ es el valor central entre estos límites y 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 $\bar{x} - E$, 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 En el análisis de las edades de todos los polizones del Queen Mary que se listan en el conjunto de datos 15 del Apéndice B, se obtiene la pantalla del 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 aleatoriamente de una población grande.

```
95.0% CI
( 24.065, 27.218)
```

SOLUCIÓN En los cálculos siguientes, los resultados se redondean a un decimal, lo cual es un espacio decimal adicional más del redondeo que se utilizó para la lista de edades original.

$$\bar{x} = \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2}$$

$$= \frac{27.218 + 24.065}{2} = 25.6 \text{ años}$$

$$E = \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2}$$

$$= \frac{27.218 - 24.065}{2} = 1.6 \text{ años}$$

Uso de los intervalos de confianza para describir, explorar o comparar datos

En algunos casos, se utiliza un intervalo de confianza para lograr el objetivo final de estimar el valor de un parámetro poblacional. Para los datos de temperatura corporal manejados en esta sección, un objetivo importante sería estimar la media de la temperatura corporal de adultos saludables; nuestros resultados sugieren con fuerza que el valor de 98.6°F, que se utiliza comúnmente, es incorrecto (puesto que tenemos un 95% de confianza de que los límites de 98.08°F y 98.32°F contienen el valor real de la media poblacional). En otros casos, un intervalo de confianza puede ser una de varias herramientas diferentes que se utilizan para describir, explorar o comparar conjuntos de datos.

Cuidado: Como en las secciones 6-2 y 6-3, es posible usar los intervalos de confianza de manera informal para comparar diferentes conjuntos de datos, pero el traslape de intervalos de confianza no debe emplearse para hacer conclusiones formales ni finales acerca de la igualdad de las medias. Los últimos capítulos incluirán procedimientos para decidir si dos poblaciones tienen medias iguales; esos métodos no enfrentarán las dificultades que se asocian con las comparaciones que se basan en el traslape de intervalos de confianza.

No utilice el traslape de intervalos de confianza como base para hacer conclusiones formales acerca de la igualdad de las medias.

Considere tres conjuntos de datos diferentes que consisten en puntuaciones de facilidad de lectura de Flesch para 12 páginas que se seleccionaron al azar de cada uno de estos tres libros: *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. Las puntuaciones de Flesch están en una escala de 1 a 100, donde los trabajos que son más fáciles de leer recibieron las puntuaciones más altas. (Véase el conjunto de datos 14 en el Apéndice B para las listas de puntuaciones muestrales). Los histogramas y las gráficas cuantilares normales sugieren que las tres distribuciones no están tan lejos de ser distribuciones normales. Los estadísticos descriptivos que se anexan se utilizan para encontrar los intervalos de confianza del 95%. (En cada caso, $n = 12$; entonces el valor crítico $t_{\alpha/2} = 2.201$ se calcula con 11 grados de libertad, en tanto que el margen de error es de $E = 2.201s/\sqrt{12}$). En la figura 6-7 graficamos los tres intervalos de confianza para compararlos mejor.

Autor	Estadísticos descriptivos	Intervalo de confianza del 95%
Clancy	$n = 12, \bar{x} = 70.73, s = 11.33$	$63.53 < \mu < 77.93$
Rowling	$n = 12, \bar{x} = 80.75, s = 4.68$	$77.78 < \mu < 83.72$
Tolstoi	$n = 12, \bar{x} = 66.15, s = 7.86$	$61.16 < \mu < 71.14$

Al comparar los estadísticos descriptivos de las tres muestras, vemos que las medias parecen ser muy diferentes. Sin embargo, la figura 6-7 nos indica que hay algún traslape entre los intervalos de confianza. Puesto que los intervalos de confianza para Tolstoi y para Rowling no se traslapan, parece que dichos autores tienen niveles de escritura muy diferentes, siendo Rowling el más fácil de leer. Clancy y Rowling apenas se traslapan, pero el traslape sugiere que sus medias poblacionales no son significantemente diferentes; entonces, no deberíamos concluir que Rowling tiene una media de puntuación Flesch más alta que la media de Clancy. Sin embargo, *todas las conclusiones que se basaron en el traslape de intervalos de confianza deben considerarse indicaciones tentativas, no conclusiones definitivas*. Los últimos capítulos introducirán métodos mejores y más confiables para determinar si las medias de la población son iguales.

Método alternativo (no se utiliza en este libro) En esta sección presentamos un método para construir un estimado del intervalo de confianza de la media

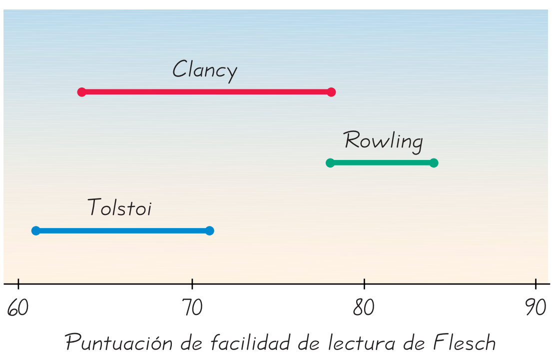


FIGURA 6-7 Comparación de intervalos de confianza

poblacional μ; este método supone que el valor de σ no se conoce. Un método alternativo que *no se utiliza en este libro* es sustituir s por σ siempre y cuando $n > 30$, para luego proceder como si σ se conociera (como en la sección 6-3). Entonces, el criterio para escoger entre las distribuciones normal y t se basa en las importantes consideraciones siguientes:

1. Con el mismo criterio que se usa en el mundo real, se estudiaron cientos de artículos de revistas profesionales.
2. Con σ desconocida, la distribución de $\dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ es una distribución t, no una distribución normal. Para tamaños de muestra muy grandes, las diferencias entre las distribuciones normal y t son despreciables, aunque el uso de la distribución t por lo general proporciona mejores resultados.
3. Después de tomar un curso introductorio de estadística, algunos estudiantes siguen adelante y toman cursos más avanzados que suelen utilizar la distribución t cuando σ no se conoce. Sería mejor que aprendieran un procedimiento que pudieran emplear nuevamente en un curso posterior que aprender uno que deba cambiarse después.
4. Trabajar con la distribución t no es mucho más difícil que hacerlo con la distribución normal, especialmente si se dispone de programas de cómputo o de una calculadora TI-83 Plus. Además, el uso de la tabla A-3 ayuda a fortalecer habilidades para emplear tablas que son importantes para actividades como determinar cantidades de impuestos a partir de tablas de impuestos de ingresos.

Utilizando la tecnología

Los procedimientos siguientes, que se aplican a intervalos de confianza para estimar una media μ, incluyen los intervalos de confianza descritos en la sección 6-3 y 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 primero que los supuestos requeridos se satisfagan. Véase los supuestos que se listan cerca del principio de esta sección y de la sección 6-3.

STATDISK Primero debe encontrar el tamaño de muestra n, la media muestral $\bar{x}$ y la desviación estándar muestral s. (Véase el procedimiento del STATDISK descrito en la sección 2-4). Seleccione **Analysis** de la barra del menú principal, elija **Confidence Intervals** y, después, seleccione **Population Mean**. Proceda a ingresar los elementos en el cuadro de diálogo; entonces, haga clic en el botón **Evaluate**. El intervalo de confianza aparecerá en la pantalla.

Minitab Minitab requiere que usted ingrese una lista de los valores muestrales originales. Minitab no realiza cálculos utilizando sólo los estadísticos que se resumen de n, $\bar{x}$ y s. El *Minitab Student Laboratory Manual and Workbook*, un suplemento de este libro de texto, describe un truco para trabajar en torno a esta limitación del Minitab. Si tiene una lista de los valores muestrales originales, ingrésela en la columna C1, luego seleccione **Stat** y **Basic Statistics**. Si no se conoce σ, elija **1-sample t** e ingrese **C1** en el cuadro de **Variables**. (Si σ se conoce, seleccione **1-sample Z**, ingrese **C1** en el cuadro de variables y luego el valor de σ en el cuadro "Sigma"). Haga clic en el botón **OK**. Para más detalles, véase la sección 5-5 del libro de trabajo de Minitab.

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 icono con forma de lápiz e ingrese el rango de datos, con el formato A1:A12, sólo si 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 es recomendable el uso de la herramienta para calcular intervalos de confianza de Excel. Ésta supone que se conoce σ; usted debe encontrar primero el tamaño de muestra n y la desviación estándar muestral s (que es posible calcular usando ***fx***, **Statistical**, **STDEV**). En lugar de generar el intervalo de confianza completo con límites específicos, tal herramienta calcula sólo el margen de error E. Entonces debe restar dicho resultado a $\bar{x}$ y sumarlo a $\bar{x}$ con la finalidad de identificar los límites reales del intervalo de confianza. Para utilizar esta herramienta cuando se conoce σ, haga clic en ***fx***, seleccione la categoría de funciones **Statistical** y el elemento de **CONFIDENCE**. En el cuadro de diálogo, ingrese el valor de α (que se llama nivel de significancia), la desviación estándar y el tamaño de muestra. El resultado será el valor del margen de error E.

TI-83 Plus **TI-83 Plus** La calculadora TI-83 Plus se puede utilizar para generar intervalos de confianza para valores muestrales originales que se guardaron en una lista; lo mismo ocurre con los estadísticos resumidos $n, \bar{x}$ y s. Ingrese los datos en la lista L1 o tenga disponibles los estadísticos resumidos; luego, presione la tecla **STAT**. Ahora seleccione **TESTS** y escoja **TInterval** si no se conoce σ. (Escoja **ZInterval** si se conoce σ). Después de efectuar los ingresos que se requieren, la pantalla de la calculadora incluirá el intervalo de confianza en el formato $(\bar{x} - E, \bar{x} + E)$.

6-4 Destrezas y conceptos básicos

Uso de la distribución correcta. *En los ejercicios 1 a 8, realice una de las siguientes acciones, según sea lo apropiado: a) calcule el valor crítico $z_{\alpha/2}$, b) calcule el valor crítico $t_{\alpha/2}$, c) establezca que no se aplican ni la distribución normal ni la distribución t.*

1. 95%; $n = 5$; σ no se conoce; la población parece distribuirse normalmente.

2. 95%; $n = 10$; σ no se concoe; la población parece distribuirse normalmente.

3. 99%; $n = 15$; σ se conoce; la población parece estar muy sesgada.

4. 99%; $n = 45$; σ se conoce; la población parece estar muy sesgada.

5. 90%; $n = 92$; σ no se conoce; la población parece distribuirse normalmente.

6. 90%; $n = 9$; $\sigma = 4.2$; la población parece estar muy sesgada.

7. 98%; $n = 7$; $\sigma = 27$; la población parece distribuirse normalmente.

8. 98%; $n = 37$; σ no se conoce; la población parece distribuirse normalmente.

Cálculo de intervalos de confianza. *En los ejercicios 9 y 10, utilice el nivel de confianza dado y los datos muestrales para calcular a) el margen de error, y b) el intervalo de confianza para la media poblacional μ. Suponga que la población tiene una distribución normal.*

9. Calificaciones del SAT en matemáticas para mujeres: 95% de confianza; $n = 15, \bar{x} = 496, s = 108$

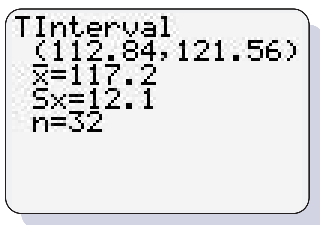
10. Longitud del codo a la punta del dedo de los hombres: 99% de confianza; $n = 32, \bar{x} = 14.50$ pulgadas, $s = 0.70$ pulgadas.

Interpretación de pantalla de calculadora. *En los ejercicios 11 y 12, utilice los datos dados y la imagen de la pantalla de la calculadora TI-83 Plus correspondiente para expresar el intervalo de confianza en el formato de $\bar{x} - E < \mu < \bar{x} + E$. Además, escriba una afirmación que interprete el intervalo de confianza.*

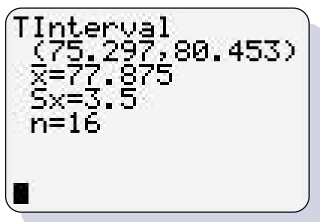
11. Puntuaciones de CI de estudiantes de estadística: 95% de confianza; $n = 32, \bar{x} = 117.2, s = 12.1$

12. Estaturas de jugadores de la NBA: 99% de confianza; $n = 16, \bar{x} = 77.875$ pulgadas, $s = 3.50$ pulgadas.

TI-83 Plus **Ejercicio 11**



TI-83 Plus **Ejercicio 12**



Construcción de intervalos de confianza. *En los ejercicios 13 a 24, construya el intervalo de confianza.*

13. Destrucción de Vipers de Dodge En la prueba destructiva, los elementos muestrales se destruyen en el proceso de probarlos. La prueba de choque de automóviles es un ejemplo muy costoso de prueba destructiva. Doce automóviles deportivos Dodge Viper (precio de lista: \$59,300) se prueban por choque en una variedad de condiciones que simulan colisiones típicas. El análisis de los 12 automóviles que se dañaron resulta en costos de reparación con una distribución que parece tener forma de campana, con una media de $\bar{x} = \$26{,}227$ y una desviación estándar de $s = \$15{,}873$ (de acuerdo con datos del Highway Loss Data Institute). Calcule el estimado del intervalo del 95% de μ, la media del costo de reparación para todos los Dodge Viper que se ocuparon en colisiones e interprete el resultado.

14. Costos hospitalarios por choque Se realizó un estudio para estimar los costos hospitalarios para víctimas de accidente que usaban cinturones de seguridad. Veinte casos que se seleccionaron aleatoriamente presentan una distribución que parece tener forma de campana, con una media de \$9004 y una desviación estándar de \$5629 (según datos del Departamento del Transporte de Estados Unidos).

a. Construya el intervalo de confianza del 99% para la media de todos los costos de este tipo.

b. Si usted es director de una compañía de seguros que ofrece tarifas más bajas para conductores que usan cinturones de seguridad, y desea un estimado conservador para la peor situación posible, ¿qué cantidad debe aplicar como posible costo hospitalario para una víctima de accidente que utiliza cinturón de seguridad?

15. Pronóstico y temperaturas reales El conjunto de datos 10 en el 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 del pronóstico para tres días de la temperatura máxima real, el resultado es una lista de 31 valores con una media de $-0.419°$ y una desviación estándar de 3.704°.

a. Construya un estimado del intervalo de confianza del 99%, de la media de la diferencia entre todas las temperaturas máximas reales y las temperaturas máximas del pronóstico para tres días.

b. ¿Incluye 0° el intervalo de confianza? 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 $-0.419°$, ¿parece ser válida esa afirmación? ¿Por qué?

16. Estaturas de padres El conjunto de datos 2 del Apéndice B incluye las estaturas de padres de 20 hombres. Si se calcula la diferencia en este dato para cada pareja de padres, restando la estatura de la madre de la estatura del padre, el resultado es una lista de 20 valores con una media de 4.4 pulgadas y una desviación estándar de 4.2 pulgadas. Un histograma y una gráfica cuantilar normal sugieren que la población presenta una distribución que no está lejos de la normal.

a. Construya un estimado del intervalo de confianza del 99%, de la media de la diferencia entre las estaturas de las madres y la de los padres.

b. ¿El intervalo de confianza incluye 0 pulgadas? Si un sociólogo afirma que las mujeres tienden a casarse con hombres que son más altos que ellas, ¿fundamenta el intervalo de confianza esta afirmación? ¿Por qué?

17. Estimación de contaminación por automóviles En una muestra de siete automóviles, cada uno se verificó para emisiones de óxido nitroso (en gramos por milla); de esto, se obtuvieron los resultados siguientes: 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 sea representativa de los automóviles en circulación, construya un estimado del intervalo de confianza del 98% de la cantidad media de emisiones de óxido nitroso para todos los automóviles. Si la agencia de protección ambiental requiere que las emisiones de óxido nitroso sean menores que 0.165 gramos/milla, ¿sería posible concluir con seguridad que se está cumpliendo tal requisito?

18. Control de plomo en el aire En la lista de abajo se incluyen cantidades medidas de plomo (en microgramos por metro cúbico o $\mu g/m^3$) en el aire. La agencia de protección ambiental estableció un estándar de calidad del aire para el plomo: 1.5 $\mu g/m^3$. Las mediciones que se muestran 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, había 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 cantidad media de plomo en el aire. ¿Hay algo acerca de este conjunto de datos que sugiera que el intervalo de confianza puede no ser muy bueno? Explique.

5.40	1.10	0.42	0.73	0.48	1.10

19. Ritmos cardiacos al trabajar con pala Ya que las muertes por deficiencias cardiacas parecen incrementarse después de las fuertes nevadas, se diseñó un experimento para comparar las demandas cardiacas de quienes remueven la nieve con una pala contra las de aquellos que utilizan un aparato eléctrico para retirarla. Diez sujetos despejaron de nieve el terreno usando ambos métodos; en consecuencia, se registraron sus frecuencias cardiacas máximas (en pulsos por minuto) durante ambas actividades. Se obtuvieron los resultados siguientes (según datos de "Cardiac Demands of Heavy Snow Shoveling", de Franklin *et al*., *Journal of the American Medical Association*, vol. 273, núm. 11):

Frecuencias cardiacas máximas de paleo de nieve manual: $n = 10, \bar{x} = 175, s = 15$

Frecuencias cardiacas máximas con aparato eléctrico para retirar nieve: $n = 10$, $\bar{x} = 124, s = 18$

a. Calcule el estimado del intervalo de confianza del 95% de la media poblacional para aquellas personas que palean nieve de manera manual.
b. Calcule el estimado del intervalo de confianza del 95% de la media poblacional de aquellas personas que usan el aparato eléctrico para retirar nieve.
c. Si fuese un doctor que se preocupa por las muertes a consecuencia de deficiencias cardiacas, que se fomentan por el paleo manual de nieve, ¿qué valor individual del intervalo de confianza del inciso *a* sería de mayor preocupación?
d. Compare los intervalos de confianza de los incisos *a* y *b*; interprete lo que encontró.

20. Pulso Una doctora quiere desarrollar criterios para determinar si el pulso de un paciente es anormal y determinar si hay diferencias significativas entre hombres y mujeres. Utilizando los pulsos muestrales del conjunto de datos 1 en el Apéndice B, los pulsos de hombres se resumen con los estadísticos $n = 40, \bar{x} = 69.4, s = 11.3$. Para las mujeres, los estadísticos son $n = 40, \bar{x} = 76.3, s = 12.5$.
a. Construya un estimado del intervalo de confianza del 95% de la media del pulso para hombres.
b. Construya un estimado del intervalo de confianza del 95% de la media del pulso para mujeres.
c. Compare los resultados anteriores. ¿Es posible concluir que las medias poblacionales para hombres y para mujeres son diferentes? ¿Por qué?

21. Amplitud craneana Amplitud de muestras de cráneos de hombres egipcios del 4000 a. C. y 150 d. C. (datos que se tomaron 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 los tamaños de la cabeza a través del tiempo sugieren una transculturación con personas de otras regiones. Utilice intervalos de confianza para determinar si los tamaños de la cabeza cambiaron del 4000 a. C. al 150 d. C. Explique su resultado.

T **22.** Circunferencias de la cabeza Para diagnosticar correctamente el trastorno de la hidrocefalia, un pediatra investiga las circunferencias de la cabeza de niños y niñas de dos años de edad. Utilice los datos muestrales del conjunto de datos 3 para construir intervalos de confianza, luego determine si hay una diferencia entre los dos géneros.

T **23.** Comparación de la Pepsi regular y la de dieta Remítase al conjunto de datos 17 en el Apéndice B y utilice los datos muestrales.

a. Construya un estimado del intervalo de confianza del 95% del peso medio del líquido en latas de Pepsi regular.

b. Construya un estimado del intervalo de confianza de 95% del peso medio del líquido en las latas de Pepsi de dieta.

c. Compare los resultados de los incisos *a* y *b*; luego interprételos.

T **24.** Índice de masa corporal Remítase al conjunto de datos 1 en el Apéndice B y utilice los datos muestrales.

a. Construya un estimado del intervalo de confianza de 99% de la media del índice de masa corporal de los hombres.

b. Construya un estimado del intervalo de confianza de 99% de la media del índice de masa corporal de las mujeres.

c. Compare e interprete los resultados. Sabemos que los hombres tienen una media de peso mayor que la media de las mujeres, así como que la estatura media de los hombres es mayor que la estatura media de las mujeres, pero ¿los hombres también tienen una media del índice de masa corporal mayor que la media del índice de masa corporal de las mujeres?

6-4 Más allá de lo básico

25. Efecto de un dato distante Pruebe el efecto de un dato distante como sigue: utilice los datos muestrales del ejercicio 17 para calcular un estimado del intervalo de confianza del 95%, de la media poblacional; después, cambie el primer valor de 0.06 gramos/milla a 60 gramos/milla. Dicho valor no es realista, pero un error de este tipo puede ocurrir fácilmente durante un proceso de captura de datos. Compare los dos intervalos de confianza. ¿Los límites del intervalo de confianza son sensibles a los datos distantes? ¿Cómo debe manejar los datos distantes cuando se encuentran en conjuntos de datos muestrales que se usarán para la construcción de intervalos de confianza?

26. Uso de la distribución incorrecta Suponga que se selecciona una muestra aleatoria simple pequeña de una población distribuida normalmente, para la que σ no se conoce. La construcción de un intervalo de confianza debe utilizar la distribución *t*, pero ¿cómo se afecta el intervalo de confianza incorrectamente, si se usa la distribución normal en lugar de la distribución *t*?

27. Efectos de unidades de medida Se construye un intervalo de confianza para una muestra aleatoria simple pequeña de temperaturas (en grados Fahrenheit) seleccionada de una población que se distribuye normalmente, para la cual σ no se conoce.

a. ¿Cómo se afecta el margen de error *E* si cada temperatura se convierte a la escala Celsius? $\left[C = \frac{5}{9}(F - 32)\right]$

b. Si los límites del intervalo de confianza se denotan por *a* y *b*, encuentre expresiones para los límites del intervalo de confianza después de que las temperaturas originales se conviertan a la escala Celsius.

c. Con base en los resultados del inciso *b*, ¿pueden calcularse los límites del intervalo de confianza para las temperaturas Celsius, convirtiendo simplemente los límites del intervalo de confianza de la escala Fahrenheit a la escala Celsius?

28. Intervalo de confianza para muestra de tamaño $n = 1$ Cuando un solo extraterrestre llega a la Tierra, se le mide y se encuentra que tiene una estatura de 3.2 pies. ¿Es razonable esperar que la estatura de todos los extraterrestres de este tipo se distribuya normalmente?

continúa

a. Los métodos de este capítulo requieren información acerca de la variación de una variable. Si sólo está disponible un valor muestral, ¿puede darnos alguna información acerca de la variación de la variable?

b. 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%?

c. Con base en el artículo "An Effective Confidence Interval for the Mean with Samples of Size One and Two" (de Wall, Boen y Tweedie, *The 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 $n = 1$ que se seleccionó aleatoriamente de una población que se distribuye normalmente y se expresa como $x \pm 9.68|x|$. Utilice dicho resultado para construir un intervalo de confianza del 95% con el valor muestral individual de 3.2 pies; expréselo en la forma de $\bar{x} - E < \mu < \bar{x} + E$. Con base en el resultado, ¿parece que algún otro extraterrestre seleccionado aleatoriamente puede medir 50 pies?

6-5 Estimación de la varianza de una población

En esta sección consideramos los mismos tres conceptos que ya se introdujeron en este capítulo: **1.** el estimado puntual, **2.** el intervalo de confianza, y **3.** la determinación del tamaño de muestra que se requiere. Mientras que en las secciones anteriores se aplicaron dichos conceptos a estimaciones de proporciones y medias, en esta sección se aplican a la varianza poblacional σ^2 o a la desviación estándar σ. He aquí los principales objetivos de esta sección:

1. Dados los valores muestrales, estimar la desviación estándar poblacional σ o la varianza poblacional σ^2.

2. Determinar el tamaño de muestra que se requiere para estimar la desviación estándar o la varianza poblacionales.

Muchas situaciones reales, como el control de calidad en procesos de fabricación, piden que estimemos valores de varianza o desviaciones estándar poblacionales. Además, para hacer productos con medidas dentro de una media que se desea, el fabricante debe hacer productos de calidad *consistente* que no se recorran de la gama de extremadamente buenos a extremadamente pobres. Como dicha consistencia puede medirse con frecuencia por medio de la varianza o la desviación estándar, éstas se vuelven estadísticos vitales en el mantenimiento de la calidad de productos y servicios.

Supuestos

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 que se distribuye normalmente se mencionó en secciones anteriores, pero aquí este requisito es más crítico. Por los métodos de la sección, los alejamientos de una distribución normal llegan a producir errores muy graves. En consecuencia, el requisito de tener una distribución normal es mucho más estricto, por lo que hay que revisar la distribución de los datos construyendo histogramas y gráficas cuantilares normales, como se describe en la sección 5-7.

Cuando consideramos estimados de proporciones y medias, utilizamos las distribuciones normal y t de Student. Cuando desarrollamos estimados de varianzas o desviaciones estándar, trabajamos con otra distribución, que se refiere 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 que se distribuye normalmente con varianza σ^2, seleccionamos aleatoriamente muestras independientes de tamaño n y calculamos la varianza muestral s^2 (véase la fórmula 2-5) para cada muestra. El estadístico muestral $\chi^2 = (n - 1)s^2/\sigma^2$ tiene una distribución llamada **distribución chi cuadrada**.

Distribución chi cuadrada

Fórmula 6-7 $$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

donde n = tamaño de la muestra
s^2 = varianza muestral
σ^2 = varianza poblacional

Denotamos chi cuadrada por χ^2, pronunciada "ji cuadrada". (Las ecuaciones matemáticas específicas que se utilizan para definir tal distribución no se darán aquí, ya que están más allá del alcance del libro). 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; en dicho sentido, en este capítulo usamos $n - 1$ grados de libertad.

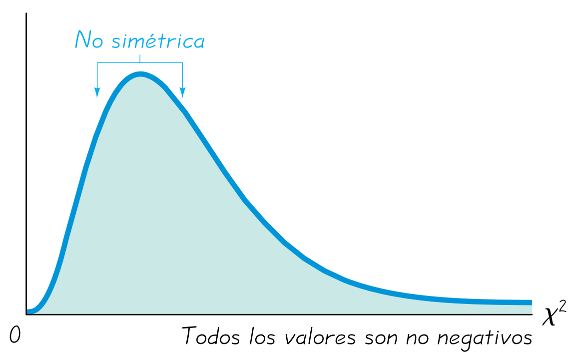
grados de libertad = $n - 1$

En capítulos posteriores encontraremos situaciones en las cuales los grados de libertad no son $n - 1$; por lo tanto, 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 6-8). (Conforme el número de grados de libertad se incrementa, la distribución se vuelve más simétrica, como ilustra la figura 6-9).

FIGURA 6-8 Distribución chi cuadrada



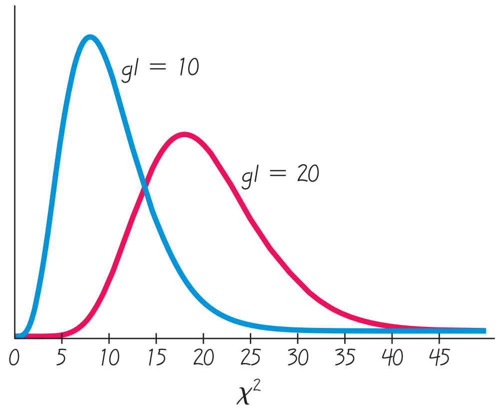


FIGURA 6-9 **Distribución chi cuadrada para gl = 10 y gl = 20**

2. Los valores de chi cuadrada pueden ser cero o positivos, pero no negativos (véase figura 6-8).
3. La distribución chi cuadrada es diferente para cada número de grados de libertad (véase figura 6-9); en esta sección el número de grados de libertad fue dado por gl $= n - 1$. Conforme el número de grados de libertad se incrementa, la distribución chi cuadrada se aproxima a la distribución normal.

Puesto que la distribución chi cuadrada es sesgada en lugar de simétrica, el intervalo de confianza no se ajusta al formato de $s^2 \pm E$, por lo que debemos hacer cálculos separados para los límites de confianza superior e inferior. Hay un procedimiento diferente para calcular valores críticos, que se ilustra en el siguiente ejemplo. Observe la característica esencial siguiente de la tabla A-4:

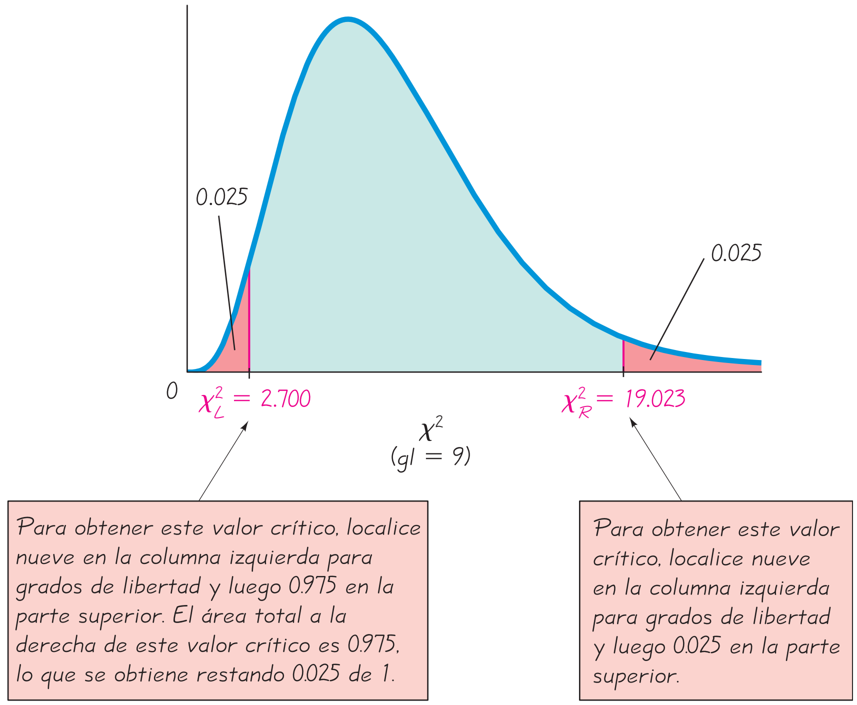
En la tabla A-4, cada valor crítico de χ^2 corresponde a un área que se encuentra en la fila superior de la tabla, en tanto que esa área representa la *región total que se localiza a la derecha* del valor crítico.

La tabla A-2 para la distribución normal estándar proporciona áreas acumulativas de la *izquierda*, mientras que la tabla A-4 para la distribución chi cuadrada provee áreas acumulativas de la *derecha*.

EJEMPLO **Valores críticos** Calcule los valores críticos de χ^2 que determinan las regiones críticas que contienen un área de 0.025 en cada cola. Suponga que el tamaño de muestra relevante es 10, de modo que el número de grados de libertad es $10 - 1$, o 9.

SOLUCIÓN Véase la figura 6-10 y remítase a la tabla A-4. El valor crítico para la derecha ($\chi^2 = 19.023$) se obtiene de manera directa localizando nueve en la columna de grados de libertad de la izquierda y 0.025 a través de la parte superior. El valor crítico de $\chi^2 = 2.700$ a la izquierda otra vez corresponde a nueve en la columna de grados de libertad, pero debemos localizar 0.975 (que se encuentra al restar 0.025 de 1) en la parte superior, puesto que los valores en la fila superior son siempre *áreas a la derecha* del valor crítico. Remítase a la figura 6-10 y véase que el área total a la derecha de $\chi^2 = 2.700$ es 0.975. La figura 6-10 nos indica que, para una muestra de 10 valores que se toman de una población que se distribuye 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.

FIGURA 6-10 Valores críticos de la distribución chi cuadrada



Cuando se obtienen valores críticos de χ^2 de la tabla A-4, note que 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 trabaje con 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 que 100, use la ecuación que se presenta en el ejercicio 22, una tabla con más detalles o un programa de cómputo de estadística.

Estimadores de σ^2

En la sección 5-4 mostramos que las varianzas muestrales s^2 (que se calculan utilizando la fórmula 2-5) tienden a apuntar (o centrarse en) al valor de la varianza poblacional σ^2. Entonces 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 a subestimar σ^2. En lugar 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. Por dichas razones, el valor de s^2 es generalmente el mejor valor individual (o estimado puntual) de los diversos estadísticos posibles que podríamos usar para estimar σ^2.

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 el caso (véase la sección 5-4). Sin embargo, si el tamaño de la muestra es grande, el sesgo es tan pequeño que podemos utilizar s como un estimado de σ razonablemente bueno. Aunque es un estimado sesgado, s 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 hay una indicación de qué tan bueno es realmente. Para compensar tal 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

$$\frac{(n-1)s^2}{\chi_D^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_I^2}$$

Tal expresión se utiliza para calcular un intervalo de confianza para la varianza σ^2, pero un intervalo de confianza (o estimado de intervalo) para la desviación estándar σ se calcula tomando la raíz cuadrada de cada componente, como se muestra 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 lugar de χ_D^2, y $\chi_{1-\alpha/2}^2$ en lugar de χ_I^2).

Notación

Con un área total de α que se dividió 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 el valor crítico de la cola derecha (como se ilustra en la figura 6-11).

Con base en los resultados precedentes, resumimos 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 supuestos que se requieren se satisfagan. (La muestra es aleatoria simple, en tanto que un histograma o una 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 que se desea.

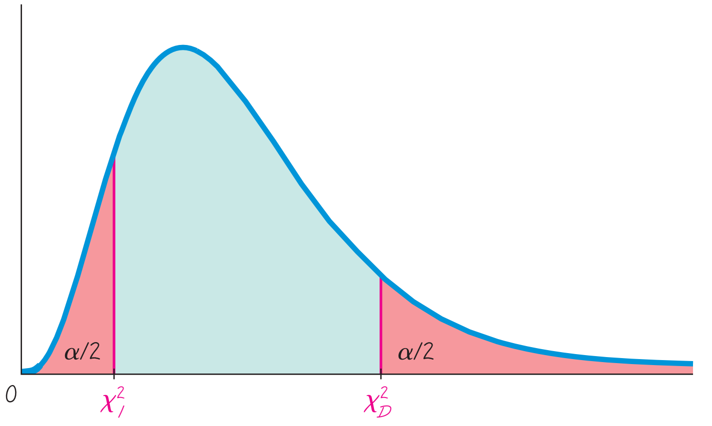


FIGURA 6-11 Distribución chi cuadrada con valores críticos χ^2_I y χ^2_D
Los valores críticos χ^2_I y χ^2_D separan las áreas extremas correspondientes a varianzas muestrales que son improbables (con probabilidad de α).

3. Evalúe los límites del intervalo de confianza superior e inferior utilizando este formato para el intervalo de confianza:

$$\frac{(n-1)s^2}{\chi^2_D} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_I}$$

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, luego cambie σ^2 a σ.
5. Redondee los límites del intervalo de confianza resultantes. Si se utiliza el conjunto de datos original, redondee a un decimal más del que se usa para el conjunto de datos original. Si se utiliza la desviación estándar o varianza muestrales, redondee los límites del intervalo de confianza al mismo número de espacios decimales.

Cuidado: Los intervalos de confianza se llegan a usar de manera informal para comparar conjuntos diferentes de datos, pero el traslape de intervalos de confianza no debe usarse para sacar conclusiones formales ni finales acerca de la igualdad de las varianzas o las desviaciones estándar. Los últimos capítulos 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 sacar conclusiones definitivas acerca de la igualdad de varianzas o desviaciones estándar.

EJEMPLO Temperaturas corporales El conjunto de datos 4 en el Apéndice B lista 106 temperaturas corporales (a las 12:00 h del día 2) que obtuvieron investigadores de la Universidad de Maryland. Utilice las siguientes características del conjunto de datos para construir un estimado del intervalo de confianza del 95% de σ, la desviación estándar de las temperaturas corporales de la población completa:

a. Como revela un histograma de los datos muestrales, la población parece tener una distribución normal.
b. La media muestral es 98.20°F.
c. La desviación estándar muestral es $s = 0.62$°F.

d. El tamaño de la muestra es $n = 106$.

e. No hay datos distantes.

SOLUCIÓN Comenzamos calculando los valores críticos de χ^2. Con una muestra de 106 valores, tenemos 105 grados de libertad. Lo anterior no está muy alejado de los 100 grados de libertad que se encuentran en la tabla A-4, entonces nos iremos con eso. (Véase el ejercicio 22 para un método que proveerá valores críticos más precisos). Para un intervalo de confianza del 95%, dividimos $\alpha = 0.05$ por igual entre las dos colas de la distribución chi cuadrada, luego buscamos los valores de 0.975 y 0.025 en la fila de la parte superior en la tabla A-4. Los valores críticos de χ^2 son $\chi_I^2 = 74.222$ y $\chi_D^2 = 129.561$. Utilizando estos valores críticos, la desviación estándar muestral de $s = 0.62$ y el tamaño de muestra de 106, construimos el intervalo de confianza del 95% evaluando lo siguiente:

$$\frac{(106 - 1)(0.62)^2}{129.561} < \sigma^2 < \frac{(106 - 1)(0.62)^2}{74.222}$$

Esto se convierte en $0.31 < \sigma^2 < 0.54$. El cálculo de la raíz cuadrada de cada parte (antes de redondear) proporciona $0.56°\text{F} < \sigma < 0.74°\text{F}$.

INTERPRETACIÓN Con base en este resultado, obtenemos el 95% de confianza de que los límites de 0.56°F y 0.74°F contienen el valor real de σ. Tenemos una confianza del 95% de que la desviación estándar de las temperaturas corporales de todas las personas saludables está entre 0.56°F y 0.74°F.

El intervalo de confianza $0.56 < \sigma < 0.74$ también se expresaría como (0.56, 0.74), pero el formato de $s \pm E$ *no puede* usarse, puesto que el intervalo de confianza no tiene a s en su centro.

En lugar de aproximar los valores críticos utilizando 100 grados de libertad, utilizamos programas de cómputo o el método descrito en el ejercicio 22; en consecuencia, el intervalo de confianza se vuelve $0.55°\text{F} < \sigma < 0.72°\text{F}$, que es muy cercano al resultado que se obtuvo aquí.

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 6-11. Para una muestra aleatoria simple, hay una probabilidad de $1 - \sigma$ 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 $1 - \sigma$ 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:

$$\frac{(n - 1)s^2}{\chi_D^2} < \sigma^2 \quad \text{y} \quad \frac{(n - 1)s^2}{\chi_I^2} > \sigma^2$$

Meta-análisis

El término *meta-análisis* se refiere a una técnica para realizar un estudio que en esencia combina resultados de otros estudios. Dicha técnica tiene la ventaja de que muestras separadas más pequeñas se pueden combinar en una gran muestra, lo que hace más significativos los resultados colectivos. También tiene la ventaja de manejar trabajo que ya se realizó. El meta-análisis tiene la desventaja de que sólo es tan bueno como los estudios que se utilicen. Si los estudios anteriores tienen defectos, ocurre el fenómeno de "entra basura, sale basura". El empleo del meta-análisis es actualmente popular en investigaciones médicas y psicológicas. Como un ejemplo, un estudio de tratamientos de dolor de cabeza por migraña 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).

Estas últimas dos desigualdades pueden combinarse en una desigualdad:

$$\frac{(n-1)s^2}{\chi_D^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_I^2}$$

Hay una probabilidad de $1-\alpha$ de que tales límites del intervalo de confianza contengan la varianza poblacional σ^2. Recuerde que debemos ser muy cuidadosos cuando interpretamos intervalos de confianza como éstos. Es un error decir que existe una probabilidad de $1-\alpha$ de que σ^2 quedará entre los dos límites del intervalo de confianza. En lugar de ello, habremos de decir que tenemos una confianza de $1-\alpha$ de que los límites contienen a σ^2. También recuerde que los supuestos que se requieren son muy importantes. Si los datos muestrales se reúnen de una forma no muy propia, el intervalo de confianza resultante será incorrecto.

Determinación del tamaño de la muestra

Los procedimientos para calcular el tamaño de muestra necesario para estimar σ^2 son mucho más complejos que los procedimientos que se vieron antes para las medias y las proporciones. En lugar de utilizar procedimientos muy complicados, usaremos la tabla 6-2. El STATDISK también provee tamaños de muestra. Con STATDISK, seleccione **Analysis**, **Sample Size Determination** y luego **Estimate St Dev**. El Minitab, Excel y la calculadora TI-83 Plus no proveen tamaños de muestra de este tipo.

Tabla 6-2

Tamaño de muestra para σ^2		Tamaño de muestra para σ	
Para tener una confianza del 95% de que s^2 está dentro	del valor de σ^2, el tamaño de muestra n debe ser al menos	Para tener una confianza del 95% de que s está dentro	del valor de σ, el tamaño de muestra 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 valor de σ^2, el tamaño de muestra n debe ser al menos	Para tener una confianza del 99% de que s está dentro	del valor de σ, el tamaño de muestra 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

EJEMPLO Queremos estimar σ, la desviación estándar de todas las temperaturas corporales, con una confianza del 95% de que nuestro estimado cae dentro del 10% del valor real de σ. ¿Qué tan grande debe ser la muestra? Suponga que la población se distribuye normalmente.

SOLUCIÓN En la tabla 6-2, vemos que un 95% de confianza y un error de 10% para σ corresponde a una muestra de tamaño 191. Debemos seleccionar aleatoriamente 191 valores de la población de temperaturas corporales.

Utilizando 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 cuantilar normal. Después, seleccione **Analysis** del menú principal, luego **Confidence Intervals** y **Population St-Dev**. Proceda a ingresar los datos que se requieren.

Minitab Primero ingrese los datos en la columna C1, entonces seleccione **Editor**, seguido por **Enable Command Language**, y aplique el comando **%DESCRIBE C1** para obtener una salida que incluye los intervalos de confianza de 95% para μ y σ. El nivel de confianza predeterminado de 95% puede cambiarse.

Excel Excel no provee intervalos de confianza para σ ni para σ^2.

TI-83 Plus La calculadora TI-83 Plus no proporciona intervalos de confianza para σ ni para σ^2.

6-5 Destrezas y conceptos básicos

Cálculo de valores críticos. *En los ejercicios 1 a 4, encuentre los valores críticos χ^2_I y χ^2_D correspondientes al nivel de confianza y tamaño de muestra dados.*

1. 95%; $n = 16$

2. 95%; $n = 51$

3. 99%; $n = 80$

4. 90%; $n = 40$

Cálculo de intervalos de confianza. *En los ejercicios 5 a 8, utilice el nivel de confianza y los datos muestrales dados para calcular el intervalo de confianza para la desviación estándar poblacional σ. En cada caso, suponga que se selecciona una muestra aleatoria simple de una población que tiene una distribución normal.*

5. Salarios de profesores de estadística: 95% de confianza; $n = 20$, $\bar{x} = \$95,000$, $s = \$12,345$

6. Edades de conductores que ocupan el carril para rebasar mientras conducen a 25 millas/hr, con la luz intermitente direccional izquierda funcionando: 99% de confianza; $n = 27$, $\bar{x} = 80.5$ años, $s = 4.6$ años

7. Tiempos entre la operación de un control remoto de televisión por hombres durante comerciales: 90% de confianza; $n = 30$, $\bar{x} = 5.24$ seg, $s = 2.50$ seg.

8. Salarios iniciales de graduados universitarios que tomaron un curso de estadística: 95% de confianza; $n = 51$, $\bar{x} = \$45{,}678$, $s = \$9900$

Determinación del tamaño de muestra. *En los ejercicios 9 a 12, suponga que cada muestra es una muestra aleatoria simple obtenida de una población que se distribuye normalmente.*

9. Calcule el tamaño de muestra mínimo que se necesita para lograr una confianza del 95% de que la desviación estándar muestral s está dentro del 10% de σ.

10. Calcule el tamaño de muestra mínimo que se necesita para lograr una confianza del 95% de que la desviación estándar muestral s está dentro del 30% de σ.

11. Calcule el tamaño de muestra mínimo que se necesita para lograr una confianza del 99% de que la varianza muestral está dentro del 1% de la varianza poblacional. ¿Resulta práctico un tamaño de muestra como éste para la mayoría de los casos?

12. Calcule el tamaño de muestra mínimo que se necesita para lograr una confianza del 95% de que la varianza muestral está dentro del 20% de la varianza poblacional.

Cálculo de intervalos de confianza. *En los ejercicios 13 a 20, suponga que cada muestra es una muestra aleatoria simple que se obtuvo de una población con una distribución normal.*

13. Destrucción de Vipers de Dodge Con la prueba destructiva, los elementos de la muestra se destruyen en el proceso de probarlos. La prueba de automóviles por choque es un ejemplo de prueba destructiva muy costosa. Los 12 automóviles deportivos Viper de Dodge (precio de lista: $59,300) se prueban por choque en una variedad de condiciones que simulan colisiones típicas. El análisis de los 12 automóviles que se dañaron resulta en costos de reparación con una distribución que parece tener forma de campana, con una media de $\bar{x} = \$26{,}227$ y una desviación estándar de $s = \$15{,}873$ (según datos del Highway Loss Data Institute). Calcule un estimado de intervalo del 95% de σ, la desviación estándar de los costos de reparación para todos los Viper de Dodge que se ocuparon en colisiones e interprete el resultado.

14. Anticongelante automotriz Se supone que un recipiente de anticongelante automotriz contiene 3785 ml de líquido. Dándose cuenta de que las fluctuaciones son inevitables, un gerente de control de calidad quiere estar completamente seguro de que la desviación estándar es menor que 30 ml. De lo contrario, algunos recipientes podrían sobrellenarse mientras que otros no tendrían suficiente refrigerante. Él selecciona una muestra aleatoria simple, con los resultados que se dan aquí. Utilice estos resultados muestrales para construir un intervalo de confianza del 99% para el valor real de σ. ¿Sugiere este intervalo de confianza que las fluctuaciones están en un nivel aceptable?

3761	3861	3769	3772	3675	3861	$n = 18$
3888	3819	3788	3800	3720	3748	$\bar{x} = 3787.0$
3753	3821	3811	3740	3740	3839	$s = 55.4$

15. Control de plomo en el aire En la lista de abajo se incluyen cantidades de plomo que se midieron en el aire (en microgramos por metro cúbico o $\mu g/m^3$). La Environmental Protection Agency estableció un estándar de plomo para la calidad del aire: 1.5 $\mu g/m^3$. 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 acerca de 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

16. Control de calidad de donas La panadería Hudson Valley hace donas que se empacan en cajas con etiquetas que dicen contener 12 donas y pesan un total de 42 oz. Si la variación entre las donas es muy grande, algunas cajas contendrán menos peso (estafando a los consumidores) y otras más (disminuyendo las ganancias). Un consumidor no estaría contento con una dona muy pequeña que pueda verse sólo con microscopio electrónico y tampoco con una dona tan grande que parezca una llanta de tractor. El supervisor de control de calidad encontró que es posible resolver el problema si las donas tienen una media de 3.50 onzas y una desviación estándar de 0.06 onzas o menor. Se seleccionan al azar 12 donas de la línea de producción y se pesan, con los resultados que se dan aquí (en onzas). Construya un intervalo de confianza del 95% para σ y luego determine si el supervisor de control de calidad está en problemas.

3.43 3.37 3.58 3.50 3.68 3.61 3.42 3.52 3.66 3.50 3.36 3.42

17. Ritmos cardiacos al trabajar con pala Ya que las muertes por deficiencias cardiacas parecen incrementarse después de nevadas abundantes, se diseñó un experimento para comparar las demandas cardiacas al remover la nieve con una pala contra las que se produjeron por el uso de un aparato eléctrico para retirar nieve. Diez sujetos despejaron de nieve el terreno con ambos métodos y sus frecuencias cardiacas máximas (en latidos por minuto) se registraron durante las dos actividades. Se obtuvieron los resultados siguientes (datos que se tomaron de "Cardiac Demands of Heavy Snow Shoveling", de Franklin *et al.*, *Journal of the American Medical Association*, vol. 273, núm. 11):

Frecuencias cardiacas máximas de paleo de nieve manual: $n = 10, \bar{x} = 175, s = 15$

Frecuencias cardiacas máximas con aparato eléctrico para retirar nieve: $n = 10, \bar{x} = 124, s = 18$

a. Construya un estimado del intervalo de confianza del 95%, de la desviación estándar poblacional σ para aquellas personas que palean nieve de manera manual.
b. Construya un estimado del intervalo de confianza del 95%, de la desviación estándar poblacional σ de aquellas personas que usan el aparato eléctrico para retirar nieve.
c. Compare e interprete los resultados. ¿Parece que la variación es diferente para los dos grupos?

18. Pulsos Un investigador médico quiere determinar si el pulso de los hombres varía más o menos que el pulso de las mujeres. Utilizando los pulsos muestrales del conjunto de datos 1 del Apéndice B, el pulso de los hombres se resume con los estadísticos $n = 40$, $\bar{x} = 69.4$, $s = 11.3$. Para las mujeres, los estadísticos son $n = 40$, $\bar{x} = 76.3$, $s = 12.5$.
a. Construya un estimado del intervalo de confianza del 95% de la desviación estándar poblacional σ de los pulsos de los hombres.
b. Construya un estimado del intervalo de confianza del 95% de la desviación estándar poblacional σ de los pulsos de las mujeres.
c. Compare los resultados anteriores. ¿Parece que las desviaciones estándar poblacionales para hombres y mujeres son diferentes? ¿Por qué?

19. a. Comparación de filas de espera Los valores que se listan son tiempos de espera (en minutos) de clientes del banco Jefferson Valley, donde los clientes se forman en una sola fila de espera para tres ventanillas de cajero. 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 que se listan son tiempos de espera (en minutos) de clientes del Bank of Providence, donde los clientes pueden formarse en cualquiera de tres filas diferentes que se alinean a tres ventanillas de cajero. 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

continúa

c. Interprete los resultados que se encontraron en los incisos *a* y *b*. ¿Sugieren los intervalos de confianza una diferencia en la variación entre los tiempos de espera? ¿Cuál acomodo parece mejor: el sistema de una sola fila o el sistema de filas múltiples?

20. Índice de masa corporal 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.

6-5 Más allá de lo básico

21. Calcular datos faltantes El artículo de una revista incluye una gráfica que exhibe que los datos muestrales se distribuyen normalmente.

a. El nivel de confianza se omite inadvertidamente cuando se declara este intervalo de confianza: $2.8 < \alpha < 6.0$. Calcule el nivel de confianza para los estadísticos muestrales dados: $n = 20$, $\bar{x} = 45.2$ y $s = 3.8$.

b. Se declara este intervalo de confianza del 95%: $19.1 < \sigma < 45.8$. Dado $n = 12$, encuentre el valor de la desviación estándar s, que se omitió en el artículo.

22. Calcular valores críticos En la construcción de intervalos de confianza para σ y σ^2, utilizamos la tabla A-4 para encontrar los valores críticos χ_I^2 y χ_D^2, pero la tabla sólo se aplica a casos en los que $n \leq 101$, por lo cual el número de grados de libertad es 100 o menor. Para números de grados de libertad más grandes, es posible aproximar χ_I^2 y χ_D^2 utilizando

$$\chi^2 = \frac{1}{2}\left[\pm z_{\alpha/2} + \sqrt{2k - 1}\right]^2$$

donde k es el número de grados de libertad y $z_{\alpha/2}$ es la puntuación crítica z que se describió al principio de la sección 6-2. Construya el intervalo de confianza del 95% para σ utilizando los siguientes datos muestrales: las estaturas medidas de 772 hombres de 18 a 24 años de edad tienen una desviación estándar de 2.8 pulgadas (datos que se tomaron de la National Health Survey).

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, además de desarrollar procedimientos para calcular cada uno de los siguientes puntos:

- Estimado puntual
- Intervalo de confianza
- Tamaño de muestra requerido

Analizamos el estimado puntual (o estimado de un solo valor) y sacamos 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 se utilizan, por lo general, intervalos de confianza (o estimados de intervalo) como estimados más reveladores y útiles. También consideramos formas para determinar los tamaños de muestra necesarios para estimar parámetros dentro de márgenes de error dados. Este capítulo también introdujo las distribuciones t de Student y chi cuadrada. Debemos tener cuidado de utilizar la distribución de probabilidad correcta para cada conjunto de circunstancias. Este capítulo utilizó los siguientes criterios para seleccionar la distribución apropiada:

Intervalo de confianza para la proporción p:	Utilice la distribución *normal* (considerando que los supuestos que se requieren se satisfacen y que $np \geq 5$ y $nq \geq 5$ para que se use la distribución normal como aproximación de la distribución binomial).
Intervalo de confianza para μ:	Véase la figura 6-6 o la tabla 6-1 para elegir entre las distribuciones *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 que se requieren se satisfacen).

Para aplicar los procedimientos del intervalo de confianza y el tamaño de muestra de este capítulo, es muy importante verificar que los supuestos que se requieren se satisfagan. Si no, no será posible utilizar los métodos de este capítulo y tal vez necesitemos otros métodos, como el *bootstrap*, que se describe en el proyecto tecnológico que viene al final de este capítulo, o métodos no paramétricos, como los que se analizan en el capítulo 12.

Ejercicios de repaso

1. Estimación de asistencia a parques temáticos Cada año se gastan millones de dólares en parques temáticos propiedad de Disney, Universal Studios, Sea World, Busch Gardens y otros. Una encuesta de 1233 personas que viajaron reveló que 111 de ellos incluyeron una visita a un parque temático (datos de la Travel Industry Association of America).

a. Calcule el estimado puntual del *porcentaje* de todas las personas que visitaron un parque temático cuando hicieron un viaje.

b. Calcule un estimado del intervalo de confianza del 95% del *porcentaje* de todas las personas que visitaron un parque temático cuando hicieron un viaje.

c. La encuesta se realizó con personas que hicieron viajes, pero no se proporcionó 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, ¿cuántas personas debe entrevistar si quiere lograr una confianza del 99% de que su porcentaje muestral está dentro de 2.5 puntos porcentuales del porcentaje correcto de la población?

2. Estimación de tiempo de propiedad de automóviles Un distribuidor de partes automotrices de la NAPA quiere información acerca de cuánto tiempo planean conservar sus vehículos los propietarios de automóviles. A este respecto, una muestra aleatoria simple de 25 propietarios de automóviles resulta en $\bar{x} = 7.01$ años y $s = 3.74$ años (de acuerdo con datos de una encuesta de Roper). Suponga que la muestra se obtuvo de una población que se distribuye normalmente.

continúa

a. Calcule un estimado del intervalo de confianza del 95% de la media poblacional.
b. Calcule un estimado del intervalo de confianza del 95% de la desviación estándar poblacional.
c. Si pasaron varios años y usted quiere realizar una nueva encuesta para estimar la cantidad media del tiempo que planean conservar sus autos los propietarios de automóviles, ¿cuántos propietarios de automóviles que se seleccionarían al azar debe encuestar? Suponga que se quiere una confianza del 99% de que la media muestral esté dentro de 0.25 años (o tres meses) de la media poblacional y que $\sigma = 3.74$ años (con base en el último resultado).
d. Cuando se realiza la encuesta descrita en el inciso *c*, descubre que el proceso de encuesta puede simplificarse con un costo sustancialmente reducido si utiliza una base de datos disponibles que consiste en personas que compraron un automóvil de la General Motors durante los 10 últimos años. ¿Se obtendrían buenos resultados de esta población?

3. Estimaciones de encuestas a votantes En una elección presidencial reciente, se encuestó a 611 votantes, de los cuales 308 dijeron que votaron por el candidato que ganó (según datos del ICR Survey Research Group).
a. Calcule el estimado puntual del *porcentaje* de votantes que, según ellos, votaron por el candidato que ganó.
b. Calcule un estimado del intervalo de confianza del 98% del *porcentaje* de votantes que, según ellos, votaron por el candidato que ganó.
c. De los que votaron, el 43% realmente votó por el candidato que ganó. ¿Es consistente este resultado con los resultados de la encuesta? ¿Cómo se explicaría una discrepancia?

4. Estimaciones de facilidad de lectura Remítase al conjunto de datos 14 en el Apéndice B para las calificaciones de nivel de Flesch-Kincaid para 12 páginas seleccionadas aleatoriamente que se tomaron de los libros de Tom Clancey, J. K. Rowling y León Tolstoi.
a. Construya un estimado del intervalo de confianza del 95% de la media de la calificación de nivel de Flesch-Kincaid para la población de todas las páginas de *El oso y el dragón*, de Tom Clancey.
b. Construya un estimado del intervalo de confianza del 95% de la media de la calificación de nivel de Flesch-Kincaid para la población de todas las páginas de *Harry Potter y la piedra filosofal*, de J. K. Rowling.
c. Construya un estimado del intervalo de confianza del 95% de la media de la calificación de nivel de Flesch-Kincaid para la población de todas las páginas de *La guerra y la paz*, de León Tolstoi.
d. Compare los intervalos de confianza anteriores. ¿Qué concluye acerca de las calificaciones de nivel?

5. Estimar facilidad de lectura El conjunto de datos 14 incluye calificaciones de nivel de Flesch-Kincaid para obras de Tom Clancey, J. K. Rowling y León Tolstoi. Si usted quiere estimar la media de la calificación de nivel de Flesch-Kincaid para las páginas de *El señor de los anillos*, de J. R. R. Tolkien, ¿cuántas páginas debe seleccionar aleatoriamente si quiere tener una confianza del 90% de que la media muestral está dentro de 0.5 de la media poblacional? Puesto que las muestras de páginas de Clancey, Rowling y Tolstoi, en el conjunto de datos 14 obtienen calificaciones de nivel de Flesch-Kincaid con desviaciones estándar de 2.45, 1.17, y 2.01, suponga que $\sigma = 2.45$ para *El señor de los anillos*.

6. Estimación de variación El conjunto de datos 14 en el Apéndice B incluye las calificaciones de nivel de Flesch-Kincaid para 12 páginas que se seleccionaron al azar de *Harry Potter y la piedra filosofal*, de J. K. Rowling. Las 12 calificaciones tienen una desviación estándar de 1.17, y parecen provenir de una población que se distribuye normalmente. Construya un estimado del intervalo de confianza del 95% de la desviación estándar σ de las calificaciones de nivel de Flesch-Kincaid para todas las páginas de *Harry Potter y la piedra filosofal*.

7. Determinación del tamaño de muestra Quiere estimar el porcentaje de estudiantes de estadística de Estados Unidos que obtienen calificaciones de B o mayores. ¿Cuántos estudiantes debe encuestar si busca obtener un nivel de confianza del 97% de que el porcentaje muestral se desvió por no más de dos puntos porcentuales?

8. Política de servicio de alcohol: determinación del tamaño de muestra En una encuesta de Gallup de 1004 adultos, el 93% indicaron que los restaurantes y los bares deberían negar el servicio a los clientes que bebieron mucho. Si usted planea realizar una nueva encuesta para confirmar que el porcentaje continúa siendo correcto, ¿cuántos adultos seleccionados al azar debe encuestar si desea obtener un nivel de confianza del 98% de que el margen de error es de cuatro puntos porcentuales?

Ejercicios de repaso acumulativo

1. Análisis de pesos de supermodelos Algunas veces las supermodelos son criticadas porque sus bajos pesos fomentan hábitos alimenticios no saludables entre las mujeres jóvenes. Abajo se listan los pesos (en libras) de nueve supermodelos que se seleccionaron al azar.

125 (Taylor)	119 (Auermann)	128 (Schiffer)	128 (MacPherson)
119 (Turlington)	127 (Hall)	105 (Moss)	123 (Mazza)
115 (Hume)			

Resuelva para cada uno de los incisos siguientes:

a. Media **b.** Mediana
c. Moda **d.** Mitad 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, intervalo, 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 de muestra necesario para estimar la media del peso de todas las modelos, con una confianza del 99% de que la media muestral sea errónea por no más de 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 se distribuyen normalmente con una media de 143 lb y una desviación estándar de 29 lb (según datos de la National Health and Examination Survey). Con base en los valores muestrales dados, ¿parece que los pesos de las supermodelos son sustancialmente menores que los pesos de mujeres que se seleccionaron al azar? Explique.

2. Trastorno recesivo del cromosoma X Un experto en genética determinó que, para ciertas parejas, hay un 0.25 de probabilidad de que cualquier hijo presente un trastorno recesivo del cromosoma X.

a. Calcule la probabilidad de que entre 200 de estos hijos, al menos 65 presenten el trastorno recesivo del cromosoma X.
b. Un estudio subsiguiente 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 estos hijos que presentan el trastorno.
c. Con base en los incisos a y b, ¿parece ser correcta la determinación del experto de un 0.25 de probabilidad? Explique.

3. Análisis de resultados de encuesta En una encuesta de Gallup, a los sujetos adultos encuestados se les preguntó: "¿Tiene usted una pistola en su casa?". De las personas que respondieron, 413 dijeron que "sí", y 646 que "no" o que no tenían opinión.

a. ¿Que porcentaje de los que respondieron contestaron "sí"?

b. Construya un estimado del intervalo de confianza del 95%, del porcentaje de todos los adultos que respondieron "sí" cuando se les preguntó si tenían una pistola en su casa.

c. ¿Podemos concluir con seguridad que menos del 50% de los adultos respondieron "sí" cuando se les preguntó si tenían una pistola en su casa? ¿Por qué?

d. ¿Cuál sería una respuesta sensible a la crítica de que la encuesta de Gallup no puede ofrecer buenos resultados puesto que el tamaño de la muestra es sólo de 1059 adultos, que se seleccionaron de una enorme población con más de 200 millones de adultos?

Actividades cooperativas 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 puede levantar una ceja sin levantar la otra. [Dichos resultados muestrales son fáciles de obtener ya que los sujetos que se encuestaron tienden a levantar una ceja (si pueden) cuando los aborda alguien haciendo preguntas].
- La media de la edad de automóviles que conducen estudiantes de estadística y/o la media de automóviles que conducen universitarios.
- La media de la edad de los libros de matemáticas y la media de la edad de los libros de ciencia en la biblioteca de su universidad (con base en las fechas de los derechos de autor).
- La media de la longitud de las palabras en los editoriales del *New York Times* y la media de la longitud de las palabras de los editoriales de su periódico local.
- La media del tamaño de las palabras en la revista *Time*, la revista *Newsweek* y la revista *People*.
- La proporción de estudiantes de su universidad capaces de identificar correctamente al presidente, al vicepresidente y al secretario de Estado de Estados Unidos.
- La proporción de estudiantes de su universidad que son mayores de 18 años de edad y se registraron en el padrón electoral.
- La media de la edad de los estudiantes de tiempo completo en su universidad.
- La proporción de vehículos con motor en su región que son automóviles.

2. Actividad en clase Forme grupos de tres o cuatro estudiantes. Pídales que examinen una revista actual como *Time* o *Newsweek*, y que calculen la proporción de páginas que incluyen anuncios comerciales. Con base en los resultados, deben construir un estimado del intervalo de confianza de 95%, del porcentaje de todas las páginas que contienen anuncios comerciales. Comparen los resultados con otros grupos.

3. Actividad en clase Forme grupos de dos estudiantes. Primero pídales que calculen el tamaño de muestra que se requiere para estimar la proporción de veces que una moneda cae en cara cuando se lanza, suponiendo que usted busca un nivel de confianza del 80% de que la proporción muestral está dentro de 0.08 de la proporción poblacional real. Luego, dígales que lancen una moneda el número requerido de veces y que registren sus resultados. ¿Qué porcentaje del intervalo de confianza podría contener realmente el valor verdadero de la proporción de la población, que sabemos que es $p = 0.5$? Verifique este último resultado comparando su intervalo de confianza con los intervalos de confianza que se encontraron en otros grupos.

Proyecto tecnológico

Muestreo repetido *bootstrap* Se puede utilizar el método *bootstrap* 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 aleatoriamente de una población con una distribución que se aleja mucho de la normal; por lo tanto, no se puede utilizar ningún método que requiera una distribución normal.

2.9 564.2 1.4 4.7 67.6 4.8 51.3 3.6 18.0 3.6

Al querer manejar los datos muestrales de arriba para la construcción del estimado de un intervalo de confianza de la media poblacional μ, notamos que la muestra es pequeña y que hay un dato distante. 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 sustraer de la muestra con reemplazo, creando así una aproximación de la población original. De esta forma, estiramos la muestra "con sus propios *bootstraps*" para simular la población original. Utilizando los datos muestrales que se dieron antes, construya un estimado del intervalo de confianza del 95% de la media poblacional μ, con el método *bootstrap* como se describe en los siguientes pasos de Minitab.

a. Desarrolle 500 muestras nuevas, cada una de tamaño 10, seleccionando 10 valores con reemplazo de los 10 valores muestrales que se dieron antes. Con Minitab, primero ingrese los valores muestrales en la columna C1, luego ingrese las probabilidades de 0.1, 0.1, . . . , 0.1 (10 veces) en la columna C2. Ahora seleccione **Calc** de la barra del menú principal, luego **Random Data**, seguida por **Discrete**. Proceda a generar 500 filas de datos, para guardarse en las columnas C11-C20, con los valores en C1 y las probabilidades en C2; para concluir, haga clic en OK.

b. Encuentre las medias de las 500 muestras *bootstrap* que se generaron en el inciso *a*. Seleccione **Calc**, **Row Statistics** y **Mean**, ingrese las variables de entrada de C11-C20 con los resultados a guardarse en C21 y haga clic en OK.

c. Ordene las 500 medias. Seleccione **Manip** de la barra del menú principal, escoja la opción de **Sort** y proceda a ordenar la columna C21. Guarde la columna ordenada en C21 y ordene la columna C21. Haga clic en OK.

d. Encuentre los percentiles $P_{2.5}$ y $P_{97.5}$ para las medias ordenadas que resultaron del paso anterior. ($P_{2.5}$ es la media de los valores 12o y 13o de la lista que se clasificó en la columna C21; $P_{97.5}$ es la media de los valores 487o y 488o en la columna C21). 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 encontrar un intervalo de confianza del 95% para la desviación estándar poblacional σ. [Utilice los mismos pasos que se listan arriba, pero especifique *desviación estándar* (Standard Deviation) en lugar de media (Mean) en el inciso *b*)]. Compare su resultado con el intervalo $318.4 < \sigma < 1079.6$, que se obtuvo utilizando incorrectamente los métodos descritos en la sección 6-5. (El uso de los métodos de la sección 6-5 es incorrecto porque estos métodos requieren que los valores muestrales provengan de una población que se distribuya normalmente, pero la población no tiene una distribución normal). Este intervalo de confianza incorrecto para σ no contiene el valor real de σ, que es 232.1. ¿Ofrece el procedimiento *bootstrap* un intervalo de confianza para σ que contiene a 232.1, verificando que este método es efectivo?

Un método alternativo al uso de Minitab es utilizar programas de cómputo que se diseñen específicamente para métodos de muestreo repetido *bootstrap*. El autor recomienda Resampling Stats, disponible en Resampling Stats, Inc., 612 N. Jackson St., Arlington, VA, 22201. Teléfono: (703) 522-2713.

de los DATOS a la DECISIÓN

Pensamiento crítico: rediseño del teclado estándar

La mayoría de los teclados que se utilizan normalmente tienen las teclas configuradas en un patrón estándar del arreglo Qwerty, que se llama así por la posición de las letras QWERTY en la fila superior. Con una fecha de desarrollo de 1872, la configuración QWERTY se supuso que obligaba a los mecanógrafos a escribir más lentamente para que sus máquinas de escribir se trabaran con menos frecuencia. El teclado Dvorak se desarrolló en 1936 como una configuración más eficiente con teclas que se acomodaron de acuerdo con su frecuencia de uso. Un artículo en la revista *Discover* sugiere que usted puede medir la facilidad de escritura utilizando este sistema de calificación por puntos: asigne a cada letra en la fila superior de letras el valor de 1, a cada letra en la fila media o "de casa" el de 0 y a cada letra de la fila inferior el de 2. (Véase "Typecasting", de Scott Kim, *Discover*). Aplicando este sistema de calificación a cada una de las 52 palabras del Preámbulo a la Constitución de Estados Unidos, obtenemos estos estadísticos para cada una de las configuraciones de teclado:

Configuración QWERTY: $n = 52$, $\bar{x} = 4.4$, $s = 2.8$
Configuración Dvorak: $n = 52$, $\bar{x} = 1.7$, $s = 1.8$

a. Utilice los datos muestrales que se dieron con los métodos de este capítulo para demostrar que la configuración Dvorak tiene calificaciones significativamente más bajas, indicando que la configuración de teclado Dvorak es más fácil de usar.

b. ¿Hay algún aspecto del sistema de calificación o de la elección de palabras para la muestra que pueda afectar la conclusión acerca de cuál configuración de teclado es más fácil de usar?

c. Escriba un breve reporte de lo que encontró y sus conclusiones.

d. Si la configuración de teclado Dvorak es realmente más fácil de usar, ¿por qué no se adopta por casi todos los que usan ahora un teclado? ¿Cómo se pueden vencer los obstáculos para adoptar la configuración Dvorak con la finalidad de que los que usamos teclados nos volvamos más eficientes?

PROYECTO DE INTERNET

Intervalos de confianza

Los intervalos de confianza en este capítulo ilustran un punto importante de la ciencia de la estimación estadística. A saber, las estimaciones que se basan en datos muestrales se hacen con ciertos grados de confianza. En el proyecto de Internet para este capítulo, usted utilizará intervalos de confianza para hacer una afirmación acerca de la temperatura de donde usted vive. Vaya al sitio de Internet de este libro de texto:

http://www.pearsoneducacion.net/triola

Localice el proyecto para este capítulo. Ahí encontrará las instrucciones sobre cómo utilizar Internet para localizar datos de temperatura que recolectó la estación meteorológica más cercana a su casa. Con estos datos a la mano, construirá intervalos de confianza para las temperaturas durante diferentes periodos e intentará concebir 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 sirve en muchos comités ambientales nacionales e internacionales que tratan con especies en peligro de extinción, contaminantes en la fauna, los efectos de químicos en el comportamiento animal y los efectos de la gente en los ecosistemas.

¿Qué conceptos de la estadística utiliza en su trabajo?

Utilizo una variedad de métodos estadísticos que incluyen métodos paramétricos y no paramétricos. Sin un firme entendimiento de la estadística, no sería capaz de probar si los factores ambientales afectan a los eventos reproductivos. Utilizo la estadística para probar hipótesis que genero observando animales dentro de sus medios naturales. Mientras que la observación nos conduce a establecer 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 usó la estadística en el pasado?

La estadística es muy útil en la identificación de factores que influyen en el comportamiento animal. Los pájaros anidan en hábitat particulares, pero nos preguntamos si anidan aleatoriamente o seleccionan sitios específicos para sus nidos. Esto es importante puesto que la conservación requiere conocer las necesidades de los animales para crear, proteger y/o manejar ese hábitat. Probé la hipótesis de que las golondrinas marinas comunes seleccionaban islas de pantanos salados particulares. Comparando estadísticamente un amplio rango de factores ambientales (como son la altura de la isla, el tamaño de la isla, y el tipo y densidad de la vegetación) en todas las islas con el mismo conjunto de factores en las que anidan las golondrinas marinas, demostramos que estas aves en realidad seleccionan un conjunto de características muy específicas. Aunque hay más de 250 islas en la bahía donde este estudio se realizó, sólo 36 reúnen el criterio que usan las golondrinas. Las aves seleccionan islas que son suficientemente altas para evitar las mareas en las tormentas de verano, pero suficientemente bajas para que los depredadores no puedan sobrevivir durante el invierno. Las islas que son bastante altas como para evitar las mareas de las tormentas de invierno a menudo tienen poblaciones viables de depredadores, tales como los zorros y los mapaches, que se comerán los huevos y los pollos 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 el de la pesca y el consumo de las personas, el saqueo de las aves costeras y la construcción de nidos de las aves marinas.

En términos de estadística, ¿qué recomendaría a los aspirantes de empleo en su campo?

Cualquiera que desee estar en el campo de la biología de conservación, la ecotoxicología, el comportamiento animal o ecológico necesita un amplio rango de habilidades estadísticas. Dos o tres cursos estarían mejor, incluyendo estadística *general* de regresión y estadística no paramétrica. La naturaleza de cada problema y las características de los datos determinarán la estadística que se requiere; uno no debería limitarse por una carencia de conocimiento de la estadística.

7

Prueba de hipótesis

7-1 Panorama general

7-2 Fundamentos de la prueba de hipótesis

7-3 Prueba de una aseveración respecto de una proporción

7-4 Prueba de una aseveración respecto de una media: σ conocida

7-5 Prueba de una aseveración respecto de una media: σ desconocida

7-6 Prueba de una aseveración respecto de una desviación estándar o de una varianza

PROBLEMA DEL CAPÍTULO

¿Nos pasamos la luz roja la mayoría de nosotros?

En el capítulo 6 utilizamos los resultados de encuestas para estimar la proporción de habitantes de Minnesota que se oponen al sistema de "cámara vigilante", que implica el uso de cámaras para multar a conductores que se pasan la luz roja de los semáforos. Los datos muestrales consistieron en 829 adultos de Minnesota, seleccionados al azar; el 51% de ellos se opusieron a una ley que implementara el sistema de cámara vigilante en su estado. Aun cuando el 51% de los 829 sujetos se opusieron a la ley de la cámara vigilante, el periódico *Star Tribune* publicó el encabezado "La opinión de los encuestados respecto a la propuesta de la 'cámara vigilante' está dividida". El encabezado del periódico afirmaba que los encuestados estaban *divididos*, pero el 51% se opuso, entonces ¿por qué no podemos decir que la *mayoría* de los ciudadanos de Minnesota se oponen?

En otro estudio realizado en Estados Unidos a nivel nacional se encuestó a 880 conductores seleccionados al azar, y el 56% admitió pasarse la luz roja del semáforo. En un artículo distribuido por la Associated Press, la reportera Sonja Barisic escribió lo siguiente: "Una encuesta reveló que casi todos los conductores estadounidenses coinciden en que pasarse la luz roja es peligroso, pero más de la mitad de ellos admiten haberlo hecho, principalmente porque tenían prisa". Esta aseveración incluye la afirmación de que la *mayoría* (más del 50%) de los estadounidenses se pasan la luz roja. ¿Apoyan en realidad los resultados de la encuesta dicha aseveración?

En este capítulo presentamos métodos estándar para probar aseveraciones como las dos siguientes, basadas en la información anterior:

- ¿Existe suficiente evidencia muestral que apoye la aseveración de que una proporción mayor al 0.5 de los adultos de Minnesota se oponen a la ley de la cámara vigilante? Es decir, ¿será suficiente evidencia una muestra de $n = 829$ adultos de Minnesota, seleccionados al azar, donde el 51% (o $\hat{p} = 0.51$) se opone a la ley de la cámara vigilante, para sustentar la aseveración de que $p > 0.5$?
- ¿Existe suficiente evidencia para sustentar la aseveración de que una proporción mayor al 0.5 de los adultos estadounidenses admiten haberse pasado la luz roja? Es decir, ¿será suficiente evidencia una muestra de $n = 880$ conductores adultos estadounidenses, seleccionados al azar, donde el 56% (o $\hat{p} = 0.56$) admite haberse pasado la luz roja, para sustentar la aseveración de que $p > 0.5$?

Existe un procedimiento estándar para probar este tipo de aseveraciones, y en este capítulo se describe dicho procedimiento.

7-1 Panorama general

Este capítulo describe el procedimiento estadístico para probar hipótesis, que es el procedimiento estándar usado comúnmente por los profesionales en una gran variedad de disciplinas. Las publicaciones científicas, tales como el *Journal of the American Medical Association, American Journal of Psychiatry* e *International Journal of Advertising*, por rutina, incluyen los mismos procedimientos básicos presentados en este capítulo. Como consecuencia, el trabajo realizado al estudiar los métodos de este capítulo encuentra aplicación en todas las disciplinas y no sólo en la estadística.

Dos actividades importantes de la estadística inferencial son la estimación de los parámetros de población (introducidos en el capítulo 6) y la prueba de hipótesis (introducida en este capítulo). Una prueba de hipótesis es un procedimiento estándar para probar alguna *aseveración*.

Definiciones

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.

Las siguientes afirmaciones son típicas de las hipótesis (aseveraciones) que se prueban usando procedimientos estudiados en este capítulo.

- Un reportero asevera que la mayoría de los conductores estadounidenses se pasan la luz roja.
- Investigadores médicos aseveran que la temperatura corporal media de adultos sanos no es igual a 98.6°F.
- Cuando se utiliza equipo nuevo para fabricar altímetros de aviones, los altímetros nuevos resultan mejores ya que se reduce la variación en los errores, de manera que las lecturas son más consistentes.

Antes de empezar el estudio de este capítulo, usted debe recordar y comprender claramente la siguiente regla básica, que se introdujo en la sección 3-1.

Regla del suceso poco común para la estadística inferencial

Si, bajo un supuesto dado, la probabilidad de un suceso observado particular es excepcionalmente pequeña, concluimos que el supuesto probablemente no sea correcto.

Siguiendo esta regla, probamos una aseveración analizando datos muestrales en un intento por distinguir entre resultados que pueden *ocurrir fácilmente por el azar* y resultados que es *extremadamente improbable que sucedan por el azar*. Podemos explicar la ocurrencia de resultados extremadamente improbables al decir que en realidad ha ocurrido un suceso poco común 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 sus posibilidades de tener un niño hasta en un 85%, y de tener una niña hasta en un 80%". Gender Choice estaba disponible en paquetes azules para parejas que deseaban niño y (ya lo adivinó) paquetes rosas para parejas que deseaban una niña. Suponga que realizamos un experimento con 100 parejas que desean tener niñas, y todas ellas siguen el "sistema casero fácil de usar" de Gender Choice, descrito en el paquete rosa. Con el propósito de probar la aseveración del incremento de posibilidades de tener niñas, suponemos que Gender Choice no tiene efecto alguno. Basados en el sentido común y sin 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 conformados por

a. 52 niñas

b. 97 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. Si las 100 parejas no hubiesen utilizado métodos especiales de selección del género, el resultado de 52 niñas podría ocurrir fácilmente por azar. El supuesto de que Gender Choice no tiene efecto alguno parece ser correcto. No existe evidencia suficiente para decir 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 mediante una de dos maneras: se trata de un evento *extremadamente* poco común que ha ocurrido por azar, o Gender Choice es eficaz. La probabilidad extremadamente baja de que resulten 97 niñas es una fuerte evidencia en contra del supuesto de que Gender Choice no tiene efecto alguno. Parece ser eficaz.

El punto 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 de la media", el resultado de 52 niñas no es significativo, mientras que 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 estándar incorporadas en un procedimiento organizado. Le sugerimos que inicie el estudio de este capítulo con la lectura de las secciones 7-2 y 7-3, de manera informal, para tener una idea general de estos conceptos, y que después lea de nuevo la sección 7-2 con mayor atención para familiarizarse con la terminología.

7-2 Fundamentos de la prueba de hipótesis

En esta sección describimos los componentes formales utilizados en la prueba de hipótesis: hipótesis nula, hipótesis alternativa, estadístico de prueba, región crítica, nivel de significancia, valor crítico, valor P, error tipo I y error tipo II. *En esta*

Explotación de datos

El término *explotación de datos* se utiliza comúnmente para describir la ahora popular práctica de analizar un gran conjunto de datos existentes, con el propósito de encontrar relaciones, patrones o cualquier resultado interesante que no se haya obtenido en los estudios originales del conjunto de datos. Algunos estadísticos expresan su preocupación por la inferencia *ad hoc*, una práctica en la que un investigador va a una expedición de pesca a través de datos viejos, encuentra algo significativo y después identifica una pregunta importante que ya ha sido contestada. Robert Gentleman, editor columnista de la revista *Chance*, escribe que "existen algunos temas estadísticos fundamentales e interesantes que la explotación de datos llega a producir. Sencillamente esperamos que su éxito y auge actuales no hagan demasiado daño a nuestra disciplina (la estadística), antes de que se discutan sus limitaciones".

sección el enfoque se centra en los componentes individuales de la prueba de hipótesis, en tanto que en las siguientes secciones se combinarán estos componentes en extensos procedimientos. He aquí los objetivos de esta sección.

Objetivos de esta sección

- Dada una aseveración, identificar la hipótesis nula y la hipótesis alternativa, y expresar ambas de forma simbólica.
- Dados una aseveración y datos muestrales, calcular el valor del estadístico de prueba.
- Dado un nivel de significancia, identificar el (los) valor(es) crítico(s).
- Dado un valor del estadístico de prueba, identificar el valor de *P*.
- Establecer la conclusión de una prueba de hipótesis en términos simples y sin tecnicismos.
- Identificar los errores tipo I y tipo II que pueden cometerse al probar una aseveración dada.

El lector debe estudiar el siguiente ejemplo hasta comprenderlo exhaustivamente. Una vez que lo logre, ya habrá captado el principal concepto de la estadística.

EJEMPLO Selección y probabilidad del género Refirámonos nuevamente al producto Gender Choice que alguna vez distribuyó ProCare Industries. En la sección 7-1 señalamos que los paquetes rosa de Gender Choice estaban elaborados para ayudar a las parejas a incrementar la posibilidad de tener una niña. ProCare Industries aseveraba que las parejas que utilizaran los paquete rosa de Gender Choice tendrían niñas en una proporción mayor al 50% o 0.5. Consideremos nuevamente un experimento en el que 100 parejas usan Gender Choice en un intento por tener una niña; supongamos que los 100 bebés incluyen exactamente 52 niñas y formalicemos parte del análisis.

En circunstancias normales, la proporción de niñas es de 0.5, de modo que la aseveración de que Gender Choice es eficaz se expresa como $p > 0.5$. El resultado de 52 niñas sustenta dicha aseveración si la probabilidad de tener *al menos 52 niñas* es pequeña, tal como menos que o igual a 0.05. [*Nota importante*: La probabilidad de tener *exactamente* 52 niñas o cualquier otro número específico de niñas es relativamente pequeña, pero nosotros necesitamos la probabilidad de obtener un resultado que es *al menos tan extremo* como el resultado de 52 niñas. Si este punto resulta confuso, revise el apartado "Uso de las probabilidades para determinar si los resultados son infrecuentes", en la sección 4-2, donde señalamos que "x éxitos en n ensayos es un número *excepcionalmente alto* de éxitos si $P(x$ o más) es muy pequeña (como 0.05 o menos)". Con este criterio, el resultado de 52 niñas en 100 nacimientos sería un número extremadamente alto de niñas si P(52 o más niñas) ≤ 0.05].

Si usamos la distribución normal como aproximación de la distribución binomial (véase sección 5-6), encontramos que P(52 o más niñas en 100 nacimientos) $= 0.3821$. Puesto que necesitamos determinar si un resultado de al menos 52 niñas tiene una baja probabilidad en circunstancias normales, suponemos que la probabilidad de una niña es 0.5. La figura 7-1 muestra que, con una probabilidad de 0.5, el resultado de 52 niñas en 100 nacimientos no es poco frecuente, de manera que *no* rechazamos el azar como una explicación razonable. Concluimos que la proporción de niñas nacidas de parejas que usan

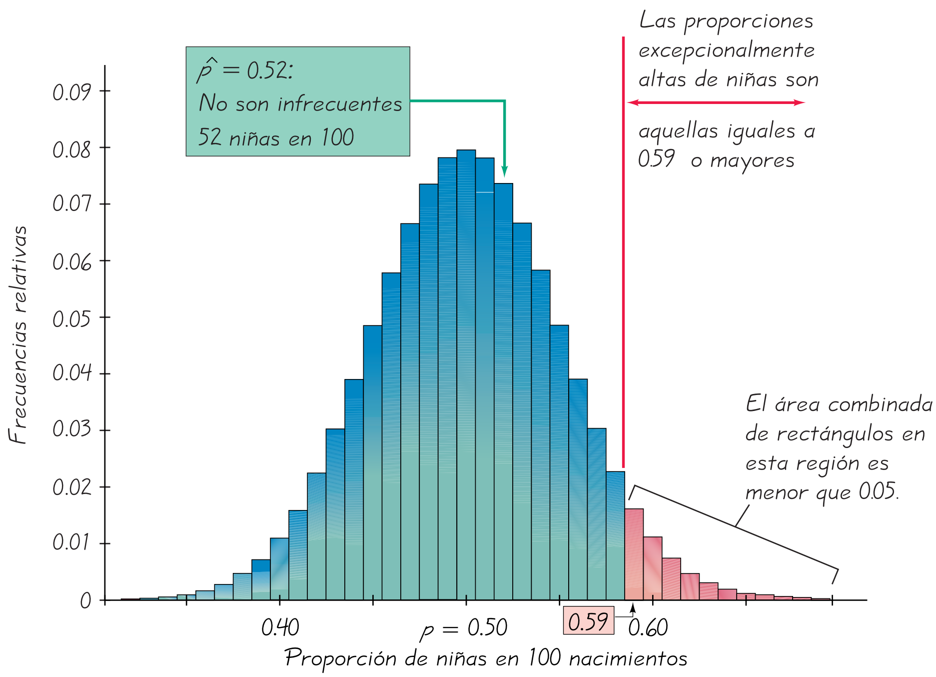


FIGURA 7-1 Distribución muestral de proporciones de niñas en 100 nacimientos

Gender Choice no es significativamente mayor que el número que esperaríamos por el azar. A continuación los puntos clave:

- 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 de Gender Choice).
- La muestra resultó en 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(al menos 52 niñas en 100 nacimientos) $= 0.3821$. (Si utilizamos los métodos de la sección 5-6, con la distribución normal como aproximación de la distribución binomial, tenemos $n = 100$, $p = 0.5$. El valor observado de 52 niñas se modifica a 51.5 por la corrección por continuidad, y 51.5 se transforma a $z = 0.30$).
- Existen dos explicaciones posibles del resultado de 52 niñas en 100 nacimientos: ha ocurrido 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. Gracias a la probabilidad de obtener al menos 52 niñas por el 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 dar a luz 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).

Detectores de mentiras

¿Por qué no requerir que todos los sospechosos de un crimen sean sometidos a la prueba del detector de mentiras y prescindir de los juicios? El Council of Scientific Affairs de la American Medical Association afirma que "está establecido que la clasificación de culpable se realiza con un 75% a 97% de precisión, aunque la tasa de falsos positivos suele ser lo suficientemente alta como para excluir el uso de esta prueba (del polígrafo) como único criterio de culpabilidad o inocencia". Un "falso positivo" es una indicación de culpabilidad cuando el sujeto es en realidad inocente. Incluso con una precisión tan alta como del 97%, el porcentaje de resultados falsos positivos puede ser del 50%, de modo que la mitad de los sujetos inocentes aparecerían incorrectamente como culpables.

El ejemplo anterior ilustra bien el método básico de razonamiento que emplearemos a lo largo de este capítulo. Enfoque su atención en el uso de la regla del suceso infrecuente de la estadística inferencial: **si, bajo un supuesto dado, la probabilidad de un suceso observado particular es excepcionalmente pequeña, concluimos que el supuesto probablemente no sea correcto.** Pero si la probabilidad de un resultado muestral particular observado *no* es muy pequeña, entonces *no* contamos con evidencia suficiente para rechazar el supuesto.

En la sección 7-3 describiremos los casos específicos que se utilizan en la prueba de hipótesis, aunque primero 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 críticas del tipo considerado en este capítulo:

  $$H_0: p = 0.5 \qquad H_0: \mu = 98.6 \qquad H_0: \sigma = 15$$

  La hipótesis nula se aprueba en forma directa, en el sentido de que asumimos 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) 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 $>$ o $\neq$. A continuación se incluyen 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 igual 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 equidad. 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. (En los lugares en que este libro de texto emplea una expresión como $p = 0.5$ para una hipótesis nula, algunos otros libros de texto podrían usar $p \leq 0.5$ o $p \geq 0.5$, en su lugar).

Nota sobre la elaboración 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 $>$ o $\neq$. No puede utilizar una prueba de hipótesis para sustentar la aseveración de que algún parámetro es *igual a* algún valor especificado.

Por ejemplo, suponga que ha creado una poción mágica que incrementa las puntuaciones de CI, de modo que la media se vuelve mayor que 100. Si desea ofrecer evidencia sobre la eficacia de la poción, debe establecer la aseveración

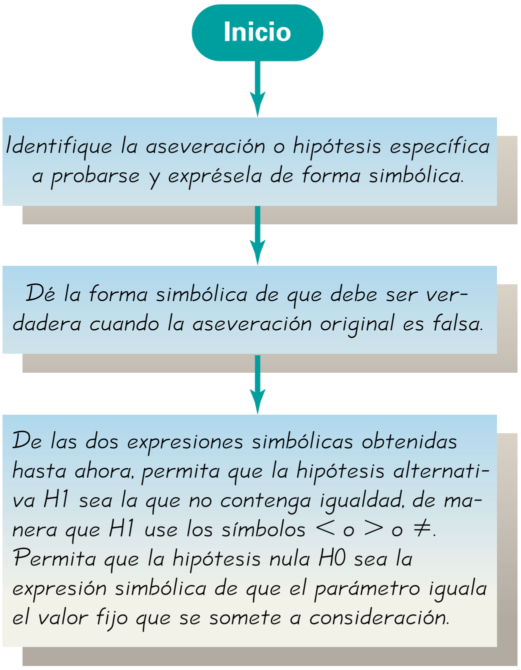


FIGURA 7-2 **Identificación de H_0 y H_1**

como $\mu > 100$. (En el contexto del intento de sustentar la meta de la investigación, la hipótesis alternativa en ocasiones se conoce como la *hipótesis de investigación*. También en este contexto, se asume que la hipótesis nula de $\mu = 100$ es verdadera con el propósito de realizar la prueba de hipótesis, pero se espera que la conclusión incluya el rechazo de la hipótesis nula, de manera que se sustente la aseveración de $\mu > 100$).

Nota sobre la identificación de H_0 y H_1: La figura 7-2 resume los procedimientos para identificar las hipótesis nula y alternativa. Observe que la afirmación original puede convertirse en la hipótesis nula, en la hipótesis alternativa o podría no corresponder 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 se expresa en símbolos tales como $\mu \geq 12$. En la figura 7-2 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 determinar 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 7-2 y utilice las aseveraciones para expresar las hipótesis nula y alternativa de forma simbólica.

a. La proporción de conductores que admiten pasarse la luz roja es mayor que 0.5.

b. La estatura media de jugadores de basquetbol profesional es de al menos siete pies.

c. La desviación estándar de las puntuaciones de actores es igual a 15.

continúa

El tamaño de muestra grande no es 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 estar integrada mayoritariamente por mujeres que tienen fuertes sentimientos acerca de los temas abordados. 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.

SOLUCIÓN Consulte la figura 7-2, que incluye el procedimiento de los tres pasos.

a. En el paso 1 de la figura 7-2, 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 tres, vimos que la expresión $p > 0.5$ no contiene igualdad, por lo que permitimos que la hipótesis alternativa H_1 sea $p > 0.5$, y permitimos que H_0 sea $p = 0.5$.

b. En el paso 1 de la figura 7-2, expresamos "una media de al menos siete pies" en símbolos como $\mu \leq 7$. En el paso 2 observamos que si $\mu \leq 7$ es falso, entonces $\mu > 7$ debe ser verdadero. En el paso 3 vemos que la expresión $\mu > 7$ no contiene igualdad, por lo que permitimos que la hipótesis alternativa H_1 sea $\mu > 7$ y que H_0 sea $\mu = 7$.

c. En el paso 1 de la figura 7-2 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 H_1 sea $\sigma \neq 15$ y que H_0 sea $\sigma = 15$.

Estadístico de prueba

- El **estadístico de prueba** es un valor calculado a partir de datos muestrales, que se utiliza para tomar la decisión sobre el rechazo de la hipótesis nula. El estadístico de prueba 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. El estadístico de prueba sirve, por lo tanto, para determinar si existe evidencia significativa en contra de la hipótesis nula. En este capítulo, consideramos las pruebas de hipótesis que incluyen proporciones, medias y desviaciones estándar (o varianzas). Con base en los resultados de capítulos previos acerca de las distribuciones muestrales de proporciones, medias y desviaciones estándar, empleamos los siguientes estadísticos de prueba:

Estadístico de prueba para proporciones
$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

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 anterior estadístico de prueba para proporciones se basa en los resultados dados en la sección 5-6, pero no incluye la corrección por continuidad que solemos emplear cuando aproximamos una distribución binomial con una distribución normal. Al trabajar con proporciones en este capítulo, utilizaremos muestras grandes, de manera que la corrección por continuidad pueda ignorarse debido a que su efecto es pequeño. Además, el estadístico de prueba para medias puede basarse en la distribución normal o distribución t de Student, dependiendo de las condiciones satisfechas. Al elegir entre las distribuciones normal y t de Student, en este capítulo usaremos los mismos criterios descritos en la sección 6-4. (Véase la figura 6-6 y la tabla 6-1).

EJEMPLO Cálculo del estadístico de prueba Una encuesta de $n = 880$ conductores adultos, seleccionados aleatoriamente, mostró que el 56% (o $\hat{p} = 0.56$) de dichos individuos admitieron pasarse la luz roja de los semáforos. Calcule el valor del estadístico de prueba para la aseveración de que la mayoría de los conductores adultos admiten pasarse la luz roja. (En la sección 7-3 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 dada genera 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:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

INTERPRETACIÓN De capítulos previos sabemos que la puntuación z de 3.56 es excepcionalmente grande. Parece que, además de ser "más que la mitad", el resultado muestral de 56% es *significativamente* mayor que el 50%. Observe la figura 7-3, donde demostramos que la proporción muestral de 0.56 (del 56%) cae dentro del rango de valores considerados significativos, es decir, aquellos valores que se encuentran tan por encima de 0.5, que no suelen suceder por el 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 hacer que rechacemos la hipótesis nula. Por ejemplo, observe la región roja sombreada en la figura 7-3.

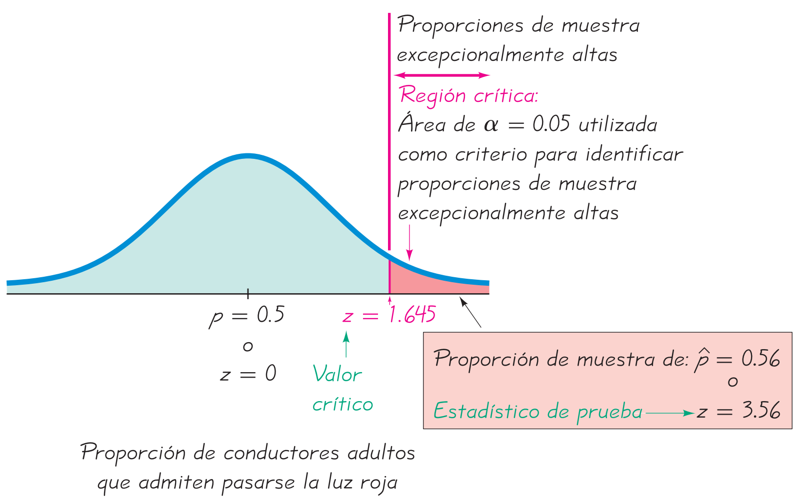


FIGURA 7-3 Región crítica, valor crítico, estadístico de prueba

- 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 modo que α es la probabilidad de cometer el error de rechazar la hipótesis nula cuando es verdadera. Se trata de la misma α introducida en la sección 6-2, 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 el 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 de muestreo que se aplique y del nivel de significancia α. Observe la figura 7-3, donde el valor crítico de $z = 1.645$ corresponde a un nivel de significancia de $\alpha = 0.05$. (Los valores críticos también se estudiaron en el capítulo 6).

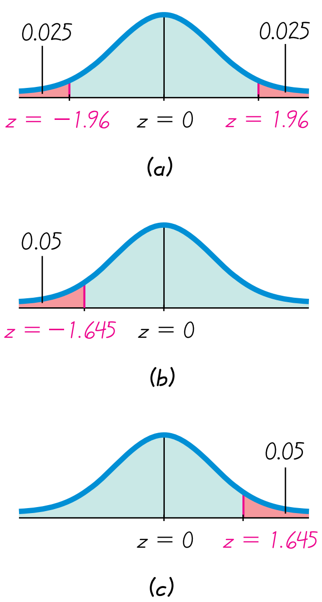


FIGURA 7-4 Cálculo de valores críticos

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):

a. $p \neq 0.5$ (de manera que la región crítica esté en *ambas* colas de la distribución normal)

b. $p < 0.5$ (de manera que la región crítica esté en la cola *izquierda* de la distribución normal)

c. $p > 0.5$ (de manera que la región crítica esté en la cola *derecha* de la distribución normal)

SOLUCIÓN

a. Observe la figura 7-4a. Las colas sombreadas contienen un área total de $\alpha = 0.05$, por lo que cada cola contiene un área de 0.025. Empleando los métodos de la sección 5-2, 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 $z = -1.96$.

b. Observe la figura 7-4b. Con una hipótesis alternativa de $p < 0.5$, la región crítica se encuentra en la cola izquierda. Con un área de cola izquierda de 0.05, se obtiene que el valor crítico es $z = -1.645$ (empleando los métodos de la sección 5-2).

c. Observe la figura 7-4c. Con una hipótesis alternativa de $p > 0.5$, la región crítica está en la cola derecha. Con un área de cola derecha de 0.05, se obtiene que el valor crítico es $z = 1.645$ (empleando los métodos de la sección 5-2).

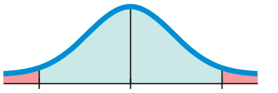
Dos colas, cola izquierda, cola derecha Las *colas* en una distribución son las regiones extremas limitadas por los valores críticos. Algunas pruebas de hipótesis incluyen dos colas, otras la cola derecha y otras la cola izquierda.

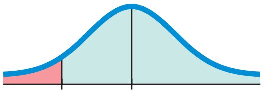
- **Prueba de dos colas:** La región crítica se encuentra en dos regiones extremas (colas) bajo la curva.

- **Prueba de cola izquierda:** La región crítica se encuentra en la región extrema izquierda (cola) bajo la curva.
- **Prueba de cola derecha:** La región crítica se encuentra en la región extrema derecha (cola) bajo la curva.

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 un 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 7-4).

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 7-5), que 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, tal como $p \neq 0.5$, corresponde a una prueba de dos colas.


Signo usado en H_1: $\neq$
Prueba de dos colas


Signo usado en H_1: $<$
Prueba de cola izquierda

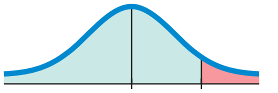

Signo usado en H_1: $>$
Prueba de cola derecha

FIGURA 7-5 Pruebas de dos colas, cola izquierda, cola derecha

- El **valor *P*** (o **valor de *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 7-6 de la siguiente página.

Decisiones y conclusiones

Hemos visto que la aseveración original en ocasiones se convierte en la hipótesis nula y en otras en la hipótesis alternativa. Sin embargo, nuestro procedimiento estándar de prueba de hipótesis requiere que siempre probemos la hipótesis nula, de modo 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, basar la decisión en intervalos de confianza. En años recientes ha disminuido el uso del método tradicional.

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 de *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 de $P > \alpha$.

Otra opción: En lugar de usar un nivel de significancia como $\alpha = 0.05$, simplemente identifique el valor de P y deje la decisión al lector.

Intervalo de confianza: Puesto que un estimado del intervalo de confianza de un parámetro de población contiene los valores posibles de dicho parámetro, rechace la aseveración de que el parámetro de población tiene un valor que no está incluido en el intervalo de confianza.

Muchos estadísticos consideran buena la práctica de seleccionar siempre un nivel de significancia *antes* de hacer una prueba de hipótesis. Éste es un procedimiento particularmente bueno cuando se utiliza el método del valor P, ya que podemos vernos tentados a ajustar el nivel de significancia con base en los resultados. Por ejemplo, con un nivel de significancia de 0.05 y un valor P de 0.06, no deberíamos rechazar la hipótesis nula, pero en ocasiones es tentador decir que la probabilidad de 0.06 es lo suficientemente pequeña para rechazar la hipótesis nula. Otros estadísticos argumentan que la selección previa de un nivel de significancia reduce la utilidad de los valores P. Ellos sostienen que no debe especificarse ningún nivel de significancia y que la conclusión debe dejarse al lector. Utilizaremos el criterio de decisión que incluye la comparación de un nivel de significancia y del valor P.

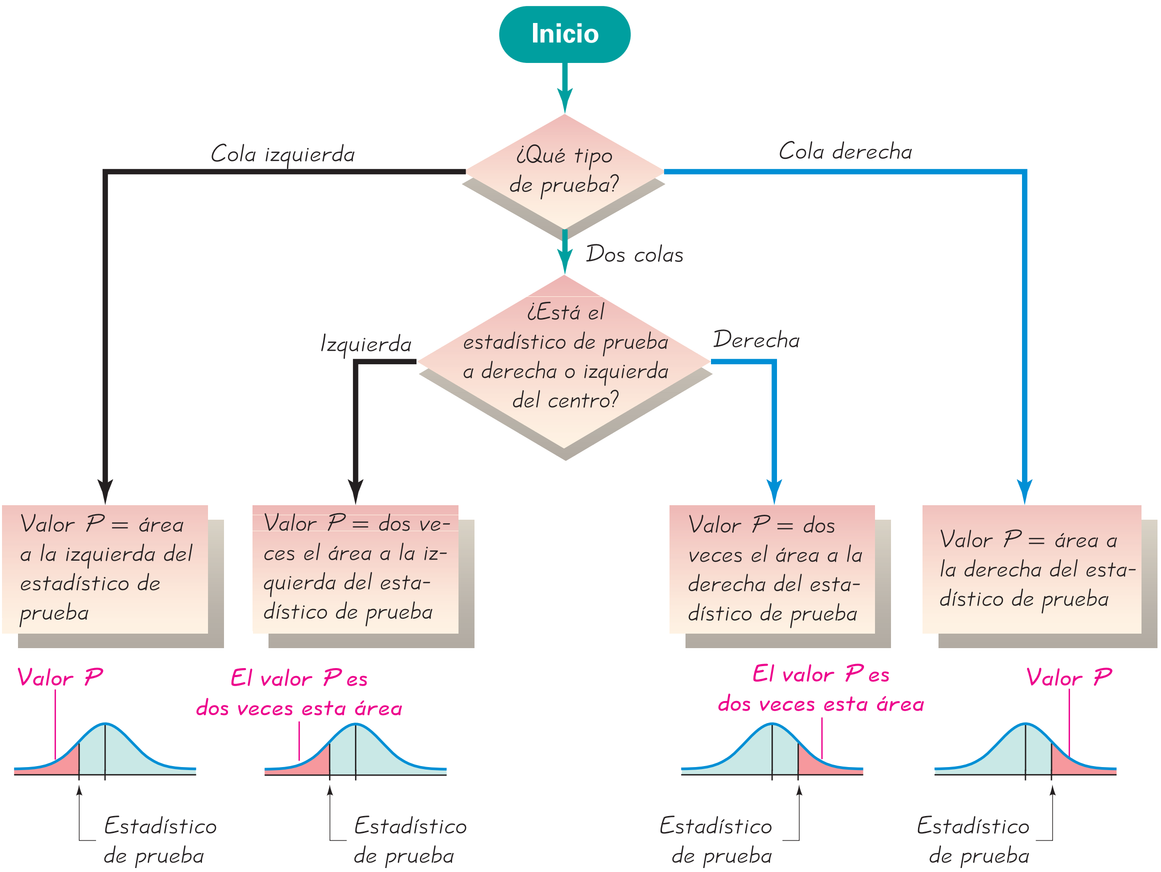


FIGURA 7-6 Procedimiento para el cálculo de los valores *P*

EJEMPLO Cálculo de valores *P* Primero determine si las condiciones dadas resultan en una prueba de cola derecha, de cola izquierda o de dos colas; después utilice la figura 7-6 para calcular el valor de *P*, luego saque una conclusión acerca de la hipótesis nula.

a. Se utiliza un nivel de significancia de $\alpha = 0.05$ para probar la aseveración de que $p > 0.25$, y los datos muestrales producen un estadístico de prueba de $z = 1.18$.

b. Se utiliza un nivel de significancia de $\alpha = 0.05$ para probar la aseveración de que $p \neq 0.25$, y los datos muestrales producen un estadístico de prueba de $z = 2.34$.

SOLUCIÓN

a. Con la aseveración de que $p > 0.25$, se trata de una prueba de cola derecha (véase la figura 7-5). Podemos calcular el valor *P* utilizando la figura 7-6. Como la prueba es de cola derecha, la figura 7-6 indica que el valor *P* es el área a la derecha del estadístico de prueba $z = 1.18$. Si empleamos los métodos de la sección 5-2, 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 el azar.

b. Con la aseveración de $p \neq 0.25$, se trata de una prueba de dos colas (véase la figura 7-5). Podemos calcular el valor *P* por medio de la figura 7-6. Como la prueba es de dos colas, ya que el estadístico de prueba $z = 2.34$ se encuentra a la derecha del centro, la figura 7-6 indica que el valor *P* es *dos veces* el área a la derecha de $z = 2.34$. Si empleamos los métodos de la sección 5-2, 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 de $P = 2 \times 0.0096 = 0.0192$. 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.

Conclusión final: La conclusión de rechazar o no la hipótesis nula es adecuada para aquellos que tenemos la inteligencia de tomar un curso de estadística, pero debemos emplear términos simples y sin tecnicismos al establecer el verdadero significado de la conclusión. La figura 7-7 de la siguiente página resume el procedimiento para plantear 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 la aseveración de alguien, indíquelo de manera tal que se convierta en la hipótesis alternativa, y después espere que la hipótesis nula sea rechazada. Por ejemplo, para sustentar la aseveración de que la temperatura corporal media difiere de 98.6°, plantee la aseveración de que $\mu \neq 98.6°$. Esta aseveración será una hipótesis alternativa que se sustentará si usted rechaza la hipótesis nula H_0: $\mu = 98.6°$. Si, por otro lado, usted asevera que $\mu = 98.6°$, rechazará o no dicha aseveración; en cualquier caso, nunca *sustentará* la aseveración de que $\mu = 98.6°$.

Aceptación/no rechazo: Algunos libros de texto dicen “aceptar la hipótesis nula” en lugar 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 suficiente-

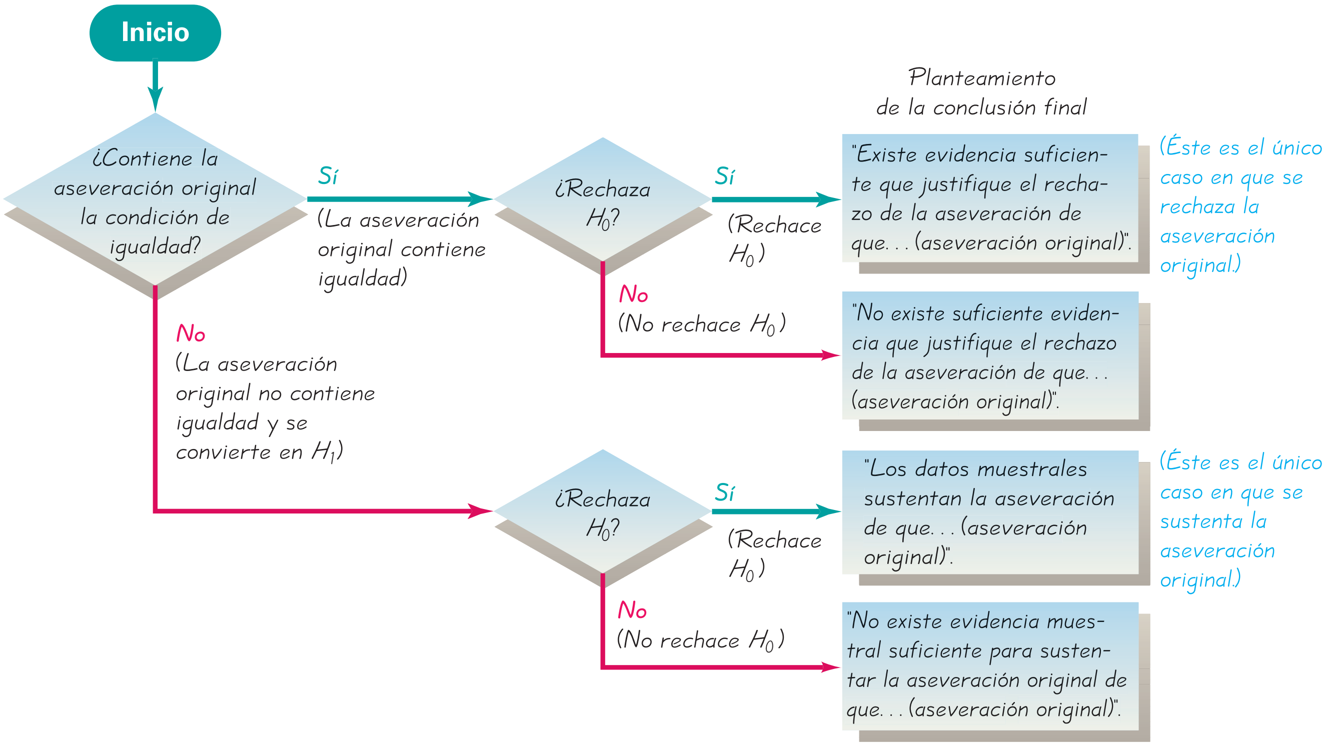


FIGURA 7-7 Conclusión final

mente fuerte como para justificar el rechazo de la hipótesis nula. Es similar a un jurado que afirma que no existe evidencia suficiente para sentenciar a un sospechoso. El término *aceptar* es un poco confuso, ya que parece indicar incorrectamente que la hipótesis nula ha sido probada. (Es confuso decir que "existe evidencia suficiente para aceptar la hipótesis nula"). La frase *no rechazar* indica correctamente 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 lugar de *aceptar la hipótesis nula*.

Múltiples negativos: Cuando se establece la conclusión final en términos no técnicos, es posible establecer 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 suelen ser confusas, por lo que sería bueno volver a redactarlas en una forma comprensible, pero teniendo cuidado de no cambiar el significado. Por ejemplo, en lugar 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", las siguientes serían mejores afirmaciones:

- No se rechaza la aseveración de que la proporción poblacional es igual a 0.5.
- Hasta no obtener evidencia más fuerte, se supone que la proporción poblacional es igual a 0.5.

EJEMPLO Conclusión final Suponga que un reportero asevera que "más de la mitad" de todos los adultos estadounidenses que conducen admiten pasarse la luz roja. 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 $p = 0.5$. Plantee la conclusión en términos simples y sin tecnicismos.

SOLUCIÓN Remítase a la figura 7-7. La aseveración original no incluye la condición de igualdad, y rechazamos la hipótesis nula. Por lo tanto, el planteamiento de la conclusión final debe ser el siguiente: "Los datos muestrales sustentan la aseveración de que más de la mitad de los conductores adultos estadounidenses admiten pasarse la luz roja".

Errores tipo I y tipo II Cuando se prueba una hipótesis nula llegamos a la conclusión de rechazarla o no rechazarla. Dichas conclusiones pueden ser correctas o incorrectas (incluso cuando hacemos todo correctamente). La tabla 7-1 resume los dos distintos tipos de errores que llegan a 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.

Tabla 7-1 Errores tipo I y tipo II

		Estado verdadero de la naturaleza	
		La hipótesis nula es verdadera.	La hipótesis nula es falsa.
Decisión	Decidimos rechazar la hipótesis nula.	**Error tipo I** (rechazar una hipótesis nula siendo verdadera) α	Decisión correcta
	No rechazamos la hipótesis nula.	Decisión correcta	**Error tipo II** (no rechazar una hipótesis nula siendo falsa) β

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)

Como los estudiantes suelen considerar difícil recordar cuál error es el tipo I y cuál es el error tipo II, recomendamos una técnica mnemónica, como podría ser "revisión no refinada" (**R**e**V**isió**N** **N**o **R**e**F**i**N**ada". 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).

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 cuando es verdadera, por lo tanto 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, por lo tanto 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 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 magnífico si tuviéramos siempre $\alpha = 0$ y $\beta = 0$, pero en realidad eso no es posible, por lo que debemos intentar manejar las probabilidades de los errores α y β. Matemáticamente, se demuestra que α, β y el tamaño de muestra n están relacionados, de manera que cuando se elige o determina cualquiera de los 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 determina 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 α. Después elija un tamaño de muestra n lo razonablemente grande, con base en consideraciones de tiempo, costo y otros factores relevantes (Las determinaciones del tamaño de muestra se estudiaron en el capítulo 6). Las siguientes consideraciones prácticas resultan relevantes:

1. Para cualquier α fija, un incremento en el tamaño de 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 causará un decremento en β.
3. Para disminuir tanto α como β, incremente 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 0.9085 g (para conformar el peso impreso en la etiqueta del empaque).
- Se supone que las tabletas Bufferin tienen un peso medio de 325 mg de aspirina.

Como los dulces M&M se consumen para disfrutarlos, mientras que las tabletas Bufferin son fármacos utilizados para el tratamiento de problemas de salud, tratamos con dos niveles muy diferentes de gravedad. Si los dulces M&M no tienen un peso medio de 0.9085 g, las consecuencias no son muy graves, pero si las tabletas Bufferin no contienen una media de 325 mg de aspirina, las consecuencias serían muy graves, incluyendo posiblemente demandas de los consumidores y acciones por parte de la Federal Drug Administration. En consecuencia, para probar la aseveración de que $\mu = 0.9085$ g de los M&M, podríamos elegir $\alpha = 0.05$ y un tamaño demuestra 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 $n = 500$. (El tamaño de muestra más grande nos permite disminuir β, mientras disminuimos también α). Se elige un nivel de significancia α menor y un tamaño de muestra n más grande por las consecuencias más graves asociadas con la prueba de un fármaco comercial.

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 estadísticos se refieren a esta probabilidad como la *potencia* de una prueba, y con frecuencia la utilizan cuando quieren 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, que se calcula utilizando un nivel de significancia α particular y un valor específico del parámetro de población que representa una alternativa al valor asumido 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.

Suponga que estamos usando un nivel de significancia de 0.05 para probar la hipótesis nula de que la estatura media de los hombres es de 6 pies (o 72 pulgadas).

FIGURA 7-8 Método tradicional

Método tradicional

Inicio

1. Identifique la aseveración o hipótesis específica que será probada y póngala en forma simbólica.
2. Dé la forma simbólica que debe ser verdadera cuando la aseveración original es falsa.
3. Dé las dos expresiones simbólicas obtenidas hasta este momento, permita que la hipótesis alternativa H_1 sea la que no contenga igualdad, de manera que H_1 emplee el símbolo $>$ o $<$ o $\neq$. Permita que la hipótesis nula H_0 sea la expresión simbólica de que el parámetro iguala el valor fijo considerado.
4. Elija el nivel de significancia α de acuerdo con la gravedad de cometer un error tipo I. Disminuya α si las consecuencias de rechazar una H_0 verdadera son graves. Los valores 0.05 y 0.01 son muy comunes.
5. Identifique el estadístico que sea relevante para esta prueba y determine su distribución muestral (tal como normal, t, de chi cuadrada).
6. Calcule el estadístico de prueba, los valores críticos y la región crítica. Dibuje una gráfica e incluya el estadístico de prueba, el (los, valor(es) crítico(s) y la región crítica.
7. Rechace H_0 si el estadístico de prueba se encuentra en la región crítica. No rechace H_0 si el estadístico de prueba no se encuentra en la región crítica.
8. Establezca esta decisión previa en términos simples y sin tecnicismos y retome la aseveración original.

Alto

FIGURA 7-9 Método del valor *P*

Método del valor *P*

Inicio

1. Identifique la aseveración o hipótesis específica que será probada y póngala de forma simbólica.
2. Dé la forma simbólica que debe ser verdadera cuando la aseveración original es falsa.
3. Dé las dos expresiones simbólicas obtenidas hasta este momento, permita que la hipótesis alternativa H_1 sea la que no contenga igualdad, de manera que H_1 emplee el símbolo $>$ o $<$ o $\neq$. Permita que la hipótesis nula H_0 sea la expresión simbólica de que el parámetro iguala el valor fijo considerado.
4. Elija el nivel de significancia α de acuerdo con la gravedad de cometer un error tipo I. Disminuya α si las consecuencias de rechazar una H_0 verdadera son graves. Los valores 0.05 y 0.01 son muy comunes.
5. Identifique el estadístico que sea relevante para esta prueba y determine su distribución muestral (tal como normal, t, de chi cuadrada).
6. Calcule el estadístico de prueba y el valor P (véase la figura 7-6). Dibuje una gráfica y muestre el estadístico de prueba y el valor P.
7. Rechace H_0 si el valor P es menor o igual que el nivel de significancia α. No rechace H_0 si el valor P es mayor que α.
8. Establezca esta decisión previa en términos simples y sin tecnicismos y retome la aseveración original.

Alto

Método del intervalo de confianza

Construya un intervalo de confianza con un nivel de confianza seleccionado de la misma forma que en la tabla 7-2. **Como un estimado del intervalo de confianza de un parámetro de población contiene los posibles valores de dicho parámetro, rechace la aseveración de que el parámetro de población tiene un valor que no está incluido en el intervalo de confianza.**

Tabla 7-2 Nivel de confianza del intervalo de confianza

		Prueba de dos colas	Prueba de una cola
Nivel de significancia para la prueba de hipótesis	0.01	99%	98%
	0.05	95%	90%
	0.10	90%	80%

A partir de los datos muestrales y de la estatura alternativa de 69 pulgadas, podemos calcular la potencia de la prueba para rechazar $\mu = 72$. Si nuestra muestra consiste en unas cuantas observaciones, la potencia será baja, pero si consta de cientos de observaciones la potencia será mucho más alta. (Además de incrementar el tamaño de muestra, existen otras formas para incrementar la potencia, tales como el aumento del nivel de significancia, el uso de valores más extremos para la media poblacional o el decremento de la desviación estándar). Así como 0.05 suele ser una opción común para un nivel de significancia, una potencia de al menos 0.80 es un requisito común para determinar si una prueba de hipótesis es efectiva. (Algunos estadísticos argumentan que la potencia debe ser más alta, como 0.85 o 0.90). Puesto que los cálculos de la potencia son realmente arduos, sólo el ejercicio 46 tiene que ver con la potencia.

Prueba profunda de hipótesis En esta sección describimos los componentes individuales utilizados en una prueba de hipótesis, pero las siguientes secciones combinarán dichos componentes en procedimientos más profundos. Probamos aseveraciones sobre parámetros de población con el uso del método tradicional que se resume en la figura 7-8, el método del valor *P* incluido en la figura 7-9, o emplear un intervalo de confianza (descrito en el capítulo 6). 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 7-2 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 dicho 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. *Cuidado*: En algunos casos, una conclusión basada en un intervalo de confianza es diferente de una conclusión basada en una prueba de hipótesis. Consulte los comentarios en las secciones individuales.**

Los ejercicios de esta sección incluyen componentes aislados de la prueba de hipótesis, pero las siguientes secciones incluirán pruebas de hipótesis completas y profundas.

7-2 Destrezas y conceptos básicos

Conclusiones sobre aseveraciones. *En los ejercicios 1 a 4, ¿qué concluye? (No emplee procedimientos formales ni cálculos exactos. Utilice sólo la regla del suceso infrecuente descrita en la sección 7-1, y haga estimados objetivos para determinar si los sucesos son posibles).*

1. Aseveración: Un método de selección del género es eficaz para ayudar a que las parejas tengan niñas y, de 50 bebés, 26 son niñas.

2. Aseveración: Un método de selección del género es eficaz para ayudar a que las parejas tengan niñas y, de 50 bebés, 49 son niñas.

3. Aseveración: A la mayoría de los adultos estadounidenses les gusta la pizza, y una encuesta de 500 adultos estadounidenses seleccionados al azar muestra que a 475 de ellos les gusta la pizza.

4. Aseveración: Las personas nacidas el 29 de febrero tienen puntuaciones de CI que varían menos que la población general, para la cual $\sigma = 15$, y una muestra aleatoria de 50 personas nacidas el 29 de febrero registra puntuaciones de CI con $s = 14.99$.

Identificación de H_0 y H_1. *En los ejercicios 5 a 12, 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.*

5. El ingreso anual medio de trabajadores con estudios de estadística es mayor que $50,000.

6. El CI medio de estudiantes de estadística es de al menos 110.

7. Más de la mitad de los usuarios de Internet realiza compras en línea.

8. El porcentaje de hombres que ven el golf por televisión no es el 70%, como afirma la Madison Advertising Company.

9. La estatura de las mujeres tiene una desviación estándar menor que 2.8 pulgadas, que es la desviación estándar de la estatura de los hombres.

10. El porcentaje de televidentes que sintoniza *60 minutos* es igual al 24%.

11. La cantidad media de Coca Cola en lata es de al menos 12 onzas.

12. Los salarios de mujeres analistas de negocios tienen una desviación estándar mayor que $ 3000.

Cálculo de valores críticos. *En los ejercicios 13 a 20, calcule los valores z críticos. En cada caso, suponga que se aplica la distribución normal.*

13. Prueba de dos colas: $\alpha = 0.05$.

14. Prueba de dos colas; $\alpha = 0.01$.

15. Prueba de cola derecha; $\alpha = 0.01$.

16. Prueba de cola izquierda; $\alpha = 0.05$.

17. $\alpha = 0.10$; H_1 es $p \neq 0.17$.

18. $\alpha = 0.10$; H_1 es $p > 0.18$.

19. $\alpha = 0.02$; H_1 es $p < 0.19$.

20. $\alpha = 0.005$; H_1 es $p \neq 0.20$.

Cálculo de estadísticos de prueba. *En los ejercicios 21 a 24, calcule el valor del estadístico de prueba z utilizando*

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

21. Encuesta de Gallup La aseveración es que la proporción de adultos que compra a través de Internet es menor que 0.5 (o 50%), y los estadísticos de muestra incluyen $n = 1025$ sujetos, de los cuales el 29% dice que utiliza Internet para realizar compras.

22. Experimento de genética La aseveración es que la proporción de chícharos con vainas amarillas es igual a 0.25 (o 25%), y los estadísticos de muestra incluyen $n = 580$ chícharos, de los cuales el 26.2% presenta vainas amarillas.

23. Estudio sobre seguridad La aseveración es que la proporción de muertes infantiles por ahogamiento, atribuibles a los globos, es mayor que 0.25, y los estadísticos muestrales incluyen $n = 400$ muertes infantiles por ahogamiento; el 29.0% de ellas puede atribuirse a los globos.

24. Prácticas policiacas La aseveración es que la proporción de conductores detenidos por la policía en un año difiere de la tasa del 10.3% reportada por el Departamento de Justicia de Estados Unidos. Los estadísticos de muestra incluyen $n = 800$ conductores seleccionados aleatoriamente; el 12% de ellos fueron detenidos durante el año anterior.

Cálculo de valores P*. En los ejercicios 25 a 32, utilice la información dada para calcular el valor P. (*Sugerencia: *consulte la figura 7-6).*

25. El estadístico de prueba en una prueba de cola derecha es $z = 0.55$.

26. El estadístico de prueba en una prueba de cola izquierda es $z = -1.72$.

27. El estadístico de prueba en una prueba de dos colas es $z = 1.95$.

28. El estadístico de prueba en una prueba de dos colas es $z = -1.63$.

29. Con H_1: $p > 0.29$, el estadístico de prueba es $z = 1.97$.

30. Con H_1: $p \neq 0.30$, el estadístico de prueba es $z = 2.44$.

31. Con H_1: $p \neq 0.31$, el estadístico de prueba es $z = 0.77$.

32. Con H_1: $p < 0.32$, el estadístico de prueba es $z = -1.90$.

Conclusiones. *En los ejercicios 33 a 36, establezca la conclusión final en términos simples y sin tecnicismos. Asegúrese de enfatizar la aseveración original. (*Sugerencia: *Consulte la figura 7-7).*

33. Aseveración original: La proporción de mujeres casadas es mayor que 0.5. Conclusión inicial: Rechazar la hipótesis nula.

34. Aseveración original: La proporción de graduados universitarios que fuman es menor que 0.27. Conclusión inicial: Rechazar la hipótesis nula.

35. Aseveración original: La proporción de accidentes fatales de aviación comercial difiere de 0.038. Conclusión inicial: No rechazar la hipótesis nula.

36. Aseveración original: La proporción de M&M azules es igual a 0.10.
Conclusión inicial: Rechazar la hipótesis nula.

Identificación de errores tipo I y tipo II*. En los ejercicios 37 a 40, identifique el error tipo I y el error tipo II correspondiente a la hipótesis dada.*

37. La proporción de mujeres casadas es mayor que 0.5.

38. La proporción de graduados universitarios que fuma es menor que 0.27.

39. La proporción de accidentes fatales de aviación comercial difiere de 0.038.

40. La proporción de M&M azules es igual a 0.10.

7-2 Más allá de lo básico

41. Prueba innecesaria Para probar la aseveración de que la mayoría de los estadounidenses adultos están en contra de aplicar la pena de muerte a una persona sentenciada por homicidio, se obtiene una muestra aleatoria de 491 adultos, y 27% de ellos se manifiestan en contra de la pena de muerte (según datos de una encuesta de Gallup). Calcule el valor de P. ¿Por qué no es necesario seguir los pasos para realizar una prueba formal de hipótesis?

42. 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?

43. Valor P Suponga que acaba de crear un nuevo proceso de fabricación que usted considera que reduce la tasa de defectos en la producción de microchips. Planea justificar su aseveración de una tasa más baja de defectos por medio de una prueba de hipótesis. ¿Que valor P preferiría, 0.10, 0.05, 0.01? ¿Por qué?

44. Prueba de aseveraciones Usted es el gerente de control de calidad de Mars, Inc. y desea probar la aseveración de la compañía de que el 10% de los dulces M&M son azules. ¿Es posible probar esa aseveración utilizando métodos de prueba de hipótesis? ¿Por qué?

45. ¿Por qué no permitir que $\alpha = 0$? Alguien sugiere que para probar hipótesis usted puede eliminar un error tipo I haciendo que $\alpha = 0$. En una prueba de dos colas, ¿qué valores críticos corresponden a $\alpha = 0$? Si $\alpha = 0$, ¿será rechazada alguna vez la hipótesis nula?

46. 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 $n = 1998$, con $\hat{p} = 0.48$.

a. Calcule β, la probabilidad de cometer un error tipo II, dado que la proporción poblacional es en realidad 0.45. (*Sugerencia:* Primero calcule los valores de las proporciones muestrales que no conducen al rechazo de H_0. Después, suponiendo que $p = 0.45$, calcule la probabilidad de obtener una proporción muestral con uno de dichos valores).

b. Calcule $1 - \beta$, que es la *potencia* de la prueba. Si β es la probabilidad de *no* rechazar la hipótesis nula falsa, describa la probabilidad de $1 - \beta$.

7-3 Prueba de una aseveración respecto de una proporción

En la sección 7-2 presentamos los componentes aislados de una prueba de hipótesis, pero en esta sección combinamos esos componentes en pruebas de hipótesis profundas de aseveraciones hechas acerca de proporciones poblacionales. Las proporciones también representan probabilidades o los equivalentes decimales de porcentajes. Los siguientes son ejemplos de los tipos de aseveraciones que es factible probar.

- Menos de 1/4 de todos los graduados universitarios fuman.
- Los sujetos que toman el fármaco Lipitor, que reduce el colesterol, experimentan dolores de cabeza en una proporción mayor que el 7% de las personas que no toman Lipitor.
- El porcentaje de televidentes nocturnos que ven *The Late Show with David Letterman* es igual al 18%.
- Con base en encuestas tempranas de salida, el candidato republicano a la presidencia ganará la mayoría (más del 50%) de los votos.

A continuación se presentan los supuestos, notación y estadístico de prueba requeridos. Básicamente, las aseveraciones sobre una proporción poblacional suelen probarse al utilizar una distribución normal como aproximación de la distribución binomial, tal como lo hicimos en la sección 5-6. En lugar de utilizar exactamente los mismos métodos de dicha sección, 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 encuestas grandes). Si no se satisfacen los supuestos dados, hay otros métodos que no se describen en esta sección. Aquí todos los ejemplos y ejercicios incluyen casos en que los supuestos se satisfacen, de manera que la distribución muestral de proporciones de muestra se aproxima usando la distribución normal.

Prueba de aseveraciones sobre una proporción poblacional p

Supuestos

1. Las observaciones muestrales son una muestra aleatoria simple. (Nunca olvide la importancia fundamental de los métodos adecuados de muestreo).
2. Se satisfacen las condiciones para una *distribución binomial.* (Hay 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 5-6).

Notación

n = tamaño de muestra o número de ensayos

$\hat{p} = \frac{x}{n}$ (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 7-6.

Valores críticos: Utilice la distribución normal estándar (tabla A-2).

La ética en los reportes

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 evaluación que considera el riesgo para los sujetos encuestados, el mérito científico de la encuesta y la garantía del consentimiento de los sujetos para participar.

PROBLEMA 7 DEL CAPÍTULO

EJEMPLO Encuesta de conductores En el problema del capítulo señalamos que un artículo distribuido por la Associated Press incluía los siguientes resultados de una encuesta nacional: de 880 conductores seleccionados aleatoriamente, el 56% admitió haberse pasado la luz roja. La reportera Sonja Barisic escribió esto: "Casi todos los conductores estadounidenses coinciden en que pasarse la luz roja es peligroso, pero más de la mitad de ellos admite haberlo hecho, . . . , encontró una encuesta". Esta afirmación incluye la aseveración de que la *mayoría* (más de la mitad) de todos los estadounidenses se pasan la luz roja. A continuación se incluye un resumen de la aseveración y de los datos muestrales:

Aseveración: Más de la mitad (de todos los estadounidenses) admite pasarse la luz roja. Es decir, $p > 0.5$.

Datos muestrales: $n = 880$ y $\hat{p} = 0.56$

continúa

Ejemplificaremos la prueba de hipótesis con el uso del método tradicional, el popular método del valor P y los intervalos de confianza. Sin embargo, antes de proceder debemos verificar que se satisfagan los supuestos requeridos. Se trata de una muestra aleatoria simple, existe un número fijo (880) de ensayos independientes con dos categorías (el sujeto admite o no admite pasarse la luz roja) y se satisfacen $np \geq 5$ y $nq \geq 5$, con $n = 880$, $p = 0.5$ y $q = 0.5$. (Técnicamente los ensayos no son independientes, aunque pueden tratarse como independientes al utilizar el siguiente lineamiento presentado en la sección 4-3: "Cuando se realiza 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. Es decir, $n \leq 0.05N$"). Una vez satisfechos todos los supuestos requeridos, ahora podemos proceder a realizar una prueba formal de hipótesis. El método tradicional, el método del valor P y el uso de intervalos de confianza se ejemplifican en la siguiente explicación.

El método tradicional

El método tradicional de prueba de hipótesis se resume en la figura 7-8. Cuando se prueba la aseveración $p > 0.5$, dada en el ejemplo anterior, los siguientes pasos corresponden al procedimiento de la figura 7-8:

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 es igual al valor fijo de 0.5. Por lo tanto, podemos expresar H_0 y H_1 como sigue:

$$H_0\colon p = 0.5$$
$$H_1\colon p > 0.5$$

Paso 4: En la ausencia de circunstancias especiales, seleccionamos $\alpha = 0.05$ para el nivel de significancia.

Paso 5: Como estamos probando una aseveración sobre una proporción poblacional p, el estadístico muestral $\hat{p}$ es relevante para esta prueba y la distribución muestral de proporciones de muestra $\hat{p}$ se aproxima por medio de una *distribución normal*.

Paso 6: El estadístico de prueba se evalúa utilizando $n = 880$ y $\hat{p} = 0.56$. En la hipótesis nula estamos suponiendo que $p = 0.5$, de modo que $q = 1 - 0.5 = 0.5$. El estadístico de prueba es

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

Se trata de una prueba de cola derecha, por lo que la región crítica es un área de $\alpha = 0.05$ en la cola derecha. Si nos remitimos a la tabla A-2 y aplicamos los métodos de la sección 5-2, encontramos que el valor crítico de $z = 1.645$ se localiza en el límite de la región crítica. Consulte la figura 7-3 (página 75), que presenta la región crítica, el valor crítico y el estadístico de prueba.

Paso 7: Ya que el estadístico de prueba cae dentro de la región crítica, rechazamos la hipótesis nula.

Paso 8: Concluimos que existe suficiente evidencia muestral que sustenta la aseveración de que la mayoría de los estadounidenses admiten pasarse la luz roja. (Véase la figura 7-7 para la redacción de esta conclusión final).

El método del valor *P*

El método del valor *P* para prueba de hipótesis se resume en la figura 7-9, y requiere del valor *P*, que se obtiene utilizando el procedimiento resumido en la figura 7-6. La comparación de las figuras 7-8 y 7-9 indica que los primeros cinco pasos del método tradicional son iguales a los primeros cinco pasos del método del valor *P*. Para la prueba de hipótesis descrita en el ejemplo anterior, los primeros cinco pasos del método del valor *P* son iguales a los que se presentan en el método tradicional anterior, por lo que ahora continuamos con el paso 6.

Paso 6: El estadístico de prueba es $z = 3.56$, tal como se muestra en el método tradicional anterior. Ahora calculamos el valor *P* (en lugar del valor crítico) utilizando el siguiente procedimiento, que se presenta en la figura 7-6:

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 $P =$	*2 veces* el área de la región extrema limitada por el estadístico de prueba z

Puesto que la prueba de hipótesis que estamos considerando es de cola derecha, con un estadístico de prueba de $z = 3.56$, el valor *P* es el área a la derecha de $z = 3.56$. Al remitirnos a la tabla A-2, observamos que para valores de $z = 3.50$ y más altos, utilizamos 0.9999 para el área acumulativa a la *izquierda* del estadístico de prueba. El área a la derecha de $z = 3.56$ es, por lo tanto, $1 - 0.9999 = 0.0001$. Ahora sabemos que el valor *P* es 0.0001. La figura 7-10 incluye el estadístico de prueba y el valor *P* para este ejemplo.

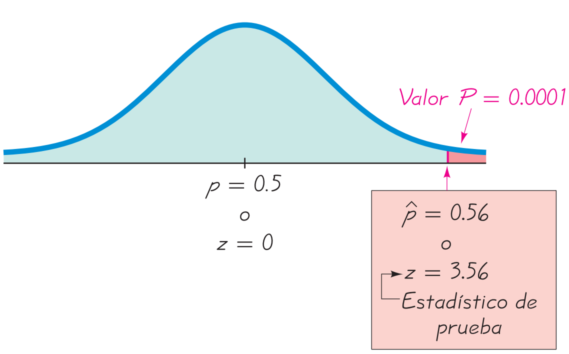


FIGURA 7-10 Método del valor *P*

Paso 7: Puesto que el valor P de 0.0001 es menor o igual que el nivel de significancia de $\alpha = 0.05$, rechazamos la hipótesis nula.

Paso 8: Igual que en el método tradicional, concluimos que existe suficiente evidencia muestral para sustentar la aseveración de que la mayoría de los estadounidenses admiten pasarse la luz roja. (Véase la figura 7-7 sobre la redacción de esta conclusión final).

Método de intervalos de confianza

Para las 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 nivel de significancia α, construya un intervalo de confianza con un nivel de confianza de $1 - 2\,\alpha$. (Véase la tabla 7-2 para los casos comunes). Por ejemplo, la aseveración de $p > 0.5$ se prueba con un nivel de significancia de 0.05, construyendo un intervalo de confianza del 90 por ciento.

Ahora utilicemos el método del intervalo de confianza para probar la aseveración de $p > 0.5$, con datos muestrales que consisten en $n = 880$ y $\hat{p} = 0.56$ (de los ejemplos al inicio de esta sección). 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 6-2 para obtener este resultado: $0.533 < p < 0.588$. Puesto que tenemos una confianza del 90% de que el valor verdadero de p está contenido dentro de los límites de 0.533 y 0.588, tenemos evidencia suficiente para sustentar la aseveración de que $p > 0.5$.

Cuidado: Cuando se prueban aseveraciones acerca de una proporción poblacional, el método tradicional y el método del valor P son equivalentes en el sentido de que siempre producen los mismos resultados, aunque 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 21). Si se obtienen conclusiones diferentes, comprenda que los métodos tradicional y del valor P emplean una desviación estándar *exacta*, con base en el supuesto de que la proporción poblacional contiene el valor dado en la hipótesis nula. Sin embargo, el intervalo de confianza se construye utilizando una desviación estándar basada en un valor *estimado* de la proporción poblacional. Si se desea estimar una proporción poblacional, hágalo construyendo un intervalo de confianza, pero si desea probar una hipótesis utilice el método del valor P o el método tradicional.

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 $\hat{p}$**
10% de los automóviles deportivos observados son rojos.	$\hat{p}$ está dada directamente: $\hat{p} = 0.10$
96 hogares encuestados tienen televisión por cable y 54 no la tienen.	$\hat{p}$ debe calcularse utilizando $\hat{p} = x/n$. $\hat{p} = \frac{x}{n} = \frac{96}{96 + 54} = 0.64$

Cuidado: Cuando la representación visual de $\hat{p}$ de una calculadora o computadora resulta con muchos decimales, utilice todos estos decimales al evaluar el estadístico de prueba z. Llegan a generarse grandes errores al redondear demasiado a $\hat{p}$.

EJEMPLO Experimentos genéticos de Mendel Cuando Gregorio Mendel realizó sus famosos experimentos de hibridación con chícharos, uno de esos experimentos produjo vástagos que consistieron en 428 chícharos con vainas verdes y 152 chícharos con vainas amarillas. Según la teoría de Mendel, $1/4$ de los chícharos vástagos 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 chícharos con vainas amarillas es igual a $1/4$.

SOLUCIÓN Una vez que se verificó que los supuestos se satisfacen, iniciamos con el método del valor P, resumido en la figura 7-9 de la sección 7-2. Observe que $n = 428 + 152 = 580$, $\hat{p} = 152/580 = 0.262$ y, para propósitos de la prueba, suponemos que $p = 0.25$.

Paso 1: La aseveración original dice que la proporción de chícharos 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: $p = 0.25$ (hipótesis nula y aseveración original)

H_1: $p \neq 0.25$ (hipótesis alternativa)

Paso 4: El nivel de significancia es $\alpha = 0.05$.

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 (siempre y cuando los supuestos requeridos se satisfagan). (Los requisitos $np \geq 5$ y $nq \geq 5$ se satisfacen, con $n = 580$, $p = 0.25$ y $q = 0.75$.)

Paso 6: El estadístico de prueba de $z = 0.67$ se calcula de la siguiente manera:

$$z = \frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} = \frac{0.262 - 0.25}{\sqrt{\dfrac{(0.25)(0.75)}{580}}} = 0.67$$

Remítase a la figura 7-6, para el procedimiento del cálculo del valor P. La figura 7-6 indica que para esta prueba de dos colas, con el estadístico de prueba localizado a la derecha del centro (debido a 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 un 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.

continúa

Encuestas y psicólogos

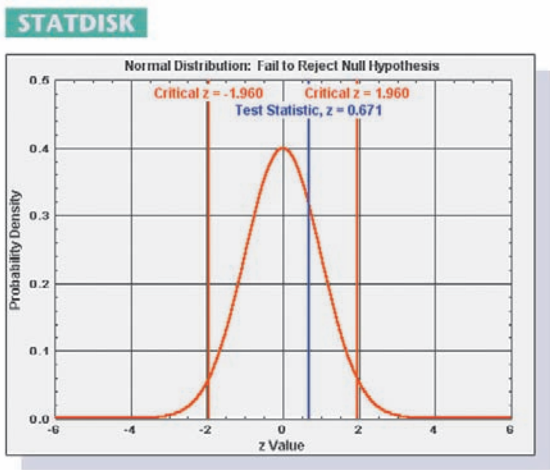
Los resultados de encuestas pueden verse seriamente afectados por la redacción de las preguntas. Las distintas personas interpretan de manera diferente una frase como "durante los últimos años". Durante los últimos años (en realidad desde 1980), los investigadores de encuestas y los psicólogos han trabajado juntos para mejorar las encuestas, disminuyendo los sesgos e incrementando la precisión. En un caso, los psicólogos estudiaron el hallazgo de que de un 10% a un 15% de los encuestados afirmaron haber votado en la última elección, cuando en realidad no lo hicieron. Ellos consideraron teorías de problemas de memoria, el deseo de ser considerado responsable y la tendencia de quienes generalmente votan para decir que votaron en la elección más reciente, aun cuando no lo hayan hecho. Se encontró que sólo esta última teoría era en realidad parte del problema.

Prueba de la terapia de contacto

Cuando tenía nueve años, Emily Rosa participó en una feria escolar de ciencias con un proyecto diseñado para probar la terapia de contacto. En lugar de tocar realmente a los sujetos, el terapeuta de contacto mueve sus manos a unas cuantas pulgadas de distancia del cuerpo del sujeto, de modo que pueda incrementar el campo humano de energía. Emily Rosa probó a 21 terapeutas de contacto, sentándose de un lado de un escudo de cartón, mientras los terapeutas colocaban sus manos a través del escudo de cartón. Emily colocó su mano por encima de una de las manos de un terapeuta (seleccionada con el lanzamiento de una moneda), y después el terapeuta intentó identificar la mano seleccionada sin ver las manos de Emily. Se esperaría un 50% de éxitos con adivinaciones al azar, pero los terapeutas de contacto sólo fueron exitosos el 44% del tiempo. Emily Rosa se convirtió en la autora más joven del *Journal of the American Medical Association* cuando su artículo se publicó: "A Close Look at Therapeutic Touch", de L. Rosa, E. Rosa, L. Sarner y S. Barrett, vol. 279, núm. 1005.

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 chícharos 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 chícharos vástagos tienen vainas amarillas.

Método tradicional: Si fuésemos a repetir 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 representación visual 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 chícharos vástagos tienen vainas amarillas.



Método del intervalo de confianza: Si repitiésemos 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 chícharos vástagos 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 5-6, junto con una corrección por continuidad, pero cuando se prueban aseveraciones sobre una proporción poblacional, hacemos dos modificaciones. Primero, no utilizamos la corrección por continuidad porque su efecto suele ser muy pequeño para las muestras grandes que estamos considerando. Además, en lugar 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, sencillamente, la misma puntuación estándar (de la sección 2-5) de $z = (x - \mu)/\sigma$, pero modificado para la notación binomial.

Utilizando la tecnología

STATDISK Seleccione **Analysis, Hypothesis Testing, Proportion-One Sample,** después proceda a introducir los datos en el cuadro de diálogo.

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, después haga clic en **Opciones** y proceda a introducir los datos en el cuadro de diálogo.

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 seleccione **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 icono del lápiz en "Num Successes" e introduzca A1. Haga clic en el icono 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.

TI-83 Plus Presione **STAT**, seleccione **TEST** y luego seleccione **1-PropZTest**. Introduzca el valor aseverado de la proporción de población para p0, luego introduzca los valores de x y n y después seleccione los tipos de pruebas. Resalte **Calculate** y luego presione la tecla **ENTER.**

7-3 Destrezas y conceptos básicos

1. Experimentos de hibridación de Mendel En uno de los famosos experimentos de Mendel sobre la hibridación se obtuvieron 8023 chícharos vástagos, de los cuales el 24% presentaba 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 los chícharos con flores verdes se presentan en una proporción del 25%.
   a. ¿Cuál es el estadístico de prueba?
   b. ¿Cuáles son los valores críticos?
   c. ¿Cuál es el valor *P*?
   d. ¿Cuál es la conclusión?
   e. ¿Se podría utilizar una hipótesis para "probar" que el porcentaje de chícharos con flores verdes es del 25%, como se aseveró?

2. Encuesta sobre bebidas alcohólicas En una encuesta de Gallup se preguntó a 1087 adultos seleccionados al azar: "¿Consume en ocasiones bebidas alcohólicas como licor, vino o cerveza, o es completamente abstemio?". El 62% de los sujetos afirmaron consumir bebidas alcohólicas. Considere una prueba de hipótesis que utiliza un nivel de significancia de 0.05 para probar la aseveración de que la mayoría (más del 50%) de los adultos consumen bebidas alcohólicas.
   a. ¿Cuál es el estadístico de prueba?
   b. ¿Cuál es el valor crítico?

continúa

c. ¿Cuál es el valor *P*?
d. ¿Cuál es la conclusión?
e. Con base en los resultados anteriores, ¿podemos concluir que el 62% es significativamente mayor que el 50% para todo este tipo de pruebas de hipótesis? ¿Por qué?

Prueba de aseveraciones sobre proporciones. *En los ejercicios 3 a 20, pruebe la aseveración dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor P o valor(es) crítico(s), la conclusión sobre 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.*

3. Encuesta de la revista Glamour La revista *Glamour* financió una encuesta de 2500 novias por casarse y encontró que el 60% de ellas gastaron menos de $750 en su traje de novia. Use 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 traje de novia. ¿De qué manera se verían afectados los resultados si supiéramos que los datos se obtuvieron de lectores de la revista que decidieron responder la encuesta a través de una página de Internet?

4. Delitos federales por drogas En un año reciente, de los 109,857 arrestos por delitos federales en Estados Unidos, el 29.1% fueron delitos por drogas (según datos de su Departamento de Justicia). Utilice un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de delitos por drogas es igual al 30%. ¿Cómo podría explicarse el resultado, dado que el 29.1% parece acercarse mucho al 30%?

5. Porcentaje de usuarios de correo electrónico La tecnología está cambiando de forma drástica nuestras comunicaciones. En 1997, una encuesta de 880 hogares estadounidenses reveló que 149 de ellos emplean el correo electrónico (de acuerdo con 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 actualmente? ¿Por qué?

6. Porcentaje de usuarios del teléfono Una encuesta reciente de 4276 hogares seleccionados al azar, reveló que 4019 de ellos tenían teléfonos (según datos del Census Bureau de Estados Unidos). Use estos resultados muestrales para probar la aseveración de que el porcentaje de hogares ahora es mayor que la tasa del 35% que se encontró en 1920. Utilice un nivel de significancia de 0.01. La tasa actual de 4019/4276 (o 94%) parece ser significativamente mayor que la tasa del 35% de 1920, pero ¿existe evidencia suficiente para sustentar dicha aseveración?

7. Legislación de la cámara vigilante El problema del capítulo incluyó esta pregunta: "¿Existirá suficiente evidencia muestral que sustente la aseveración de que la proporción de todos los adultos de Minnesota que se oponen a la ley de la cámara vigilante es mayor que 0.5? Utilice un nivel de significancia de 0.10 para probar la aseveración de que la proporción es mayor que 0.5. La evidencia muestral consiste en $n = 829$ adultos de Minnesota, seleccionados al azar, con un 51% que se opone a la ley de la cámara vigilante. Como la muestra incluye únicamente a ciudadanos de Minnesota, ¿se aplica la conclusión a todos los estadounidenses adultos?

8. Encuesta sobre la clonación En una encuesta de Gallup de 1012 adultos, seleccionados al azar, el 9% opinó que debería permitirse la clonación humana. Utilice un nivel de significancia de 0.05 para probar la aseveración de que menos del 10% de todos los adultos opinan que debe permitirse la clonación humana. Entonces, ¿un periódico publicaría un encabezado que afirme que "menos del 10% de todos los adultos se oponen a la clonación humana"?

9. 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 (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 precio, 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, ¿parecería que los verificadores de precios ayudan a los consumidores a evitar los sobreprecios?

10. 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 consume drogas?

11. Porcentaje de “strikes” marcados por árbitros En un año reciente, algunos jugadores profesionales de béisbol se quejaron de que los árbitros estaban marcando más *strikes* que el porcentaje promedio del 61.0% del año anterior. En cierto momento de esta temporada, el árbitro Dan Morrison marcó *strike* en 2231 de 3581 lanzamientos (según datos de *US Today*). Utilice un nivel de significancia de 0.05 para probar la aseveración de que el porcentaje de *strikes* es mayor que el 61.0 por ciento.

12. Prueba del Lipitor para reducción del colesterol En una prueba clínica del fármaco Lipitor (nombre genérico, atorvastatín), 863 pacientes se trataron con dosis de 10 mg de atorvastatín, y 19 de ellos experimentaron síntomas de gripe (según datos de Parke-Davis). Utilice un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de pacientes tratados con síntomas de gripe es mayor que el porcentaje del 1.9% de pacientes que no recibieron el tratamiento. ¿Parecería que los síntomas de gripe son una reacción adversa del tratamiento?

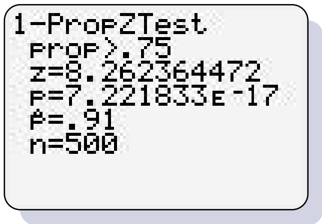
13. Teléfonos celulares y cáncer En un estudio de 420,095 usuarios daneses de teléfonos celulares, 135 sujetos desarrollaron cáncer cerebral o del sistema nervioso (según datos del *Journal of the National Cancer Institute*). Pruebe la aseveración, antes popular, de que estos tipos de cáncer se ven afectados 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 del 0.0340% de las personas que no utilizan teléfonos celulares. Ya que este tema es de gran importancia, utilice un nivel de significancia del 0.005. ¿Deberían preocuparse los usuarios de teléfonos celulares del cáncer cerebral o del sistema nervioso?

14. Prueba de la eficacia de los parches de nicotina En un estudio realizado a fumadores que intentaron dejar el cigarrillo con terapia de parches de nicotina, un año después del tratamiento 39 de ellos continuaban fumando y 32 ya no fumaban (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 del tratamiento. ¿Sugieren estos resultados que la terapia de parches de nicotina es ineficaz?

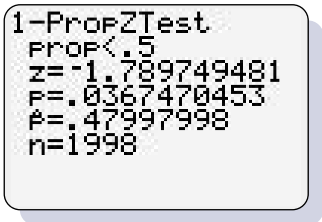
15. Tabaquismo y educación universitaria Una encuesta reveló que de 785 sujetos seleccionados aleatoriamente y que completaron cuatro años de estudios universitarios, 144 fuman y 641 no fuman (según datos de la American Medical Association). Utilice un nivel de significancia de 0.01 para probar la aseveración de que el porcentaje de fumadores que tienen cuatro años de estudios universitarios es menor que el porcentaje del 27% de la población general. ¿Por qué los graduados universitarios que fuman tienen una tasa menor del resto?

16. Audiencia televisiva Una muestra aleatoria de hogares con televisores encendidos, revela que 1024 de ellos estaban sintonizando *60 minutos*, mientras que 3836 estaban sintonizando algún otro programa. Utilice un nivel de significancia de 0.025 para probar la aseveración de un ejecutivo de la CBS de que “*60 minutos* tiene una audiencia mayor que 20”, lo que significa que más del 20% de los televisores en uso sintonizan *60 minutos*. Si usted fuese un anunciante comercial y tratara de negociar costos más bajos, ¿cuál sería su argumento?

TI-83 Plus



TI-83 Plus



17. Interpretación de la representación visual de una calculadora La Federal Aviation Administration financiará investigaciones sobre la desorientación espacial de los pilotos, si existe suficiente evidencia muestral (con un nivel de significancia de 0.01) para concluir que entre los accidentes aéreos relacionados con este tipo de desorientación, más de tres cuartas partes provocan muertes. Un estudio de 500 accidentes aéreos relacionados con desorientación espacial del piloto, reveló que el 91% de estos accidentes provocaron muertes (según datos del Departamento del Transporte de Estados Unidos). Se obtiene la representación visual de la calculadora TI-83 Plus que se presenta al margen. Interprétela. Con base en estos resultados muestrales, ¿se aprobará el financiamiento?

18. Interpretación de la representación visual de una calculadora Un ejecutivo de televisión asevera que "menos de la mitad de todos los adultos están preocupados por la violencia que se muestra en la televisión". Datos muestrales de una encuesta de Roper mostraron que el 48% de 1998 adultos encuestados indicaron preocupación por la violencia televisiva. Se obtiene la representación visual de la calculadora TI-83 Plus que se incluye al margen. ¿Sustentan los datos muestrales la aseveración del ejecutivo?

19. Uso de datos de M&M Remítase al conjunto de datos 19 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 10% de sus dulces M&M son azules.

20. Consumo de alcohol y tabaco en películas infantiles de dibujos animados Utilice los resultados listados en el conjunto de datos 7 del Apéndice B para probar la aseveración de que la mayoría de las películas infantiles de dibujos animados muestran consumo de alcohol o tabaco (o ambos). Utilice un nivel de significancia de 0.05.

7-3 Más allá de lo básico

21. 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 aleatoriamente, 119 son ceros. Si los números se seleccionan aleatoriamente, la proporción de ceros debe ser de 0.1.

a. 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.

b. 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.

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 el método del valor *P* y con el método del intervalo de confianza. ¿Conducen todos a la misma conclusión?

22. Uso de la corrección por continuidad Repita el ejercicio 20, pero incluya la corrección por continuidad que se estudió en la sección 5-6. ¿De qué manera se ven afectados los resultados al incluir la corrección por continuidad?

23. Prueba de aseveraciones En el artículo de *USA Today* "Power Lines Not a Cancer Risk for Kids", la primera oración dice que "reportes médicos afirman, en el estudio más extenso realizado sobre uno de los temas más polémicos, que los niños que viven cerca de líneas eléctricas de alto voltaje no parecen estar más expuestos a padecer leucemia que el resto de los niños". Represente el porcentaje de leucemia en niños que no viven cerca de líneas eléctricas de alto voltaje con la constante *c*, redacte la aseveración de manera simbólica y después identifique las hipótesis nula y alternativa sugeridas por esta afirmación. Como rechazamos o no rechazamos la hipótesis nula, ¿qué posibles conclusiones se obtienen respecto de la aseveración original? ¿Los datos muestrales sustentan la aseveración de que los niños que viven cerca de líneas eléctricas de alto voltaje no son más propensos a padecer leucemia, que el resto de los niños?

24. Método alternativo para probar una aseveración acerca de p En un estudio sobre percepción se prueba a 80 hombres y resulta que 7 de ellos son daltónicos (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 daltonismo es mayor que el porcentaje del 0.25% de las mujeres.

a. ¿Por qué no podemos utilizar los métodos de esta sección?

b. Suponiendo que el porcentaje de daltonismo 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 tenga este tipo de ceguera al color. Describa el método utilizado para calcular esa probabilidad.

c. Con base en los resultados del inciso *b*, ¿qué concluye?

25. Manejo de no éxitos En una muestra aleatoria simple de 50 dulces M&M, 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. ¿Son útiles los métodos de esta sección? Si es así, pruebe la aseveración; si no, explique por qué no.

26. Estadísticos confusos Chemco, distribuidor de contenedores de desperdicios químicos, descubre que el 3% de una muestra de 500 unidades presenta defectos. Siendo una persona básicamente deshonesta, el gerente de producción de Chemco desea plantear la aseveración de que el porcentaje de unidades defectuosas no supera un porcentaje especificado, y no quiere que esta aseveración se rechace al nivel de significancia de 0.05, si se utilizan los datos muestrales. ¿Cuál es el porcentaje de defectos *más bajo* que puede aseverar en estas condiciones?

27. Aseveración falsa Un investigador aseveró que al tratar 20 ratones, el porcentaje de éxitos fue igual al 47%. ¿Cuál es la base para rechazar esta aseveración?

28. Probabilidad del error tipo II 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. Remítase al ejercicio 20. Suponiendo que el valor verdadero de p es 0.45, calcule β, la probabilidad de un error tipo II. Utilice el siguiente procedimiento. [*Sugerencia:* En el paso 3 utilice los valores $p = 0.45$ y $pq/n = (0.45)(0.55)/50$].

Paso 1: Calcule el (los) valor(es) del estadístico muestral $\hat{p}$ que corresponde al (los) valor(es) crítico(s). En

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

sustituya el (los) valor(es) crítico(s) para z, sustituya los valores de p (de la hipótesis nula) y q, después resuelva para $\hat{p}$.

Paso 2: Dado un valor particular p alternativo al valor dado en la hipótesis nula, dibuje la curva normal con este nuevo valor p alternativo al centro. También grafique el (los) valor(es) de $\hat{p}$ calculado(s) en el paso 1.

Paso 3: Remítase a la gráfica en el paso 2 y calcule el área de la nueva región crítica limitada por el (los) valor(es) de $\hat{p}$ obtenido(s) en el paso 1. (Asegúrese de emplear la desviación estándar basada en el nuevo valor p). Ésta es la probabilidad de rechazar la hipótesis nula, en tanto que el nuevo valor p sea correcto.

Paso 4: El valor β es 1 menos el área del paso 3. Ésta es la probabilidad de no rechazar la hipótesis nula, en tanto que el nuevo valor p sea correcto.

Los pasos anteriores le permitirán calcular la probabilidad de no rechazar H_0 cuando es falsa. Usted está determinando el área bajo la curva que excluye la región crítica en la que rechaza H_0; esta área corresponde a no rechazar una H_0 falsa, y sabemos que H_0 es falsa debido a que estamos empleando un valor alternativo que se supone es la proporción poblacional correcta.

7-4 Prueba de una aseveración respecto de una media: σ conocida

En esta sección consideramos métodos para probar aseveraciones hechas acerca de una media poblacional μ y suponemos que se conoce la desviación estándar poblacional σ. En muy raras ocasiones las circunstancias nos permitirían conocer σ sin conocer μ, pero la sección 7-5 trata con casos en los que no conocemos σ. Aun cuando esta sección incluye casos menos realistas que los de la sección 7-5, es importante porque describe el mismo método general empleado en la siguiente sección. Además, existen casos en los que se desconoce el valor específico de σ, aunque se emplee alguna información acerca de ella. El ejemplo que se presenta en esta sección incluye el supuesto poco realista de que sabemos que σ es igual a 0.62°F. El estadístico de prueba en ese ejemplo es $z = -6.64$, que conduce al rechazo de la creencia común de que la temperatura corporal media es igual a 98.6°F. Si analizamos la variación de las temperaturas corporales, queda claro que σ no puede ser tan alta como 2°F en las temperaturas corporales; además, emplear $\sigma =$ 2°F generaría un estadístico de prueba $z = -2.05$, que nuevamente conduce al rechazo de la aseveración de que $\mu = 98.6$°F. Puesto que σ debe ser menor que 2°F para las temperaturas corporales, el estadístico de prueba debe ser al menos tan extremo como $z = -2.05$. (Consulte el ejercicio 17). Esto demuestra que, aun cuando no conozcamos un valor específico de σ, existen casos en los que el uso de valores muy conservadores de σ nos permitirá sacar algunas conclusiones importantes.

Los supuestos, 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 conocida)

Supuestos

1. La muestra es aleatoria simple. (Recuerde este punto muy importante, planteado en el capítulo 1: *Los datos reunidos con descuido serían tan inútiles que ninguna cantidad de tortura estadística pueda salvarlos).*
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 7-6.

Valores críticos: Utilice la distribución normal estándar (tabla A-2).

Antes de iniciar el procedimiento de prueba de hipótesis, debemos explorar primero el conjunto de datos. Con los métodos descritos en el capítulo 2, 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 dato distante. Debemos verificar que los supuestos requeridos se satisfagan. Para la muestra de 106 temperaturas corporales del siguiente ejemplo, un histograma indica que los datos muestrales parecen provenir de una población que se distribuye normalmente. Además, no hay datos distantes. El aspecto de la normalidad no es demasiado importante para este ejemplo debido a que la muestra es muy grande, pero sí es importante saber que no existen datos distantes que afectarían de forma drástica los resultados.

EJEMPLO Método del valor *P* El conjunto de datos 4 del Apéndice B incluye la lista de una muestra de 106 temperaturas corporales, con una media de 98.20°F. Suponga que la muestra es aleatoria simple y que se sabe que la desviación estándar poblacional σ es 0.62°F. Utilice un nivel de significancia de 0.05 para probar la creencia común de que la temperatura corporal media de adultos sanos es igual a 98.20°F. Aplique el método del valor *P*, siguiendo el procedimiento descrito en la figura 7-9.

SOLUCIÓN Remítase a la figura 7-9 y siga estos pasos:

Paso 1: La aseveración de que la media es igual a 98.6 se expresa en forma simbólica como $\mu = 98.6$.

Paso 2: La alternativa (en forma simbólica) a la aseveración original es $\mu \neq 98.6$.

Paso 3: Puesto que la afirmación $\mu \neq 98.6$ no contiene la condición de igualdad, se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de que $\mu = 98.6$.

H_0: $\mu = 98.6$ (aseveración original)
H_1: $\mu \neq 98.6$

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

continúa

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.

$\bar{x} = 98.20$. Puesto que se supone que conocemos σ (0.62) y $n > 30$, 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:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{98.20 - 98.6}{\frac{0.62}{\sqrt{106}}} = -6.64$$

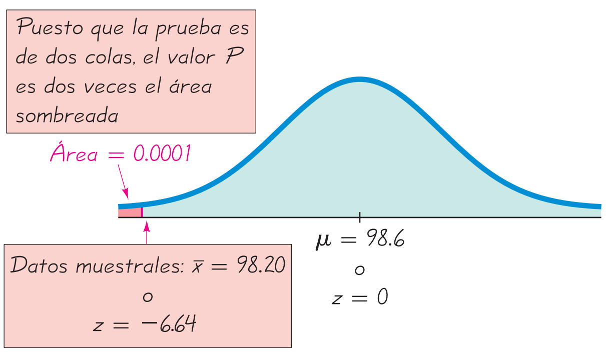
Utilizando el estadístico de prueba $z = -6.64$, ahora procedemos al cálculo del valor *P*. Observe el diagrama de flujo de la figura 7-9 que resume el procedimiento para el cálculo de los valores *P*. Se trata de una prueba de dos colas y el estadístico de prueba se encuentra a la izquierda del centro (puesto que $z = -6.64$ es menor que $z = 0$), de modo que el valor *P* es *dos veces* el área a la izquierda de $z = -6.64$. Ahora nos remitimos a la tabla A-2 para encontrar que el área a la izquierda de $z = -6.64$ es 0.0001, de manera que el valor *P* es $2(0.0001) = 0.0002$. (Resultados más precisos muestran que el valor *P* en realidad es mucho menor que 0.0002). Consulte la figura 7-11.

Paso 7: Puesto que el valor *P* de 0.0002 es menor que el nivel de significancia de $\alpha = 0.05$, rechazamos la hipótesis nula.

INTERPRETACIÓN El valor *P* de 0.0002 es la probabilidad de obtener una media muestral tan extrema como 98.20°F (con un tamaño de muestra de $n = 106$) por el azar, suponiendo que $\mu = 98.6$°F y $\sigma = 0.62$°F. Ya que esta probabilidad es muy pequeña, rechazamos al azar como una posible explicación y concluimos que el supuesto de $\mu = 98.6$°F debe ser incorrecto. Nos remitimos a la figura 7-7, en la sección 7-2, para establecer de forma correcta la conclusión final. Estamos rechazando la hipótesis nula, que es la aseveración original, de manera que concluimos que existe evidencia suficiente para justificar el rechazo de la aseveración de que la temperatura corporal media de adultos sanos es 98.6°F. Existe evidencia suficiente para concluir que la temperatura corporal media de todos los adultos sanos es diferente de 98.6°F.

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 los valores críticos de $z = -1.96$ y $z = 1.96$, en lugar de calcular el valor *P*. Nuevamente rechazaríamos la hipótesis nula, ya que el estadístico de prueba $z = -6.64$ caería en la región crítica. La conclusión final sería la misma.

FIGURA 7-11 Método de prueba del valor *P*
H_0: $\mu = 98.6$



Método del intervalo de confianza Ahora podemos emplear un intervalo de confianza para probar una aseveración acerca de μ cuando conocemos σ. Para una prueba de hipótesis de dos colas, con un nivel de significancia de 0.05, construimos un intervalo de confianza del 95%. Si utilizamos los datos muestrales del ejemplo anterior ($n = 106$ y $\bar{x} = 98.20$), y suponemos que conocemos $\sigma = 0.62$, podemos probar la aseveración de que $\mu = 98.6$ aplicando los métodos de la sección 6-3 para construir este intervalo de confianza del 95%: $98.08 < \mu < 98.32$. Puesto que el valor aseverado de $\mu = 98.6$ no está incluido dentro del intervalo de confianza, rechazamos esta aseveración. Tenemos una confianza del 95% de que el valor verdadero de μ se encuentra dentro de los límites de 98.08 y 98.32, por lo que parece que 98.6 no puede ser el valor verdadero de μ.

En la sección 7-3 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 dicha diferencia y los tres métodos son equivalentes.

Cuidado: Cuando se prueba una aseveración sobre μ empleando un intervalo de confianza, asegúrese de utilizar el nivel de confianza apropiado para un nivel de significancia específico. En el caso de pruebas de dos colas, es fácil ver que un nivel de significancia de 0.05 corresponde a un nivel de confianza del 95%, pero con las pruebas de una cola se vuelve confuso. Para probar la aseveración de que $\mu < 98.6$, con un nivel de significancia de 0.05, construya un intervalo de confianza del 90%. Para probar la aseveración de que $\mu > 98.6$, con un nivel de significancia de 0.01, construya un intervalo de confianza del 98%.

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 los fundamentos que sustentan la prueba de hipótesis. La clave para comprenderlos radica en la regla del evento poco común 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, hacemos una suposición (hipótesis nula) de igualdad. Después comparamos el supuesto y los resultados muestrales para llegar a una de las siguientes conclusiones:

- Si los resultados muestrales (o resultados más extremos) ocurren 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, con la conclusión de que el supuesto no es verdadero, por lo que rechazamos el supuesto.

Método alternativo (no utilizado en este libro) Un método alternativo, que no se aplica este libro, implica el uso de s, estimado de σ desconocida, siempre y cuando la muestra sea grande ($n > 30$). Es decir, si el tamaño de muestra n es mayor que 30, sustituya la σ desconocida con la desviación estándar muestral s, después utilice los métodos de esta sección, procediendo como si conociera σ. La sección 7-5 lista las razones por las que este método alternativo no se emplea en el presente libro.

Utilizando 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 2-4. 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 trabaja únicamente con la lista de los datos originales. (Para saber cómo superar esta restricción, véase *Minitab Student Laboratory Manual and Workbook*, que complementa a este libro). Primero introduzca los datos en la columna C1, después seleccione **Stat, Basic Statistics** y **1-Sample z** del menú e introduzca los datos requeridos. El cuadro denominado "alternative" se utiliza para seleccionar la forma de la hipótesis alternativa y puede incluir **not equal, less than** o **grater than.**

Excel La función ZTEST creada de Excel es extremadamente confusa, debido a 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 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 icono del lápiz e introduzca el rango de valores de datos, tal como A1:A106, si tiene 106 valores listados en la columna A. Haga clic en **OK**. Siga los cuatro pasos del cuadro de diálogo. Después de hacer clic en **Compute** del paso cuatro, obtendrá el valor P, el estadístico de prueba y la conclusión.

TI-83 Plus Si utiliza la calculadora TI-83 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 proporcionar las entradas indicadas en la representación visual de la ventana. Los primeros tres elementos de los resultados de la TI-83 Plus incluirán la hipótesis alternativa, el estadístico de prueba y el valor P.

7-4 Destrezas y conceptos básicos

Verificación de supuestos. *En los ejercicios 1 a 4, 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.*

1. El tamaño de muestra es $n = 25$, $\sigma = 6.44$ y la población original se distribuye de manera normal.

2. El tamaño de muestra es $n = 7$, se desconoce σ y la población original se distribuye de manera normal.

3. El tamaño de muestra es $n = 11$, se desconoce σ y la población original se distribuye de manera normal.

4. El tamaño de muestra es $n = 47$, $\sigma = 12.6$ y la población original no se distribuye de manera normal.

Cálculo de los componentes de prueba. *En los ejercicios 5 a 8, calcule el estadístico de prueba, el valor P, el (los) valor(es) crítico(s) y establezca la conclusión final.*

5. Aseveración: La puntuación media del CI de profesores de estadística es mayor que 118. Datos muestrales: $n = 50$, $\bar{x} = 120$. Suponga que $\sigma = 12$ y que el nivel de significancia es $\alpha = 0.05$.

6. Aseveración: La temperatura corporal media de adultos sanos es menor que 98.6°F. Datos muestrales: $n = 106$, $\bar{x} = 98.20$°F. Suponga que $\sigma = 0.62$ y que el nivel de significancia es $\alpha = 0.01$.

7. Aseveración: El tiempo medio que transcurre para que los hombres vuelvan a usar el control remoto del televisor, durante los comerciales, es igual a 5.00 segundos.

continúa

Datos muestrales: $n = 80$, $\bar{x} = 5.25$ s. Suponga que $\sigma = 2.50$ s y que el nivel de significancia es $\alpha = 0.01$.

8. Aseveración: El salario medio inicial de estudiantes universitarios graduados que han tomado un curso de estadística es igual a \$46,000.

Datos muestrales: $n = 65$, $\bar{x} = \$45{,}678$. Suponga que $\sigma = \$9900$ y que el nivel de significancia es $\alpha = 0.05$.

Prueba de hipótesis. *En los ejercicios 9 a 12 pruebe la aseveración dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor* P *o el (los) valor(es) crítico(s), la conclusión sobre 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.*

9. Temperaturas de los Everglades Para verificar la salud ecológica de los Everglades de Florida, se registran varias mediciones en momentos diferentes. Las temperaturas más bajas se registran en la estación Garfield Bight, y se obtiene la media de 30.4°C para las 61 temperaturas registradas. Suponiendo que $\sigma = 1.7$°C, pruebe la aseveración de que la media poblacional es mayor que 30.0°C. Utilice un nivel de significancia de 0.05.

10. Pesos de osos La salud de una población de osos en el Yellowstone National Park se verifica por medio de mediciones periódicas, tomadas de osos anestesiados. Una muestra de 54 osos tiene un peso medio de 182.9 libras. Suponiendo que sabemos que σ es igual a 121.8 libras, utilice un nivel de significancia de 0.10 para probar la aseveración de que la media poblacional de todos los pesos de osos es menor que 200 libras.

11. Niveles de cotinina de fumadores Cuando las personas fuman, la nicotina que absorben se convierte en cotinina, que es susceptible de medición. Una muestra de 40 fumadores tiene un nivel medio de cotinina de 172.5. Suponiendo que sabemos que σ es igual a 119.5, utilice un nivel de significancia de 0.01 para probar la aseveración de que el nivel medio de cotinina de todos los fumadores es igual a 200.0.

12. Circunferencia de cabezas Se obtiene una muestra aleatoria de 100 bebés y se descubre que la circunferencia media de las cabezas es de 40.6 cm. Suponiendo que sabemos que la desviación estándar poblacional es de 1.6 cm, utilice un nivel de significancia de 0.05 para probar la aseveración de que la circunferencia media de las cabezas de todos los bebés de dos meses de edad es igual a 40.0 cm.

Interpretación de resultados de computadora y calculadora. *En los ejercicios 13 a 16, utilice los resultados de una computadora o una calculadora para sacar una conclusión.*

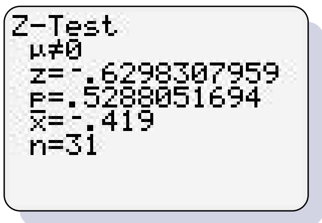
13. Peso medio de dulces M&M Un paquete de dulces M&M dice contener 1361 g y tiene 1498 dulces, de modo que el peso medio de los dulces individuales debe ser 1361/1498, o 0.9085 g. La Mars Company desea producir dulces M&M con pesos que no engañen a los consumidores, pero tampoco desean desperdiciar dinero de producción en una media significativamente mayor de lo necesario. Para probar la aseveración de que $\mu \neq 0.9085$ g, se selecciona al azar una muestra de 100 M&M. (Consulte el conjunto de datos 19 del Apéndice B). Cuando se utiliza Minitab con los 100 pesos, el despliegue de los resultados es igual al que se muestra a continuación (suponiendo que sabemos que σ es igual a 0.03691 g). Interprete estos resultados. ¿Se está engañando a los consumidores? ¿Se está desperdiciando dinero al hacer M&M más pesados de lo necesario?

```
Test of mu = 0.9085 vs mu not = 0.9085
The assumed sigma = 0.03691

Variable       N       Mean      StDev    SE Mean
M&M          100    0.91470    0.03691    0.00369

Variable          95.0% CI           Z          P
M&M        ( 0.90747, 0.92193)    1.68      0.093
```

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 precisa. Cuando se seleccionan aleatoriamente 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 cincos. Se utilizan los últimos dígitos de las longitudes (en pies) de los "home runs" anotados por Barry Bonds en el 2001 para probar la aseveración de que provienen de una población con una media de 4.50 (según datos de *USA Today*). Cuando se utiliza Minitab para probar esa aseveración, resulta la representación visual mostrada a continuación. Con un nivel de significancia de 0.05, interprete los resultados de Minitab. ¿Parece que las distancias se midieron con precisión?

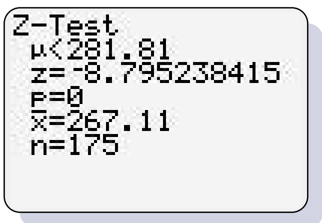
```
Test of mu = 4.5 vs mu not = 4.5
The assumed sigma = 2.87

Variable       N      Mean     StDev    SE Mean
BONDS         73     1.753     2.650      0.336

Variable         95.0% CI            Z        P
BONDS       ( 1.095,   2.412)    −8.18    0.000
```

TI-83 Plus



15. Diferencias entre las altas temperaturas pronosticadas y las reales El conjunto de datos 10 del Apéndice B incluye las altas temperaturas reales y las altas temperaturas pronosticadas para tres días. Una forma para investigar la precisión de las temperaturas pronosticadas es el cálculo de las diferencias entre las temperaturas reales y las pronosticadas. Las 31 diferencias (altas reales − altas pronosticadas para tres días) tienen una media de −0.419°. Suponiendo que $\sigma = 3.704°$, obtenemos los resultados mostrados en la siguiente representación de la pantalla de la calculadora TI-83 Plus. Interprete los resultados. ¿Parece la diferencia media acercarse a 0°? ¿O parece existir una diferencia significativa? ¿Qué sugieren estos resultados respecto a la precisión de las altas temperaturas pronosticadas para tres días?

TI-83 Plus



16. ¿Son más débiles las latas de aluminio delgado? El conjunto de datos 20 del Apéndice B incluye las cargas axiales medidas (en libras) de 175 latas de refresco que utilizan aluminio de 0.0109 pulgadas de grosor. Antes de obtener estos resultados muestrales, las latas estándar tenían un grosor de 0.0111 pulgadas, y la carga axial media era de 281.81 libras. Cuando se utilizaron las cargas axiales de las latas más delgadas en la prueba de la aseveración de que la carga axial media es menor que 281.81 libras, la calculadora TI-83 Plus proporcionó los resultados que aparecen al margen. (Los resultados se basan en el supuesto de que sabemos que σ es de 22.11 libras). Suponga que estamos empleando un nivel de significancia de 0.01 e interprete los resultados. ¿Parece que las latas más delgadas tienen una carga axial media menor que 281.81 libras?

7-4 Más allá de lo básico

17. Prueba de la σ supuesta En el ejemplo incluido en esta sección, rechazamos H_0: $\mu = 98.6$ y sustentamos H_1: $\mu \neq 98.6$, bajo el supuesto de que $\sigma = 0.62$ y los datos muestrales consistentes de $n = 106$ valores con $\bar{x} = 98.20$.

a. ¿Qué aspecto de este ejemplo no es realista?

b. Calcule el valor más grande de σ que da como resultado la misma conclusión planteada al asumir que $\sigma = 0.62$.

c. En tanto que las 106 temperaturas corporales tienen una desviación estándar de 0.62, ¿hay alguna posibilidad razonable de que el verdadero valor de σ sea mayor que el valor calculado en el inciso *b*? ¿Qué implica esto para el supuesto de que $\sigma = 0.62$?

18. Cálculo de la desviación estándar Un artículo de una revista reportó que una hipótesis nula de $\mu = 100$ fue rechazada debido a que el valor P fue menor que 0.01. El tamaño de muestra era de 62 y la media muestral de 103.6. Calcule la desviación estándar más grande posible.

19. Cálculo de la probabilidad de un error tipo II Para una prueba de hipótesis con un nivel de significancia α dado, la probabilidad de un error tipo I es el valor fijo α, pero la probabilidad β de un error tipo II depende del valor particular de μ que se utilice como alternativa a la hipótesis nula. Para pruebas de hipótesis del tipo estudiado en esta sección, podemos calcular β de la siguiente manera:

Paso 1: Calcule el (los) valor(es) de $\bar{x}$ que corresponde(n) al (los) valor(es) crítico(s). En

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

sustituya el (los) valor(es) para z, introduzca los valores de $\mu_{\bar{x}}$, σ y n, y después resuelva para $\bar{x}$.

Paso 2: Se tiene un valor particular de μ que es una alternativa al valor dado en la hipótesis nula. Dibuje la curva normal con este nuevo valor de μ al centro. También grafique el (los) valor(es) calculado(s) en el paso 1.

Paso 3: Remítase a la gráfica del paso 2 y calcule el área de la nueva región crítica limitada por el (los) valor(es) de $\bar{x}$, calculados en el paso 1. Ésta es la probabilidad de rechazar la hipótesis nula, en tanto que el nuevo valor de μ sea correcto y el valor de μ dado en la hipótesis nula sea falso.

Paso 4: El valor de β es 1 menos el área del paso 3. Ésta es la probabilidad de no rechazar la hipótesis nula, en tanto que el nuevo valor de μ sea correcto.

Estos pasos le permiten calcular la probabilidad de no rechazar la hipótesis nula cuando es falsa. Usted está determinando el área bajo la curva que excluye la región crítica en que rechaza H_0; esta área corresponde al no rechazo de H_0 falsa, ya que utilizamos un valor particular de μ que va en contra de H_0. Remítase al ejemplo de las temperaturas corporales presentado en esta sección y calcule β (la probabilidad de un error tipo II) correspondiente a lo siguiente:

a. $\mu = 98.7$
b. $\mu = 98.4$

20. Potencia de una prueba La *potencia* de una prueba, expresada como $1 - \beta$, es la probabilidad de rechazar una hipótesis nula falsa. Suponga que al probar la aseveración de que $\mu < 98.6$, los datos muestrales son $n = 106$ y $\bar{x} = 98.20$. Suponga que $\sigma = 0.62$ y un nivel de significancia de 0.05. Si la prueba de la aseveración $m < 98.6$ tiene una potencia de 0.8, calcule la media μ que se está empleando como alternativa al valor dado en H_0 (véase el ejercicio 19).

7-5 Prueba de una aseveración respecto de una media: σ desconocida

Una de las grandes ventajas de aprender los métodos de prueba de hipótesis, descritos en las secciones anteriores de este capítulo, es que esos mismos métodos se modifican fácilmente para aplicarse en muchas otras circunstancias, tales como las que estudiaremos en esta sección. El principal objetivo de esta sección es el de desarrollar la habilidad de probar aseveraciones hechas sobre medias poblacionales, cuando se desconoce la desviación estándar poblacional σ.

Mejores resultados con clases más pequeñas

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 grupos grandes que oscilaban entre 150 y 200 estudiantes. En un curso de cálculo, los porcentajes de fracaso fueron del 19% en los grupos pequeños, en comparación con un 50% en los grupos grandes. Los porcentajes de calificación A fueron del 24% para los grupos pequeños y del 3% para los grupos grandes. Estos resultados sugieren que los estudiantes se benefician de los grupos reducidos, que permiten una interacción más directa entre los alumnos y los maestros.

La sección 7-4 presentó métodos de prueba de aseveraciones acerca de μ cuando se conoce σ, pero en pocas ocasiones desconocemos el valor de μ y conocemos el valor de σ. Los métodos de esta sección son mucho más prácticos y realistas porque suponen que se desconoce σ, como generalmente sucede. Los supuestos, el estadístico de prueba, el valor P y los valores críticos se resumen de la siguiente manera.

Prueba de aseveraciones acerca de una media poblacional (σ desconocida)

Supuestos

1. La muestra es una muestra 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 aprobar 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 7-6 para los procedimientos del cálculo del valor P).

El requisito de una población con distribución normal no es estricto y generalmente podemos considerar que la población se distribuye normalmente después de utilizar datos muestrales que confirmen que no existen datos distantes y cuando el histograma tiene una forma no muy diferente de una distribución normal. Además, utilizamos el criterio simplificado de $n > 30$ como justificación para tratar la distribución de medias de muestra como una distribución normal, pero el tamaño de muestra mínimo en realidad depende de qué tanto la distribución de la población se aparta de una distribución normal. Como desconocemos el valor de σ, la estimamos con el valor de la desviación estándar muestral s, aunque esto introduce otra fuente de baja confiabilidad, en especial con muestras pequeñas. Para compensar esta baja confiabilidad añadida, calculamos los valores P y los valores críticos empleando una distribución t, en lugar de la distribución normal que se empleó en la sección 7-4, con σ conocida. He aquí 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 (consulte la figura 6-5 en la sección 6-4).
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 al tamaño de la muestra y es mayor que 1 (a diferencia de la distribución normal estándar que tiene $\sigma = 1$).
5. Conforme aumenta el tamaño de muestra 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 el capítulo 12; la técnica *bootstrap* de muestreo se describe en el "Proyecto tecnológico" que está al final del capítulo 6.) Revise las páginas 336 y 337, donde la figura 6-6 y la tabla 6-1 resumen las decisiones a tomarse al elegir entre las distribuciones normal y t de Student. En ellas se observa que cuando se prueban aseveraciones acerca de medias poblacionales, la distribución t de Student se aplica en tales 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 Temperaturas corporales A un estudiante del propedéutico de la carrera de medicina se le pide realizar un proyecto en clase. Intrigado por las temperaturas corporales del conjunto de datos 4 del Apéndice B, planea recolectar su propio conjunto de datos para probar la aseveración de que la temperatura corporal media es menor que 98.6°F, como suele pensarse. Por limitación del tiempo impuesto por otros cursos y al deseo de mantener una vida social que vaya más allá de hablar en sueños, se da cuenta de que tiene tiempo para reunir datos únicamente de 12 personas. Después de planear cuidadosamente un procedimiento para obtener una muestra aleatoria simple de 12 adultos sanos, mide sus temperaturas corporales y obtiene los resultados listados abajo. 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

SOLUCIÓN Antes de llegar a la prueba de hipótesis, primero exploremos los datos muestrales. No se presentan datos distantes y, con base en un histograma y una gráfica cuantilar normal, podemos suponer que los datos provienen de una población con una distribución normal. Utilizamos los datos muestrales para calcular los siguientes estadísticos: $n = 12$, $\bar{x} = 98.39$, $s = 0.535$. La media muestral de $\bar{x} = 98.39$ es menor que 98.6, pero necesitamos determinar si es *significativamente* menor que 98.6. Procedamos con una prueba de hipótesis formal. Emplearemos el método tradicional de prueba de hipótesis resumido en la figura 7-8.

continúa

La pena de muerte como correctivo

Un argumento utilizado comúnmente para sustentar la pena de muerte es que ésta desanima a otros individuos para cometer asesinatos. Jeffrey Grogger, de la Universidad de California, analizó los datos sobre los homicidios diarios en California 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 lo 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". La pena capital 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 modo que tengamos información precisa que nos permita abordar temas como éste.

Paso 1: La aseveración original de que “la temperatura corporal media es menor que 98.6°F” se expresa de manera simbólica como $\mu < 98.6$.

Paso 2: El opuesto de la aseveración original es $\mu \geq 98.6$.

Paso 3: De las dos expresiones simbólicas obtenidas hasta ahora, la expresión $\mu < 98.6$ no contiene igualdad, por lo tanto se convierte en la hipótesis alternativa H_1. La hipótesis nula es el supuesto de que $\mu = 98.6$.

$$H_0: \mu = 98.6$$
$$H_1: \mu < 98.6 \quad \text{(aseveración original)}$$

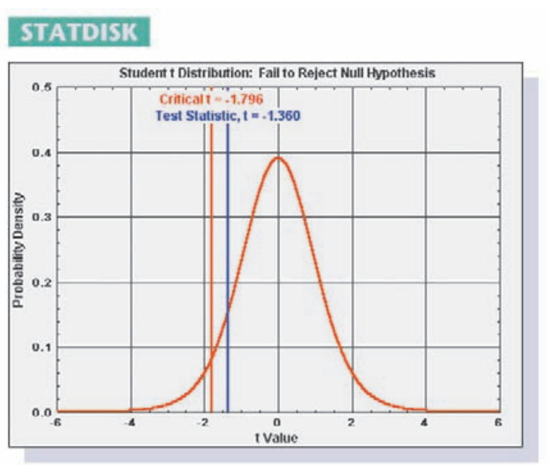
Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: En esta prueba de una aseveración acerca de la media poblacional, el estadístico más relevante es la media muestral. Para seleccionar la distribución correcta, nos remitimos a la figura 6-6 o a la tabla 6-1. Seleccionamos la distribución t de Student por las siguientes condiciones: tenemos una muestra aleatoria simple, desconocemos el valor de σ y los datos muestrales parecen provenir de una población con una distribución normal.

Paso 6: El estadístico de prueba es

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{98.39 - 98.6}{\frac{0.535}{\sqrt{12}}} = -1.360$$

El valor crítico de $t = -1.796$ se calcula consultando la tabla A-3. Primero localice $n - 1 = 11$ grados de libertad en la columna de la izquierda. Como la prueba es de cola izquierda, con $\alpha = 0.05$, remítase a la columna que indica un área de 0.05 en una cola. El estadístico de prueba y el valor crítico se presentan en la siguiente representación de la pantalla de STATDISK.



Paso 7: Puesto que el estadístico de prueba $t = -1.360$ no cae en la región crítica, no rechazamos H_0.

INTERPRETACIÓN (Remítase a la figura 7-7 para saber cómo redactar la conclusión final). No existe evidencia suficiente para sustentar la aseveración de que la muestra proviene de una población con una media menor que 98.6°F. Esto no "prueba" que la media sea 98.6°F. De hecho, μ bien puede ser menor que 98.6°F, pero los dos valores muestrales no proporcionan una evidencia suficientemente fuerte para sustentar esa aseveración. Si utilizáramos las 106 temperaturas corporales incluidas en el conjunto de datos 4 del Apéndice B, encontraríamos que existe evidencia suficiente para sustentar la aseveración de que la temperatura corporal media es menor que 98.6°F, pero los 12 valores muestrales incluidos en este ejemplo no sustentan dicha aseveración.

El valor crítico en el ejemplo anterior fue $t = -1.796$, pero si se hubiese utilizado la distribución normal, el valor crítico habría sido $z = -1.645$. El valor crítico de la t de Student se encuentra más cargado a la izquierda, 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

El ejemplo anterior siguió el método tradicional de prueba de hipótesis, pero STATDISK, Minitab, la calculadora TI-83 Plus y muchos artículos de revistas científicas presentan valores P. Para el ejemplo anterior, STATDISK presenta un valor P de 0.1023, Minitab y Excel presentan un valor P de 0.102 y la calculadora TI-83 Plus muestra un valor P de 0.1022565104. 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 Plus, utilice 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 Plus.
2. Si no dispone de la tecnología, consulte la tabla A-3 para identificar un rango de valores P. (Observe el siguiente ejemplo).

EJEMPLO **Cálculo de valores *P*** Suponiendo que no disponemos de un programa de cómputo o de una calculadora TI-83 Plus, consultamos la tabla A-3 para obtener un rango de valores para el valor P, correspondientes a los resultados dados.

a. En una prueba de hipótesis de cola izquierda, el tamaño de la muestra es $n = 12$ y el estadístico de prueba es $t = -2.007$.

b. En una prueba de hipótesis de cola derecha, el tamaño de la muestra es $n = 12$ y el estadístico de prueba es $t = 1.222$.

c. En una prueba de hipótesis de dos colas, el tamaño de la muestra es $n = 12$ y el estadístico de prueba es $t = -3.456$.

SOLUCIÓN De la figura 7-6, recuerde que el valor P es el área que se determina de la siguiente manera:

continúa

Prueba de cola izquierda:	El valor P es el área que se ubica a la *izquierda* del estadístico de prueba.
Prueba de cola derecha:	El valor P es el área que se ubica a la *derecha* del estadístico de prueba.
Prueba de dos colas:	El valor P es *dos veces* el área en la cola limitada por el estadístico de prueba.

En cada uno de los incisos *a*, *b* y *c*, el tamaño de la muestra es $n = 12$, de manera que el número de grados de libertad es gl $= n - 1 = 11$. Observe la porción de la tabla A-3 que se presenta a continuación, para 11 grados de libertad junto con los recuadros que describen los procedimientos para el cálculo de los valores P.

a. Se trata de una prueba de cola izquierda, con estadístico de prueba $t = -2.007$, por lo que el valor P es el área ubicada a la izquierda de -2.007. Por la simetría de la distribución t, es igual al área ubicada a la derecha de $+\ 2.007$. Observe la siguiente ilustración que muestra que cualquier

Tabla A-3 Cálculo de valores P con la tabla A-3

	Área en una cola				
	0.005	0.01	0.025	0.05	0.10
	Área en dos colas				
Grados de libertad	0.01	0.02	0.05	0.10	0.20
• • •					
11	3.106	2.718	2.201	1.796	1.363
	• • •				

Para un estadístico de prueba positivo mayor que 3.106:
- La prueba de cola dere–cha tiene un valor P menor que 0.005.
- La prueba de dos colas tiene un valor P menor que 0.01.

Para un estadístico de prueba positivo que esté entre 2.201 y 1.796:
- La prueba de cola dere–cha tiene un valor P entre 0.025 y 0.05.
- La prueba de dos colas tiene un valor P entre 0.05 y 0.10.

Para un estadístico de prueba positivo menor que 1.363:
- La prueba de cola dere–cha tiene un valor P mayor que 0.10.
- La prueba de dos colas tiene un valor P mayor que 0.20.

Nota: Si el estadístico de prueba es *negativo* elimine el signo negativo cuando...
- La prueba de cola izquierda tiene el mismo valor P descrito arriba para una prueba de cola derecha.
- La prueba de dos colas tiene el mismo valor P descrito arriba para una prueba de dos colas.
- La prueba de cola derecha tiene un valor P mayor que 0.5.

estadístico de prueba que esté entre 2.201 y 1.796 posee un valor P de cola derecha que se encuentra entre 0.025 y 0.05. Concluimos que $0.025 <$ valor $P < 0.05$. (El valor P exacto calculado por medio de un programa de cómputo es 0.0350).

b. Se trata de una prueba de cola derecha, con estadístico de prueba $t = 1.222$, de modo que el valor P es el área ubicada a la derecha de 1.222. Observe la ilustración que indica que cualquier estadístico de prueba menor que 1.363 tiene un valor P de cola derecha que es mayor que 0.10. Concluimos que el valor P es > 0.10. (El valor P exacto calculado con un programa de cómputo es 0.124).

c. Se trata de una prueba de dos colas, con estadístico de prueba $t = -3.456$. El valor P es *dos* veces el área ubicada la izquierda de –3.456, pero con la simetría de la distribución t, que es igual al doble del área ubicada a la derecha de $+3.456$. Observe la ilustración que indica que cualquier estadístico de prueba mayor que 3.106 tiene un valor P de dos colas que es menor que 0.01. Concluimos que el valor P es < 0.01. (El valor P exacto calculado con un programa de cómputo es 0.00537).

Una vez que se ha comprendido el formato de la tabla A-3, no es difícil calcular un rango de números para los valores P. Verifique sus resultados para asegurarse de que siguen los mismos patrones presentados en la tabla A-3. De izquierda a derecha, las áreas *se incrementan* conforme los valores de *t disminuyen*. Por ejemplo, en el inciso *b*, el estadístico de prueba $t = 1.222$ es menor que 1.363, de manera que el área de cola derecha es *mayor que* 0.10.

Recuerde, los valores P se calculan con facilidad si se utiliza un programa de cómputo o una calculadora TI-83 Plus. Además, se puede utilizar el método tradicional de prueba de hipótesis en lugar 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 7-2). Utilizando los datos muestrales del primer ejemplo de esta sección ($n = 12$ y $\bar{x} = 98.39$, $s = 0.535$), sin conocer σ y utilizando un nivel de significancia de 0.05, podemos probar la aseveración de que $\mu < 98.6$ por medio del método del intervalo de confianza. Construya este intervalo de confianza del 90%: $98.11 < \mu < 98.67$ (Consulte la sección 6-4). Como el valor supuesto de $\mu = 98.6$ está contenido dentro del intervalo de confianza, no podemos rechazar dicho supuesto. Con base en los 12 valores muestrales dados en el ejemplo, no tenemos evidencias suficientes para sustentar la aseveración de que la temperatura corporal media es menor que 98.6°F. Con base en el intervalo de confianza es probable que el valor verdadero de μ esté entre 98.11 y 98.67, incluyendo 98.6.

En la sección 7-3 aprendimos que cuando probamos una aseveración acerca de una proporción poblacional, el método tradicional y el método del valor P son equivalentes, pero el método del intervalo de confianza es un poco diferente. Cuando se prueba una aseveración acerca de una media poblacional no existe dicha diferencia y los tres métodos son equivalentes.

Método alternativo (no se usa en este libro) Cuando se prueba una aseveración acerca de la media poblacional μ, utilizando una muestra aleatoria simple que proviene de una población que se distribuye normalmente, con σ desconocida, un método alternativo (no utilizado en este libro) es el de aplicar 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 7-4). Este método alternativo no se utiliza en este libro por las siguientes razones (también citadas en la sección 6-4): **1.** Los criterios para elegir entre las distribuciones normal y t empleadas en este libro son los mismos que se usan en el mundo real. **2.** Cuando se desconoce σ, la distribución de $(\bar{x} - \mu) \div (s/\sqrt{n})$ es una distribución t, no una distribución normal; para tamaños de muestra muy grandes, las diferencias entre las distribuciones normal y t son despreciables, pero el uso de la distribución t suele producir mejores resultados. **3.** Para aquellos estudiantes que tomarán más cursos de estadística, es mejor que aprendan un procedimiento que puedan aplicar posteriormente y no un procedimiento que deban cambiar después. **4.** No es mucho más difícil trabajar con la distribución t que con la distribución normal, especialmente si se dispone de un programa de cómputo o de una calculadora TI-83 Plus.

Utilizando la tecnología

STATDISK Si se trabaja con la lista de los valores muestrales originales, primero calcule el tamaño de la muestra, la media muestral y la desviación estándar muestral por medio del procedimiento de STATDISK descrito en la sección 2-4. 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 trabaja únicamente con la lista de los datos originales. (Para saber cómo superar esta restricción, véase *Minitab Student Laboratory Manual and Workbook*, que complementa a este libro.) Primero introduzca los datos en la columna C1, después seleccione **Stat, Basic Statistics** y **1-Sample t** del menú e introduzca los datos requeridos. El cuadro denominado "alternative" se utiliza para seleccionar la forma de la hipótesis alternativa, y puede incluir **not equal, less than** o **grater than.**

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 Test**. 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, tal 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 Plus Si utiliza la calculadora TI-83 Plus, presione **STAT**, luego seleccione **TESTS** y elija la segunda opción **T-Test**. Usted puede utilizar los datos originales o un resumen de los estadísticos **(Stats)** al proporcionar las entradas indicadas en la representación visual de la ventana. Los primeros tres elementos de los resultados de la TI-83 Plus incluirán la hipótesis alternativa, el estadístico de prueba y el valor P.

7-5 Destrezas y conceptos básicos

Uso de la distribución correcta. *En los ejercicios 1 a 4, determine si la prueba de hipótesis incluye una distribución muestral de medias con distribución normal, distribución t de Student o ninguna de ellas. (Sugerencia: Consulte la figura 6-6 y la tabla 6-1.)*

1. Aseveración: $\mu = 100$. Datos muestrales: $n = 15$, $\bar{x} = 102$, $s = 15.3$. Los datos muestrales parecen provenir de una población que se distribuye normalmente, con μ y σ desconocidas.

2. Aseveración: $\mu = 75$. Datos muestrales: $n = 25$, $\bar{x} = 102$, $s = 15.3$. Los datos muestrales parecen provenir de una población con una distribución muy alejada de lo normal, con σ desconocida.

3. Aseveración: $\mu = 980$. Datos muestrales: $n = 5$, $\bar{x} = 950$, $s = 27$. Los datos muestrales parecen provenir de una población que se distribuye normalmente, con $\sigma = 30$.

4. Aseveración: $\mu = 2.80$. Datos muestrales: $n = 150$, $\bar{x} = 2.88$, $s = 0.24$. Los datos muestrales parecen provenir de una población que no se distribuye normalmente, con σ desconocida.

***Cálculo de valores* P.** *En los ejercicios 5 a 8, utilice la información dada para calcular un rango de números para el valor* P. *(Sugerencia: Véase el ejemplo y su representación visual en el apartado de "cálculo de valores* P *con la distribución* t *de Student").*

5. Prueba de cola derecha con $n = 12$ y estadístico de prueba $t = 2.998$

6. Prueba de cola izquierda con $n = 12$ y estadístico de prueba $t = -0.855$

7. Prueba de dos colas con $n = 16$ y estadístico de prueba $t = 4.629$

8. Prueba de dos colas con $n = 9$ y estadístico de prueba $t = -1.577$

Cálculo de los componentes de prueba. *En los ejercicios 9 a 12, suponga que se seleccionó una muestra aleatoria simple, de una población distribuida de manera normal. Calcule el estadístico de prueba, el valor* P*, el (los) valor(es) crítico(s) y establezca la conclusión final.*

9. Aseveración: La puntuación media del CI de profesores de estadística es mayor que 118. Datos muestrales: $n = 20$, $\bar{x} = 120$, $s = 12$. El nivel de significancia es $\alpha = 0.05$.

10. Aseveración: La temperatura corporal media de adultos sanos es menor que 98.6°F. Datos muestrales: $n = 35$, $\bar{x} = 98.20$°F, $s = 0.62$. El nivel de significancia es $\alpha = 0.01$.

11. Aseveración: El tiempo medio que transcurre para que los hombres vuelvan a utilizar el control remoto del televisor, durante los comerciales, es igual a 5.00 segundos. Datos muestrales: $n = 81$, $\bar{x} = 5.25$ s, $s = 2.50$ s. El nivel de significancia es $\alpha = 0.01$.

12. Aseveración: El salario medio inicial de estudiantes universitarios graduados que han tomado un curso estadística es igual a $46,000.

Datos muestrales: $n = 27$, $\bar{x} = \$45,678$, $s = \$9900$. El nivel de significancia es $\alpha = 0.05$.

Prueba de hipótesis. *En los ejercicios 13 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.*

13. Harry Potter y nivel de lectura El conjunto de datos 14 del Apéndice B incluye medidas del nivel de lectura de 12 páginas seleccionadas al azar del libro *Harry Potter y la piedra filosofal*, de J. K. Rowling. Las medidas del nivel de Flesch-Kincaid se resumen en los siguientes estadísticos: $n = 12$, $\bar{x} = 5.075$, $s = 1.168$. Los maestros en West Park School District no utilizarán el libro a menos que se demuestre que el nivel de lectura de una página típica esté por encima del cuarto grado. Utilice un nivel de significancia de 0.05 para aprobar la aseveración de que la media es mayor que 4. ¿Utilizarán los maestros el libro?

14. Azúcar en el cereal El conjunto de datos 16 del Apéndice B lista el contenido de azúcar (gramos de azúcar por gramo de cereal) de una muestra de distintos cereales.

Estas cantidades se resumen en los siguientes estadísticos: $n = 16$, $\bar{x} = 0.295$ g, $s = 0.168$ g. Utilice un nivel de significancia de 0.05 para aprobar la aseveración de que la media de todos los cereales es menor que 0.3 g.

15. Temperaturas reales y pronosticadas El conjunto de datos 10 del Apéndice B incluye una lista de altas temperaturas reales y la lista correspondiente de altas temperaturas pronosticadas a tres días. Si la diferencia para cada día se calcula restando el pronóstico a tres días de altas temperaturas de las altas temperaturas reales, el resultado es una lista de 31 valores con una media de $-0.419°$ y una desviación estándar de $3.704°$. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la diferencia media es distinta de 0°. Con base en el resultado, ¿parece razonablemente preciso el pronóstico de altas temperaturas para tres días?

16. Estatura de los padres El conjunto de datos 2 del Apéndice B incluye las estaturas de padres de 20 varones. Si la diferencia de la estatura de cada conjunto de padres se calcula restando la estatura de la madre de la estatura del padre, el resultado es una lista de 20 valores con una media de 4.4 pulgadas y una desviación estándar de 4.2 pulgadas. Utilice un nivel de significancia de 0.01 para probar la aseveración de que la diferencia media es mayor que 0. ¿Sustentan los resultados la aseveración de un sociólogo de que las mujeres tienden a casarse con hombres más altos que ellas?

17. Prueba de la precisión de relojes de pulso Los estudiantes del autor seleccionaron al azar a 40 personas y midieron la precisión de sus relojes de pulso. Los errores positivos representan relojes que están adelantados; y los errores negativos, relojes que están retrasados. Los 40 valores tienen una media de 117.3 s y una desviación estándar de 185.0 s. Utilice un nivel de significancia de 0.01 para probar la aseveración de que la población de todos los relojes tiene una media igual a 0 s. ¿Qué se concluye acerca de la precisión de los relojes de pulso de las personas?

18. Precios de libros de texto Heather Carielli es una ex alumna del autor que obtuvo el grado de maestría en estadística en la Universidad de Massachusetts. Al seleccionar al azar 16 libros de texto nuevos en la librería de la universidad, descubrió que la media de los precios era de \$70.41 y la desviación estándar era de \$19.70. ¿Existirá evidencia suficiente para justificar el rechazo de la aseveración que aparece en el catálogo de la universidad de que el precio medio de un libro de texto ahí es menor que \$75?

19. 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 hombres 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 hombres tienen un periodo medio de vida mayor que 69.5 años. ¿Parecería que los directores de orquesta hombres viven más que los hombres de la población general? ¿Por qué la experiencia de ser un director de orquesta hace que los hombres vivan más tiempo? (*Sugerencia:* ¿Los directores de orquesta nacen, o se convierten en directores a una edad mucho más tardía?)

20. Pelotas de béisbol En pruebas previas, se dejaron caer pelotas de béisbol 24 pies sobre una superficie de concreto, y rebotaron un promedio de 92.84 pulgadas. En una prueba realizada a una muestra de 40 pelotas nuevas, rebotaron un promedio de 92.67 pulgadas, con una desviación estándar de 1.79 pulgadas (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 pulgadas. ¿Parecería que las pelotas son diferentes?

21. Prueba de choques de BMW Por el costo que implican, las pruebas de choques de automóviles suelen utilizar muestras pequeñas. Cuando se chocan cinco automóviles BMW en condiciones estándar, se emplean los costos de reparación (en dólares) para probar la aseveración de que el costo medio de reparación de todos los automóviles BMW es menor que \$1000. Los resultados de Minitab de esta prueba de hipótesis se presentan abajo. Con base en los resultados de esta prueba de hipótesis, ¿se justificaría que BMW anunciara que, en condiciones estándar, el costo promedio de reparación es menor que \$1000?

```
Test of mu = 1000 vs mu < 1000

Variable       N      Mean      StDev      SE Mean
Cost           5       767        285          127

Variable   95.0% Upper Bound        T            P
Cost                  1039      -1.83        0.071
```

22. Confiabilidad de radios de aeronaves El tiempo medio que transcurre entre las fallas (en horas) de un radio de la Telektronic Company, utilizado en aeronaves ligeras, es de 420 h. Después de que se modificaron 15 radios nuevos, en un intento por mejorar su confiabilidad, se realizaron pruebas para medir los tiempos transcurridos entre las fallas. Cuando se utilizó Minitab para probar la aseveración de que los radios modificados tienen una media mayor que 420 h, se obtuvieron los resultados que se presentan a continuación. ¿Parecería que las modificaciones incrementaron la confiabilidad?

```
Test of mu = 420 vs mu > 420

Variable       N      Mean      StDev      SE Mean
Time          15     442.2       44.0         11.4

Variable   95.0% Lower Bound        T            P
Time                 422.2       1.95        0.035
```

23. Efecto de un complemento vitamínico en el peso al momento de nacer Se registran los pesos al nacimiento (en kg) de una muestra de bebés hombres nacidos de madres que tomaron un complemento vitamínico especial (según datos del New York State Department of Health). Al probar la aseveración de que el peso medio al nacimiento de todos los bebés cuyas madres tomaron vitaminas es igual a 3.39 kg, que es la media de la población de todos los varones, la calculadora TI-83 Plus produjo los resultados al margen. Con base en esos resultados, ¿parecería que el complemento vitamínico tiene un efecto sobre el peso al momento de nacer?

TI-83 Plus

```
T-Test
 μ≠3.39
 t=1.73432095
 p=.1033573367
 x̄=3.675
 Sx=.6573177821
 n=16
```

24. Pulso En el momento más intenso de un programa de ejercicio, el autor aseveró que su pulso era menor que el pulso medio de estudiantes de estadística. La medida del pulso del autor fue de 60 latidos por minuto, y se midió el pulso de los 20 estudiantes de su clase. Al probar la aseveración de que los estudiantes de estadística tenían un pulso medio mayor que 60 latidos por minuto, en la calculadora TI-83 Plus se obtuvieron los resultados presentados al margen. Con base en esos resultados, ¿existe evidencia suficiente para sustentar la aseveración de que el pulso medio de los estudiantes de estadística es mayor que 60 latidos por minuto?

TI-83 Plus

```
T-Test
 μ>60
 t=6.393083002
 p=1.9683974E-6
 x̄=74.35
 Sx=10.03821645
 n=20
```

25. Verificación de plomo en el aire Más adelante se listan cantidades medidas de plomo (en microgramos por metro cúbico o $\mu g/m^3$) en el aire. La Environmental Protection Agency (EPA) ha establecido un estándar de calidad del aire para el plomo: 1.5 $\mu g/m^3$. Las mediciones presentadas abajo se registraron en el edificio cinco 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 estándar de la EPA, de 1.5 $\mu g/m^3$. ¿Existe algo en estos datos que sugiera que el supuesto de una población que se distribuye normalmente podría no ser válido?

5.40	1.10	0.42	0.73	0.48	1.10

26. Tratamiento del síndrome de fatiga crónica Se probaron pacientes con síndrome de fatiga crónica, luego se trataron con fludrocortisona y después se probaron nuevamente. Abajo se presentan los cambios en la fatiga después del tratamiento (datos tomados de "The Relationship 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). Se utilizó una escala estándar de -7 a $+7$, donde 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?

6 5 0 5 6 7 3 3 2 6 5 5 0 6 3 4 3 7 0 4 4

27. 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 aleatoriamente de la población de todos los juegos olímpicos pasados y futuros, pruebe la aseveración de que el tiempo medio es menor que 10.5 segundos. ¿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 10.5 segundos? ¿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	

28. Nicotina en cigarrillos La Carolina Tobacco Company anunció que sus cigarrillos sin filtro más vendidos contienen a lo sumo 40 mg de nicotina; sin embargo, la revista *Consumer Advocate* realizó pruebas a 10 cigarrillos seleccionados al azar y descubrió las cantidades (en mg) de la lista que se presenta a continuación. Es grave acusar a la compañía de que su anuncio sea incorrecto, por lo que el editor de la revista elige un nivel de significancia de $\alpha = 0.01$ para probar su creencia de que el contenido medio de nicotina es mayor que 40 mg. Empleando un nivel de significancia de 0.01, pruebe 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

T **29.** Nivel de lectura de Tom Clancy Remítase al conjunto de datos 14 del Apéndice B y utilice las mediciones del nivel de Flesch-Kincaid para la obra *El oso y el dragón*, de Tom Clancy. Un maestro de preparatoria desea asignar el libro para una tarea de lectura, pero requiere de un libro con un nivel de lectura por encima del sexto grado. ¿Hay evidencia suficiente para sustentar la afirmación de que el libro de Clancy cumple este requisito?

T **30.** Consumo de tabaco en películas infantiles Remítase al conjunto de datos 7 del Apéndice B y utilice únicamente las películas que presentan algún consumo de tabaco. Pruebe la aseveración de un crítico de cine de que "entre las películas que muestran el consumo de tabaco, el tiempo medio de exposición es de dos minutos". Dados los datos muestrales, ¿son engañosos los datos?

T **31.** Volúmenes de Coca Cola El conjunto de datos 17 del Apéndice B incluye los volúmenes (en onzas) de la Coca Cola regular en una muestra de 36 latas diferentes etiquetadas con 12 onzas. Un gerente de línea afirma que la cantidad media de Coca Cola clásica es mayor que 12 onzas, lo que causa menores ganancias a la compañía. Con un nivel de significancia de 0.01, pruebe la aseveración del gerente de que la media es mayor que 12 onzas. ¿Deberá ajustarse el proceso de producción?

T **32.** Sodio en el cereal Remítase al conjunto de datos 16 del Apéndice B y pruebe la aseveración de un nutriólogo de que "la caja promedio de cereal contiene más de 6 mg de sodio por gramo de cereal". Si se considera que 6 mg de sodio por gramo de cereal es excesivo, ¿podemos decir que el cereal no es sano debido a su alto contenido de sodio?

7-5 Más allá de lo básico

33. Uso de resultados de computadora Remítase a los resultados de Minitab del ejercicio 22. Si la aseveración se cambia de "mayor que 420 h" a "no igual a 420 h", ¿de qué manera se ven afectados el estadístico de prueba, el valor *P* y la conclusión?

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 de emplear la distribución *t* de Student para calcular los valores críticos y/o un valor *P*. Si, en su lugar, se utiliza de forma incorrecta una distribución normal estándar, ¿este error lo hace más propenso a rechazar la hipótesis nula, o no hace ninguna diferencia? Explique.

35. Efecto de un dato distante Repita el ejercicio 25 después de cambiar primero el valor 5.40 por 540. Con base en los resultados, describa el efecto de un dato distante en una prueba *t*.

36. Cálculo de los valores críticos *t* Al calcular 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

$$t = \sqrt{\text{gl} \cdot (e^{A^2/\text{gl}} - 1)}$$

donde $\text{gl} = n - 1$, $e = 2.718$, $A = z(8 \cdot \text{gl} + 3)/(8 \cdot \text{gl} + 1)$, 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. Probabilidad de un error tipo II Remítase al ejercicio 28 y suponga que está probando la aseveración de que $\mu > 40$ mg. Calcule β, la probabilidad de un error tipo II, si el valor real de la media poblacional es $\mu = 45.0518$ mg. (Véase el ejercicio 19 en la sección 7-4).

7-6 Prueba de una aseveración respecto de una desviación estándar o de una varianza

El mundo industrial comparte esta meta común: mejorar la calidad reduciendo la variación. Los ingenieros de control de calidad desean asegurarse de que un producto tiene una media aceptable, pero también desean producir artículos de calidad *consistente*, de modo que se presenten pocos defectos. Por ejemplo, la consistencia de altímetros de aeronaves está determinada por la regla 91.36 de la Federal Aviation, la cual requiere que los altímetros de aeronaves se prueben y calibren para dar una lectura "dentro de 125 pies (con una base de probabilidad del 95%)". Aun cuando la lectura de la altitud media sea exactamente correcta, resultará una desviación estándar excesivamente grande en lecturas individuales que son peligrosamente bajas o altas. La consistencia se mejora al reducir la desviación estándar. En las secciones anteriores de este capítulo describimos métodos para probar aseveraciones acerca de medias y proporciones poblacionales. Esta sección se enfoca en la variación, que es sumamente importante en muchas aplicaciones, incluyendo el control de calidad. El objetivo principal de esta sección es presentar métodos para probar aseveraciones acerca de una desviación estándar poblacional σ o varianza poblacional σ^2. Los supuestos, el estadístico de prueba, el valor *P* y los valores críticos se resumen de la siguiente manera.

Prueba de aseveraciones acerca de σ o σ^2

Supuestos

1. La muestra es una muestra 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 7-4 y 7-5).

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*).

En las secciones 7-4 y 7-5 vimos que los métodos de prueba de aseveraciones acerca de medias requieren de una población distribuida de forma normal, y que estos métodos trabajan razonablemente 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*, lo que quiere decir que los resultados pueden ser muy confusos si la población no tiene una distribución normal. Por consiguiente, la condición de una población que se distribuye normalmente es un requisito mucho más estricto en esta sección. Si la población tiene una distribución que se aleja mucho de lo normal y usted utiliza los métodos de esta sección para rechazar una hipótesis nula, en realidad no sabrá si la desviación estándar no es como se supuso o si el rechazo se debe a la carencia de normalidad.

No se confunda cuando nos refiramos a las distribuciones normal y chi cuadrada. Después de verificar que los datos muestrales parecen provenir de una población distribuida normalmente, entonces debemos pensar en términos de la distribución chi cuadrada. La distribución chi cuadrada se introdujo en la sección 6-5, 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 7-12).
2. Existe una distribución χ^2 diferente para cada número de grados de libertad (véase la figura 7-13).
3. Todos los valores críticos se encuentran en la tabla A-4, utilizando

$$\textbf{grados de libertad} = \mathbf{n - 1}$$

La tabla A-4 se basa en áreas acumulativas de la zona *derecha* (a diferencia de los datos de la tabla A-4 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 $= n - 1$). A

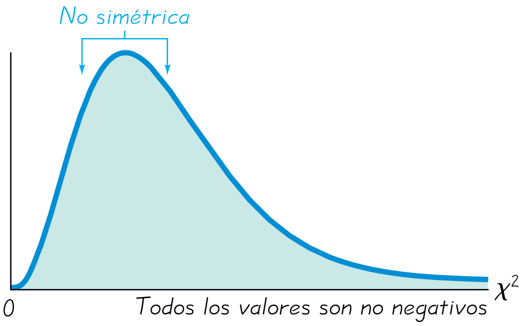


FIGURA 7-12 Propiedades de la distribución chi cuadrada

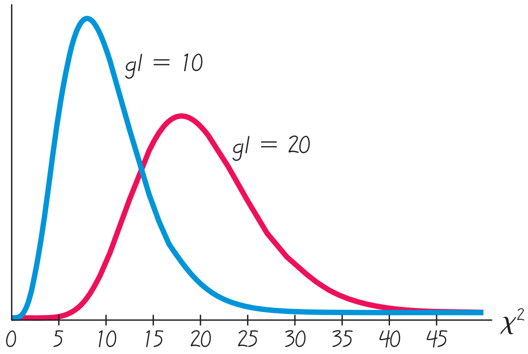


FIGURA 7-13 Distribución chi cuadrada para 10 y 20 grados de libertad

continuación, se utiliza el nivel de significancia α para determinar la columna correcta. Los siguientes ejemplos están basados en un nivel de significancia de $\alpha = 0.05$, aunque se puede emplear cualquier otro nivel de significancia de manera similar. Observe que en cada caso el área clave es la región que se encuentra a la *derecha* del (los) valor(es) crítico(s).

Prueba de cola derecha:	Considerando 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 un área de cola izquierda de 0.05, el área a la *derecha* del valor crítico es 0.95, por lo tanto 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 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. (Consulte la figura 6-10 y el ejemplo en las páginas 349-350).

EJEMPLO Puntuaciones de CI de profesores de estadística

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 desviación estándar y no se da cuenta de que ésta debe ser menor que 15 (ya que los profesores de estadística tienen una variación menor que la población general). En su lugar, él asevera que los profesores de estadística tienen puntuaciones de CI con una desviación estándar igual a 15, como la población general. Suponga que las puntuaciones de CI de los 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 sobre la desviación estándar de las puntuaciones de CI de los profesores de estadística?

continúa

La ética en experimentos

Con frecuencia es posible obtener datos muestrales con sólo observar o encuestar a miembros seleccionados de la población. Muchas otras situaciones requieren que manipulemos circunstancias, de alguna manera, para obtener datos muestrales. En ambos casos llegan a surgir dilemas éticos. Investigadores en Tuskegee, Alabama, aplicaron el tratamiento de penicilina eficaz a víctimas de sífilis para poder estudiar la enfermedad. ¡Este experimento continuó por un periodo de 27 años!

SOLUCIÓN Emplearemos el método tradicional de prueba de hipótesis, tal como se describe en la figura 7-8.

Paso 1: La aseveración se expresa en forma simbólica como $\sigma = 15$.

Paso 2: Si la aseveración original es falsa, entonces $\sigma \neq 15$.

Paso 3: La expresión $\sigma \neq 15$ no contiene igualdad, por lo que se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de que $\sigma = 15$.

H_0: $\sigma = 15$ (aseveración original)
H_1: $\sigma \neq 15$

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Puesto que la aseveración se refiere a σ, utilizamos la distribución chi cuadrada.

Paso 6: El estadístico de prueba es

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(13-1)(7.2)^2}{15^2} = 2.765$$

Los valores críticos de 4.404 y 23.337 se localizan en la tabla A-4, en el 12o renglón (grados de libertad $= n - 1 = 12$) en las columnas correspondientes a 0.975 y 0.025. Observe el estadístico de prueba y los valores críticos que aparecen en la figura 7-14.

Paso 7: Puesto que el estadístico de prueba se encuentra en la región crítica, rechazamos la hipótesis nula.

INTERPRETACIÓN Existe suficiente evidencia para justificar el rechazo de la aseveración de que la desviación estándar es igual a 15. Parece que los profesores de estadística tienen puntuaciones de CI con una desviación estándar que es significativamente diferente de la desviación estándar de 15 de la población general.

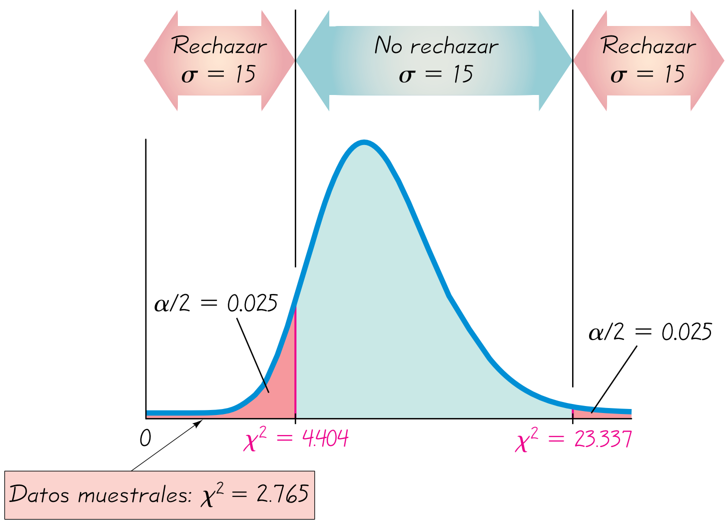


FIGURA 7-14 **Prueba de la aseveración de que $\sigma = 15$**

Método del valor *P*

En lugar de utilizar el método tradicional de prueba de hipótesis en el ejemplo anterior, también podemos utilizar el método del valor *P* que se resume en las figuras 7-6 y 7-9. Si se usa STATDISK en el ejemplo anterior, se obtendrán el valor *P* de 0.0060. Si empleamos la tabla A-4, generalmente no obtendremos valores *P exactos*, ya que la tabla de la distribución chi cuadrada únicamente incluye valores selectos de α. (Por esta limitación, es más fácil probar aseveraciones acerca de σ o σ^2 con la tabla A-4 utilizando el método tradicional, que utilizando el método del valor *P*). Si empleamos la tabla A-4 podemos identificar los límites que contienen al valor *P*. El estadístico de prueba del último ejemplo es $\chi^2 = 2.765$, y sabemos que la prueba es de dos colas con 12 grados de libertad. Remítase al 12o. renglón de la tabla A-4 y observe que el estadístico de prueba de 2.765 es menor que cualquier dato en ese renglón, lo que significa que el área a la izquierda del estadístico de prueba es menor que 0.005. El valor *P* para una prueba de dos colas es *dos veces* el área de la cola limitada por el estadístico de prueba, de modo que duplicamos 0.005 para concluir que el valor *P* es menor que 0.01. Puesto que el valor *P* es menor que el nivel de significancia de $\alpha = 0.05$, rechazamos la hipótesis nula. Nuevamente, 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. Utilizando los métodos descritos en la sección 6-5, podemos emplear los datos muestrales ($n = 13$, $s = 7.2$) para construir el siguiente intervalo de confianza del 95%: $5.2 < \sigma < 11.9$. Como el valor aseverado de $\sigma = 15$ *no* está contenido dentro del intervalo de confianza, rechazamos la aseveración de que $\sigma = 15$ y sacamos la misma conclusión que con los métodos tradicional y del valor *P*.

Utilizando 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.** El STATDISK desplegará el estadístico de prueba, los valores críticos, el valor *P*, la conclusión y el intervalo de confianza.

Minitab **Excel** **TI-83 Plus** Estas herramientas tecnológicas aún no están diseñadas para probar aseveraciones acerca de σ o σ^2.

7-6 Destrezas y conceptos básicos

Cálculo de valores críticos. *En los ejercicios 1 a 4 calcule el estadístico de prueba, después utilice la tabla A-4 para obtener el (los) valor(es) crítico(s) de χ^2 y los límites que contienen al valor P; luego determine si existe suficiente evidencia para sustentar la hipótesis alternativa dada.*

1. H_1: $\sigma \neq 15$, $\alpha = 0.05$, $n = 20$, $s = 10$.

2. H_1: $\sigma > 12$, $\alpha = 0.01$, $n = 5$, $s = 18$.

3. H_1: $\sigma < 50$, $\alpha = 0.01$, $n = 30$, $s = 30$.

4. H_1: $\sigma \neq 4.0$, $\alpha = 0.05$, $n = 81$, $s = 4.7$.

Prueba de aseveraciones sobre variación. *En los ejercicios 5 a 16 pruebe la aseveración dada. Suponga que se selecciona una muestra aleatoria simple de una población que se distribuye normalmente. Utilice el método tradicional de prueba de hipótesis, a menos que su profesor indique otra cosa.*

5. 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 0.04 g. Una muestra de 40 dulces M&M de cacahuate tiene pesos con una desviación estándar de 0.31 g. ¿Por qué tendrán los dulces de cacahuate pesos que varían más que los dulces sencillos?

6. Variación de pistones Al diseñar un pistón para una bomba de transferencia de soluciones líquidas, los ingenieros especificaron una media de 0.1 pulgadas para el radio del pistón. La desviación estándar máxima se especificó en 0.0005 pulgadas (según datos de Taylor Industries). Cuando se selecciona al azar 12 pistones de la línea de producción y se miden, sus radios tienen una desviación estándar de 0.00047 pulgadas. ¿Existe suficiente evidencia para sustentar la aseveración de que los pistones se están fabricando con radios que tienen una desviación estándar menor que el mínimo especificado de 0.0005 pulgadas? Utilice un nivel de significancia de 0.05.

7. 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 $s = 52.3$ pies. Al 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 pies, que era la desviación estándar del antiguo método de producción. Parece que la desviación estándar ha cambiado, ¿parece ser mejor o peor el nuevo método de producción en comparación con el anterior?

8. Puntuaciones de exámenes de estadística Los exámenes de clases anteriores de estadística del autor tienen calificaciones con una desviación estándar igual a 14.1. Una de sus clases recientes incluye 27 calificaciones de examen con una desviación estándar de 9.3. Utilice un nivel de significancia de 0.01 para probar la aseveración de que las clases actuales tienen menor variación que las clases anteriores. ¿La desviación estándar menor sugiere que les va mejor a las clases actuales?

9. Tiempos de espera de clientes bancarios El banco Jefferson Valley, que utiliza filas individuales en las distintas ventanillas, encontró que la desviación estándar de los tiempos de espera los viernes en la tarde, distribuidos normalmente, era de 6.2 min. El banco experimentó con una fila única y observó que para una muestra aleatoria simple de 25 clientes, los tiempos de espera tenían una desviación estándar de 3.8 min. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la fila única causa una menor variación en los tiempos de espera. ¿Por qué los clientes preferirían tiempos de espera con menor variación? ¿Resulta en una espera menor el uso de una fila única?

10. Temperaturas corporales En la sección 7-4 probamos la aseveración de que la temperatura corporal media es igual a 98.6°F, y utilizamos datos muestrales del conjunto de datos 4 del Apéndice B. Las temperaturas corporales tomadas a las 12:00 AM del día dos pueden resumirse con los siguientes estadísticos: $n = 106$, $\bar{x} = 98.20$°F, $s = 0.62$°F, y un histograma muestra que los valores tienen una distribución aproximadamente normal. En la sección 7-4 asumimos que $\sigma = 0.62$°F, que es un supuesto poco realista. Sin embargo, el estadístico de prueba causará el rechazo de $\mu = 98.6$°F, siempre y cuando la desviación estándar sea menor que 2.11°F. Utilice los estadísticos muestrales y un nivel de significancia de 0.005 para probar la aseveración de que $\sigma < 2.11$°F.

11. 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 libras. A continuación se listan los pesos (en libras) de 9 supermodelos seleccionadas al azar.

125 (Taylor)	119 (Auermann)	128 (Schiffer)	128 (MacPherson)
119 (Turlington)	127 (Hall)	105 (Moss)	123 (Mazza)
115 (Hume)			

12. 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 pulgadas. A continuación se listan las estaturas (en pulgadas) de supermodelos seleccionadas al azar (Taylor, Harlow, Mulder, Goff, Evangelista, Auermann, Schiffer, MacPherson, Turlington, Hall, Crawford, Campbell, Herzigova, 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

T **13.** Volúmenes de Pepsi Un nuevo gerente de producción asevera que los volúmenes de latas de Pepsi normal tienen una desviación estándar menor que 0.10 onzas. Utilice un nivel de significancia de 0.05 para probar la aseveración con los resultados muestrales incluidos en el conjunto de datos 17 del Apéndice B. ¿Qué problemas se causan por una media que no es de 12 onzas? ¿Qué problemas surgen por una desviación estándar demasiado alta?

T **14.** Presión sanguínea sistólica de mujeres La presión sanguínea sistólica resulta de las contracciones del corazón. Con base en resultados pasados del National Health Survey, se asevera que las mujeres tienen presiones sanguíneas sistólicas con una media y una desviación estándar de 130.7 y 23.4, respectivamente. Use las presiones sanguíneas sistólicas de mujeres que se listan en el conjunto 1 de datos del Apéndice B y pruebe la aseveración de que la muestra proviene de una población con una desviación estándar de 23.4.

T **15.** Pesos de hombres Se emplean datos de una encuesta antropométrica para publicar valores que sirven en el diseño de productos adecuados para que los adultos los utilicen. Según Gordon, Churchill *et al*., los hombres tienen pesos con una media de 172.0 libras y una desviación estándar de 28.7 libras. Con la muestra de pesos de hombres del conjunto de datos 1 del Apéndice B, pruebe la aseveración de que la desviación estándar es de 28.7 libras. Emplee un nivel de significancia de 0.05. Al diseñar elevadores, ¿cuál sería una consecuencia de la creencia de que los pesos de hombres varían menos de lo que realmente varían?

T **16.** Estaturas de mujeres Se emplean datos de una encuesta antropométrica para publicar valores que pueden usarse en el diseño de productos adecuados para que los adultos los utilicen. Según Gordon, Churchill *et al*., las mujeres tienen estaturas con una media de 64.1 pulgadas y una desviación estándar de 2.52 pulgadas. Con la muestra de estaturas de mujeres del conjunto de datos 1 del Apéndice B, pruebe la aseveración de que la desviación estándar es de 2.52 pulgadas. Utilice un nivel de significancia de 0.05. Al diseñar asientos de automóvil para mujeres, ¿cuál sería la consecuencia de creer que las estaturas de mujeres varían menos de lo que realmente varían?

7-6 Más allá de lo básico

17. Control de la variación en latas de Pepsi Remítase al ejercicio 13 y, para una muestra de tamaño $n = 36$ y con un nivel de significancia de 0.05, calcule la desviación estándar muestral más grande que pueda utilizarse para sustentar la aseveración de que $\sigma < 0.10$ onzas.

18. 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 150, permitimos que $k = 149$ con $z = -1.96$, seguido por $k = 149$ y $z = 1.96$.

a. Utilice esta aproximación para estimar los valores críticos de χ^2 en una prueba de hipótesis de dos colas con $n = 101$ y $\alpha = 0.05$. Compare los resultados con los obtenidos en la tabla A-4.

b. Use esta aproximación para estimar los valores críticos de χ^2 en una prueba de hipótesis de dos colas, con $n = 150$ y $\alpha = 0.05$.

19. Cálculo de los valores críticos de χ^2 Repita el ejercicio 18 aplicando esta aproximación (con k y z descritas en el ejercicio 18):

$$\chi^2 = k\left(1 - \frac{2}{9k} + z\sqrt{\frac{2}{9k}}\right)^3$$

20. Efectos de un dato distante Al utilizar el procedimiento de prueba de hipótesis de esta sección, ¿se vería muy afectado el resultado con la presencia de un dato distante? Describa cómo llegó a su respuesta.

21. Análisis del último dígito En ocasiones los últimos dígitos de datos muestrales se emplean para determinar si los datos fueron medidos o simplemente reportados por el sujeto. Los datos reportados suelen tener últimos dígitos con un exceso de ceros y cincos. Los datos medidos tienden a tener últimos dígitos con una media de 4.5, una desviación estándar de alrededor de 3 y los dígitos suelen presentarse casi con la misma frecuencia.

a. ¿De qué manera se afecta la desviación estándar cuando existe un exceso de ceros y cincos?

b. ¿Por qué no podemos utilizar los métodos de esta sección para probar que los últimos dígitos de los datos muestrales tienen una desviación estándar igual a 3?

22. Probabilidades de un error tipo II Remítase al ejercicio 9. Suponiendo que σ es en realidad igual a 4.0, calcule β, que denota la probabilidad de un error tipo II. Revise el ejercicio 19 de la sección 7-4 y modifique el procedimiento de tal modo que se aplique a una prueba de hipótesis que incluya σ en lugar de μ.

Repaso

En este capítulo se presentaron métodos básicos para probar 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, tal como se ilustra en muchas de sus revistas científicas.

En la sección 7-2 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 tradicional (resumido en la figura 7-8)

2. El método del valor P (resumido en la figura 7-9)

3. Los intervalos de confianza (estudiados en el capítulo 6)

Tabla 7-3 Pruebas de hipótesis

Parámetro	Condiciones	Distribución y estadístico de prueba	Valores P y críticos
Proporción	$np \geq 5$ y $nq \geq 5$	Normal: $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$	Tabla A-2
Media	σ conocida y población que se distribuye normalmente o σ conocida y $n > 30$	Normal: $z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\dfrac{\sigma}{\sqrt{n}}}$	Tabla A-2
	σ desconocida y población distribuida normalmente o σ desconocida y $n > 30$	t de Student: $t = \dfrac{\bar{x} - \mu_{\bar{x}}}{\dfrac{s}{\sqrt{n}}}$	Tabla A-3
	Población no distribuida normalmente y $n \leq 30$	Usar método no paramétrico o *bootstrap*.	
Desviación estándar o varianza	Población distribuida normalmente	Chi cuadrada: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$	Tabla A-4

En las secciones 7-3 a 7-6 estudiamos métodos específicos para manejar distintos parámetros. Puesto que es tan importante seleccionar correctamente la distribución y el estadístico de prueba, presentamos la tabla 7-3, que resume los procedimientos de este capítulo para la prueba de hipótesis.

Ejercicios de repaso

1. **a.** Usted acaba de reunir una muestra muy grande ($n = 2575$) de respuestas obtenidas de adultos estadounidenses que enviaron por correo las respuestas a un cuestionario impreso en la revista *Fortune*. Una prueba de hipótesis realizada con un nivel de significancia de 0.01 conduce a la conclusión de que la mayoría (más del 50%) de los adultos se oponen a los impuestos estatales. ¿Concluiremos que la mayoría de los adultos estadounidenses se oponen a los impuestos estatales? ¿Por qué?

   b. Al probar un fármaco para el control del peso, una prueba de hipótesis basada en 5000 sujetos seleccionados al azar revela que la pérdida media de peso de 0.2 libras es significativa al nivel 0.01. ¿Deberán utilizar este fármaco los sujetos que desean perder peso? ¿Por qué?

   c. Usted acaba de inventar 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?

   d. Al probar la aseveración de que la cantidad media de refresco de cola en latas es mayor que 12 onzas, usted no rechaza la hipótesis nula. Plantee la conclusión final que retoma la aseveración original.

   e. Complete la afirmación: "Un error tipo I es el error de..."

2. Identificación de hipótesis y distribuciones Con base en las condiciones acordadas, identifique la hipótesis alternativa y la distribución muestral (normal, *t*, chi cuadrada) del estadístico de prueba.

a. Aseveración: El ingreso anual medio de estudiantes universitarios de tiempo completo está por debajo de $10,000. Datos muestrales: para 750 estudiantes universitarios seleccionados aleatoriamente, la media es $3662 y la desviación estándar es $2996.

b. Aseveración: Con el ensamble manual de las partes de teléfonos, los tiempos de ensamble varían más que los tiempos del ensamble automatizado, del cual se sabe que tiene una media de 27.6 segundos y una desviación estándar de 1.8 segundos.

c. Aseveración: La mayoría de los estudiantes universitarios son mujeres. Datos muestrales: de 500 estudiantes universitarios seleccionados aleatoriamente, el 58% son mujeres.

d. Aseveración: Cuando se selecciona al azar a un grupo de adultos encuestados, su CI medio es igual a 100. Datos muestrales: $n = 150$ y $\bar{x} = 98.8$. Es razonable suponer que $\sigma = 15$.

3. Generación aleatoria de datos La calculadora TI-83 Plus se utiliza para generar datos aleatorios a partir de una población que se distribuye normalmente. El comando **rand-Norm(100,15,50)** genera 50 valores de una población distribuida normalmente, con $\mu = 100$ y $\sigma = 15$. Una muestra generada de este tipo de 50 valores tiene una media de 98.4 y una desviación estándar de 16.3.

a. 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 $\sigma = 15$.

b. Repita el inciso *a*, suponiendo que se desconoce σ.

c. Utilice un nivel de significancia de 0.10 para probar la aseveración de que esta muestra en realidad proviene de una población con una desviación estándar igual a 15. ¿Qué dice este resultado acerca de la variación entre los valores muestrales generados?

d. Con base en los resultados anteriores, ¿parecería que el generador de números aleatorios de la calculadora está funcionando correctamente?

4. Errores de entrevista Una encuesta de Accountmeps, 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.

5. Pesos de monedas de 25 centavos de dólar Si nos remitimos a los pesos (en gramos) de monedas de 25 centavos de dólar, listados en el conjunto de datos 29 del Apéndice B, encontramos 50 pesos con una media de 5.622 g y una desviación estándar de 0.068 g. El Departamento del Tesoro de Estados Unidos asevera que el procedimiento utilizado para acuñar estas monedas produce un peso medio de 5.670 g. Con un nivel de significancia de 0.01, pruebe la aseveración de que el peso medio de las monedas de 25 centavos de dólar en circulación es de 5.670 g. Si se rechaza la aseveración, ¿cuál sería una explicación posible para la discrepancia?

6. Pesos de dulces M&M azules Con los pesos de los dulces M&M *azules* listados en el conjunto de datos 19 del Apéndice B, pruebe la aseveración de que la media es de al menos 0.9085 g, el valor medio necesario para que los 1498 M&M produzcan un total de 1361 g, tal como lo indica la envoltura. Utilice un nivel de significancia de 0.05. Para los M&M azules, $\bar{x} = 0.9014$ g y $s = 0.0573$ g. Con base en el resultado, ¿podemos concluir que el contenido de los paquetes no coincide con el peso aseverado impreso en la envoltura?

7. Porcentaje de visitas a parques temáticos Cada año se gastan miles de millones de dólares en los parques temáticos propiedad de Disney, Universal Studios, Sea World, Busch Gardens y otros. Una encuesta a 1233 personas que hicieron viajes, reveló que 111 de ellos incluyeron una visita a un parque temático (según datos de la Travel

Industry Association of America). Con base en esos resultados de encuesta, la consultora gerencial Laura Croft asevera que menos del 10% de los viajes incluyen una visita a un parque temático. Utilice un nivel de significancia de 0.05 para probar su aseveración. ¿Sería adecuado que ella utilizara esta afirmación para tratar de convencer a la gerencia de los parques temáticos de que incremente sus gastos en publicidad?

8. Votos para el candidato ganador En una elección presidencial reciente, se entrevistó a 611 votantes, y 308 de ellos dijeron haber votado por el candidato ganador (de acuerdo con datos del ICR Survey Research Group). Utilice un nivel de significancia de 0.04 para probar la aseveración de que, entre todos los votantes, el 43% dijo haber votado por el candidato ganador. (Los registros de votos revelaron que el porcentaje real que votó por el candidato ganador fue del 43%). ¿Qué sugiere el resultado acerca de las percepciones de los votantes?

9. ¿Están siendo engañados 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 onzas de cerveza de raíz en sus latas. Al seleccionar y medir aleatoriamente 24 latas, se descubre que la cantidad media es de 11.4 onzas y la desviación estándar es de 0.62 onzas. El presidente de la compañía, Harry Windsor, asevera que la muestra es demasiado pequeña para tener algún significado. Utilice los datos muestrales para probar la aseveración de que los consumidores están siendo engañados. ¿Tiene alguna validez el argumento de Harry Windsor?

10. Porcentaje de personas que piensa que Elvis está vivo *USA Today* publicó un reporte acerca de una encuesta de la Universidad de Carolina del Norte, realizada a 1248 adultos del sur de Estados Unidos. Se reportó que el 8% de los encuestados creen que Elvis Presley aún vive. El artículo comenzaba con la aseveración de que "casi uno de 10" sureños cree que Elvis aún está vivo. Con un nivel de significancia de 0.01, pruebe la aseveración de que el verdadero porcentaje es menor que el 10%. Con base en el resultado, determine si el resultado muestral del 8% justifica la frase "casi uno de cada diez".

11. ¿Es mejor la nueva máquina? La Medassist Pharmaceutical Company utiliza una máquina para llenar botellas con medicina para el resfriado, de tal modo que la desviación estándar de los pesos es de 0.15 onzas. Se prueba una nueva máquina en 71 botellas y la desviación estándar de esta muestra es de 0.12 onzas. La Dayton Machine Company, que fabrica una nueva máquina, asevera que ésta llena las botellas con menor variación. Con un nivel de significancia de 0.05, pruebe la aseveración hecha por la Dayton Machine Company. Si la máquina de Dayton se utiliza por ensayos, ¿debe considerarse su compra?

T **12.** Pesos de paquetes de azúcar Remítase al conjunto de datos 28 del Apéndice B y pruebe la aseveración de que los pesos de los paquetes de azúcar tienen una media igual a 3.5 g, ¿qué concluye?

Ejercicios de repaso acumulativos

1. Verificación de dioxina A continuación se listan las cantidades medidas de dioxina en el aire en el World Trade Center, los días posteriores a los ataques terroristas del 11 de septiembre de 2001. La dioxina incluye un grupo de químicos producidos por los incendios y ciertos tipos de manufactura. Las cantidades listadas se miden en nanogramos por metro cúbico (ng/m^3) y se presentan en orden, de tal modo que los valores registrados inicialmente se encuentran a la izquierda. Los datos fueron proporcionados por la Environmental Protection Agency de Estados Unidos.

0.161 0.175 0.176 0.032 0.0524 0.044 0.018 0.0281 0.0268

a. Calcule la media de esta muestra.
b. Calcule la mediana.
c. Calcule la desviación estándar.

continúa

d. Calcule la varianza.
e. Calcule el rango.
f. Construya un estimado del intervalo de confianza del 95% de la media poblacional.
g. La EPA (Agencia de Protección Ambiental) emplea un "nivel de verificación" de 0.16 ng/m^3, que se "establece como protección en contra del incremento significativo de los riesgos de cáncer y de otros efectos nocivos para la salud". Utilice un nivel de significancia de 0.05 para probar la aseveración de que esta muestra proviene de una población con una media menor que 0.16 ng/m^3.
h. ¿Existe alguna característica importante en los datos que no se señale en los resultados anteriores? Si es así, ¿cuál es?

2. Calificaciones de mujeres en el área de matemáticas del SAT 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.

3. PES Un estudiante de psicología diseña un experimento para probar la percepción extrasensorial (PES). En este experimento se selecciona al azar un naipe de un mazo mezclado y un sujeto vendado de los ojos debe adivinar el palo (diamantes, tréboles, corazones, espadas) del naipe seleccionado. El experimento se repite 25 veces, reemplazando el naipe y mezclando el mazo cada vez.
a. Para los sujetos que adivinan sin PES, calcule el número medio de respuestas correctas.
b. Para los sujetos que adivinan sin PES, calcule la desviación estándar del número de respuestas correctas.
c. Para los sujetos que adivinan sin PES, calcule la probabilidad de obtener más de 12 respuestas correctas.
d. Si un sujeto tiene más de 12 respuestas correctas, pruebe la aseveración de que trató de adivinar. Utilice un nivel de significancia de 0.05.
e. Usted desea realizar una encuesta para estimar el porcentaje de adultos estadounidenses que creen que algunas personas tienen PES. ¿A cuántas personas debe encuestar si busca, con un nivel de confianza del 90%, que el error en su porcentaje muestral no sobrepase cuatro puntos porcentuales?

Actividades de cooperación en equipo

1. Actividad en clase Cada estudiante debe estimar la longitud del salón de clases. Los valores deben basarse en estimados visuales, sin tomar mediciones reales. Una vez que se han reunido las estimaciones, mida la longitud de la habitación, después pruebe 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 sea aproximadamente igual a la longitud real de la habitación?

2. 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 establezca que "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?

3. Actividad en clase En un grupo de tres o cuatro personas, realice un experimento de PES, seleccionando a uno de los miembros del grupo como sujeto. Dibuje un círculo en un pequeño pedazo de papel y dibuje un cuadrado en otro papel del mismo tamaño. Repita el siguiente experimento 20 veces: seleccione aleatoriamente el círculo o el cuadrado y colóquelo en la mano del sujeto a sus espaldas, de manera que no pueda verlo; después pida al sujeto que identifique la figura (sin verla); registre si la respuesta es correcta. Pruebe la aseveración de que el sujeto tiene PES, debido a que la proporción de respuestas correctas es mayor que 0.5.

4. Actividad en clase Después de formar grupos con tamaños entre 10 y 20 individuos, cada miembro debe registrar su número de latidos cardiacos por minuto. Después de calcular $\bar{x}$ y s, que cada grupo proceda a probar 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!)

5. Actividad fuera de clase Como parte de una encuesta de Gallup, a unos sujetos se les preguntó: "¿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 6-2 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. Después divida el tamaño de muestra 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.

Proyecto tecnológico

Se dispone de STATDISK, Minitab, Excel, la calculadora TI-83 Plus y muchas otras herramientas para generar datos aleatoriamente a partir de una población que se distribuye normalmente, con una media y una desviación estándar dadas.

a. Utilice una de estas herramientas para generar cinco valores al azar de una población que se distribuye normalmente, con una media de 100 y una desviación estándar de 15 (los parámetros de una prueba típica del CI).

b. Con los cinco valores muestrales generados en el inciso *a*, pruebe la aseveración de que la muestra proviene de una población con una media igual a 100. Utilice un nivel de significancia de 0.10.

c. Repita el inciso *a* y el inciso *b* nueve veces más, de modo que se generen un total de 10 muestras diferentes y se realicen 10 pruebas de hipótesis distintas.

d. Con un nivel de significancia de 0.10, existe una probabilidad de 0.10 de cometer un error tipo I (rechazar la hipótesis nula cuando es verdadera). Por la forma en que se generan los datos muestrales, sabemos que 100 es la verdadera media poblacional, de manera que cometemos un error tipo I en este experimento cada vez que rechazamos la hipótesis nula de $\mu = 100$. Para los 10 ensayos de este experimento, ¿cuántas veces se rechazó realmente la hipótesis nula? Cuando realizamos 10 ensayos de este tipo, ¿cuántas veces esperamos que se rechace la hipótesis nula? ¿Son consistentes los resultados reales con los resultados teóricos? Explique.

A continuación se incluyen instrucciones de los pasos *a* y *b* para el uso de STATDISK, Minitab, Excel y la calculadora TI-83 Plus.

STATDISK

a. Haga clic en **Data**, después en **Normal Generator**. En el cuadro de diálogo introduzca un tamaño de muestra de 5, una media de 100, una desviación estándar de 15 y anote 0 en el número de lugares decimales. Haga clic en **Generate**.

Ahora proceda a calcular los valores de la media y la desviación estándar muestrales. Cuando aparezcan los cinco valores generados, haga clic en **Copy**, después en **Data** de la barra del menú principal. Haga clic en **Descriptive Statistics** del menú y luego en **Paste**. Los datos muestrales generados deben aparecer. Haga clic en **Evaluate**. Registre los valores de la media y la desviación estándar muestrales.

b. Haga clic en **Analysis** del menú principal y luego en **Hypothesis Testing**. Seleccione **Mean-One Sample.** En el cuadro de diálogo introduzca un nivel de significancia de 0.10, una media aseverada de 100, un tamaño de muestra de 5 e introduzca los valores de la media y la desviación estándar muestrales que registramos en el inciso *a*. Haga clic en **Evaluate** y registre el resultado (rechace la hipótesis nula o no lo haga).

Minitab

a. Haga clic en **Calc** del menú principal, después en **Random Data** y luego en **Normal**. En el cuadro de diálogo introduzca 5 para el número de renglones generados, introduzca C1 para la columna en que se almacenarán los datos, introduzca una media de 100 y una desviación estándar de 15. Haga clic en **OK**.

b. Haga clic en **Stat** del menú principal, seleccione **Basic Statistics**, después **1-Sample t.** En el cuadro de diálogo, introduzca C1 para el nombre de la variable, haga clic en el botón de Test Mean, anote 100 en el cuadro adyacente y luego haga clic en **OK**. Interprete los resultados y registre la conclusión (rechace la hipótesis nula o no lo haga).

Excel

a. 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, 5 para el número de números aleatorios, seleccione la opción de distribución Normal, introduzca una media de 100 y una desviación estándar de 15, después haga clic en OK.

b. Haga clic en **DDXL** y seleccione **Hypothesis Test** y luego **1 Var t Test.** Haga clic en el icono del lápiz e introduzca el rango de celdas que contienen los datos muestrales generados. Por ejemplo, introduzca A1:A5 para cinco valores en los renglones 1 al 5 de la columna A. Haga clic en **OK**. En el cuadro de diálogo, haga clic en la barra en Step 1 y proceda a introducir la media aseverada de 100. Proceda a hacer clic en las barras, en los pasos restantes. Después de hacer clic en **Compute**, registre la conclusión (rechace la hipótesis nula o no lo haga).

TI-83 Plus

a. 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 100,15,5 y presione **ENTER**. Guarde los datos muestrales generados presionando **STO L1**, seguido por la tecla **ENTER**.

b. Presione **STAT**, seleccione **TESTS** y luego **2:T-Test** y presione **ENTER**. Seleccione **Data** (debido a que tenemos los datos generados en la lista L1), introduzca 100 para la media aseverada y proceda a obtener los resultados de la prueba de hipótesis. Interprete los resultados y registre la conclusión (rechace la hipótesis nula o no lo haga).

de los DATOS a la DECISIÓN

Pensamiento crítico: cuestionamiento de los resultados de encuesta

Como las encuestas se han extendido tanto en nuestra sociedad, cada uno de nosotros debería desarrollar la capacidad para pensar críticamente acerca de ellas. Debemos cuestionarnos los procesos de selección de los sujetos encuestados, la redacción de la pregunta, la significancia de los resultados, la objetividad del patrocinador de la encuesta y del grupo que la realiza, así como otros aspectos. Considere el siguiente artículo hipotético de un periódico.

> **"Una encuesta prueba que la mayoría de los estadounidenses no hacen trampa en los impuestos"** En *Every Day is Saturday*, una revista para personas jubiladas, se reportó que una encuesta realizada a 250 sujetos seleccionados al azar, incluyó a un 55% que dijo no hacer trampa al pagar los impuestos. La revista publicó la encuesta en su ejemplar de marzo, y las 250 respuestas se seleccionaron aleatoriamente de las encuestas que los lectores enviaron por correo.

La primera pregunta de la encuesta decía: "¿Hace usted trampa en los impuestos o es honesto?". La revista afirmó que "es motivante saber que la mayoría de los estadounidenses son honestos, al menos cuando se trata de llenar las formas de impuestos".

Análisis de los resultados

a. Utilice los métodos de este capítulo para probar la aseveración de que "la mayoría de los estadounidenses no hacen trampa en los impuestos". ¿Qué concluye? ¿Será posible que los resultados de una encuesta prueben que la mayoría de los estadounidenses no hacen trampa en sus impuestos?

b. Además de los resultados de la prueba de hipótesis, existen al menos otros cuatro aspectos importantes que afectan la validez de los resultados de las encuestas. Identifique los otros aspectos y describa de qué manera afectan la validez de los resultados.

PROYECTO DE INTERNET

Prueba de hipótesis

Este capítulo introdujo la metodología para la prueba de hipótesis, una técnica esencial para la estadística inferencial. Este proyecto de Internet requerirá que realice pruebas con el uso de 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 al sitio Web 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, y 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, que prueba productos y servicios, y proporciona calificaciones y recomendaciones a los consumidores en la revista Consumer Reports.

Nota del autor: El autor se reunió con Mike Saccucci y los otros estadísticos 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 estadísticos 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.

¿Qué conceptos y procedimientos estadísticos utiliza en Consumers Union?

Cada día, los estadísticos tienen que utilizar diversos 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 estuviesen 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 estadístico puede necesitar construir un diseño completamente aleatorizado, un diseño aleatorizado por bloques o algún otro tipo de diseño experimental para asegurarse de que nuestros resultados son precisos y sin sesgos. Durante la fase de análisis, el estadístico emplea diversas técnicas, tales como el análisis de varianza, el análisis de regresión, el análisis de series de tiempo, el análisis categórico y/o análisis no paramétricos.

¿Qué hacen los estadísticos en Consumers Union?

Los estadísticos realizan gran variedad de tareas. En las primeras etapas de un proyecto, el estadístico trabaja con el equipo del proyecto para desarrollar el protocolo de prueba y ayudar a seleccionar los productos que van a probarse. Después, el estadístico ayuda a crear un diseño experimental adecuado para emplearse durante la prueba. Una vez que se han obtenido los datos de prueba, el estadístico analiza los resultados y presenta los hallazgos en un reporte estadístico. El estadístico también se incorpora 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 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 localizados a lo largo de todo 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 es práctico, los artículos que se prueban se codifican de forma ciega, de tal manera que los probadores no saben qué marcas están evaluando.

¿Las recomendaciones y calificaciones de la revista Consumer Reports sólo están basadas 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 que no hay una diferencia importante en los resultados de la prueba, no consideramos una marca mejor que otra. Por ejemplo, al probar selladores de agua, podríamos encontrar que existe una diferencia estadísticamente significativa entre las cantidades de agua que se escurre a través de dos marcas diferentes de sellador. Sin embargo, si la diferencia son unas cuantas gotas de agua, calificaríamos de forma similar a los productos respecto a dicha 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, creo que el conocimiento básico de la estadística se considera favorablemente en casi cualquier campo de estudio. Esto es verdad sobre todo en las áreas cuantitativas, tales como ciencias, ingeniería y 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 para llevar a cabo sus responsabilidades con excelencia?

La misión de Consumers Union es adelantarse a los intereses de los consumidores al proporcionar 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 en extremo valioso para ayudar a que Consumers Union logre esta misión.

Cuando era estudiante universitario, ¿esperaba utilizar la estadística en su trabajo?

Inicié mi carrera en matemáticas y realmente no me interesé en la estadística sino hasta el último año de la carrera. Fue en el posgrado, mientras trabajaba bajo la dirección del profesor Hoerl en la Universidad de Delaware, que me di cuenta de lo interesante que sería una carrera en estadística. A pesar de los sentimientos negativos que muchos estudiantes tienen por la estadística, yo creo tener uno de los trabajos más interesantes. Nunca sé qué esperar durante el día. Un día puedo estar sentado en una sesión de entrenamiento sobre cata de vinos para aprender acerca de los procedimientos de prueba. Otro día puedo estar discutiendo diversas formas para probar pinturas. Sin embargo, la mayor parte de los días paso gran parte del tiempo frente una computadora para diseñar un futuro estudio o buscando grandes cantidades de datos que se utilizarán finalmente como base de las evaluaciones de productos.

8

Inferencias a partir de dos muestras

Uso de la estadística para identificar la discriminación racial

La *discriminación racial* es la práctica polémica de señalar que alguien manifiesta una conducta criminal con base en la raza, nación de procedencia o grupo étnico al que pertenece esa persona. Ocurren ejemplos de discriminación racial cuando se detiene desproporcionadamente a más negros que blancos para recibir multas de tránsito, o cuando se detiene desproporcionadamente en los aeropuertos a más gente procedente de Medio Oriente para efectuar revisiones meticulosas.

Considere los datos seleccionados al azar de la tabla 8-1 para conductores detenidos por la policía en un año reciente (de acuerdo con datos del Departamento de Justicia de Estados Unidos). Podría argumentarse que se detuvo a muchos más blancos que negros. Sin embargo, la población incluye muchos más blancos que negros, por lo tanto no tiene mucho sentido comparar los 147 conductores blancos detenidos con los 24 conductores negros. Cualquier comparación debería tomar en cuenta las *proporciones* en las que se detiene a blancos y a negros.

También podría aseverarse que la tasa del 12.0% de los negros no es *significativamente* mayor que la tasa del 10.5% de los blancos. Esta afirmación será puesta a prueba en este capítulo. Probaremos la aseveración de que parece que la proporción de negros detenidos es *mayor* que la proporción de blancos detenidos, con base en las proporciones muestrales de 24/200 para los negros y 147/1400 para los blancos. Utilizaremos procedimientos estadísticos que son muy importantes para temas como la discriminación racial.

Tabla 8-1 Datos de discriminación racial

	Raza y grupo étnico	
	Negros y no hispanos	Blancos y no hispanos
Conductores detenidos por la policía	24	147
Número total de conductores observados	200	1400
Porcentaje detenido por la policía	**12.0%**	**10.5%**

8-1 Panorama general

El capítulo 6 introdujo una importante actividad de la estadística inferencial: se utilizaron datos muestrales para construir *estimados* de intervalos de confianza de parámetros poblacionales. El capítulo 7 introdujo una segunda actividad importante de la estadística inferencial: se utilizaron datos muestrales para *probar hipótesis* acerca de parámetros poblacionales. En los capítulos 6 y 7, todos los ejemplos y los ejercicios implicaron el uso de *una* muestra para hacer una inferencia acerca de *una* población. En la realidad, sin embargo, existen muchas situaciones importantes en las que es necesario comparar *dos* conjuntos de datos muestrales. Los siguientes son ejemplos típicos de los que se incluyen en este capítulo, que presentan métodos para utilizar datos muestrales a partir de dos poblaciones de modo que puedan hacerse inferencias acerca de éstas.

- Cuando se prueba una aseveración de discriminación racial para determinar si la proporción de conductores negros detenidos por la policía es mayor que la proporción de conductores blancos detenidos por la policía.
- Cuando se prueba la eficacia de la vacuna de Salk en la prevención de la poliomielitis paralítica para determinar si el grupo de tratamiento tiene una menor incidencia de la enfermedad que el grupo que recibió un placebo.
- Cuando se investiga la precisión de la estatura reportada por personas para determinar si existe una diferencia significativa entre las estaturas reportadas y las estaturas reales medidas.

Los capítulos 6 y 7 incluyen métodos que se aplicaron a proporciones, medias y medidas de variación (desviación estándar y varianza). Este capítulo abordará estos mismos parámetros y aplicará los mismos métodos introducidos en los capítulos 6 y 7 a situaciones que requieren realizar comparaciones entre dos muestras, en lugar de estudiar una sola.

8-2 Inferencias acerca de dos proporciones

Existen muchas situaciones importantes y reales en las que es necesario utilizar datos muestrales para comparar dos proporciones poblacionales. De hecho, podría argumentarse enfáticamente que esta sección es una de las más importantes secciones en el libro puesto que es donde describimos métodos para tratar con dos proporciones muestrales. Si bien esta sección se basa en proporciones, podemos 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 hecha acerca de dos proporciones poblacionales o cuando se construye un intervalo de confianza para la diferencia entre dos proporciones poblacionales, partimos de los siguientes supuestos y utilizamos la siguiente notación.

Supuestos

1. Tenemos proporciones de dos muestras aleatorias simples que son *independientes*, lo cual quiere decir que los valores muestrales seleccionados de una población no están relacionados ni apareados de ninguna forma con los valores muestrales seleccionados de la otra población.
2. Para ambas muestras, las condiciones $np \geq 5$ y $nq \geq 5$ se satisfacen. Esto es, existen al menos cinco éxitos y cinco fracasos en cada una de las dos muestras. (En muchos casos, probaremos la aseveración de que dos poblaciones tienen proporciones iguales para que $p_1 - p_2 = 0$. Puesto que asumimos que $p_1 - p_2 = 0$, no es necesario especificar el valor particular que p_1 y p_2 tienen en común. En casos como éste, las condiciones $np \geq 5$ y $nq \geq 5$ se revisan reemplazando p con el estimado apareado de la proporción $\overline{p}$, lo cual se analizará después).

Notación para dos proporciones

Para la población 1 permitimos que

p_1= proporción *poblacional*

n_1= tamaño de la muestra

x_1= número de éxitos en la muestra

$\hat{p} = \dfrac{x_1}{n_1}$ (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.

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 otras es necesario calcular los valores de x_1 y x_2.

Por ejemplo, considere la afirmación de que "cuando 734 hombres fueron tratados con Viagra, el 16% de ellos experimentaron dolores de cabeza". A partir de esta afirmación podemos ver que $n_1 = 734$ y $\hat{p}_1 = 0.16$, pero no está dado el número real de éxitos x_1. Sin embargo, de $\hat{p}_1 = x_1/n_1$, sabemos que

$$x_1 = n_1 \cdot \hat{p}_1$$

de modo que $x_1 = 734 \cdot 0.16 = 117.44$. Pero usted no puede tener 117.44 hombres que experimentaron dolor de cabeza, puesto que cada uno o experimentó un dolor de cabeza o no lo hizo y, por lo tanto, el número de éxitos x_1 debe ser un número entero. Podemos redondear 117.44 a 117. Ahora podemos utilizar $x_1 = 117$ en los cálculos que requieran de este valor. En realidad es bastante sencillo: el 16% de 734 quiere decir 0.16×734, que da como resultado 117.44, que redondeamos a 117.

Pruebas de hipótesis

En la sección 7-2 analizamos pruebas de hipótesis acerca de una sola proporción poblacional. Ahora consideraremos pruebas de hipótesis acerca de dos proporciones poblacionales, pero *sólo probaremos la aseveración de que* $p_1 = p_2$, y utilizaremos el siguiente estimado agrupado (o combinado) del valor que p_1 y p_2 tienen en común. (Para aseveraciones de que la diferencia entre p_1 y p_2 es igual a una constante que no sea cero, véase el ejercicio 34 en esta sección). Por la forma del estimado apareado $\bar{p}$ usted puede ver que éste básicamente combina las dos muestras diferentes en una gran muestra.

Estimado apareado de p_1 y p_2

El **estimado apareado de p_1 y p_2** se denota por $\bar{p}$ y está dado por

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Denotamos el complemento de $\bar{p}$ por $\bar{q}$, entonces $\bar{q} = 1 - \bar{p}$.

Prueba estadística para dos proporciones (con H_0: $p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\,\bar{q}}{n_1} + \dfrac{\bar{p}\,\bar{q}}{n_2}}}$$

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. (Utilice el valor calculado del estadístico de prueba z y obtenga el valor P siguiendo el procedimiento resumido en la figura 7-6).

Valores críticos: Utilice la tabla A-2. (Con base en el nivel de significancia α, obtenga valores críticos utilizando los procedimientos introducidos en la sección 7-2).

Una vez más, el estadístico de prueba se ajusta al formato común de

$$\frac{(\text{estadístico muestral}) - (\text{valor aseverado del parámetro})}{(\text{desviación estándar de los estadísticos muestrales})}$$

El ejemplo siguiente ayudará a aclarar los papeles de x_1, n_1, $\hat{p}_1$ $\bar{p}$, etcétera. En particular, usted debe reconocer que bajo el supuesto de proporciones iguales, el mejor estimado de la proporción común se obtiene agrupando ambas muestras en una gran muestra, para que $\bar{p}$ se vuelva un estimado más evidente de la proporción poblacional común.

EJEMPLO Discriminación racial Para los datos muestrales listados en la tabla 8-1, utilice un nivel de significancia de 0.05 para probar la aseveración de que la proporción de conductores negros detenidos por la policía es mayor que la proporción de conductores blancos detenidos.

SOLUCIÓN Para fines de notación estipulamos que la muestra 1 es el grupo de conductores negros y la muestra 2 es el grupo de conductores blancos. Podemos resumir los datos muestrales como sigue.

Conductores negros	Conductores blancos
$n_1 = 200$	$n_2 = 1400$
$x_1 = 24$	$x_2 = 147$
$\hat{p}_1 = \dfrac{x_1}{n_1} = \dfrac{24}{200} = 0.120$	$\hat{p}_2 = \dfrac{x_2}{n_2} = \dfrac{147}{1400} = 0.105$

Ahora utilizaremos el método de prueba de hipótesis del valor *P*, como se resumió en la figura 7-9.

Paso 1: La aseveración de una proporción mayor de conductores negros se representa por $p_1 > p_2$.

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

$$H_0: p_1 = p_2 \qquad H_1: p_1 > p_2 \qquad \textbf{(aseveración original)}$$

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. Tenemos dos muestras independientes, y las condiciones $np \geq 5$ y $nq \geq 5$ se satisfacen para cada una de las dos muestras. Para corroborar esto señalamos que al realizar esta prueba suponemos que $p_1 = p_2$, cuyo valor común es el estimado apareado $\bar{p}$ calculado como se muestra abajo, con espacios decimales extras utilizados para minimizar los errores de redondeo en cálculos posteriores.

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{24 + 147}{200 + 1400} = 0.106875$$

Con $\bar{p} = 0.106875$, se sigue que $\bar{q} = 1 - 0.106875 = 0.893125$.

continúa

Experimento de 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. A aproximadamente 200,000 niños se les inyectó una solución salina inocua, y a otros 200,000 se les inyectó la vacuna. El experimento fue "doble ciego" porque los niños inyectados no sabían si estaban recibiendo la vacuna real o el placebo, y 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 la enfermedad. Un análisis estadístico de estos y otros resultados llevó a la conclusión de que la vacuna de Salk realmente era efectiva contra la poliomielitis paralítica.

Verificamos que $np \geq 5$ y $nq \geq 5$ para ambas muestras como se indica abajo, con p estimada por medio de $\bar{p}$ y q estimada por medio de $\bar{q}$.

Muestra 1	Muestra 2
$n_1 p = (200)(0.106875) = 21.375 \geq 5$	$n_2 p = (1400)(0.106875) = 149.625 \geq 5$
$n_1 q = (200)(0.893125) = 178.625 \geq 5$	$n_2 q = (1400)(0.893125) = 1250.375 \geq 5$

Paso 6: Ahora podemos calcular el valor del estadístico de prueba.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\,\bar{q}}{n_1} + \frac{\bar{p}\,\bar{q}}{n_2}}}$$

$$= \frac{\left(\frac{24}{200} - \frac{147}{1400}\right) - 0}{\sqrt{\frac{(0.106875)(0.893125)}{200} + \frac{(0.106875)(0.893125)}{1400}}} = 0.64$$

El valor P de 0.2611 se calcula como sigue: se trata de una prueba de cola derecha, entonces el valor P es el área ubicada a la derecha del estadístico de prueba $z = 0.64$ (véase la figura 7-6). Remítase a la tabla A-2 y encuentre que el área a la derecha del estadístico de prueba $z = 0.64$ es 0.7389, entonces el valor P es $1 - 0.7389 = 0.2611$. (Los programas de cómputo demuestran que un valor P más exacto es 0.2603). El estadístico de prueba y el valor P se incluyen en la figura 8-1a.

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 de $p_1 = p_2$.

INTERPRETACIÓN Debemos retomar la afirmación original de que los conductores negros son detenidos en una mayor proporción que los conductores blancos. Puesto que no rechazamos la hipótesis nula, concluimos que no existe

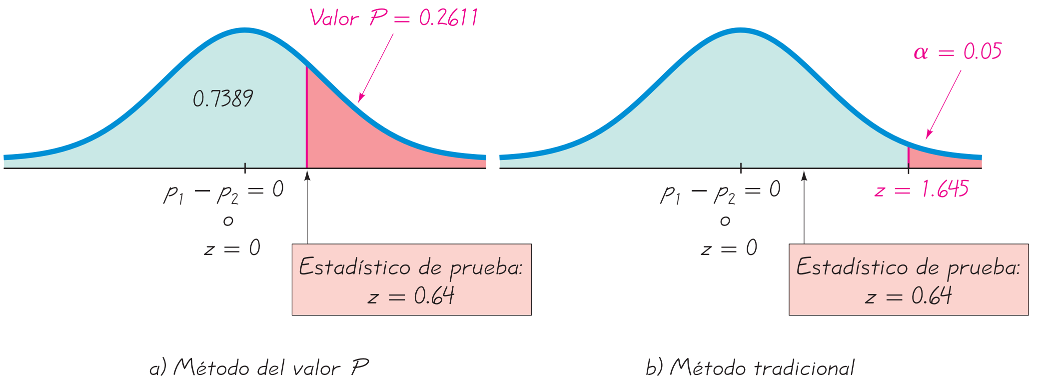


FIGURA 8-1
Prueba de la aseveración de que $p_1 > p_2$

evidencia suficiente para fundamentar la aseveración de que la proporción de conductores negros detenidos por la policía es mayor que la de los conductores blancos. (Véase la figura 7-7 para la redacción de la conclusión final). Esto *no* significa que queda probado que no hay discriminación racial; significa únicamente que la evidencia todavía no es lo suficientemente fuerte como para concluir que la proporción del 12.0% de detención de conductores negros es significativamente mayor que la proporción del 10.5% de detención de conductores blancos. Con más datos, la evidencia podría ser lo suficientemente fuerte (véase el ejercicio 33). De hecho, conjuntos de datos más grandes que los utilizados en el ejemplo sugieren que la discriminación racial sucedió.

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 lugar 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 un área de $\alpha = 0.05$ en la cola derecha corresponde al valor crítico de $z = 1.645$. Véase la figura 8-1b donde podemos observar que el estadístico de prueba no cae en la región crítica limitada por el valor crítico de $z = 1.645$. Una vez más no rechazamos la hipótesis nula y concluimos que no hay suficiente evidencia para sustentar la aseveración de que los conductores negros son detenidos en una mayor proporción que los conductores blancos.

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 que aparece en la página 444. 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, recomendamos no utilizar un estimado del intervalo de confianza de $p_1 - p_2$ como base para probar la aseveración de que $p_1 = p_2$, por las siguientes razones.

No utilice un intervalo de confianza para probar la aseveración de que $p_1 = p_2$ (puesto que la desviación estándar que se usa para los intervalos de confianza es diferente de la desviación estándar que se usa para la prueba de hipótesis que emplea el método del valor P o el método tradicional). Cuando se prueban aseveraciones acerca de la diferencia entre dos proporciones poblacionales, el método tradicional y el método del valor P son equivalentes en el sentido de que siempre proporcionan los mismos resultados, pero el estimado del intervalo de confianza de la diferencia puede sugerir una conclusión diferente (véase el ejercicio 32). Si se obtienen diferentes conclusiones, comprenda 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. Utilice esta estrategia: si usted quiere 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.

¿Ayuda la aspirina a prevenir ataques cardiacos?

En un estudio reciente realizado a 22,000 médicos, a la mitad de ellos se les administraron dosis normales de aspirina mientras que a la otra mitad se le dieron placebos. El estudio duró seis años y tuvo un costo de 4.4 millones de dólares. Entre aquellos que tomaron la aspirina, 104 sufrieron ataques cardiacos. Entre aquellos que tomaron los placebos, 189 sufrieron ataques cardiacos. (Estas cifras están basadas en datos de *Time* y *New England Journal of Medicine*, vol. 318, núm. 4). Éste es un experimento clásico que incluye 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 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 cardiacos en el grupo muestral que tomó aspirina.

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, *The American Statistician*, vol. 55, núm. 3). Véase el ejercicio 31.

Estimado del intervalo de confianza de $p_1 - p_2$

El estimado del intervalo de confianza de la diferencia $p_1 - p_2$ 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 $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

EJEMPLO Discriminación racial Utilice los datos muestrales que se presetnan en la tabla 8-1 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 prueba de hipótesis de cola derecha anterior. Véase la tabla 7-2 en la sección 7-2).

SOLUCIÓN 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 muestra.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 1.645\sqrt{\frac{\left(\frac{24}{200}\right)\left(\frac{176}{200}\right)}{200} + \frac{\left(\frac{147}{1400}\right)\left(\frac{1253}{1400}\right)}{1400}} = 0.040$$

Con $\hat{p}_1 = 24/200 = 0.120$, $\hat{p}_2 = 147/1400 = 0.105$ y $E = 0.040$, el intervalo de confianza se evalúa como sigue.

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

$$(0.120 - 0.105) - 0.040 < (p_1 - p_2) < (0.120 - 0.105) + 0.040$$

$$-0.025 < (p_1 - p_2) < 0.055$$

INTERPRETACIÓN Los límites del intervalo de confianza contienen a 0, lo que sugiere que no existe una diferencia significativa entre las dos proporciones. Sin embargo, si el objetivo es probar la igualdad de las dos proporciones poblacionales, debemos usar el valor P o el método tradicional de prueba de hipótesis; no debemos basar la decisión en el intervalo de confianza.

El autor como testigo

El autor fue requerido para testificar en la Suprema Corte del estado de Nueva York por un ex alumno que impugnaba una reelección perdida a la oficina del Dutchess County Clerk. 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 preguntó acerca de 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, defendiendo la afirmación del ex alumno, y ordenó una nueva elección en el distrito impugnado. Este juicio después fue derribado por la corte de apelación con base en que las irregularidades de la votación debían haber sido impugnadas antes de la elección, no después.

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 $n_1p_1 \geq 5$ y $n_1q_1 \geq 5$, la distribución de $\hat{p}_1$ puede aproximarse con una distribución normal con media p_1, desviación estándar $\sqrt{p_1q_1/n_1}$ y varianza p_1q_1/n_1. Estas conclusiones se basan en las secciones 5-6 y 6-2, 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 aproximará por medio de una distribución normal con media $p_1 - p_2$ y varianza

$$\sigma^2_{(\hat{p}_1-\hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2} = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}$$

   (El resultado de arriba se basa en esta propiedad: la varianza de las *diferencias* entre dos variables aleatorias independientes es la suma de sus varianzas individuales. Véase el ejercicio 38).
3. Puesto que los valores de p_1, q_1, p_2, y q_2 suelen ser desconocidos, y a partir de la hipótesis nula suponemos que $p_1 = p_2$, podemos agrupar (o combinar) los datos muestrales. El estimado agrupado 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 muestra 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 ese supuesto (como en la construcción del intervalo de confianza), estimamos la varianza de $p_1 - p_2$ como

$$\sigma^2_{(\hat{p}_1-\hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2} = \frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}$$

y la desviación estándar se vuelve

$$\sqrt{\frac{\hat{p}_1\hat{q}_2}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

En el estadístico de prueba

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}}$$

utilice los valores positivo y negativo de z (para dos colas) y resuelva para $p_1 - p_2$. Los resultados son los límites del intervalo de confianza que se dieron antes.

El margen de error líder

Los autores Stephen Ansolabehere y Thomas Belin escribieron en su artículo "Poll Faulting" (revista *Chance*) lo siguiente: "Nuestra mayor crítica al reporte de los resultados de la encuesta es para el margen de error de una proporción individual (±3%), tomando en cuenta que la atención de los medios de comunicación está claramente dirigida al *liderazgo* de un candidato". Ellos señalan que el liderazgo es realmente la *diferencia* entre dos proporciones ($p_1 - p_2$) y proceden a explicar 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 ±3% se convierte en alrededor de ±5% por el liderazgo de un candidato sobre el otro. Ellos afirman que debe reportarse el margen de error para el liderazgo.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego seleccione **Hypothesis Testing**, luego **Proportion-Two Samples**. Introduzca los elementos requeridos en el cuadro de diálogo. Los límites del intervalo de confianza se incluyen con los resultados de la prueba de hipótesis.

Minitab Ahora Minitab maneja estadísticos resumidos para dos muestras. Seleccione **Stat** de la barra del menú principal, luego seleccione **Basic Statistics**, luego **2 Proportions**. Haga clic en el botón para **Summarize Data**. Haga clic en la barra de **Options**. Introduzca el nivel de confianza deseado, introduzca 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 es un suplemento de este libro. Primero haga estas entradas: en la celda A1 indique 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 anote 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 Plus La calculadora TI-83 Plus se utiliza para pruebas de hipótesis e intervalos de confianza. Oprima **STAT** y seleccione **TESTS**. Luego escoja 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 Plus mostrará en la pantalla un valor P en lugar de valores críticos, por lo tanto se utiliza el método del valor P para prueba de hipótesis.

8-2 Destrezas y conceptos básicos

Cálculo del número de éxitos. *En los ejercicios 1 a 4, calcule el número de éxitos x sugeridos por la afirmación dada.*

1. Del Arizona Department of Weights and Measures: De 37 inspecciones a las tiendas de NAPA Auto Parts, el 81% falló.

2. Del *New York Times*: De 240 guantes de vinilo sujetos a pruebas de tensión, el 63% de ellos presentaron filtración de virus.

3. De *Sociological Methods and Research*: Cuando se encuestó a 294 residentes de una ciudad central, el 28.9% se rehusó a responder.

4. De una encuesta de Time/CNC: El 24% de 205 mujeres solteras dijeron que ellas "definitivamente quieren contraer matrimonio".

Cálculos para probar aseveraciones. *En los ejercicios 5 y 6, 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.*

5.

Trabajadores	Jefes
$n_1 = 436$	$n_2 = 121$
$x_1 = 192$	$x_2 = 40$

6.

Actividad baja	Actividad alta
$n_1 = 10{,}239$	$n_2 = 9877$
$x_1 = 101$	$x_2 = 56$

7. Correo electrónico y privacidad Una encuesta de 436 trabajadores mostró que 192 dijeron que vigilar el correo electrónico de los empleados era un grave atentado contra la ética. Cuando 121 jefes de alto nivel fueron encuestados, 40 dijeron que vigilar el correo electrónico de los empleados era un grave atentado contra la ética (según datos de una encuesta de Gallup). Utilice un nivel de significancia de 0.05 para probar que en la aseveración de aquellos que dijeron que vigilar el correo electrónico es un grave atentado contra la ética, la proporción de empleados es mayor que la proporción de jefes.

8. Correo electrónico y privacidad Remítase a los datos muestrales dados en el ejercicio 7 y construya un estimado del intervalo de confianza del 90% de la diferencia entre las dos proporciones poblacionales. ¿Existe una diferencia sustancial entre los empleados y los jefes?

9. El ejercicio y el trastorno cardiaco coronario En un estudio de mujeres y el trastorno cardiaco coronario se obtuvieron los siguientes resultados muestrales: de 10,239 mujeres con un bajo nivel de actividad física (menos de 200 kcal/semana), hubo 101 casos de trastorno cardiaco coronario. De 9877 mujeres con actividad física medida de entre 200 y 600 kcal/semana, hubo 56 casos de trastorno cardiaco coronario (según datos de "Physical Activity and Coronary Heart Disease in Women", de Lee, Rexrode *et al*., *Journal of the American Medical Association*, vol. 285, núm. 11). Construya un estimado del intervalo de confianza del 90% para la diferencia entre las dos proporciones. ¿La diferencia parece ser sustancial? ¿Parece que la actividad física corresponde a una menor proporción de trastorno cardiaco coronario?

10. El ejercicio y el trastorno cardiaco coronario Remítase a los datos muestrales del ejercicio 9 y utilice un nivel de significancia de 0.05 para probar la aseveración de que el porcentaje de trastornos cardiacos coronarios es más alto para las mujeres que tienen los niveles más bajos de actividad física. ¿Qué sugiere la conclusión?

11. Repetición instantánea en fútbol En la temporada de fútbol del 2000, los oficiales revisaron 247 juegos, utilizando repeticiones instantáneas de video, y en 83 de ellos se revocó la decisión original. En la temporada de fútbol del 2001, se revisaron 258 juegos y 89 de ellos se revocaron (datos tomados de "Referees Turn to Video Aid More Often", de Richard Sandomir, *New York Times*). ¿Existe una diferencia significativa en las dos tasas de revocación? ¿Parece que la tasa de revocación fue la misma en ambos años?

12. Eficacia de las prohibiciones de fumar La Joint Commission on Accreditation of Healthcare Organizations mandó que se prohibiera fumar en los hospitales en 1994. En un estudio de los efectos de esta prohibición, se seleccionaron al azar sujetos fumadores de dos poblaciones diferentes. De 843 empleados fumadores en hospitales con la prohibición de fumar, 56 dejaron de fumar un año después de la prohibición. De 703 empleados fumadores en lugares de trabajo sin una prohibición de fumar, 27 dejaron de fumar un año después de 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 proporción de quienes dejaron de fumar?

13. Prueba de eficacia de vacuna En un artículo de *USA Today* acerca de una vacuna experimental en aerosol nasal para niños, se presentó la siguiente afirmación: "En un ensayo 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 refiere un estudio que afirma que el aerosol nasal experimental "corta las posibilidades de que los niños contraigan gripe". ¿Existe suficiente evidencia muestral que fundamente la aseveración establecida?

14. Daltonismo En un estudio de daltonismo en hombres y mujeres se seleccionaron aleatoriamente y se examinaron 500 hombres y 2100 mujeres. Entre los hombres, 45 tenían daltonismo. Entre las mujeres, 6 tenían daltonismo (según datos del *USA Today*).

continúa

a. ¿Existe evidencia suficiente que sustente la aseveración de que los hombres tienen una tasa más alta de daltonismo que las mujeres? Utilice un nivel de significancia de 0.01.

b. Construya el intervalo de confianza del 98% para la diferencia entre las tasas de daltonismo de hombres y mujeres. ¿Parece haber una diferencia sustancial?

c. ¿Por qué sería el tamaño de la muestra para las mujeres mucho más grande que el tamaño de la muestra para los hombres?

15. Cinturones de seguridad y tiempo de hospitalización Se realizó un estudio de 413 niños hospitalizados a causa de choques de vehículos motorizados. De 290 niños que no utilizaban cinturones de seguridad, 50 se hirieron de gravedad. De 123 niños que usaban cinturones de seguridad, 16 se hirieron de gravedad (datos tomados de “Morbidity Among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts”, de Osberg y Di Scala, *American Journal of Public Health*, vol. 82, núm. 3). ¿Existe suficiente evidencia muestral para concluir, a un nivel de significancia de 0.05, que la tasa de heridas graves es menor para los niños que usaban cinturones de seguridad? Con base en estos resultados, ¿qué acción se debería tomar?

16. Alcoholismo y crimen Karl Pearson, quien desarrolló muchos conceptos importantes en estadística, recolectó datos de crímenes en 1909. De aquellos convictos por provocar incendios, 50 eran bebedores y 43 abstemios. De aquellos convictos por fraude, 63 eran bebedores y 144 abstemios. Utilice un nivel de significancia de 0.01 para probar la aseveración de que la proporción de bebedores entre los incendiarios convictos es mayor que la proporción de bebedores entre aquellos convictos por fraude. ¿Parece razonable que el alcoholismo podría tener un efecto en el tipo de delito? ¿Por qué?

17. Interpretación de pantalla de computadora Un reporte del Departamento de Justicia de Estados Unidos (NCJ-156831) incluyó la aseveración de que “en casos de asesinato conyugal, las esposas acusadas tuvieron menor probabilidad de ser declaradas culpables que los maridos acusados”. Los datos muestrales consistieron en 277 sentencias de culpabilidad entre 318 maridos acusados, y 155 sentencias de culpabilidad entre 222 esposas acusadas. Pruebe la aseveración establecida e identifique una posible explicación para el resultado. Aquí se muestran los resultados de Minitab.

Sample	X	N	Sample p
1	277	318	0.871069
2	155	222	0.698198

```
Estimate for p(1) − p(2): 0.172871
95% lower bound for p(1) − p(2): 0.113511
Test for p(1) − p(2) = 0 (vs > 0): Z = 4.94 P-value = 0.000
```

TI-83 Plus

```
2-PropZTest
 p1<p2
 z=-6.741605792
 p=7.88E-12
 p̂1=1.65E-4
 p̂2=5.75E-4
↓p̂=3.7E-4
```

18. Eficacia de la vacuna de Salk para la poliomielitis En los experimentos preliminares de la vacuna de Salk, 33 de 200,000 niños vacunados presentaron poliomielitis después. De 200,000 niños vacunados con un placebo, 115 presentaron poliomielitis después. Aquí se muestra la pantalla de la calculadora TI-83 Plus. A un nivel de significancia de 0.01, pruebe la aseveración de que la vacuna de Salk es eficaz en la reducción de la tasa de poliomielitis. ¿Parece que la vacuna es eficaz?

19. Inspecciones con resultado no aprobatorio Efectuando revisiones a tiendas de partes de automóvil, el Arizona Department of Weights and Measures realizó 100 inspecciones a tiendas de Autozone y encontró que el 63% de esas inspecciones no aprobaron la norma. Entre 37 inspecciones a las tiendas de NAPA Auto Parts, el 81% no aprobó. Utilice un nivel de significancia de 0.05 para determinar si existe una diferencia significativa entre estas dos tasas de no aprobación. ¿Parece que alguna tienda es una mejor opción para los consumidores?

20. Factor de carga de aerolínea En un año reciente, la Southwest Airlines tenía 3,131,727 asientos de avión disponibles en todos sus vuelos, y 2,181,604 de ellos se ocuparon por pasajeros. La America West tuvo 2,091,859 asientos disponibles, y 1,448,255 de ellos se ocuparon. El porcentaje de asientos ocupados se conoce como *factor de carga*, por lo tanto estos resultados muestran que el factor de carga es del 69.7% (redondeado) para la Southwest Airlines y de 69.2% (redondeado) para la America West. (Los datos son del Departamento del Transporte de Estados Unidos). Responda lo siguiente, suponiendo que los resultados provienen de muestras seleccionadas aleatoriamente.

a. Pruebe la aseveración de que ambas aerolíneas tienen el mismo factor de carga.

b. Como el 69.7% y el 69.2% parecen ser evidentemente muy cercanos, ¿cómo explica usted los resultados del inciso *a*?

c. Generalice el punto clave de este ejemplo completando la siguiente oración: "Si dos tamaños muestrales son extremadamente grandes, incluso las aparentes pequeñas diferencias en las proporciones muestrales. . .".

21. Actitudes hacia el matrimonio En una encuesta de Time/CNN, el 24% de 205 mujeres solteras dijeron que "definitivamente querían casarse". En la misma encuesta, el 27% de 260 hombres solteros dieron esta misma respuesta. Construya un estimado del intervalo de confianza del 99% de la diferencia entre las proporciones de mujeres solteras y hombres solteros que definitivamente quieren casarse. ¿Existe una diferencia de género en este tema?

22. Actitudes hacia el matrimonio Remítase a los mismos datos muestrales del ejercicio 21 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 quieren casarse. ¿Parece existir una diferencia?

23. Crimen violento y grupo de edad El nuevo director asignado de la agencia estatal de salud mental asevera que una proporción más pequeña de los crímenes cometidos por personas menores de 21 años de edad son crímenes violentos (cuando se comparó a los crímenes cometidos por personas de 21 años de edad o mayores). De 2750 arrestos de criminales menores de 21 años de edad seleccionados al azar, el 4.25% se relacionaba con crímenes violentos. De 2200 arrestos de criminales de 21 años de edad o mayores seleccionados al azar, el 4.55% implicaba crímenes violentos (con base en datos de *Uniform Crime Reports*). Construya un intervalo de confianza del 95% para la diferencia entre las dos proporciones de crímenes violentos. ¿Indica el intervalo de confianza que no existe una diferencia significativa entre las dos tasas de crímenes violentos?

24. 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. A 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.

25. Encuesta escrita y encuesta por computadora En un estudio de 1700 jóvenes en el rango de 15 a 19 años de edad, la mitad de ellos recibieron encuestas escritas y a la otra mitad se le aplicó una encuesta utilizando un programa de computadora anónimo. Entre aquellos a los que se les dieron encuestas escritas, el 7.9% dijo que portó una pistola en los últimos 30 días. Entre aquellos a quienes se encuestó por computadora, el 12.4% dijo que portó una pistola en los últimos 30 días (según datos del Urban Institute).

a. Los porcentajes muestrales de 7.9% y 12.4% evidentemente no son iguales, pero ¿es significativa la diferencia? Explique.

b. Construya un estimado el intervalo de confianza del 99% de la diferencia entre los dos porcentajes poblacionales e interprete el resultado.

26. Reacciones adversas a los fármacos El medicamento Viagra se ha vuelto bastante conocido y ha tenido un impacto económico sustancial para su fabricante, Pfizer

Pharmaceuticals. En las pruebas preliminares para las reacciones adversas, se encontró que cuando 734 hombres fueron tratados con Viagra, el 16% experimentó dolores de cabeza. (Esto es una auténtica ironía). De 725 hombres en un grupo placebo, el 4% experimentó dolores de cabeza (con base en datos de Pfizer Pharmaceuticals).

a. Utilizando un nivel de significancia de 0.01, ¿existe evidencia suficiente para sustentar la aseveración de que entre los hombres que tomaron Viagra, los dolores de cabeza ocurrieron a una tasa mayor que en aquellos que no tomaron Viagra?

b. Construya un estimado del intervalo de confianza del 99% de la diferencia entre las tasas de dolores de cabeza de los usuarios de Viagra y la tasa de dolores de cabeza de quienes recibieron un placebo. ¿Qué sugiere el intervalo de confianza acerca de las dos tasas?

27. Tasa de rechazo a encuestas Los encuestadores profesionales están comenzando a preocuparse acerca de la tasa creciente de rechazos de potenciales sujetos de encuesta. Analizando el problema, existe una necesidad de saber si la tasa de rechazo es universal o si existe una diferencia entre las tasas de residentes de ciudades centrales y aquellos que no viven en ciudades centrales. Específicamente, se encontró que al encuestar a 294 residentes de ciudades centrales, el 28.9% se rehusó a responder. Una encuesta realizada a 1015 residentes que no viven en una ciudad central reveló una tasa de rechazo del 17.1% (datos tomados de "I Hear You Knocking But You Can't Come In", de Fitzgerald y Fuller, *Sociological Methods and Research*, vol. 11, núm. 1). Con un nivel de significancia de 0.01, pruebe la aseveración de que la tasa de rechazo de las ciudades centrales es la misma que la tasa de rechazo en otras áreas.

28. Ventaja del campo en casa Cuando se muestrearon partidos durante una temporada, se encontró que el equipo de casa ganó 127 de 198 partidos de *basquetbol* profesional, y el equipo de casa 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). Construya un intervalo de confianza del 95% para la diferencia entre las proporciones de los triunfos en casa. ¿Parece existir una diferencia significativa entre las proporciones de triunfos en casa? ¿Qué concluye usted acerca de la ventaja de jugar en casa?

29. Alcoholismo y tabaquismo en películas infantiles Pruebe la aseveración de que la proporción de 25 de 50 películas infantiles que muestran alcoholismo, seleccionadas aleatoriamente, es significativamente menor que la proporción muestral de 28 de otras 50 películas de este tipo que muestran tabaquismo. ¿Se aplican los resultados al conjunto de datos 7?

30. Encuesta de salud Remítase al conjunto de datos 1 en el 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.

8-2 Más allá de lo básico

31. Interpretación del traslape de intervalos de confianza En el artículo "On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals", de Schenker y Gentleman (*The 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 y 112 personas en la primera muestra 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% de 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 6-2 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 usted 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 usted acerca de la igualdad de p_1 y p_2? ¿Cuál de los tres métodos anteriores es el menos efectivo para probar la igualdad de p_1 y p_2?

32. Equivalencia de prueba de hipótesis e intervalo de confianza Se obtienen dos muestras aleatorias simples a partir de dos poblaciones diferentes. La primera muestra consta de 20 personas, 10 de las cuales tienen un atributo en común. La segunda muestra consta de 2000 personas con 1404 que 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 $p_1 - p_2$.

33. Las mismas proporciones con muestras más grandes Esta sección utilizó los datos muestrales de la tabla 8-1 para probar la aseveración de que $p_1 = p_2$ y para construir un estimado del intervalo de confianza de $p_1 - p_2$. ¿Cómo se ven afectados los resultados si los datos muestrales de la tabla 8-1 se modifican para que p_1 se convierta en $240/2000$, en lugar de $24/200$ y p_2 se convierta en $1470/14{,}000$, en lugar de $147/1400$? Note que ambas proporciones muestrales permanecen iguales, pero los tamaños muestrales son mayores. ¿Existe ahora suficiente evidencia para sustentar la aseveración de que la proporción de conductores negros detenidos por la policía es mayor que la proporción de conductores blancos detenidos?

34. 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

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - c}{\sqrt{\dfrac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}}$$

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 26 y utilice un nivel de significancia de 0.05 para probar la aseveración de que la tasa de dolor de cabeza de usuarios de Viagra es 10 puntos porcentuales más alta que el porcentaje de aquellos a quienes se administró un placebo.

35. La transitividad de las pruebas de hipótesis Se seleccionan al azar datos muestrales a partir de tres poblaciones independientes, cada una de tamaño 100. Las proporciones muestrales son $\hat{p}_1 = 40/100$, $\hat{p}_2 = 30/100$, y $\hat{p}_3 = 20/100$.

a. Al nivel de significancia de 0.05, pruebe H_0: $p_1 = p_2$.

b. Al nivel de significancia de 0.05, pruebe H_0: $p_2 = p_3$.

c. Al nivel de significancia de 0.05, pruebe H_0: $p_1 = p_3$.

d. En general, ¿si las pruebas de hipótesis nos llevan a las conclusiones de que $p_1 = p_2$ y $p_2 = p_3$ son razonables, se sigue que $p_1 = p_3$ es también razonable? ¿Por qué?

36. Determinación de tamaño de la muestra El tamaño de la muestra 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

$$E = z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

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 resuelva para n.

continúa

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 el 95% de confianza de que su error no sea mayor de 0.03.

37. Interpretación de resultados de prueba de fármaco El Ziac es un fármaco de Lederle Laboratories elaborado para tratar la hipertensión. Lederle Laboratories reportó que cuando 221 personas fueron tratadas con Ziac, el 3.2% experimentó mareo. También se reportó que de 144 personas en el grupo placebo, el 1.8% experimentó mareo.

a. ¿Utilizaría los métodos de esta sección para probar la aseveración de que existe una diferencia significativa entre las dos tasas de mareo? ¿Por qué?

b. ¿Es correcta la información dada? ¿Por qué?

38. Verificación de la propiedad de varianzas Cuando se analizaron los fundamentos de los métodos de esta sección se estableció que como $\hat{p}_1$ y $\hat{p}_2$ se aproximan cada una a una distribución normal, $\hat{p}_1 - \hat{p}_2$ también se aproximará a una distribución normal con media $p_1 - p_2$ y varianza $\sigma^2_{(\hat{p}_1-\hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2}$. 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 cuatro sucesos simples, luego calcule la proporción de caras en cada uno de los cuatro casos. Utilice la fórmula $\sigma^2 = \Sigma(x - \mu)^2/N$ 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.

8-3 Inferencias acerca de dos medias: muestras independientes

En esta sección consideramos métodos para utilizar datos muestrales provenientes de dos muestras independientes para probar hipótesis acerca de dos medias poblacionales o para construir estimados del intervalo de confianza de la diferencia entre dos medias poblacionales. Comenzamos por definir formalmente las muestras *independientes* y *dependientes*.

Definiciones

Dos muestras son **independientes** si los valores muestrales seleccionados a partir de una población no están relacionados, apareados o asociados de alguna manera con los valores muestrales seleccionados a partir de la otra población. Si existe alguna relación, de modo que cada valor en una muestra esté apareado con un valor correspondiente en la otra muestra, las muestras son **dependientes**. Las muestras dependientes se conocen con frecuencia como **datos apareados** o **muestras equiparadas.** (Utilizaremos el término *datos apareados*, pues 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 un segundo grupo separado de sujetos reciben un placebo. Estos dos grupos muestrales son independientes puesto que los individuos en el grupo de tratamiento no están en ninguna forma apareados o equiparados con miembros correspondientes en el grupo placebo.

Datos apareados (o muestras dependientes): La eficacia de una dieta se prueba utilizando los pesos de los sujetos medidos antes y después del tratamiento de 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 lo siguiente.

Supuestos

1. Las dos muestras son *independientes*.
2. Ambas muestras son *muestras aleatorias simples*.
3. Cualquiera o ambas de estas condiciones se satisfacen: los dos tamaños de muestra 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 en tanto que no existan datos distantes y no existan sesgos fuertes).

Estadístico de prueba de hipótesis para dos medias: muestras independientes

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

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 por lo regular dan como 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 $n_2 - 1$.

continúa

2. Los programas de cómputo de estadística por lo regular utilizan el estimado más exacto pero más difícil dado en la fórmula 8-1. (Nosotros no utilizaremos la fórmula 8-1 en los ejemplos y ejercicios de este libro).

Fórmula 8-1
$$\text{gl} = \frac{(A + B)^2}{\dfrac{A^2}{n_1 - 1} + \dfrac{B^2}{n_2 - 1}}$$

donde $A = \dfrac{s_1^2}{n_1}$ y $B = \dfrac{s_2^2}{n_2}$

Valores *P*: Remítase a la tabla A-3. Utilice el procedimiento resumido en la figura 7-6. (Véase también la subsección de "cálculo de valores *P* con la distribución *t* de Student" en la sección 7-5).

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

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

$$\text{donde} \qquad E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

y el número de grados de libertad gl es como el descrito arriba 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. Por consecuencia, la hipótesis nula de $\mu_1 = \mu_2$ (o $\mu_1 - \mu_2 = 0$) se prueba determinando si el intervalo de confianza incluye a 0. Para 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 7-2 para casos comunes). Por ejemplo, la aseveración de que $\mu_1 > \mu_2$ se prueba con un nivel de significancia de 0.05, construyendo un intervalo de confianza del 90%.

Posteriormente analizaremos los fundamentos de las expresiones anteriores en esta sección. Por ahora, observe que los supuestos listados no incluyen las condiciones de que deben conocerse las desviaciones estándar poblacionales σ_1 y σ_2 ni tampoco suponemos que las dos poblaciones tienen la misma desviación estándar. Más tarde en esta sección se analizarán métodos alternativos basados en estos supuestos adicionales.

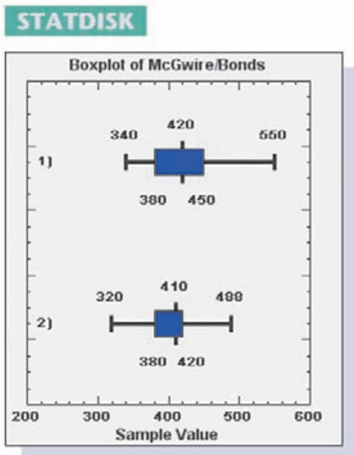
Exploración de los conjuntos de datos

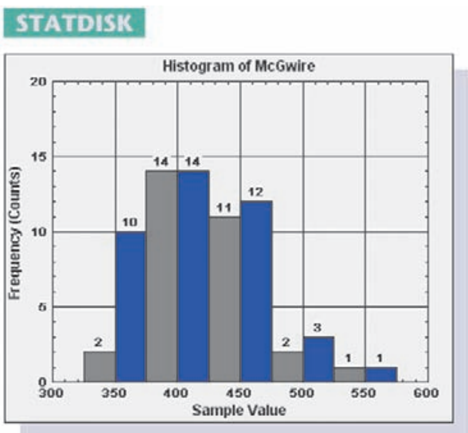
Debemos verificar los supuestos requeridos cuando utilizamos dos muestras independientes para hacer inferencias acerca de dos medias poblacionales. En lugar de realizar de inmediato una prueba de hipótesis o construir un intervalo de confianza, primero debemos *explorar* las dos muestras utilizando los métodos descritos en el capítulo 2. Para cada una de las dos muestras, debemos investigar el centro, la variación, la distribución, los datos distantes y si la población parece cambiar con el tiempo (CVDDT). Podría ser muy útil hacer lo siguiente:

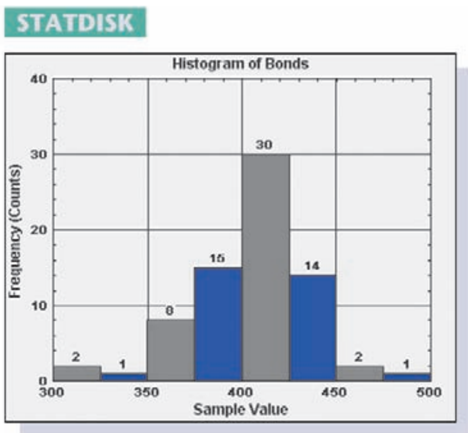
- Calcular estadísticos descriptivos para ambos conjuntos de datos, incluyendo n, $\bar{x}$, y s.
- Crear gráficas de caja de ambos conjuntos de datos, hechas en la misma escala para que puedan compararse.
- Crear histogramas de ambos conjuntos de datos, de modo que las distribuciones puedan compararse.
- Identificar cualquier dato distante.

EJEMPLO Prueba de hipótesis de distancias de "home run" de Bonds y McGwire El conjunto de datos 30 en el Apéndice B incluye las distancias de los "home runs" anotados en los récords de temporada por Mark McGwire y Barry Bonds. Los estadísticos muestrales, el histograma y las gráficas de cuadro se muestran abajo. Suponga que tenemos muestras aleatorias simples de poblaciones grandes y utilizamos un nivel de significancia de 0.05 para probar la aseveración de que las distancias provienen de poblaciones con medias diferentes.

	McGwire	Bonds
n	70	73
$\bar{x}$	418.5	403.7
s	45.5	30.6







continúa

Una 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 hacer ejercicio. Puesto que la Federal Trade Commission identificó aseveraciones que parecían no tener fundamento, se multó a la compañía con 10 millones de dólares por publicidad engañosa.

La eficacia de tratamientos como éstos puede determinarse con experimentos en los cuales un grupo de sujetos seleccionados al azar reciben el tratamiento, mientras que otro grupo de sujetos seleccionados al azar reciben un placebo. Las pérdidas de peso resultantes se comparan utilizando métodos estadísticos, como los descritos en esta sección.

El efecto placebo

Durante mucho tiempo se ha creído que los placebos realmente ayudan a algunos pacientes. De hecho, algunos estudios serios han mostrado que cuando se da un placebo (un tratamiento sin valor medicinal), muchos sujetos de prueba experimentan 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 publicado en el *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".

SOLUCIÓN Con la intención de explorar los dos conjuntos de datos, vemos que las medias muestrales son diferentes, los histogramas sugieren que las poblaciones tienen distribuciones que son aproximadamente normales, y las gráficas de cuadro parecen mostrar una diferencia. No parecen existir datos distantes. Procedamos con la prueba de la hipótesis formal para determinar si la diferencia entre las dos medias muestrales es realmente significativa. Puesto que es un poco difícil calcular el valor *P* en este ejemplo, utilizaremos el método tradicional de prueba de hipótesis.

Paso 1: La aseveración de medias diferentes se expresa simbólicamente como $\mu_1 \neq \mu_2$.

Paso 2: Si la aseveración original es falsa, entonces $\mu_1 = \mu_2$.

Paso 3: La hipótesis alternativa es la expresión que no contiene igualdad y la hipótesis nula es una expresión de igualdad, de modo que tenemos

$$H_0\colon \mu_1 = \mu_2 \qquad H_1\colon \mu_1 \neq \mu_2 \quad \textbf{(aseveración original)}$$

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:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{(418.5 - 403.7) - 0}{\sqrt{\dfrac{45.5^2}{70} + \dfrac{30.6^2}{73}}} = 2.273$$

Puesto que estamos utilizando una distribución *t*, los valores críticos de $t = \pm 1.994$ se encuentran en la tabla A-3. [Con una área de 0.05 en dos colas, queremos el valor *t* correspondiente a 69 grados de libertad, que sea el más chico de $n_1 - 1$ y $n_2 - 1$ (o el más chico de 69 y 72). La tabla A-3 no incluye 69 grados de libertad, entonces utilizamos el valor más cercano de 70 grados de libertad para obtener los valores críticos *t* de ± 1.994]. El estadístico de prueba, los valores críticos y la región crítica se muestran en la figura 8-2.

Utilizando STATDISK, Minitab, Excel o una calculadora TI-83 Plus, también podemos encontrar que el valor *P* es 0.0248 y los dos valores críticos más precisos son $t = \pm 1.995$ (con base en gl $= 69$). También podríamos utilizar la tabla A-3 para encontrar que, con gl $= 69$, el estadístico de prueba $t = 2.273$ corresponde a un valor *P* entre 0.02 y 0.05.

Paso 7: Puesto que el estadístico de prueba cae dentro de la región crítica, rechace la hipótesis nula $\mu_1 = \mu_2$ (o $\mu_1 - \mu_2 = 0$).

INTERPRETACIÓN Existe suficiente evidencia para sustentar la aseveración de que existe una diferencia entre las distancias medias de "home run" de Mark McGwire y Barry Bonds.

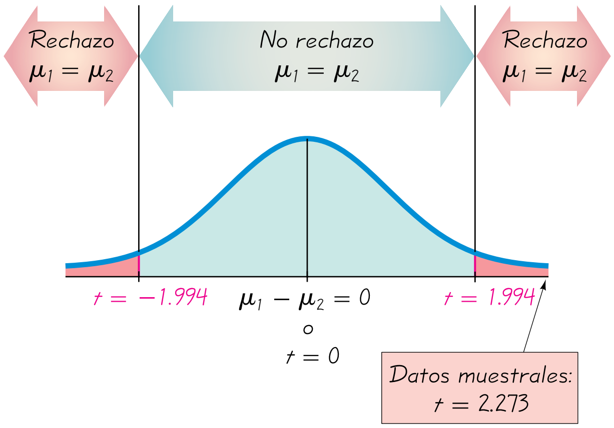


FIGURA 8-2
Distribución de valores $\bar{x}_1 - \bar{x}_2$

EJEMPLO Intervalo de confianza para la distancias de "home run" de Bonds y McGwire Utilice los datos muestrales que se dan en el ejemplo anterior y construya un estimado del intervalo de confianza del 95% de la diferencia entre la distancia media de "home run" de Mark McGwire y la distancia media de "home run" de Barry Bonds.

SOLUCIÓN Primero calculamos el valor del margen de error E. Utilizamos $t_{\alpha/2} = 1.994$, que se encuentra en la tabla A-3 como la puntuación t que corresponde a un área de 0.05 en dos colas y gl $= 70$. [Como en el ejemplo anterior, buscamos la puntuación t correspondiente a 69 grados de libertad, que es la más pequeña de $n_1 - 1$ y $n_2 - 1$ (o el más chico de 69 y 72). La tabla A-3 no incluye 69 grados de libertad, por lo tanto utilizamos el valor más cercano de 70 grados de libertad].

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_1^2}{n_2}} = 1.994\sqrt{\frac{45.5^2}{70} + \frac{30.6^2}{73}} = 13.0$$

Ahora calculamos el intervalo de confianza deseado como sigue:

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$
$$(418.5 - 403.7) - 13.0 < (\mu_1 - \mu_2) < (418.5 - 403.7) + 13.0$$
$$1.8 < (\mu_1 - \mu_2) < 27.8$$

Si utilizamos programas de cómputo o la calculadora TI-83 Plus para obtener resultados más precisos, obtenemos un intervalo de confianza de $1.9 < (\mu_1 - \mu_2) < 27.7$, por lo tanto podemos ver que el intervalo de confianza de arriba es bastante bueno.

INTERPRETACIÓN Tenemos una confianza del 95% de que los límites de 1.8 pies y 27.8 pies realmente contienen la diferencia entre las dos medias poblacionales. Este resultado podría presentarse con más claridad estableciendo que μ_1 excede a μ_2 por una cantidad que está entre 1.8 pies y 27.8 pies. Puesto que estos límites no contienen a 0, este intervalo de confianza sugiere que es muy improbable que las dos medias poblacionales sean iguales.

Comerciales

Las cadenas de televisión tienen sus propios departamentos de autorización para examinar los comerciales y verificar las aseveraciones. La National Advertising Division, una rama del Council of Better Business Bureaus, investiga las aseveraciones que se hacen en la publicidad. También participan la Federal Trade Commission y los fiscales de distrito locales. En cierta ocasión, Firestone tuvo que quitar una aseveración respecto a que sus neumáticos permitían frenar un 25% más rápido, y Warner Lambert tuvo que gastar 10 millones de dólares para informar a los consumidores que Listerine no previene ni cura los resfriados. Muchos anuncios engañosos se desechan voluntariamente y muchos otros escapan al escrutinio simplemente porque los mecanismos reguladores no son capaces de mantener el ritmo de la avalancha de comerciales.

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 periodos de tiempo comparables, las muestras de ventas tienen medias que son significativamente diferentes. La cantidad media de ventas decrece significativamente. 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 aplican los métodos para probar hipótesis e identificar indicadores como éstos. (Véase "How to Catch a Thief", de Manly y Thomson, *Chance*, vol. 11, núm. 4).

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 a $s_1^2/n_1 + s_2^2/n_2$, tomando en cuenta que $\bar{x}_1$ y $\bar{x}_2$ sean independientes. (Véase el ejercicio 31).

Método alternativo: $\boldsymbol{\sigma_1}$ y $\boldsymbol{\sigma_2}$ 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 en lugar de una distribución *t*. Véase lo siguiente.

Estadístico de prueba: $$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

Intervalo de confianza: $$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

donde $$E = z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

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 ambas muestras son grandes (con $n_1 > 30$ y $n_2 > 30$). Este método alternativo se usa con σ_1 reemplazado por s_1 y σ_2 reemplazado por s_2. Puesto que σ_1 y σ_2 en realidad rara vez se conocen, este libro no utilizará este método alternativo. Véase la figura 8-3.

Método alternativo: suponga que $\boldsymbol{\sigma_1 = \sigma_2}$ y *agrupe* las varianzas muestrales. Aun cuando los valores específicos de s_1^2 y s_2^2 no se conozcan, si se supone que tienen el *mismo* valor, las varianzas muestrales y pueden *agruparse* para obtener un estimado de la varianza poblacional σ^2 común. El **estimado agrupado de $\boldsymbol{\sigma^2}$** 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.

Supuestos

1. Las dos poblaciones tienen la misma desviación estándar. Esto es $\sigma_1 = \sigma_2$.
2. Las dos muestras son *independientes*.
3. Ambas muestras son *muestras 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. (Para muestras pequeñas, el requisito de normalidad es menos estricto en el sentido de que los procedimientos funcionan bien siempre y cuando no existan datos distantes y no existan sesgos fuertes).

Estadístico de prueba de hipótesis para dos medias: muestras independientes y $\sigma_1 = \sigma_2$

Estadístico de prueba: $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$

donde $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$ (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 $E = t_{\alpha/2}\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}$ y s_p^2 es como se dio en el estadístico de prueba anterior y el número de grados de libertad está dado por gl $= n_1 + n_2 - 2$.

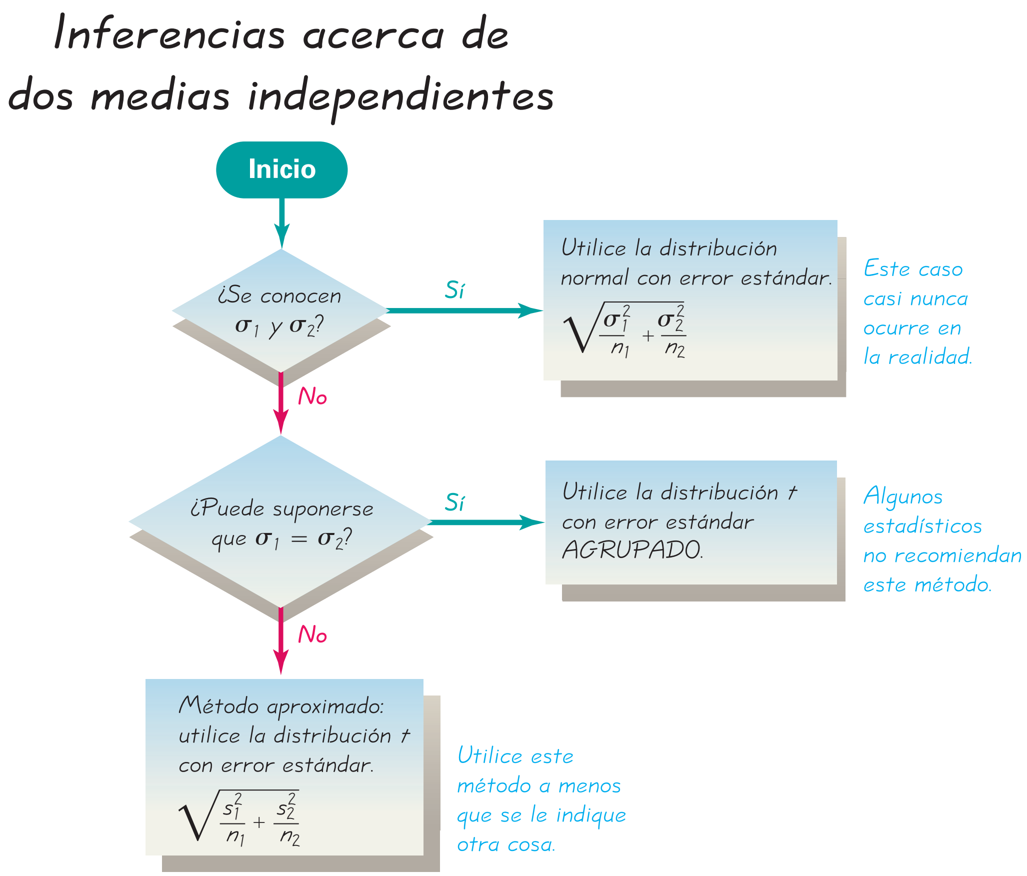


FIGURA 8-3 Métodos para inferencias acerca de dos medias independientes

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 se dio en la sección 8-5, 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, *The American Statistician*, vol. 46, núm. 1), los autores señalan que rara vez se conoce que $\sigma_1 = \sigma_2$. Ellos analizan el funcionamiento de las pruebas diferenciales considerando los tamaños muestrales y la potencia de las pruebas; concluyen que debe dedicarse más esfuerzo al aprendizaje del método presentado casi al inicio de esta sección y que debe ponerse menos énfasis en el método basado en el supuesto de que $\sigma_1 = \sigma_2$. Si no se indica de otra manera, 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 presentados casi al inicio de esta sección (véase la figura 8-3).**

Utilizando la tecnología

STATDISK Seleccione los elementos del menú **Analysis**, **Hypothesis Testing** y **Mean-Two Independent Samples**. Ingrese los valores requeridos en el cuadro de diálogo. Usted tiene las opciones de "Not Eq vars: NO POOL", "Eq vars: POOL" o "Prelim F Test". Se recomienda la opción **Not Eq vars: NO POOL.** (La prueba *F* se describe en la sección 8-5).

Minitab Minitab requiere las listas de datos muestrales originales y no funciona con un resumen de estadísticos. Si se conocen los valores muestrales originales, ingréselos en las columnas C1 y C2. (Si usted no conoce los valores muestrales originales, existe una forma de utilizar Minitab, pero es difícil; véase el *Minitab Student Laboratory Manual and Workbook*). Después de ingresar los datos muestrales en las columnas C1 y C2, seleccione las opciones **Stat**, **Basic Statistics** y **2-Sample t**, luego haga clic en **Samples in different columns** y proceda a ingresar C1 para la primera muestra y C2 para la segunda muestra. En el cuadro identificado como **alternative**, seleccione lo adecuado para la hipótesis alternativa (*no igual*, o *menor que* o *mayor que*), e ingrese el intervalo de confianza apropiado para la prueba (como 0.95 para $\alpha = 0.05$). La pantalla del Minitab también incluye los límites del intervalo de confianza.

Si las dos varianzas poblacionales parecen ser iguales, Minitab no permite el uso de un estimado agrupado de la varianza común. Aparecerá un cuadro después para **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 programa de complemento 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 icono del lápiz para la primera columna cuantitativa e ingrese el rango de valores para la primera muestra, por ejemplo A1:A14. Haga clic en el icono 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 no iguales. (Usted también puede seleccionar **Pooled** para la suposición de varianzas poblacionales iguales, pero no se recomienda este método).

Para utilizar el programa de complemento Data Analysis de Excel, haga clic en **Tools** y seleccione **Data Analysis**. Seleccione uno de los siguientes dos elementos (recomendamos la suposición de varianzas *no iguales*):

prueba *t*: Dos muestras suponiendo variables iguales
prueba *t*: Dos muestras suponiendo varianzas no iguales

Proceda a ingresar el rango de valores de la primera muestra (por ejemplo A1:A14) 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 Plus La calculadora TI-83 Plus le da la opción de utilizar varianzas "agrupadas" (si usted cree que $\sigma_1^2 = \sigma_2^2$) o de no agrupar las varianzas, pero recomendamos que las varianzas no se agrupen. Para realizar pruebas del tipo de las que se encuentran en esta sección, oprima **STAT**, luego seleccione **TESTS** y escoja **2-SampTTest** (para una prueba de hipótesis) o **2-SampTInt** (para un intervalo de confianza).

8-3 Destrezas y conceptos básicos

Muestras independientes y datos apareados. *En los ejercicios 1 a 4, determine si las muestras son independientes o si consisten en datos apareados.*

1. La eficacia del Prilosec para tratar la acidez estomacal se prueba midiendo la secreción ácida gástrica en un grupo de pacientes tratados con Prilosec, y a otro grupo de pacientes se les da un placebo.

2. La eficacia de Prilosec para tratar la acidez estomacal se prueba midiendo la secreción ácida gástrica en pacientes, antes y después del tratamiento con el fármaco. Los datos consisten en mediciones antes/después para cada paciente.

3. Se prueba la precisión de respuestas verbales en un experimento en el que los sujetos reportan sus pesos y luego se pesan en una báscula médica. Los datos consisten en el peso reportado y el peso medido en cada sujeto.

4. Se prueba el efecto del azúcar como ingrediente, con una muestra de latas de Coca Cola clásica y otra muestra de latas de Coca Cola de dieta.

En los ejercicios 5 a 24 suponga que las dos muestras son aleatorias simples independientes, seleccionadas a partir de poblaciones distribuidas normalmente. No suponga que las desviaciones estándar poblacionales son iguales.

5. 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 estos 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 (datos tomados 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 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: $n = 65, \bar{x} = 51.3, s = 4.5$

6. Intervalo de confianza del efecto del consumo de marihuana en estudiantes universitarios Remítase a los datos muestrales utilizados en el ejercicio 5 y construya un intervalo de confianza del 98% para la diferencia entre las dos medias poblacionales. ¿Incluye el intervalo de confianza a 0? ¿Qué sugiere el intervalo de confianza acerca de la igualdad de las dos medias poblacionales?

7. Intervalo de confianza para tratamiento de depresión bipolar En ensayos clínicos que incluyen 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: $n = 33, \bar{x} = 20.38, s = 3.91$

8. Prueba de hipótesis para tratamiento de depresión bipolar Remítase a los datos muestrales del ejercicio 7 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?

9. Prueba de hipótesis para tratamiento magnético del dolor La gente gasta enormes sumas de dinero (actualmente alrededor de 5000 millones de dólares 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 eficientes 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 aquellas personas tratadas con magnetos tienen una mayor reducción del dolor que quienes recibieron un tratamiento simulado (similar a un placebo). ¿Parece que los magnetos son eficientes en el tratamiento del dolor de espalda? ¿Es válido argumentar que los magnetos podrían parecer efectivos si los tamaños de muestra 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: $n = 20$, $\bar{x} = 0.49, s = 1.4$

10. Intervalo de confianza para tratamiento magnético del dolor Remítase a los datos muestrales del ejercicio 9 y construya un estimado del intervalo de confianza del 90% de la diferencia entre la reducción media del dolor para las personas tratadas con magnetos y la reducción media del dolor para quienes recibieron un tratamiento simulado. Con base en el resultado, ¿parece que los magnetos son eficientes en la reducción del dolor?

Coca Cola clásica	Coca Cola de dieta
$n_1 = 36$	$n_2 = 36$
$\bar{x}_1 = 0.81682$	$\bar{x}_2 = 0.78479$
$s_1 = 0.007507$	$s_2 = 0.004391$

11. Referencias a partir de muestras de Coca Cola clásica y Coca Cola de dieta Al utilizar el conjunto de datos 17 en el Apéndice B, encontramos los estadísticos muestrales para los pesos (en libras) de Coca Cola clásica y Coca Cola de dieta como se listan al margen.

a. Utilice un nivel de significancia de 0.01 para probar la aseveración de que las latas de Coca Cola clásica y Coca Cola de dieta tienen el mismo peso medio. Si parece existir una diferencia, trate de dar una explicación.

b. Construya un estimado del intervalo de confianza del 99% de $\mu_1 - \mu_2$, para la diferencia entre el peso medio de la Coca Cola clásica y el peso medio de la Coca Cola de dieta.

Nicotina (mg)

Largos con filtro	Largos sin filtro
$n_1 = 21$	$n_2 = 8$
$\bar{x}_1 = 0.94$	$\bar{x}_2 = 1.65$
$s_1 = 0.31$	$s_2 = 0.16$

12. Filtros de cigarrillos y nicotina Remítase a los resultados muestrales listados al margen para el contenido medido de nicotina de cigarrillos largos con filtro y sin filtro seleccionados al azar. Todas las mediciones están en miligramos y los datos son de la Federal Trade Commission.

a. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los cigarrillos largos con filtro tienen una cantidad media más baja de nicotina que la cantidad media de nicotina en cigarrillos largos sin filtro.

b. Construya un estimado del intervalo de confianza del 90% de la diferencia entre las dos medias poblacionales.

c. ¿Parece que los filtros de cigarrillo son eficaces en la reducción de nicotina?

13. Prueba de hipótesis para identificar trastornos psiquiátricos ¿Están relacionados los trastornos psiquiátricos graves con factores biológicos observables médicamente? Un estudio utilizó tomografía computarizada (TC) por rayos X 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 (datos tomados 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 del 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?

Grupo control:	$n = 10, \bar{x} = 0.45, s = 0.08$
Pacientes obsesivo compulsivos:	$n = 10, \bar{x} = 0.34, s = 0.08$

14. Intervalo de confianza para identificar trastornos psiquiátricos Remítase a los datos muestrales en el ejercicio 13 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?

15. Intervalo de confianza para efectos del alcohol Se realizó un experimento para probar los efectos del alcohol. Los errores se registraron en una prueba de destrezas visuales y motrices para un grupo de tratamiento de personas que bebieron etanol y 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 del 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?

Grupo de tratamiento	Grupo placebo
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 4.20$	$\bar{x}_2 = 1.71$
$s_1 = 2.20$	$s_2 = 0.72$

16. Prueba de hipótesis para efectos del alcohol Remítase a los datos muestrales del ejercicio 15 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 control. Si existe una diferencia significativa, ¿podemos concluir que el tratamiento causa una disminución en las destrezas visuales y motrices?

17. Prueba de hipótesis para la precisión de los testigos oculares de la policía ¿Afecta el estrés la capacidad de memoria de policías que han sido testigos oculares? Este tema se estudió en un experimento que probó la memoria de testigos oculares una semana después de un interrogatorio no estresante a un sospechoso cooperativo y un interrogatorio estresante a un sospechoso no cooperativo y beligerante. El número de detalles recordados una semana después del incidente se resumen al margen (datos de "Eyewitness Memory of Police Trainees for Realistic Role Plays", de Yuille *et al*., *Journal of Applied Psychology*, vol. 79, núm. 6). Utilice un nivel de significancia de 0.01 para probar la aseveración del artículo de que "el estrés disminuyó la cantidad de recuerdos".

Sin estrés	Con estrés
$n_1 = 40$	$n_2 = 40$
$\bar{x}_1 = 53.3$	$\bar{x}_2 = 45.3$
$s_1 = 11.6$	$s_2 = 13.2$

18. Intervalo de confianza para la precisión de los testigos oculares de la policía Utilizando los datos muestrales del ejercicio 17, construya un estimado del intervalo de confianza del 98% de la diferencia entre las dos medias poblacionales. ¿Sustenta el resultado la aseveración del artículo de que "el estrés disminuyó la cantidad de recuerdos"? ¿Por qué?

19. Polizones del *Queen Mary* El conjunto de datos 15 en el Apéndice B lista las edades de pasajeros polizones en viajes del *Queen Mary* por la costa oeste y por la costa este. Cuando se utiliza Excel con estos dos conjuntos de edades, se despliegan los resultados que se muestran abajo. ¿Existe una diferencia significativa entre las edades de los pasajeros polizones en viajes por la costa oeste del *Queen Mary* y las edades de los polizones en viajes por la costa este?

Excel

t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	*Variable 2*
Mean	26.71428571	24.84
Variance	103.2987013	67.81189189
Observations	56	75
Hypothesized Mean Difference	0	
df	104	
t Stat	1.130487967	
P(T<=t) one-tail	0.130435525	
t Critical one-tail	1.659636837	
P(T<=t) two-tail	0.26087105	
t Critical two-tail	1.983034963	

TI-83 Plus

2-SampTTest
$\mu_1 \neq \mu_2$
t=-2.831291425
p=.0128658011
df=14.65189403
$\bar{x}_1$=70.73333333
↓$\bar{x}_2$=80.75

20. Niveles de lectura Cuando se utiliza una calculadora TI-83 Plus con las puntuaciones de facilidad de lectura de Flesch para *El oso y el dragón* de Tom Clancey, y *Harry Potter y la piedra filosofal* de J. K. Rowling, se obtienen los resultados adjuntos. (Los datos muestrales están listados en el conjunto de datos 14 en el Apéndice B). ¿Existe evidencia suficiente para concluir que la media de la puntuación de facilidad de lectura de Flesch para Clancy es diferente de la media para Rowling?

T **21.** Alquitrán y cigarrillos Remítase a los datos muestrales listados abajo y utilice un nivel de significancia de 0.05 para probar 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			

T **22.** Bloqueo en exámenes Muchos estudiantes han tenido la experiencia poco placentera de pánico en exámenes porque la primera pregunta era excepcionalmente difícil. Se estudió el orden de las preguntas de exámenes para sus efectos en la ansiedad. Las siguientes puntuaciones son mediciones de "ansiedad debilitante por exámenes", que la mayoría de nosotros llamamos pánico o bloqueo (de acuerdo con datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimco, *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 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				

T **23.** IMC de hombres y mujeres Remítase al conjunto de datos 1 en el Apéndice B y pruebe la aseveración de que el índice de masa corporal (IMC) medio de los hombres es igual al índice de masa corporal medio de las mujeres.

T **24.** Corredores de maratón Remítase al conjunto de datos 8 en el Apéndice B y pruebe la aseveración de que la media de la edad de un corredor hombre en el maratón de la ciudad de Nueva York es igual a la media de la edad de una corredora mujer en ese maratón.

En los ejercicios 25 a 28 suponga que las 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.

25. Intervalo de confianza con agrupamiento Realice el ejercicio 7 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?

26. Prueba de hipótesis con agrupamiento Realice el ejercicio 8 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?

27. Prueba de hipótesis con agrupamiento Realice el ejercicio 9 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?

28. Intervalo de confianza con agrupamiento Realice el ejercicio 10 con la suposición adicional de que $\sigma_1 = \sigma_2$. ¿De qué manera se ven afectados los resultados por esta suposición adicional?

8-3 Más allá de lo básico

29. Efectos de un dato distante

a. Remítase al ejercicio 17 e incluya un dato distante consistente en un polizón de 90 años de edad en un viaje de crucero del *Queen Mary* por la costa oeste. ¿Se afecta drásticamente la prueba de hipótesis por la presencia del dato distante?

b. Remítase al ejercicio 19 e incluya un dato distante consistente en un polizón de 5000 años de edad en un viaje de crucero del *Queen Mary* por la costa oeste. ¿Por qué disminuye el estadístico de prueba t en lugar de incrementarse?

30. Efectos de las unidades de medida ¿De qué manera se ven afectados los resultados del ejercicio 12, si todas las cantidades de nicotina se convierten de miligramos a onzas? En general, ¿afecta la elección de la escala las conclusiones acerca de la igualdad de dos medias poblacionales y afecta dicha elección al intervalo de confianza?

31. Verificación de una propiedad de las varianzas

a. Calcule la varianza para esta *población* de x valores: 5, 10, 15. (Véase la sección 2-5 para la varianza σ^2 de una población).

b. Calcule la varianza para esta *población* de y valores: 1, 2, 3.
c. Haga una lista de la *población* de todas las diferencias posibles $x - y$, y calcule la varianza de esta población.
d. 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 ($\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y$). (Este principio se utiliza para derivar el estadístico de prueba y el intervalo de confianza dados en esta sección).
e. ¿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?

32. 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 de 0.05 para probar la aseveración de que los dos grupos muestrales provienen de poblaciones con la misma media. Los resultados se basan 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.

Grupo de tratamiento	Grupo placebo
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 0.049$	$\bar{x}_2 = 0.000$
$s_1 = 0.015$	$s_2 = 0.000$

33. Cálculo de grados de libertad ¿De qué manera se ve afectado el número de grados de libertad en los ejercicios 13 y 14 si se utiliza la fórmula 8-1 en lugar de seleccionar el más chico de $n_1 - 1$ y $n_2 - 1$? Si se utiliza la fórmula 8-1 para el número de grados de libertad en lugar del más pequeño de $n_1 - 1$ 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 chico de $n_1 - 1$ y $n_2 - 1$" es un estimado más conservador del número de grados de libertad que el estimado que se obtiene con la fórmula 8-1?

8-4 Inferencias a partir de datos apareados

En la sección 8-3 definimos que dos muestras son *independientes* si los valores muestrales, seleccionados a partir de una población, no están relacionados, apareados ni asociados con los valores muestrales seleccionados a partir de la otra población. La sección 8-3 trató con inferencias acerca de las medias de dos poblaciones independientes, y esta sección se enfoca en muestras dependientes, a las que nos referimos como *datos apareados*. En éstos, existe alguna relación para que cada valor en una muestra se aparee con un valor correspondiente en la otra muestra. A continuación se presentan algunos ejemplos típicos de datos apareados:

- Cuando se realiza un experimento para probar la eficacia de una dieta baja en grasas, el peso de cada sujeto se mide una vez antes de la dieta y una vez después de la dieta.
- La eficacia de un programa de entrenamiento para el SAT (prueba de aptitudes académicas) se prueba efectuando a cada sujeto un examen del SAT antes del programa y otro examen del SAT equivalente después del programa.

- La precisión de los pesos reportados se analiza con una muestra de personas cuando, para cada persona, el peso reportado se registra y el peso real se mide.

Para tratar con inferencias acerca de medias y datos apareados, abajo se incluyen resúmenes de los supuestos relevantes, 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 se prueba determinando si el intervalo de confianza incluye a 0. [Para 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 nivel de significancia α, construya un intervalo de confianza con un nivel de confianza de $1 - 2\alpha$. (Véase la tabla 7-2 para casos comunes). Por ejemplo, la aseveración de que la diferencia de la media es mayor que 0 se puede probar con un nivel de significancia de 0.05, construyendo un intervalo de confianza del 90%].

Supuestos

1. Los datos muestrales consisten en datos apareados.
2. Las muestras son muestras aleatorias simples.
3. Cualquiera o ambas de estas condiciones se satisfacen: el número de datos apareados o 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 quizá podamos utilizar los métodos no paramétricos que se analizan en el capítulo 12).

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

$\overline{d}$ = 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

$$t = \frac{\overline{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

donde los grados de libertad $= n - 1$.

Valores *P* y valores críticos: Tabla A-3 (distribución t)

continúa

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.

Exploración de los conjuntos de datos

Como siempre, debemos evitar la aplicación descuidada de cualquier procedimiento estadístico. Debemos considerar el centro, la variación, la distribución, los datos distantes y cualquier cambio que tenga lugar en el tiempo (CVDDT). Puesto que queremos ilustrar los métodos de esta sección con cálculos sencillos, los siguientes ejemplos utilizan datos muestrales consistentes en sólo cinco datos apareados. Observamos que las temperaturas mínimas reales parecen ser sustancialmente diferentes de las temperaturas mínimas pronosticadas cinco días antes. Una gráfica cuantilar normal de estas cinco diferencias muestrales sugiere que tienen una distribución que es aproximadamente normal. (Estos cinco datos apareados se tomaron del conjunto de datos 10 en el Apéndice B, y un histograma de la lista completa de las 31 diferencias indica que la población de diferencias tiene una distribución que es aproximadamente normal). Podemos ver que no existen datos distantes. Es particularmente importante considerar los datos distantes puesto que su presencia llega a afectar drásticamente los resultados.

EJEMPLO ¿Son precisos los pronósticos de temperatura? La tabla 8-2 incluye cinco temperaturas mínimas reales y las correspondientes temperaturas mínimas que se pronosticaron cinco días antes. 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.

SOLUCIÓN Seguiremos el mismo método básico de prueba de hipótesis que se introdujo en el capítulo 7, pero utilizaremos el estadístico de prueba de arriba para datos apareados.

Paso 1: La aseveración de que existe una diferencia entre las temperaturas mínimas reales y las temperaturas mínimas pronosticadas para cinco días se expresa como $\mu_d \neq 0$.

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

$$H_0: \mu_d = 0 \qquad H_1: \mu_d \neq 0 \quad \textbf{(Aseveración original)}$$

Investigación en gemelos

Los gemelos idénticos se gestan cuando un solo óvulo fertilizado se divide en dos, de manera que ambos gemelos comparten el mismo paquete genético. Actualmente existe una explosión en la investigación enfocada en este tipo de gemelos. Hablando para el Center of Study of Multiple Birth, Louis Keith señala que actualmente "tenemos mucha más capacidad de analizar los datos en gemelos utilizando computadoras con nuevos programas estadísticos instalados de fábrica". Una meta común de estudios de este tipo es explorar el tema clásico de "naturaleza contra crianza". Por ejemplo, Thomas Bouchard, quien realizó el Minnessota Study of Twins Reared Apart, encontró que el CI es heredado en un 50%–60%, mientras que el resto es el resultado de fuerzas externas.

Los gemelos idénticos son pares conjugados que proveen mejores resultados permitiéndonos reducir la variación genética inevitable con pares no relacionados de personas.

Tabla 8-2 Temperatura real y pronosticada

Mínima real	1	−5	−5	23	9
Mínima pronosticada cinco días antes	16	16	20	22	15
Diferencia d = real − pronosticada	−15	−21	−25	1	−6

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Utilizamos la distribución t de Student ya que se satisfacen los supuestos requeridos. (Estamos probando una aseveración acerca de datos apareados, tenemos dos muestras aleatorias simples y una gráfica cuantilar normal de las diferencias muestrales que indica que tienen una distribución aproximadamente normal).

Paso 6: Antes de calcular el valor del estadístico de prueba, debemos calcular los valores de $\bar{d}$ y s_d. Remítase a la tabla 8-2 y utilice las diferencias de −15, −21, −25, 1 y −6 para calcular estos estadísticos muestrales: $\bar{d} = -13.2$ y $s_d = 10.7$. Utilizando estos estadísticos muestrales y la suposición de la prueba de hipótesis de que $\mu_d = 0$, podemos ahora calcular el valor del estadístico de prueba.

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-13.2 - 0}{\frac{10.7}{\sqrt{5}}} = -2.759$$

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 de $n - 1 = 4$. La figura 8-4 nos muestra el estadístico de prueba, los valores crítico y la región crítica.

Paso 7: Puesto que el estadístico de prueba no cae en la región crítica, no rechazamos la hipótesis nula.

continúa

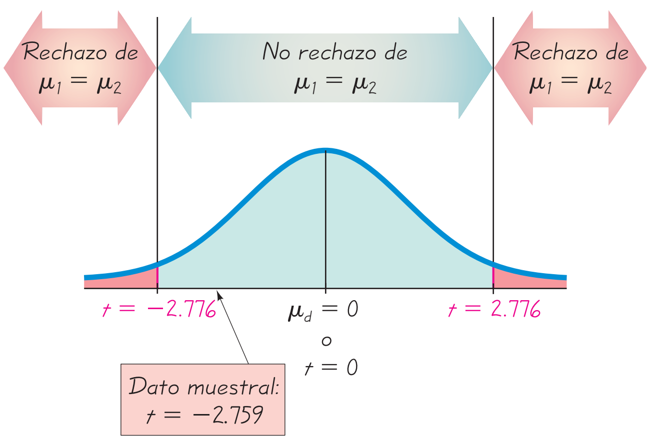


FIGURA 8-4 Distribución de diferencias d entre valores de datos apareados

Crest y muestras dependientes

A fines de la década de 1950, Procter & Gamble introdujo la pasta dental Crest como el primer producto de este tipo con fluoruro. A 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 características similares de alimentación, de cepillado y genéticas. Los resultados mostraron que los gemelos que usaron Crest tenían un número significativamente menor de caries que los que no la usaron. Este empleo de gemelos como muestras dependientes permitió a los investigadores controlar muchas de las diferentes variables que afectan las caries.

¿Salvan vidas las bolsas de aire?

La National Highway Transportation Safety Administration reportó que en un año reciente se salvaron 3448 vidas gracias a las bolsas de aire. Se reportó que para conductores de automóviles implicados en choques frontales, la tasa de fatalidad se redujo en un 31%; para pasajeros, se redujo en un 27%. Se señaló que "el cálculo de vidas salvadas se realiza con un análisis matemático de casos reales de acontecimientos fatales en vehículos con bolsas de aire comparado con vehículos sin bolsas de aire. Estudios como éste se llaman estudios de comparación de doble apareo, y son métodos de análisis estadístico ampliamente aceptados".

INTERPRETACIÓN Los datos muestrales de la tabla 8-2 no proporcionan suficiente evidencia para sustentar la aseveración de que las temperaturas mínimas reales y pronosticadas para cinco días son diferentes. Esto *no* establece que las temperaturas reales y pronosticadas sean iguales. Quizá datos muestrales adicionales podrían proporcionar la evidencia necesaria para concluir que las temperaturas mínimas reales y pronosticadas son diferentes. (Consulte el ejercicio 19 donde se utilizan resultados de 31 días).

Método del valor *P*. En el ejemplo anterior se utilizó el método tradicional, aunque se puede utilizar el método del valor P modificando los pasos 6 y 7. En el paso 6, utilice el estadístico de prueba de $t = -2.759$ y remítase al 4o. renglón de la tabla A-3 para encontrar que el estadístico de prueba (sin el signo negativo) está entre 2.776 y 2.132, indicando que el valor P está entre 0.05 y 0.10. Con STATDISK, Excel, Minitab y la calculadora TI-83 Plus, se calcula el valor P que es 0.0507. 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 ¿Son precisos los pronósticos de temperatura? Utilice los mismos datos apareados de la tabla 8-2, construya un estimado del 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 Utilizamos los valores de $\bar{d} = -13.2$, $s_d = 10.7$, $n = 5$ y $t_{\alpha/2} = 2.776$ (a partir de la tabla A-3 con $n - 1 = 4$ grados de libertad y un área de 0.05 en dos colas). Primero calculamos el valor del margen de error E.

$$E = t_{\alpha/2}\frac{s_d}{\sqrt{n}} = 2.776 \cdot \frac{10.7}{\sqrt{5}} = 13.3$$

Ahora se calcula el intervalo de confianza.

$$\bar{d} - E < \mu_d < \bar{d} + E$$
$$-13.2 - 13.3 < \mu_d < -13.2 + 13.3$$
$$-26.5 < \mu_d < 0.1$$

INTERPRETACIÓN Algunas veces el resultado se expresa como -13.2 ± 13.3 o como $(-26.5, 0.1)$. 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, indicando 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 las pronosticadas.

Utilizando la tecnología

STATDISK Seleccione **Analysis**, luego **Hypothesis Testing** y después **Mean-Matched Pairs**. En el cuadro de diálogo escoja el formato de la aseveración, ingrese un nivel de significancia, ingrese los datos muestrales y luego haga clic en **Evaluate**. El STATDISK provee automáticamente los límites del intervalo de confianza.

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 **2 Var t Interval**. En el cuadro de diálogo, haga clic en el icono del lápiz para la primera columna cuantitativa e ingrese el rango de valores para la primera muestra, por ejemplo A1:A14. Haga clic en el icono 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 la diferencia poblacional media deseada 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, y los valores críticos para una prueba de una cola y para una prueba de dos colas.

TI-83 Plus *Cuidado:* No utilice el elemento del menú **2-SampTTest,** puesto que éste se aplica a muestras *independientes*. En su lugar, ingrese los datos para la primera variable en la lista L1, ingrese los datos para la segunda variable en la lista L2, luego despeje la pantalla e ingrese **L1** − **L2** → **L3**. Después oprima **STAT**, luego seleccione **TESTS** y escoja la opción de **T-Test**. Utilizando la opción de alimentación de **Data**, ingrese los datos indicados, incluyendo la lista L3, y oprima **ENTER** cuando lo haya hecho. También es posible calcular un intervalo de confianza oprimiendo **STAT**, luego seleccionando **TESTS** y por último **TInterval**.

8-4 Destrezas y conceptos básicos

Cálculos para datos apareados. *En los ejercicios 1 y 2 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) $\overline{d}$, b) s_d, c) el estadístico de prueba t, y d) los valores críticos.*

1.

x	1	1	3	5	4
y	0	2	5	8	0

2.

x	5	3	7	9	2	5
y	5	1	2	6	6	4

3. Utilice los datos muestrales apareados del ejercicio 1 y construya un intervalo de confianza del 95% para la media poblacional de todas las diferencias $x - y$.

4. Utilice los datos muestrales apareados del ejercicio 2 y construya un intervalo de confianza del 99% para la media poblacional de todas las diferencias $x - y$.

5. Estaturas de mujeres reportadas y medidas Como parte de la National Health and Nutrition Examination Survey realizada por el Department of Health and Human Services, se obtuvieron estaturas reportadas y medidas para mujeres de 12 a 16 años de edad. Abajo se listan resultados muestrales.

a. ¿Existe evidencia suficiente para sustentar la aseveración de que existe una diferencia entre las estaturas reportadas y las estaturas medidas de mujeres de 12 a 16 años de edad? Utilice un nivel de significancia de 0.05.

b. 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	53	64	61	66	64	65	68	63	64	64	64	67
Estatura medida	58.1	62.7	61.1	64.8	63.2	66.4	67.6	63.5	66.8	63. 9	62.1	68.5

6. Estaturas de hombres reportadas y medidas Como parte de la National Health and Nutrition Examination Survey, realizada por el Department of Health and Human Services, se obtuvieron estaturas reportadas y medidas para hombres de 12 a 16 años de edad. Abajo se listan resultados muestrales.

a. ¿Existe evidencia suficiente para sustentar la aseveración de que hay 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.

b. 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 acerca de las implicaciones de 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

7. Eficacia del curso para el SAT Remítase a los datos en la tabla que lista calificaciones del SAT antes y después de que una muestra de 10 estudiantes tomara un curso preparatorio (según datos del College Board y de "An Analysis of the Impact of Commercial Test Preparation Courses on SAT Scores", de Sesnowitz, Bernhardt y Knain, *American Educational Research Jornal*, vol. 19, núm. 3).

a. ¿Existe evidencia suficiente para concluir que el curso preparatorio es efectivo en elevar las calificaciones? Utilice un nivel de significancia de 0.05.

b. Construya un estimado del intervalo de confianza del 95% de la diferencia media entre las calificaciones antes y después. Escriba un enunciado que interprete el intervalo de confianza resultante.

Estudiante	A	B	C	D	E	F	G	H	I	J
Calificación del SAT antes del curso (x)	700	840	830	860	840	690	830	1180	930	1070
Calificación del SAT después del curso (y)	720	840	820	900	870	700	800	1200	950	1080

8. Resultados antes/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 mm de mercurio) se midieron antes y después de tomar el fármaco, con los resultados que se dan en la tabla adjunta (según datos de "Essential Hypertension: Effect of an Oral Inhibitor of Angiotensin-Converting Enzyme", de MacGregor *et al*., *British Medical Journal*, vol. 2).

a. Utilice los datos muestrales para construir un intervalo de confianza del 99% para la diferencia media entre las lecturas antes y después.

b. ¿Existe suficiente evidencia para sustentar la aseveración de que el captopril es eficaz en la reducción de la presión sanguínea sistólica?

Sujeto	A	B	C	D	E	F	G	H	I	J	K	L
Antes	200	174	198	170	179	182	193	209	185	155	169	210
Después	191	170	177	167	159	151	176	183	159	145	146	177

9. 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 (basada en "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.

a. Construya un intervalo de confianza del 95% para la media de las diferencias "antes-después".

b. Utilice un nivel de significancia 0.05 para probar la afirmación de que las mediciones sensoriales son más bajas después de la hipnosis.

c. ¿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

10. Medición de inteligencia en niños Las mediciones mentales de niños pequeños con frecuencia se efectúan dándoles cubos y pidiéndoles que construyan una torre tan alta como les 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 (datos tomados de "Tower Building", de Johnson y Courtney, *Child Development*, vol. 3).

a. ¿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 primero 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

11. 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 terrenos adyacentes. Los valores listados son las cosechas de mazorcas en libras por acre.

a. Utilizando un nivel de significancia de 0.05, 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 del 95% de la diferencia media entre las cosechas de los dos tipos de semilla.

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

12. Estaturas de los padres Remítase al conjunto de datos 2 en el Apéndice B y utilice sólo los datos que corresponden a niños varones. Utilice los datos apareados que consisten en la estatura de la madre y la estatura del padre.

a. Utilice un nivel de significancia de 0.01 para probar la aseveración de que las madres de niños varones son más bajas que los padres.

b. Construya un estimado del intervalo de confianza del 98% de la media de las diferencias entre las estaturas de las madres y las estaturas de los padres.

13. Tratamiento del malestar por movimiento La siguiente pantalla de Minitab es el resultado de un experimento en el que se probó a 10 sujetos para el malestar por movimiento antes y después de tomar el fármaco astemizole. Los datos de la columna C3 del Minitab consisten en las diferencias en el 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 mientras estuviera a bordo de una embarcación de crucero?

b. Suponga que en lugar 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
```

TI-83 Plus

```
T-Test
 μ>0
 t=6.431306409
 p=3.340322E-4
 x̄=5.142857143
 Sx=2.115700942
 n=7
```

14. Dietas: interpretación de pantalla de calculadora Algunos investigadores obtuvieron datos de pérdida de peso de una muestra de personas sometidas a una dieta, utilizando las instalaciones del New World Athletic Club. Se registraron los pesos antes y después, luego se calcularon las diferencias (antes − después). Se muestran los resultados de la calculadora TI-83 Plus para la prueba de la aseveración de que la dieta es eficaz.

a. ¿Existe evidencia suficiente para fundamentar la aseveración de que la dieta es eficaz? Explique.

b. ¿Cuál es la media de la pérdida de peso? ¿Es lo suficientemente grande para que la dieta sea práctica para alguna persona que quiere perder peso?

c. Utilice los resultados de la pantalla para construir un intervalo de confianza del 95% para la media de la pérdida de peso.

15. Pesos de hombres reportados y medidos Remítase a la pantalla de Excel que muestra los resultados obtenidos cuando se prueba la aseveración de que no existe diferencia entre los pesos reportados y medidos de hombres de 12 a 16 años de edad. ¿Existe evidencia suficiente para fundamentar la aseveración de que hay una diferencia? Los datos son de la National Health and Nutrition Examination Survey, realizada por el Department of Health and Human Services.

Excel

t-Test: Paired Two Sample for Means		
	Variable 1	*Variable 2*
Mean	133.75	134.7333333
Variance	291.4772727	280.6151515
Observations	12	12
Pearson Correlation	0.919502265	
Hypothesized Mean Difference	0	
df	11	
t Stat	-0.501440942	
P(T<=t) one-tail	0.312972232	
t Critical one-tail	1.795883691	
P(T<=t) two-tail	0.625944463	
t Critical two-tail	2.200986273	

STATDISK

Sample Size, n	11
Difference Mean, $\bar{x}_d$	0.6727
Difference St Dev, s_d	0.8259
Test Statistic, t	2.7014
Critical t	1.8125
P-Value	0.0111

16. Estaturas reportadas y medidas de estudiantes de estadística hombres Se aplicó una encuesta a estudiantes de estadística hombres que incluía una pregunta pidiéndoles que reportaran su estatura en pulgadas. No se les dijo que su estatura sería medida, pero las estaturas se midieron con precisión después de que la encuesta se completó. Se mantuvo el anonimato con el uso de números codificados en lugar de nombres, de manera que ninguna información personal sería anunciada públicamente y nadie sería avergonzado por los resultados. Se presentan los resultados del STATDISK para la aseveración de que $\mu_d > 0$ para un nivel de significancia de 0.05. ¿Existe evidencia suficiente para sustentar la aseveración de que los estudiantes de estadística hombres exageran sus estaturas?

T **17.** Temperaturas corporales matutinas y nocturnas Remítase al conjunto de datos 4 en el Apéndice B. Utilice los datos apareados que consisten de temperaturas corporales de mujeres a las 8:00 AM y a las 12:00 AM del día 2.

a. Construya un intervalo de confianza del 95% de la diferencia media de las temperaturas a las 8:00 AM, menos las temperaturas a las 12:00 AM.

b. Utilizando un nivel de significancia de 0.05, pruebe la aseveración de que para estas temperaturas, la diferencia media es 0. Con base en los resultados, ¿parece que las temperaturas corporales matutinas y nocturnas son casi las mismas?

T **18.** Alcohol y tabaco en películas infantiles Remítase al conjunto de datos 7 en el Apéndice B. Utilice los datos apareados que consisten en las ocasiones en que las películas mostraron consumo de tabaco y en las que mostraron consumo de alcohol.

a. ¿Existe evidencia suficiente para concluir que el número de ocasiones difiere?

b. 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 a los niños se les muestra consumo de tabaco y el número de veces en que se les muestra consumo de alcohol?

T **19.** Temperaturas reales y pronosticadas Para los ejemplos en esta sección se utilizaron sólo cinco pares de datos muestrales con la finalidad de que los cálculos fueran sencillos. Remítase al conjunto de datos 10 en el 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?

T **20.** Temperaturas reales y pronosticadas Remítase al conjunto de datos 10 en el Apéndice B y utilice todas las temperaturas mínimas reales y las temperaturas mínimas que se pronosticaron un día 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 un día 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 un día antes.

c. Compare los resultados con los obtenidos en el ejercicio 19. ¿Parecen ser mejores los valores del pronóstico para un día que para cinco días?

8-4 Más allá de lo básico

21. Efectos de un dato distante y unidades de medida

a. Al utilizar los métodos de esta sección, ¿un dato distante tendría un efecto drástico en la prueba de hipótesis y en el intervalo de confianza?

b. Para los ejemplos en 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?

22. Intervalos de confianza y pruebas de un factor El intervalo de confianza del 95% para un conjunto de datos muestrales apareados es $0.0 < \mu_d < 1.2$. Con base en este intervalo de confianza, el método tradicional de prueba de hipótesis nos lleva a la conclusión de que se sustenta la aseveración de que $\mu_d > 0$. ¿Cuál es el menor valor posible del nivel de significancia de la prueba de hipótesis?

23. Uso del procedimiento correcto

a. Considere que los datos muestrales que se dan abajo son datos apareados y utilice un nivel de significancia de 0.05 para probar la aseveración de que $\mu_d > 0$.

b. Considere que los datos muestrales que se dan abajo son dos muestras independientes. Utilice un nivel de significancia de 0.05 para probar la aseveración de que $\mu_1 > \mu_2$.

c. Compare los resultados de los incisos *a* y *b*. ¿Es esencial 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

8-5 Comparación de la variación en dos muestras

Puesto que la característica de variación entre los datos es extremadamente importante, esta sección presenta un método del uso de dos muestras para comparar las varianzas de dos poblaciones de las que se obtienen las muestras. En la sección 2-5 vimos que la variación en una muestra se mide por medio de la desviación estándar, la varianza y otras medidas como son el rango y la desviación media absoluta. Puesto que la desviación estándar es una medida de variación muy efectiva, y ya que es más fácil de entender que la varianza, los primeros capítulos de este libro han enfatizado el uso de la desviación estándar en lugar de la varianza. Aunque el procedimiento básico de esta sección está diseñado para varianzas, podemos utilizarlo también para desviaciones estándar. Repasemos brevemente esta relación entre la desviación estándar y la varianza: la varianza es el cuadrado de la desviación estándar.

Medidas de variación

s = desviación estándar *muestral*

s^2 = varianza *muestral* (desviación estándar muestral al cuadrado)

σ = desviación estándar *poblacional*

σ^2 = varianza *poblacional* (desviación estándar poblacional al cuadrado)

Los cálculos de esta sección se simplificarán en gran medida si designamos las dos muestras de manera que s_1^2 represente a la *más grande* de las dos varianzas muestrales. Matemáticamente, en realidad no importa cuál muestra se designe como la muestra 1, así que la vida será mejor 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.

Supuestos

1. Las dos poblaciones son *independientes* una de la otra. (Recuerde de la sección 8-2 que dos muestras son independientes si la muestra seleccionada a partir de una población no está relacionada con la muestra seleccionada a partir de la otra población. Las muestras no están apareadas o asociadas).
2. Las dos poblaciones están *distribuidas normalmente*. (Este supuesto es importante puesto que los métodos de esta sección son extremadamente 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 a partir de la cual 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 prueba de hipótesis con dos varianzas

$$F = \frac{s_1^2}{s_2^2} \qquad \text{(donde } s_1^2 \text{ es } \textit{la más grande} \text{ de las dos varianzas muestrales)}$$

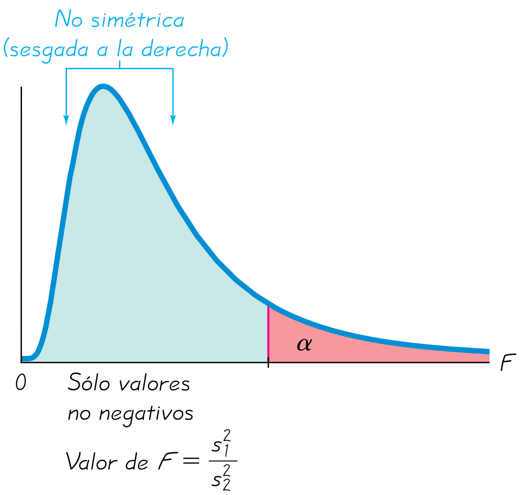
Valores críticos: Utilice la tabla A-5 para encontrar valores críticos F que se determinan por lo siguiente:

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 ilustra en la figura 8-5 con los valores críticos que se listan en la tabla 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.

FIGURA 8-5 Distribución *F*

Existe una distribución *F* diferente para cada par diferente de grados de libertad para el numerador y el denominador.



En la figura 8-5 note estas 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), entonces 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 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 lugar de esto, 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 19).

Con frecuencia tenemos números de grados de libertad que no se incluyen en la tabla A-5. Podríamos utilizar interpolación lineal para aproximar los valores que no están, pero en la mayoría de los casos esto no es necesario puesto que el estadístico de prueba F es menor que el valor crítico posible más bajo o mayor que el valor crítico posible más alto. Por ejemplo, la tabla A-5 muestra 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 mayor que 2.1952 provocará el rechazo de la hipótesis nula, y la interpolación sólo es necesaria si el estadístico de prueba F parece caer entre 2.0677 y 2.1952. El uso de una programa de cómputo de estadística como el STATDISK o el Minitab elimina este problema porque proporciona valores críticos o valores P.

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 acercarse a 1 puesto que s_1^2 y s_2^2 tienden a aproximar su valor. Pero si las dos poblaciones tienen varianzas radicalmente diferentes, s_1^2 y s_2^2 tienden a ser números muy distintos. Denotando la más grande de las varianzas muestrales por s_1^2, vemos que la proporción s_1^2/s_2^2 será un número grande, siempre y cuando 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 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.

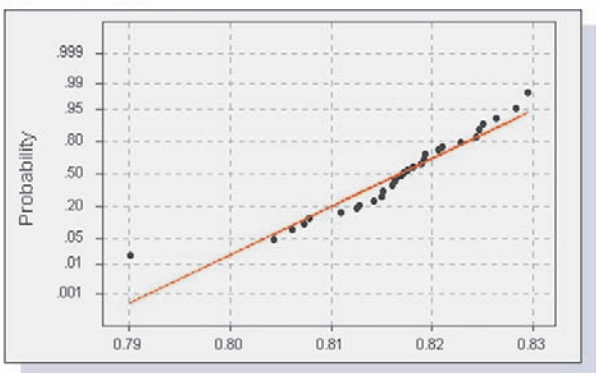
Los valores de *F grandes* son evidencia en contra de que $\sigma_1^2 = \sigma_2^2$.

Aseveraciones acerca de desviaciones estándar: El estadístico de prueba F se aplica a una aseveración hecha 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.

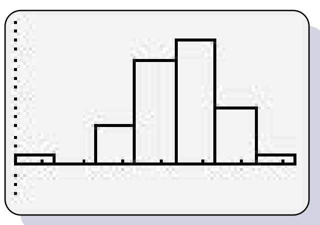
Exploración de los datos

Puesto que el requisito de distribuciones normales es muy importante y muy estricto, debemos comenzar comparando los dos conjuntos de datos muestrales a través de herramientas como los histogramas, las gráficas de cuadro y las gráficas cuantilares normales (véase la sección 5-7), y debemos buscar datos distantes (véase el ejercicio 17). Debemos 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 17 en el Apéndice B). Aquí se muestra un histograma de una calculadora TI-83 Plus y una gráfica de probabilidad normal del Minitab. El histograma indica que los datos tienen una distribución que es aproximadamente normal y que existe un valor que es un dato distante potencial. La gráfica de probabilidad normal, que puede interpretarse como si fuera una gráfica cuantilar normal, indica que los puntos se aproximan razonablemente a una línea recta, pero no ajustan a la línea recta perfectamente. Este conjunto de datos satisface claramente 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.

Minitab



TI-83 Plus



Menor variación, mayor calidad

Ford y Mazda estaban produciendo transmisiones similares que, se suponía, 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 estaban detectando variabilidad alguna entre las cajas de engranajes de las transmisiones Mazda. Ellos 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 remanufactura y de las reparaciones por garantía.

EJEMPLO Coca contra Pepsi El conjunto de datos 17 en el 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
n	36	36
$\bar{x}$	0.81682	0.82410
s	0.007507	0.005701

SOLUCIÓN En lugar de utilizar las desviaciones estándar muestrales para probar la aseveración de desviaciones estándar poblacionales iguales, utilizaremos las varianzas muestrales para probar la aseveración de varianzas poblacionales iguales. 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 perfila en la figura 7-8.

Paso 1: La aseveración de desviaciones estándar iguales es equivalente a una aseveración para varianzas iguales, lo que 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 como la hipótesis alternativa no podría contener igualdad, tenemos

$$H_0: \sigma_1^2 = \sigma_2^2 \quad \text{(Aseveración original)} \qquad H_1: \sigma_1^2 \neq \sigma_2^2$$

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Puesto que esta prueba comprende dos varianzas poblacionales, utilizamos la distribución F.

Paso 6: El estadístico de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{0.007507^2}{0.005701^2} = 1.7339$$

Para 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 del estadístico de prueba F, necesitamos encontrar sólo el valor crítico de cola derecha. En la tabla A-5 vemos que el valor crítico de F está entre 1.8752 y 2.0739, que encontramos 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 proporcionan un valor crítico de 1.9611).

Paso 7: La figura 8-6 indica que el estadístico de prueba $F = 1.7339$ no cae dentro de la región crítica, por lo tanto no rechazamos la hipótesis nula de varianzas iguales.

INTERPRETACIÓN No existe suficiente evidencia para justificar el rechazo de la aseveración de que las dos varianzas son iguales. De cualquier manera, de-

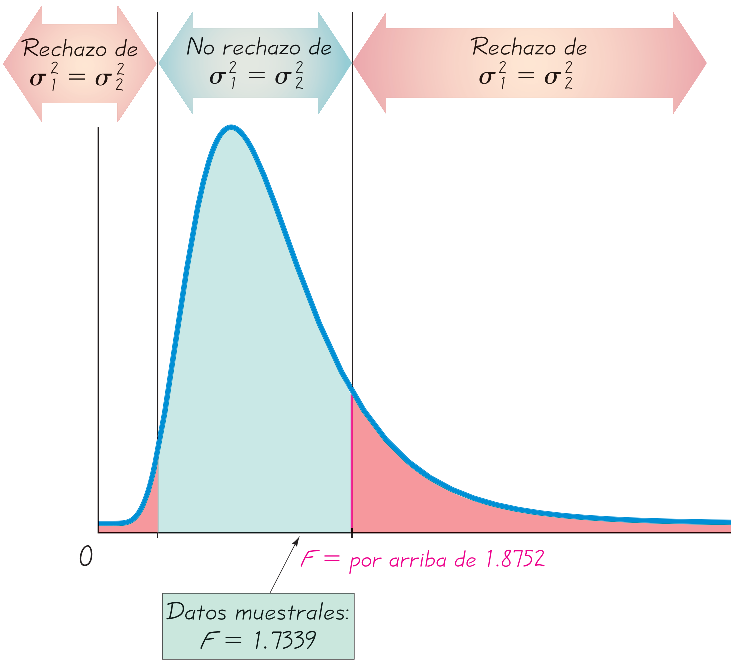


FIGURA 8-6 **Distribución de s_1^2/s_2^2 para pesos de Coca clásica y Pepsi clásica**

beríamos reconocer que la prueba F es extremadamente sensible a distribuciones que no son normales, por lo tanto esta conclusión podría hacer que parezca que no existe una diferencia significativa entre las varianzas poblacionales cuando realmente existe una diferencia que se ocultó por las distribuciones no normales.

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 de $F = 1.7339$, pero un valor crítico de F diferente.

Hemos descrito el método tradicional de prueba de hipótesis acerca de dos varianzas poblacionales. El ejercicio 18 se refiere al método del valor P, y el ejercicio 20 a la construcción de intervalos de confianza.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego seleccione **Hypothesis Testing**, y luego **StDev-Two Samples**. Ingrese los elementos requeridos en el cuadro de diálogo.

Minitab Primero ingrese todos los datos a partir de las dos muestras en la columna C1, con los valores de la primera muestra acumulados encima de los valores de la segunda muestra. En la columna C2 ingrese los "subíndices" de identificación, que consisten en un 1 junto a cada valor de la primera muestra y un 2 junto a cada valor de la segunda muestra. Ahora seleccione **Stat**, luego **ANOVA** y luego **Homogeneity of Variance**. Ingrese C1 para la variable de respuesta e ingrese C2 para los factores. Ingrese el nivel de confianza, con 0.95 correspondiente a un nivel de significancia de 0.05. Haga clic en **OK**. De los diversos resultados mostrados en la pantalla, encuentre el estadístico de prueba F y el correspondiente valor P. Si el valor P es menor que o igual al nivel de significancia, rechace la hipótesis nula de varianzas iguales.

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. Ingrese el valor del nivel de significancia en el cuadro "Alfa". Excel proporcionará 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, duplique el valor P dado por Excel.

TI-83 Plus **Plus** Oprima la tecla **STAT**, luego seleccione **TESTS** y luego **2-SampFTEST**. Usted puede utilizar los estadísticos resumidos o utilizar los datos que se han ingresado como listas.

8-5 Destrezas y conceptos básicos

Prueba de hipótesis de varianzas iguales. *En los ejercicios 1 y 2, pruebe la aseveración dada. Utilice un nivel de significancia de $\alpha = 0.05$ y suponga que todas las poblaciones están distribuidas normalmente. Utilice el método tradicional de prueba de hipótesis perfilado en la figura 7-8.*

1. Aseveración: La población de tratamiento y la población placebo tienen varianzas diferentes.

Grupo de tratamiento: $n = 25$, $\bar{x} = 98.6$, $s = 0.78$
Grupo placebo: $n = 30$, $\bar{x} = 98.2$, $s = 0.52$

2. Aseveración: Las estaturas de estudiantes de estadística hombres tienen una varianza mayor que las de estudiantes de estadística mujeres.

Hombres: $n = 16$, $\bar{x} = 68.4$, $s = 0.54$
Mujeres: $n = 12$, $\bar{x} = 63.2$, $s = 0.39$

3. Prueba de hipótesis para tratamiento magnético del dolor Ciertos investigadores realizaron un estudio para determinar si los magnetos son eficaces en el tratamiento del dolor de espalda, con los resultados que se presentan abajo (datos tomados 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 de quienes recibieron el tratamiento con magnetos.

Reducción en el nivel de dolor después del tratamiento fingido: $n = 20$, $\bar{x} = 0.44$, $s = 1.4$

Reducción en el nivel de dolor después del tratamiento magnético: $n = 20$, $\bar{x} = 0.49$, $s = 0.96$

4. Prueba de hipótesis para el efecto del consumo de marihuana en estudiantes universitarios En un estudio sobre los efectos del consumo de marihuana, en la universidad se probó la capacidad de memoria de consumidores ocasionales y frecuentes de marihuana, con los resultados que se presentan abajo (datos tomados 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: $n = 65, \bar{x} = 51.3, s = 4.5$

5. Pesos de Coca Cola clásica y Coca Cola de dieta En esta sección se incluyó un ejemplo acerca de una prueba de hipótesis para la aseveración de que los pesos de Coca clásica y Pepsi clásica tienen la misma desviación estándar. Pruebe la aseveración de que la Coca clásica y la Coca de dieta tienen pesos con desviaciones estándar diferentes. Los pesos muestrales se encuentran en el conjunto de datos 17 en el Apéndice B, pero aquí tenemos los estadísticos resumidos: la muestra de 36 pesos de Coca clásica tiene una desviación estándar de 0.007507 libras, y la muestra de 36 pesos de Coca de dieta tiene una desviación estándar de 0.004391 libras. Utilice un nivel de significancia de 0.05. Si los resultados mostraran que las desviaciones estándar son significativamente diferentes, ¿cuál sería un factor importante que podría explicar la diferencia?

6. Cargas axiales de latas de aluminio El conjunto de datos 20 en el Apéndice B incluye cargas axiales (en libras) de una muestra de 175 latas de aluminio con un espesor de 0.0109 pulgadas y otra muestra de 175 latas de aluminio con un espesor de 0.0111 pulgadas. (Una carga axial es el peso máximo que soportan los costados. Se mide utilizando una placa para aplicar presión creciente a la parte superior de la lata hasta que ésta se colapse). La muestra de latas de 0.0109 pulgadas tuvo cargas axiales con una media de 267.1 libras y desviación estándar de 22.1 libras. La muestra de latas de 0.0111 pulgadas tuvo cargas axiales con una media de 281.8 libras y desviación estándar de 27.8 libras. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las muestras provienen de poblaciones con la misma desviación estándar.

7. Filtros de cigarrillos y nicotina Remítase a los resultados muestrales que se listan al margen para los contenidos de nicotina medidos de cigarrillos largos con filtro y sin filtro seleccionados al azar. Todas las mediciones son en miligramos y los datos son de la Federal Trade Commission. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los cigarrillos largos con filtro tienen cantidades de nicotina que varían más que las cantidades de nicotina de los cigarrillos largos sin filtro.

Nicotina (mg)

Largos con filtro	Largos sin filtro
$n_1 = 21$ $\bar{x}_1 = 0.94$ $s_1 = 0.31$	$n_2 = 8$ $\bar{x}_2 = 1.65$ $s_2 = 0.16$

8. 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 para un grupo de tratamiento de personas que bebieron etanol y 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). 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 del grupo placebo.

Grupo de tratamiento	Grupo placebo
$n_1 = 22$ $\bar{x}_1 = 4.20$ $s_1 = 2.20$	$n_2 = 22$ $\bar{x}_2 = 1.71$ $s_2 = 0.72$

9. Antigüedad de automóviles de profesores y estudiantes Los estudiantes en la universidad del autor seleccionaron aleatoriamente 217 automóviles de estudiantes y encontraron que tienen antigüedades con una media de 7.89 años y una desviación estándar de 3.67 años. Ellos 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?

Grupo con suplemento de zinc	Grupo placebo
$n = 294$	$n = 286$
$\bar{x} = 3214$	$\bar{x} = 3088$
$s = 669$	$s = 728$

10. Prueba de efectos del zinc Se realizó un estudio de madres con deficiencia de zinc para determinar los efectos del suplemento de zinc durante el embarazo. Los datos muestrales están listados al margen (según datos de "The Effects of Zinc Supplementation on Pregnancy Outcome", de Goldenberg *et al*., *Journal of the American Medical Association*, vol. 274, núm. 6). Los pesos se midieron en gramos. Utilizando un nivel de significancia de 0.05, ¿existe evidencia suficiente para sustentar la aseveración de que la variación de los pesos al nacer de los bebés de la población placebo es mayor que la variación de la población tratada con suplementos de zinc?

11. Precipitación pluvial en fines de semana *USA Today* y otros periódicos reportaron un estudio que, al parecer, mostraba que llovía más durante los fines de semana. El estudio se refería a áreas en la costa este de Estados Unidos cerca del océano. El conjunto de datos 11 en el Apéndice B lista las cantidades de lluvia en Boston en un año. Las 52 cantidades de lluvia para el miércoles tienen una media de 0.0517 pulgadas y una desviación estándar de 0.1357 pulgadas. Las 52 cantidades de lluvia para el domingo tienen una media de 0.0677 pulgadas y una desviación estándar de 0.2000 pulgadas.

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 lugar de construir histogramas o gráficas cuantilares normales, examine simplemente el número de días sin lluvia. ¿Están distribuidas normalmente las cantidades de lluvia del miércoles? ¿Están distribuidas normalmente las cantidades de lluvia del domingo?

c. ¿Qué se concluye a partir de los resultados de los incisos *a* y *b*?

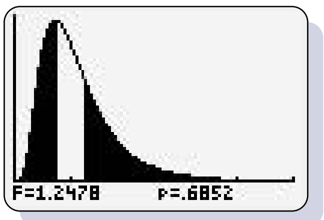
12. Consumo de tabaco y alcohol en películas infantiles de dibujos animados El conjunto de datos 7 en el Apéndice B lista tiempos (en segundos) en los que las películas de dibujos animados para niños muestran consumo de tabaco y consumo de alcohol. Los 50 tiempos de consumo de tabaco tienen una media de 57.4 segundos y una desviación estándar de 104.0 segundos. Los 50 tiempos de consumo de alcohol tienen una media de 32.46 segundos y una desviación estándar de 66.3 segundos.

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. Considere el prerrequisito de poblaciones distribuidas normalmente. En lugar de construir histogramas o gráficas cuantilares normales, examine simplemente el número de películas que no muestran consumo de tabaco o alcohol. ¿Están distribuidos normalmente los tiempos para tabaco? ¿Están distribuidos normalmente los tiempos para alcohol?

c. ¿Qué se concluye a partir de los resultados de los incisos *a* y *b*?

TI-83 Plus



13. Calcio y presión sanguínea Se reunieron datos muestrales en un estudio de suplementos de calcio y su efecto en la presión sanguínea. Se inició el estudio con las mediciones de la presión sanguínea de un grupo placebo y de un grupo de calcio (según datos de "Blood Pressure and Metabolic Effects of Calcium Supplementation in Normotensive White and Black Men", de Lyle *et al*., *Journal of the American Medical Association*, vol. 257, núm. 13). Se listan los valores muestrales y se presenta la pantalla de la TI-83 Plus. A un nivel de significancia de 0.05, pruebe la aseveración de que los dos grupos muestrales provienen de poblaciones con la misma desviación estándar. Si el experimento requiere grupos con desviaciones estándar iguales, ¿son estos dos grupos aceptables?

Placebo:	124.6	104.8	96.5	116.3	106.1	128.8	107.2	123.1
	118.1	108.5	120.4	122.5	113.6			
Calcio:	129.1	123.4	102.7	118.1	114.7	120.9	104.4	116.3
	109.6	127.7	108.0	124.3	106.6	121.4	113.2	

14. Bloqueo en exámenes Muchos estudiantes han tenido la experiencia poco placentera de sentir pánico en exámenes pues 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 otro grupo con preguntas de examen ordenadas de difícil a fácil. (Véase la lista de calificaciones de examen en el ejercicio 22 en la sección 8-3). La pantalla de Excel se muestra abajo (con base en datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimco, *Journal of Experimental Education*, vol. 52, núm. 4). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las dos muestras provienen de poblaciones con la misma varianza.

Excel

	A	B	C
1	F-Test Two-Sample for Variances		
2		*Variable 1*	*Variable 2*
3	Mean	27.1152	31.72813
4	Variance	47.01983	18.1489
5	Observations	25	16
6	df	24	15
7	F	2.590782	
8	P(F<=f) one-tail	0.029928	
9	F Critical one-tail	2.287827	

T **15.** Comparación de facilidad de lectura de J. K. Rowling y León Tolstoi Remítase al conjunto de datos 14 en el Apéndice B y utilice un nivel de significancia de 0.05 para probar la aseveración de que, respecto a sus puntuaciones en la evaluación de facilidad de lectura de Flesch, las páginas de *Harry Potter y la piedra filosofal* de J. K. Rowling tienen la misma variación que las páginas de *La guerra y la paz* de León Tolstoi.

T **16.** Comparación de edades de corredores de maratón Remítase al conjunto de datos 8 en el Apéndice B y utilice un nivel de significancia de 0.05 para probar la aseveración de que para los corredores del maratón de la ciudad de Nueva York, los hombres y las mujeres tienen edades con cantidades diferentes de variación.

8-5 Más allá de lo básico

17. Efecto de un dato distante Los métodos de esta sección tienen el requisito bastante estricto de que las dos poblaciones tengan distribuciones normales. La presencia de un dato distante es evidencia en contra de que una población esté distribuida normalmente. Repita el ejercicio 6 después de borrar el dato distante de 504 libras en la muestra de cargas axiales de latas que tienen 0.0111 pulgadas de espesor. Después de eliminar este dato distante, los 174 valores tienen una media de 280.5 libras y una desviación estándar de 22.1 libras. ¿Tiene el dato distante un gran efecto en los resultados?

18. Determinación de valores *P* Para probar una aseveración acerca de dos varianzas poblacionales utilizando el método del valor *P*, primero calcule el estadístico de prueba *F*, luego remítase a la tabla A-5 con la finalidad de determinar cómo se compara con los valores críticos listados para $\alpha = 0.025$ y $\alpha = 0.05$. Con referencia al ejercicio 5, ¿qué se concluye acerca del valor *P*?

19. 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) se calcula como sigue: primero intercambie los grados de libertad y después tome el recíproco de valor F resultante encontrado en la tabla A-5. (F_D algunas veces se denota por $F_{\alpha/2}$ y F_I algunas veces se denota por $F_{1-\alpha/2}$.) Calcule los valores críticos F_D y F_I para pruebas de hipótesis de dos colas con base 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. $n_1 = 7, n_2 = 10, \alpha = 0.05$

20. Construcción de intervalos de confianza Además de probar aseveraciones que incluyen 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.

$$\left(\frac{s_1^2}{s_2^2} \cdot \frac{1}{F_R}\right) < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2} \cdot \frac{1}{F_L}\right)$$

Aquí F_D y F_I son como se describe en el ejercicio 19. Remítase a los datos del ejercicio 13 y construya un estimado del intervalo de confianza del 95% para la proporción de la varianza del grupo placebo a la varianza del grupo con suplementos de calcio.

Repaso

En los capítulos 6 y 7 introdujimos dos conceptos importantes de la estadística inferencial: la estimación de parámetros poblacionales y los métodos para probar hipótesis acerca de parámetros poblacionales. Los capítulos 6 y 7 consideraron sólo casos que comprenden una sola población, pero este capítulo consideró dos muestras obtenidas a partir de dos poblaciones.

- La sección 8-2 consideró inferencias hechas acerca de dos proporciones poblacionales.
- La sección 8-3 consideró inferencias hechas acerca de las medias de dos poblaciones independientes. La sección 8-3 incluyó tres métodos diferentes, pero un método se utiliza en raras ocasiones puesto que requiere que se conozcan las dos desviaciones estándar poblacionales. Otro método consiste en agrupar las dos desviaciones estándar muestrales para desarrollar un estimado del error estándar, pero este método se basa en el supuesto de que se sabe que las dos desviaciones estándar poblacionales son iguales, y esta suposición es con frecuencia riesgosa. Consulte la figura 8-3 para determinar cuál método aplicar.
- La sección 8-4 consideró inferencias hechas acerca de la diferencia media para una población consistente en datos apareados.
- La sección 8-5 presentó métodos para probar aseveraciones acerca de la igualdad de dos desviaciones estándar o varianzas poblacionales.

Ejercicios de repaso

1. ¿Se recuperan mejor los pacientes quirúrgicos tibios? Un artículo publicado en el *USA Today* afirmó que "en un estudio de 200 pacientes de cirugía colorectal, a 104 se les mantuvo tibios con mantas y líquidos intravenosos; a los otros 96 se les mantuvo frescos. Los resultados indican que sólo 6 de los pacientes que se mantuvieron tibios presentaron infecciones en la herida contra 18 de los que se mantuvieron frescos".

a. Utilice un nivel de significancia de 0.05 para probar la aseveración del encabezado del artículo: "Los pacientes quirúrgicos tibios se recuperan mejor". Si estos resultados se confirman, ¿deberá entibiarse por rutina a los pacientes quirúrgicos?

b. Si se utilizara un intervalo de confianza para probar la aseveración del inciso *a*, ¿qué nivel de confianza debería utilizarse?

c. Utilice el nivel de confianza del inciso *b* y construya un estimado del intervalo de confianza de la diferencia entre las dos proporciones poblacionales.

d. 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, ¿será la conclusión basada en el intervalo de confianza siempre la misma que la conclusión de una prueba de hipótesis estándar?

2. 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 semilla de paja (común y secada al horno) que se utilizaron en terrenos adyacentes. Los valores listados son las cosechas de paja en toneladas por acre.

a. 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 semilla.

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 horno	25	24	24	28	22.5	19.5	22.25	16	17.25	15.75	17.25

3. Volumen cerebral y trastornos psiquiátricos 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 de un grupo control de personas saludables. Abajo se dan los resultados muestrales (en mL) de los volúmenes cerebrales totales (datos tomados 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).

a. 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. Suponga que las dos poblaciones tienen varianzas que *no son iguales*.

b. Suponiendo que las varianzas poblacionales *no son 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.

c. 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 control: $n = 10, \bar{x} = 1268.41, s = 137.97$

4. Variación de volúmenes cerebrales Utilice los datos muestrales que se dieron en el ejercicio 3, con un nivel de significancia de 0.05, para probar la aseveración de que las poblaciones de volúmenes cerebrales totales para pacientes obsesivo compulsivos y para el grupo control tienen diferentes cantidades de variación.

5. Monóxido de carbono y cigarrillos Remítase a los datos que se proporcionan para las cantidades medidas de monóxido de carbono (CO) de muestras de cigarrillos largos con filtro y sin filtro. Todas las mediciones son en miligramos, y los datos son de la Federal Trade Commission. Utilice un nivel de significancia de 0.05 para probar la aseveración de que la cantidad media de monóxido de carbono en los cigarrillos largos con filtro es igual a la cantidad media de monóxido de carbono para los cigarrillos largos sin filtro. Con base en este resultado, ¿son eficaces los filtros de cigarrillos en la reducción de monóxido de carbono?

Con filtro:	14 12 14 16 15 2 14 16 11 13 13 12 13 12 13 14 14 14 9 17 12
Sin filtro:	14 15 17 17 16 16 14 16

Grupo con suplemento de zinc	Grupo placebo
$n = 294$ $\bar{x} = 3214$ $s = 669$	$n = 286$ $\bar{x} = 3088$ $s = 728$

6. Zinc para madres Se realizó un estudio de madres con deficiencia de zinc para determinar si el suplemento de zinc durante el embarazo da como resultado bebés con un mayor peso al nacer. Al margen se listan los datos muestrales (según datos de "The Effects of Zinc Supplementation on Pregnancy Outcome", de Goldenberg *et al*., *Journal of the American Medical Association*, vol. 274, núm. 6). Los pesos se midieron en gramos. Utilizando un nivel de significancia de 0.05, ¿existe suficiente evidencia para sustentar la aseveración de que el suplemento de zinc da como resultado un mayor peso al nacer?

7. Personas que ayudan a otras En un estudio de personas que se detienen a ayudar a conductores con automóviles descompuestos, los investigadores formularon la hipótesis de que se detenía más gente a ayudar a alguien si primero veían a otro conductor con un automóvil averiado recibiendo ayuda. En un experimento, 2000 conductores primero vieron a una mujer recibiendo ayuda con un neumático desinflado y luego vieron a una segunda mujer que estaba sola, camino adelante, con un neumático desinflado; el 2.90% de aquellos 2000 conductores se detuvieron a ayudar a la segunda mujer. Otros 2000 conductores no vieron a la primera mujer recibiendo ayuda, y sólo el 1.75% se detuvo a ayudar (según datos de "Help on the Highway", de McCarthy, *Psychology Today*). Con un nivel de significancia de 0.05, pruebe la aseveración de que el porcentaje de personas que se detienen después de haber visto primero a un conductor con un automóvil averiado recibiendo ayuda es mayor que el porcentaje de personas que se detienen sin haber visto primero a alguien recibiendo ayuda.

8. Prueba de efectos del entrenamiento físico Se realizó un estudio para investigar algunos efectos del entrenamiento físico. Los datos muestrales se listan abajo, con todos los pesos dados en kilogramos. (Véase "Effect of Endurance Training on Possible Determinants of VO_2 During Heavy Exercise", de Casaburi *et al*., *Journal of Applied Physiology*, vol. 62, núm. 1).

a. ¿Existe suficiente evidencia para concluir que hay una diferencia entre los pesos previos al entrenamiento y posteriores al entrenamiento? ¿Qué concluye usted acerca del efecto del entrenamiento sobre el peso?

b. Construya un intervalo de confianza del 95% para la media de las diferencias entre los pesos previos al entrenamiento y posteriores al entrenamiento.

Antes del entrenamiento:	99	57	62	69	74	77	59	92	70	85
Después del entrenamiento:	94	57	62	69	66	76	58	88	70	84

Ejercicios de repaso acumulativos

	¿Multados por exceso de velocidad durante el año pasado?	
	Sí	No
Hombres	26	224
Mujeres	27	473

1. Multas por exceso de velocidad para hombres y mujeres Los datos en la tabla adjunta se obtuvieron por medio de una encuesta de sujetos seleccionados al azar (datos tomados de R. H. Bruskin Associates).

a. Si se selecciona a un sujeto de la encuesta al azar, calcule la probabilidad de obtener a alguien multado por exceso de velocidad.

b. Si se selecciona a un sujeto de la encuesta al azar, calcule la probabilidad de obtener a un hombre o alguna persona multados por exceso de velocidad.

c. Calcule la probabilidad de obtener alguna persona multada por exceso de velocidad, dado que la persona seleccionada sea un hombre.

d. Encuentre la probabilidad de obtener alguna persona multada por exceso de velocidad, dado que la persona seleccionada sea una mujer.

e. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el porcentaje de mujeres multadas por exceso de velocidad es menor que el porcentaje de hombres. ¿Podemos concluir que por lo general los hombres corren más que las mujeres?

2. Teléfonos celulares y choques: análisis de reporte de periódico En un artículo de la Associated Press, se reportó que unos investigadores "seleccionaron aleatoriamente a 100 conductores que habían estado implicados en un accidente y a 100 que no. De aquellos que estuvieron implicados en accidentes, el 13.7% tenía un teléfono celular, mientras sólo el 10.6% de los conductores sin accidentes tenían un teléfono en el automóvil". Analice estos resultados.

3. Ensayos clínicos de Viagra En ensayos clínicos de reacciones adversas al fármaco Viagra, el 4.0% de los 734 sujetos del grupo de tratamiento sufrió congestión nasal, y el 2.1% de los 725 sujetos del grupo placebo sufrió congestión nasal (con base en datos de Pfizer Pharmaceuticals).

a. Construya un estimado del intervalo de confianza del 95% de la proporción de usuarios de Viagra que sufrieron congestión nasal.

b. Construya un estimado del intervalo de confianza del 95% de la proporción de usuarios de un placebo que sufrieron congestión nasal.

c. Construya un estimado del intervalo de confianza del 95% de la diferencia entre las dos proporciones poblacionales.

d. Cuando se intenta determinar si existe una diferencia significativa entre las dos proporciones poblacionales, ¿cuál de los siguientes métodos es el mejor?

i. Determinar si los intervalos de confianza en los incisos *a* y *b* se traslapan.

ii. Determinar si el intervalo de confianza en el inciso *c* contiene el valor de cero.

iii. Realizar una prueba de hipótesis de la hipótesis nula $p_1 = p_2$, utilizando un nivel de significancia de 0.05.

iv. Los métodos de los incisos *i*, *ii* y *iii* son todos igualmente buenos.

T **4.** Finalistas de maratón Remítase a los resultados de mujeres finalistas de maratón en el conjunto de datos 8 en el Apéndice B.

a. Calcule la proporción de mujeres corredoras que finalizaron en el maratón de la ciudad de Nueva York y luego pruebe la aseveración de que la proporción es menor que 0.5.

b. Para los tiempos de las mujeres finalistas, calcule la media, la mediana, la desviación estándar, describa la naturaleza de la distribución e identifique cualquier dato distante.

c. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las mujeres finalistas tienen un tiempo medio de carrera menor que 5 horas.

continúa

d. Los corredores varones que se incluyen en el conjunto de datos 8 tienen un tiempo medio de 15,415.2 segundos y una desviación estándar de 3036.8 segundos. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el tiempo medio para hombres es diferente del tiempo medio para mujeres.

e. Identifique las proporciones de corredoras mujeres y corredores hombres. ¿Qué error se comete al utilizar estas dos proporciones muestrales con los métodos de la sección 8-2 en una prueba de la aseveración de que las proporciones poblacionales de mujeres y hombres son diferentes?

Actividades de cooperación en equipo

1. Actividad fuera de clase ¿Se ven influidos los estimados por números anclados? Remítase a la actividad de cooperación en grupos relacionada en el capítulo 2. En el capítulo 2 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 2, se les 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 2, pudimos comparar los dos conjuntos de resultados utilizando estadísticos (como la media) y gráficas (como las gráficas de caja). Ahora los métodos del capítulo 8 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 Divida al grupo de acuerdo al género, con alrededor de 10 o 12 estudiantes en cada equipo. Cada miembro del grupo deberá registrar su pulso contando el número de latidos en un minuto, y deben calcularse los estadísticos del grupo (n, $\bar{x}$ y s). Los grupos deben probar 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 convictas 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 con el género de la persona que realiza las preguntas. ¿Parece que la respuesta es 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.

Proyecto tecnológico

El STATDISK, el Minitab, Excel y la calculadora TI-83 Plus generan datos distribuidos normalmente obtenidos de una población con una media y 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 a partir 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 Plus Oprima **MATH**, seleccione **PRB**, luego seleccione **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 repite muchas veces, ¿cuál es el porcentaje esperado de ensayos que nos llevan 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 repite 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?

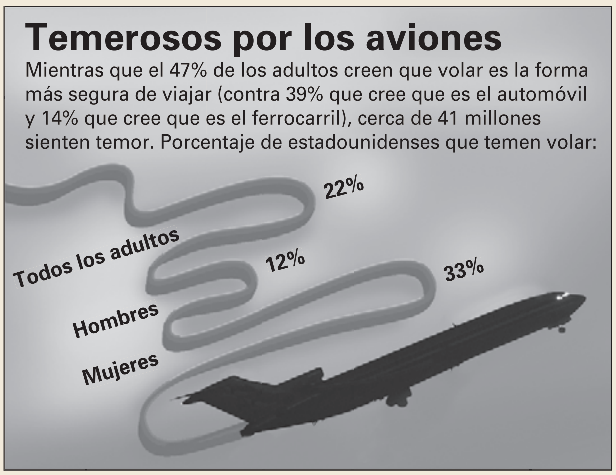
de los DATOS *a la* DECISIÓN

Pensamiento crítico: el temor de volar

Las vidas de muchas personas se ven afectadas por un temor que les impide volar. El comentarista de deportes John Madden ganó popularidad cuando cruzaba el país en ferrocarril o en un remolque vivienda, 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 se describen en la ilustración anexa publicada en *USA Today*. Los resultados de la encuesta muestran que el 12% de los hombres y el 33% de las mujeres temen volar.

Análisis de los resultados

1. ¿Existe evidencia suficiente para concluir que hay una diferencia significativa entre el porcentaje de hombres y el porcentaje de mujeres que temen volar?
2. Construya un estimado del intervalo de confianza del 95% de la diferencia entre el porcentaje de hombres y el porcentaje de mujeres que temen volar. ¿Contienen los límites del intervalo de confianza a 0, y cuál es la significancia de si lo contienen o no?



Fuente: Marist Institute for Public Opinion **Por Scott Boeck y Genevieve Lynn, *USA Today***

Fotografía de USA Today. "A look at statistics that shape the nation."

continúa

3. Construya un intervalo de confianza del 95% para el porcentaje de hombres que temen volar.
4. Con base en el resultado del ejercicio 3, 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 teme volar es del 22%, con un margen de error de ______".
5. Examine la afirmación completa del ejercicio 4. ¿Qué parte importante de la información que no está incluida debiera incluirse?
6. En una encuesta separada de Gallup, se pidió a 1001 adultos seleccionados al azar que respondieran a esta pregunta: "Si usted tuviera que volar mañana en un avión, ¿cómo describiría sus sensaciones acerca de volar? ¿Estaría usted muy temeroso, un poco temeroso, no muy temeroso, o sin temor alguno?". Aquí están 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 explicarían las discrepancias por el hecho de que la encuesta de Gallup se realizó después de los ataques terroristas del 11 de septiembre del 2001, mientras que la otra encuesta se realizó antes de esa fecha?
7. ¿Funciona bien la ilustración del *USA Today* como descripción de los resultados de la encuesta? Construya una gráfica que ilustre con mayor claridad los resultados de la encuesta.

PROYECTO DE INTERNET Comparación de poblaciones

El capítulo anterior le mostró métodos para prueba de hipótesis acerca de una sola población. Este capítulo se explaya en 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. Vaya al proyecto de Internet para este capítulo a

http://www.pearsoneducacion.net/triola

Ahí usted encontrará varios problemas de prueba de hipótesis que incluyen a múltiples poblaciones. En estos problemas, usted analizará la imparcialidad de los salarios, la demografía poblacional 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 un conocimiento funcional de estadística básica".

Mark T. Lycett

Kathleen Morrison

Mark T. Lycett y Kathleen Morrison están en la facultad del Departamento de Antropología en la Universidad de Chicago.

La investigación del doctor Lycett trata 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 el uso de la estadística en la arqueología?

Sería imposible realizar investigación arqueológica sin un conocimiento funcional de estadística básica.

¿Qué conceptos de estadística utilizan ustedes?

Los arqueólogos hacen un uso extensivo de la estadística descriptiva e inferencial a diario. El análisis exploratorio de datos que utiliza una variedad de resúmenes gráficos y numéricos cada vez es más común en la arqueología moderna. Los problemas arqueológicos incluyen de rutina estudios de asociación para variables categóricas, pruebas de hipótesis para datos de 2 muestras y k muestras, problemas de correlación y regresión y una serie de métodos no paramétricos.

Por favor, expongan 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 similar morfológicamente, los cultivos alimenticios básicos del nuevo mundo (maíz) tienen granos de polen mucho más grandes que la hierba silvestre, y 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 fósil a partir de contextos arqueológicos, hemos sido capaces de 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; y para estudiar los efectos del colonialismo y el imperialismo en las prácticas sociales, económicas y religiosas locales.

¿El uso que ustedes hacen de la probabilidad y de la estadística, aumenta, decrece 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 sofisticadas se vuelven disponibles a través del uso muy difundido de la tecnología de Geographic Information Systems.

En términos de estadística, ¿qué recomendarían ustedes a los futuros trabajadores?

Cuando éramos estudiantes universitarios, entendimos que la estadística sería parte de nuestras vidas profesionales, pero nunca imaginamos el grado en el que la utilizaríamos a diario. Los estudiantes interesados en la arqueología deben comenzar con un curso introductorio de probabilidad y estadística. Quienes tengan metas profesionales o académicas en este campo deben considerar un curso más avanzado en la licenciatura o a nivel de posgrado en análisis de datos cuantitativos.

9

Correlación y regresión

PROBLEMA DEL CAPÍTULO

¿Deberían imponerse restricciones a la navegación de barcos para reducir las muertes de manatíes?

Los manatíes, también llamados “vacas marinas”, son grandes mamíferos que viven bajo el agua, con frecuencia cerca de rutas marinas con un gran tráfico de embarcaciones. En Florida, las muertes de manatíes por encuentros con barcos han sido motivo de gran controversia entre ambientalistas y operadores de embarcaciones. Recientemente, Andrew Revkin escribió un artículo para el *The New York Times* titulado “How Endangered a Species?” El periodista afirmó que “las muertes [de manatíes] por barcos han continuado a pesar de la creación de una red de refugios y zonas de baja velocidad para las embarcaciones; el resultado es uno de los debates más intensos del país sobre especies en peligro de extinción”. El artículo incluyó dos gráficas juntas, similares a las que se presentan en la figura 9-1. (Las gráficas del *New York Times* incluían datos desde 1976 hasta el 2000). Las gráficas de la figura 9-1 reflejan los datos de la tabla 9-1 (de acuerdo con datos del Florida Department of Highway Safety and Motor Vehicles y del Florida Marine Research Institute).

He aquí algunos temas cubiertos en este capítulo:

- En comparación con la figura 9-1, ¿hay otra gráfica que ilustre mejor la relación entre el número de barcos registrados y el número de muertes de manatíes por los barcos?
- ¿De qué manera es posible utilizar métodos de estadística para determinar objetivamente si hay una relación entre dos variables, tales como el número de barcos registrados y la cantidad de manatíes muertos por los barcos?
- Si hay una relación entre el número de embarcaciones que se registran y el número de manatíes que asesinan los barcos, ¿cómo se describiría? ¿Existe alguna ecuación que sea posible utilizar para predecir las muertes de manatíes que provocan las barcos, a partir de un número particular de barcos registrados?
- Además del número de barcos registrados, ¿hay otras variables importantes que afecten el número de manatíes que asesinan los barcos?

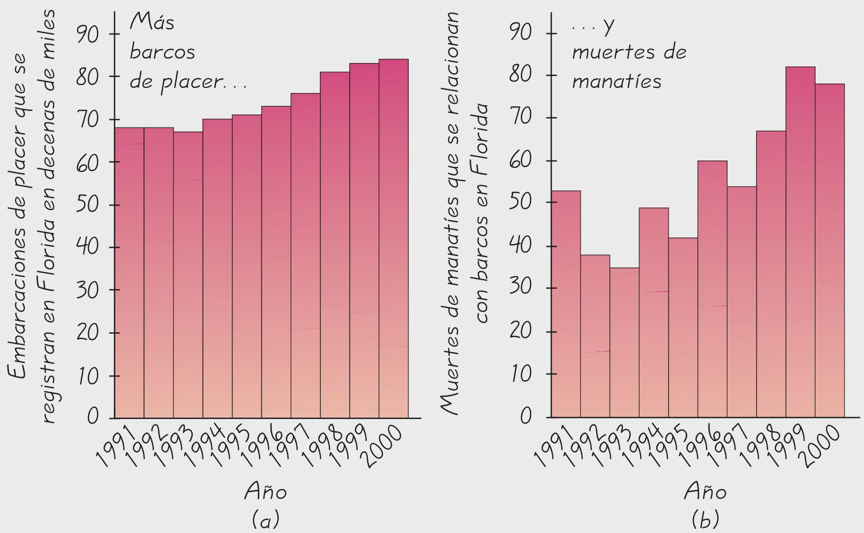


FIGURA 9-1 Embarcaciones de placer que se registran en Florida (en decenas de miles) y muertes de manatíes que se relacionan con barcos

Tabla 9-1 Embarcaciones de placer que se registraron en Florida (en decenas de miles) y muertes de manatíes que se relacionan con barcos

Año	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
x: Barcos	68	68	67	70	71	73	76	81	83	84
y: Muertes de manatíes	53	38	35	49	42	60	54	67	82	78

9-1 Panorama general

Este capítulo introduce métodos importantes para hacer inferencias que se basan en datos muestrales que se ordenan en *pares*. La sección 8-4 utilizó muestras apareadas, pero las inferencias en la sección 8-4 se referían a diferencias entre dos medias poblacionales. Este capítulo tiene el objetivo de determinar si hay una relación entre las dos variables; en caso de existir tal relación, queremos describirla con una ecuación que permita hacer predicciones.

Iniciamos en la sección 9-2 considerando el concepto de correlación, que se emplea para determinar si existe una relación estadísticamente significativa entre dos variables. Investigaremos la correlación por medio del diagrama de dispersión (una gráfica) y el coeficiente de correlación lineal (una medida de la dirección y el poder de la asociación lineal entre dos variables). En la sección 9-3 investigaremos el análisis de regresión; describiremos la relación entre dos variables con una ecuación que las relaciona y mostraremos cómo utilizar esa ecuación para predecir valores de una variable cuando conocemos los valores de la otra variable.

En la sección 9-4 analizaremos las diferencias entre los valores predichos y los valores reales que se observan de una variable. Las secciones 9-2 a 9-4 implican relaciones entre *dos* variables, pero en la sección 9-5 emplearemos conceptos de regresión múltiple para describir la relación entre tres o más variables. Finalmente, en la sección 9-6 describiremos algunos métodos básicos para crear un modelo matemático que permita describir la relación entre dos variables. Aun cuando la sección 9-3 se limita a relaciones lineales, la sección 9-6 incluye algunas relaciones no lineales comunes.

9-2 Correlación

El principal objetivo de esta sección es analizar un conjunto de datos muestrales apareados (que llamamos en ocasiones **datos bivariados**) y determinar si parece haber una relación entre las dos variables. En estadística, nos referimos a una relación como ésta como una *correlación*. (Consideraremos únicamente relaciones *lineales*, lo que significa que, cuando se grafica, los puntos se aproximan a un patrón de línea recta. Además, consideramos solamente datos cuantitativos).

Definición

Correlación: Existe entre dos variables cuando una de ellas se relaciona con la otra de alguna manera.

La tabla 9-1, por ejemplo, se compone de datos barco/manatí apareados para cada año de la década pasada. Determinaremos si hay una correlación entre la variable x (número de barcos que se registran) y la variable y (número de manatíes que asesinan los barcos).

Exploración de los datos

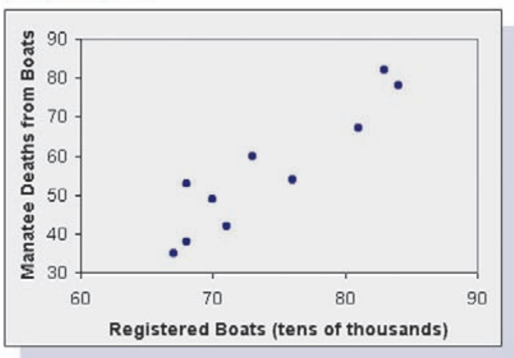
Antes de trabajar con los métodos más formales de cálculo de esta sección, primero debemos explorar el conjunto de datos para ver qué es posible aprender. Con frecuencia encontraremos una relación entre dos variables al construir una gráfica que se denomina *diagrama de dispersión*.

Definición

Diagrama de dispersión: Una gráfica en la que datos muestrales apareados (x, y) se grafican en un eje x horizontal y un eje y vertical. Cada par individual (x, y) se grafica como un solo punto.

Como ejemplo, véase el resultado de Excel de los 10 pares de datos que se listan en la tabla 9-1. Cuando examinamos un diagrama de dispersión como ése, es necesario estudiar el patrón general de los puntos graficados. Si existe un patrón, debemos señalar su dirección. Es decir, mientras una variable se incrementa, ¿la otra parece aumentar o disminuir? Tenemos que observar si hay datos distantes, que son puntos que se ubican muy lejos de todos los demás. El diagrama de dispersión que se genera en Excel no parece revelar un patrón que indique que un mayor número de barcos que se registran se asocie con una mayor cantidad de muertes de manatíes debidas a los barcos. El diagrama de dispersión sirve mucho más para visualizar la asociación entre los barcos que se registran y las muertes de manatíes que las gráficas de barras que se muestran en la figura 9-1. Como se ordenan de acuerdo con una secuencia temporal, las gráficas de barras de la figura 9-1 son muy útiles para mostrar la tendencia a largo plazo del número de barcos que se registran y el número de muertes de manatíes que ocasionan los barcos, aunque el diagrama de dispersión es útil para ilustrar la relación existente entre esas dos variables.

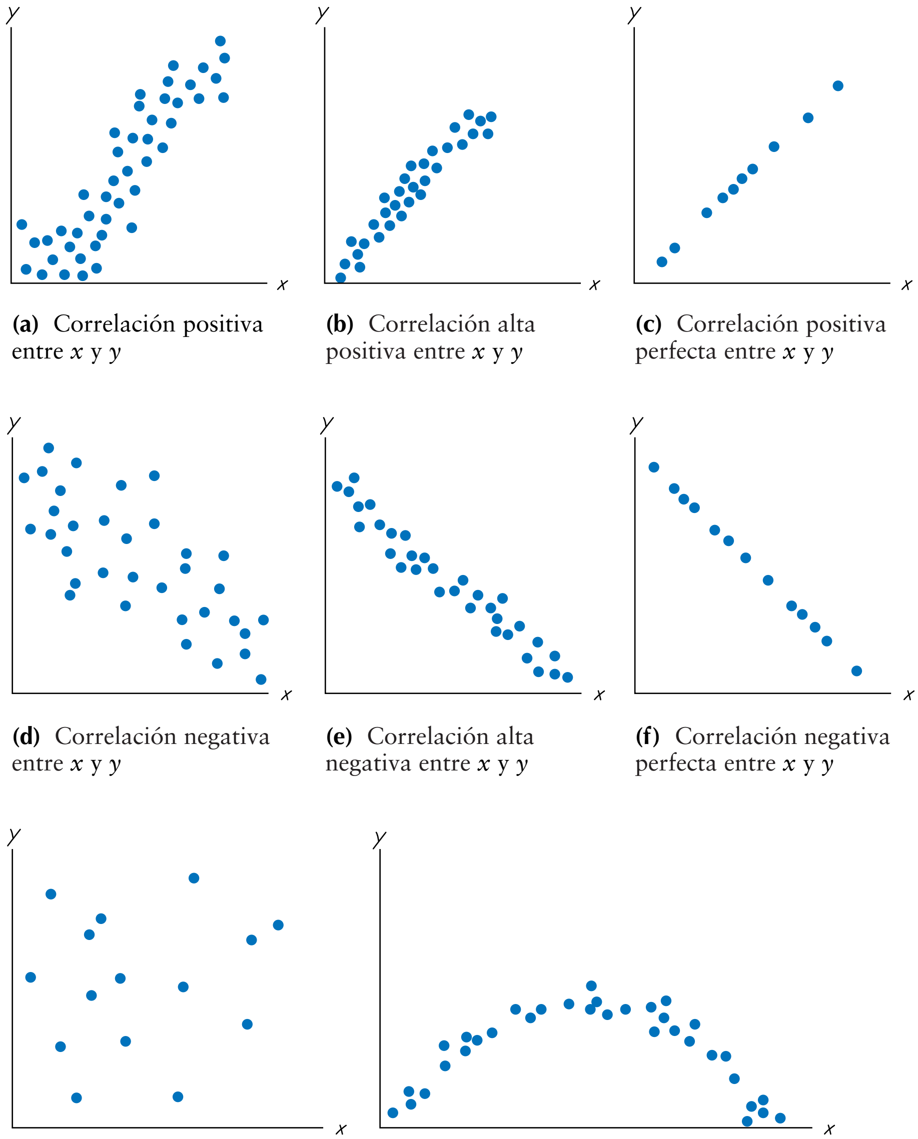
Excel



Los profesores son evaluados por estudiantes

Muchas universidades consideran que las buenas evaluaciones de los profesores, hechas por estudiantes, son equivalentes a una buena enseñanza, ecuación que se fomenta por el hecho de que las evaluaciones de los estudiantes son fáciles de administrar y medir. Sin embargo, un estudio que comparó evaluaciones de profesores hechas por alumnos con la cantidad de material que se aprende reveló una fuerte correlación *negativa* entre los dos factores. Los maestros que mejor evalúan los alumnos parecieron producir menor aprendizaje. En un estudio que se relaciona, una audiencia otorgó una evaluación alta a un conferencista que ofreció muy poca información, pero que era interesante y entretenido. (Véase "Rating the Teachers", de Miriam Rodin, *Center Magazine,* vol. VIII, núm. 5).

En la figura 9-2 se presentan otros ejemplos de diagramas de dispersión. Las gráficas en la figura 9-2a, b y c describen un patrón de valores crecientes de *y*, que corresponde a valores crecientes de *x*. Conforme vamos de *a* a *c*, el patrón de puntos se aproxima a una línea recta, lo que sugiere que la relación entre *x* y *y* se hace más fuerte. Los diagramas de dispersión en *d*, *e* y *f* describen patrones en los que los valores de *y* disminuyen mientras los de *x* aumentan. Nuevamente, conforme vamos de *d* a *f*, la relación se hace más fuerte. En contraste con las primeras seis gráficas, el diagrama de dispersión de *g* no presenta ningún patrón y



(a) Correlación positiva entre *x* y *y*

(b) Correlación alta positiva entre *x* y *y*

(c) Correlación positiva perfecta entre *x* y *y*

(d) Correlación negativa entre *x* y *y*

(e) Correlación alta negativa entre *x* y *y*

(f) Correlación negativa perfecta entre *x* y *y*

(g) Sin correlación entre *x* y *y*

(h) Relación no lineal entre *x* y *y*

FIGURA 9-2 Diagramas de dispersión

sugiere que no hay correlación alguna (o relación) entre x y y. Finalmente, el diagrama de dispersión de h indica un patrón, pero no de tipo lineal.

Coeficiente de correlación lineal

Puesto que el examen visual de los diagramas de dispersión es muy subjetivo, necesitamos medidas más precisas y objetivas. El coeficiente de correlación lineal r sirve para detectar patrones lineales.

Definición

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 9-1, que se incluye en el siguiente recuadro. [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 con datos muestrales, se trata de un estadístico muestral que se emplea para medir la fuerza de la correlación lineal entre x y y. Si tuviésemos cada par de los valores poblacionales de x y y, el resultado de la fórmula 9-1 sería un parámetro poblacional, que se representaría por ρ (rho griega). El siguiente recuadro incluye los supuestos que se requieren, la notación y la fórmula 9-1.

Supuestos

1. La muestra de datos apareados (x, y) es una muestra *aleatoria* de datos cuantitativos.
2. Los pares de datos (x, y) tienen una **distribución normal bivariada**. (Las distribuciones normales se estudiaron en el capítulo 5, pero este supuesto requiere básicamente que, para cualquier valor fijo de x, los valores correspondientes de y contengan 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, pero es posible realizar una verificación parcial determinando si las distribuciones básicamente de los valores de x y y tienen forma de campana.

Notación para el coeficiente de correlación lineal

n	representa el número de pares de datos presentes.
Σ	denota la suma de los elementos que se indican.
Σx	denota la suma de todos los valores de x.
Σx^2	indica que cada valor de x debe elevarse al cuadrado y que después dichos cuadrados se suman.
$(\Sigma x)^2$	indica que los valores de x deben sumarse y el total elevarse al cuadrado. Es sumamente importante evitar confundirse entre Σx^2 y $(\Sigma x)^2$.

continúa

LA ESTADÍSTICA EN LAS NOTICIAS

“El mercado de valores varía de acuerdo con la superstición del triunfo patriota”

Este encabezado del *New York Post* es una afirmación acerca del *pronóstico del Súper Bowl*, que establece que la victoria de un equipo de la NFL es seguida por un año cuando el índice de intercambio de la bolsa de Nueva York se incrementa; de otra manera, disminuye. (En 1970, la NFL y la AFL se unieron para formar la actual NFL). Tal indicador ha sido correcto en 29 de los últimos 35 años, principalmente por el hecho de que los equipos de la NFL ganan con mayor frecuencia y que el mercado de valores tiende a subir con el paso del tiempo. El pronóstico y las predicciones son metas importantes de los estadísticos y de los consejeros de inversiones, aunque el sentido común sugiere que nadie debe basar sus inversiones en el resultado de un juego de futbol. Otros indicadores que se utilizan para pronosticar el desempeño del mercado de valores incluye la aparición de faldas más cortas, la venta de aspirinas, las limusinas en Wall Street, las órdenes de cajas de cartón, las ventas de cerveza en relación con las de vino y el tráfico de elevadores en el mercado de valores de Nueva York.

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 *muestra*.
ρ	representa el coeficiente de correlación lineal de una *población*.

Fórmula 9-1 $$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

Interpretación de *r* por medio de la tabla A-6: Si el valor absoluto del valor que se calculó de r excede el valor de la tabla A-6, concluya que hay una correlación lineal significativa. De lo contrario, no existe evidencia suficiente para sustentar la conclusión de una correlación lineal significativa.

Redondeo del coeficiente de correlación lineal

Redondee el coeficiente de correlación lineal r a tres decimales (de manera que su valor se compare directamente con los valores críticos de la tabla A-6). Al calcular r y otros estadísticos en este capítulo, hacer un redondeo a la mitad de un cálculo suele crear errores importantes; por lo tanto, trate de utilizar la memoria de su calculadora para almacenar los resultados inmediatos y redondee sólo al final. Muchas calculadoras baratas incluyen la fórmula 9-1, por lo que evalúan automáticamente r después de introducir los datos muestrales.

EJEMPLO Cálculo de *r* Con los datos que se presentan a continuación, calcule el valor del coeficiente de correlación lineal r.

x	1	1	3	5
y	2	8	6	4

SOLUCIÓN Para la muestra dada de datos apareados, $n = 4$, ya que existen cuatro pares de datos. Los otros componentes que se requieren para la fórmula 9-1 se obtienen de los cálculos en la tabla 9-2. Note cómo este formato vertical facilita los cálculos.

Con los valores que se calcularon y la fórmula 9-1 es posible evaluar r de la siguiente forma:

$$\begin{aligned} r &= \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}} \\ &= \frac{4(48) - (10)(20)}{\sqrt{4(36) - (10)^2}\ \sqrt{4(120) - (20)^2}} \\ &= \frac{-8}{\sqrt{44}\ \sqrt{80}} = -0.135 \end{aligned}$$

Tabla 9-2 Obtención de estadísticos que se emplean para calcular r

	x	y	$x \cdot y$	x^2	y^2
	1	2	2	1	4
	1	8	8	1	64
	3	6	18	9	36
	5	4	20	25	16
Total	**10** $\uparrow$ Σx	**20** $\uparrow$ Σy	**48** $\uparrow$ Σxy	**36** $\uparrow$ Σx^2	**120** $\uparrow$ Σy^2

Con grandes conjuntos de datos, estos cálculos dan resultado muy desordenados; por fortuna, el coeficiente de correlación lineal se determina automáticamente usando calculadoras y programas de cómputo. Consulte la sección "Utilizando la tecnología" al final de esta sección para referencias de STATDISK, Minitab, Excel, y la calculadora TI-83 Plus.

Interpretación del coeficiente de correlación lineal

Necesitamos interpretar un valor calculado de r, tal como el valor de -0.135 que se obtuvo en el ejemplo anterior. Como la manera en donde la fórmula 9-1 se construyó, el valor de r siempre debe estar entre -1 y $+1$, inclusive. Si r se acerca a 0, concluimos que no hay una correlación lineal significativa entre x y y, pero si r se acerca a -1 o $+1$, concluimos que hay una correlación lineal significativa entre x y y. Interpretaciones tales como "cercano a" 0 o 1 o -1 son vagas, por lo que utilizamos el siguiente criterio de decisión muy específico:

Si el valor absoluto del valor que se calculó de r excede el valor de la tabla A-6, se concluye que hay una correlación lineal significativa. De lo contrario, no existe evidencia suficiente para sustentar la conclusión de una correlación lineal significativa.

Cuando en realidad no hay 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 A-6 nos indica que con $n = 10$ pares de datos muestrales, los valores críticos son 0.632 (para $\alpha = 0.05$) y 0.765 (para $\alpha = 0.01$). Los valores críticos y el papel de α se describen en detalle en los capítulos 6 y 7. He aquí cómo interpretamos dichos números: con 10 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 que se calcula exceda 0.632. Con $n = 10$ y sin correlación lineal, hay una probabilidad del 1% de que $|r|$ exceda 0.765.

Lectura de las manos

Algunas personas piensan que la longitud de la línea de la vida de la palma de sus manos es útil para predecir la longevidad. En una carta publicada en el *Journal of the American Medical Association*, M. E. Wilson y L. E. Mather refutaron esta creencia con un estudio de cadáveres. Se registraron las edades a la muerte, junto con las longitudes de la línea de la vida de sus palmas. Los autores concluyeron que no hay una correlación significativa entre la edad al morir y la longitud de la línea de la vida. La quiromancia pierde, bajen las manos.

EJEMPLO Barcos y manatíes Con los datos muestrales de la tabla 9-1, calcule el valor del coeficiente de correlación lineal r, después remítase a la tabla A-6 para determinar si existe una correlación lineal significativa entre el número de barcos que se registran y la cantidad de manatíes que asesinan los barcos. En la tabla A-6, utilice el valor crítico para $\alpha = 0.05$. (Con $\alpha = 0.05$, concluimos que hay 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 Al utilizar el mismo procedimiento que se ilustra en el ejemplo anterior o empleando las herramientas tecnológicas, obtenemos que los 10 pares de datos barco/manatí de la tabla 9-1 dan como resultado $r = 0.922$. A continuación se presentan los resultados de Minitab:

Minitab

```
Pearson correlation of Boats and Manatees = 0.922
P-Value = 0.000
```

Si nos remitimos a la tabla A-6, localizaremos el renglón en que $n = 10$ (porque hay 10 pares de datos). Este renglón contiene los valores críticos de 0.632 (para $\alpha = 0.05$) y 0.765 (para $\alpha = 0.01$). Con el valor crítico para $\alpha = 0.05$, observamos que hay una probabilidad menor al 5% de que, sin correlación lineal, el valor absoluto de r que se calculó exceda 0.632. Puesto que $r = 0.922$, su valor absoluto excede 0.632, por lo que concluimos que existe una correlación lineal significativa entre el número de barcos que se registran y la cantidad de muertes de manatíes a consecuencia de los barcos.

Ya señalamos que la fórmula 9-1 requiere que el valor que se calculó de r caiga siempre entre -1 y $+1$, inclusive. Listamos esa propiedad, junto con otras que también son importantes.

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 cualquier variable se convierten a una escala diferente.*
3. *El valor de r no se afecta 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 se diseñó para medir la fuerza de una relación que no sea lineal.

Interpretación de *r*: variación explicada

Si concluimos que hay una correlación lineal significativa entre x y y, obtendremos una ecuación lineal que exprese y en términos de x; la ecuación se utiliza para predecir valores de y a partir de valores dados de x. En la sección 9-3 describiremos un procedimiento para el cálculo de dichas 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, hay otros factores que afectan a y, como la variación aleatoria y otras características que no se incluyen en el estudio. En la sección 9-4 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 se explica por la relación lineal entre x y y.

PROBLEMA DEL CAPÍTULO 9

EJEMPLO Barcos y manatíes Con los datos barco/manatí de la tabla 9-1, encontramos que el coeficiente de correlación lineal es $r = 0.922$. ¿Qué proporción de la variación de las muertes de manatíes se explicaría por la variación en el número de registros de barcos?

SOLUCIÓN Con $r = 0.922$, obtenemos $r^2 = 0.850$. (Redondeando el valor de r resulta $r^2 = 0.849$).

INTERPRETACIÓN Concluimos que 0.850 (o aproximadamente el 85%) de la variación en las muertes de manatíes debidas a los barcos se explicaría por la relación lineal entre el número de registros de barcos y el número de muertes de manatíes por los barcos. Lo anterior implica que cerca del 15% de la variación de estas muertes no se explica por el número de registros de barcos. Otro factor verdaderamente importante es el tamaño de la población de manatíes; de hecho, hay evidencia que indica que su población continúa en crecimiento. Algunas personas argumentan que el incremento de las muertes de manatíes por los barcos se explica por el hecho de que la población creciente de manatíes ocasiona una mayor cantidad de estos animales en el agua, y que la tasa creciente de muertes por los barcos es un síntoma de una población saludable y en aumento de manatíes, por lo que las restricciones adicionales a los barcos son innecesarias. Puesto que los estimados de la población de manatíes se basan en observaciones aéreas, otros argumentan que los estimados del tamaño de esta población son poco confiables. El investigador Thomas Fraser sugirió en un reporte que "el Estado debe poner en marcha un programa vigoroso de captura y etiquetación, y de recaptura para obtener mejor información sobre el tamaño de la población y sus cambios". (Véase la tercera actividad de cooperación en equipo del capítulo 3).

Errores comunes en las correlaciones

Ahora identificamos tres de las fuentes más comunes de errores que se cometen al interpretar los resultados de correlaciones:

1. *Un error común es concluir que la correlación implica causalidad*. Con los datos muestrales de la tabla 9-1, concluiríamos que hay una correlación entre el número de barcos que se registran y la cantidad de manatíes que son asesinados por barcos, pero no es posible concluir que un mayor número de barcos que se registró *causa* más muertes de manatíes. Las muertes de manatíes por los barcos pueden afectarse por alguna otra variable interventora en los ante-

cedentes. (Una **variable interventora** es aquella que afecta a las variables que se estudian, pero que no está incluida en la investigación). Por ejemplo, temperaturas más cálidas llegan a afectar el número de barcos y el número de manatíes asesinados por los barcos. Por lo tanto, la temperatura sería una variable interventora.

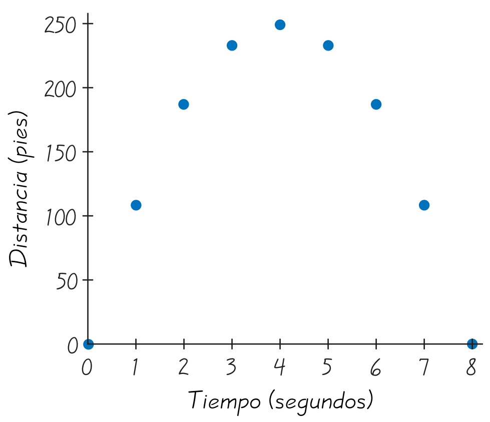
2. *Otro error proviene de los datos basados en promedios*. Los promedios eliminan la variación individual y pueden inflar el coeficiente de correlación. Un estudio produjo un coeficiente de correlación lineal de 0.4 para datos apareados que relacionaban el ingreso y la educación de individuos, pero 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 significativa. Los datos presentados en la figura 9-3 tienen un valor de $r = 0$, lo que indica que no existe una correlación *lineal* entre las dos variables. Sin embargo, al observar la figura, con facilidad podemos percibir que existe un patrón que refleja una relación *no lineal* muy fuerte. (La figura 9-3 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).

Prueba formal de hipótesis (requiere el estudio del capítulo 7)

Presentamos dos métodos (resumidos en el recuadro siguiente y en la figura 9-4) para utilizar una prueba formal de hipótesis a fin de determinar si existe una correlación lineal significativa entre dos variables. Algunos profesores prefieren el método 1 debido a que refuerza conceptos introducidos en capítulos anteriores. Otros prefieren el método 2 porque implica cálculos más fáciles. El método 1 emplea la distribución t de Student, con un estadístico de prueba con la forma $t = (r - \mu_r)/s_r$, donde μ_r y s_r denotan el valor aseverado de la media y de 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 se expresa como $\sqrt{(1 - r^2)/(n - 2)}$.

La figura 9-4 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 crí-

FIGURA 9-3 Diagrama de dispersión que muestra un patrón no lineal



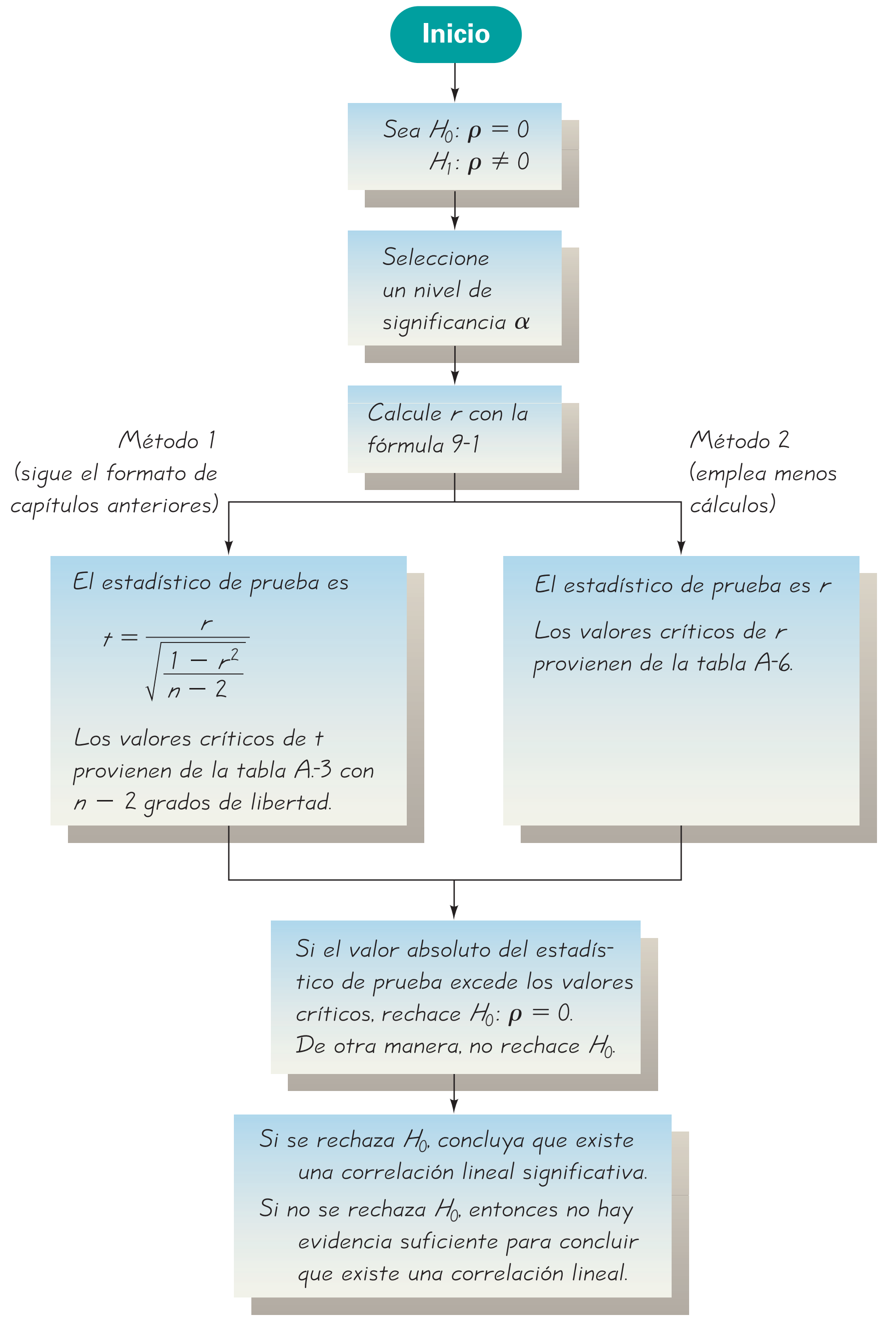


FIGURA 9-4 **Prueba de hipótesis para una correlación lineal**

ticos; 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.

Prueba de hipótesis de correlación (Véase la figura 9-4).

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: $t = \dfrac{r - \mu_r}{\sqrt{\dfrac{1 - r^2}{n - 2}}}$

donde μ_r denota el valor aseverado de la media de los valores de r. Sea $\mu_r = 0$ al probar la hipótesis nula de $\rho = 0$.
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.

PROBLEMA 9 DEL CAPÍTULO

EJEMPLO Barcos y manatíes Con los datos muestrales de la tabla 9-1, pruebe la aseveración de que existe una correlación lineal entre el número de barcos registrados y la cantidad de manatíes asesinados por barcos. Para obtener el estadístico de prueba utilice a) el método 1 y b) el método 2.

SOLUCIÓN Remítase a la figura 9-4. Aseverar que existe una correlación lineal significativa equivale a aseverar que el coeficiente de correlación lineal poblacional ρ es distinto 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 $\alpha = 0.05$.

En un ejemplo previo ya calculamos que $r = 0.922$. 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

$$t = \frac{r}{\sqrt{\dfrac{1 - r^2}{n - 2}}} = \frac{0.922}{\sqrt{\dfrac{1 - 0.922^2}{10 - 2}}} = 6.735$$

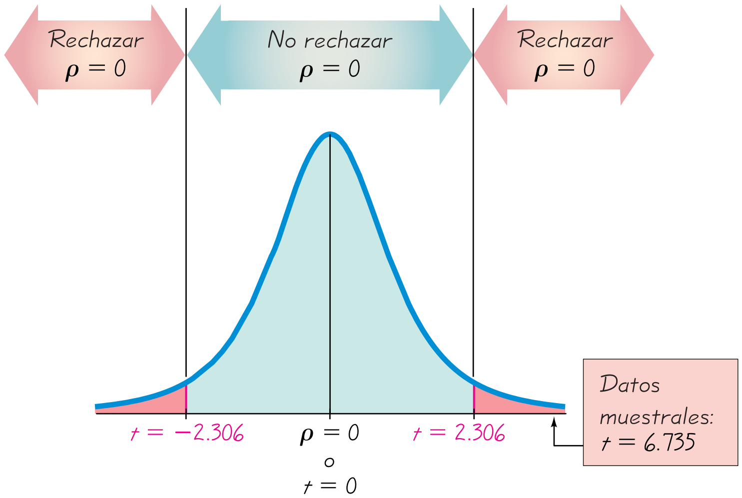


FIGURA 9-5 **Prueba de H_0: $\rho = 0$ con el método 1**

Los valores críticos de $t = \pm 2.306$ se encuentran en la tabla A-3, donde 2.306 corresponde a un área de 0.05, dividida entre dos colas, y el número de grados de libertad es $n - 2 = 8$. Observe la figura 9-5 con la gráfica que incluye el estadístico de prueba y los valores críticos.

b. *Método 2:* El estadístico de prueba es $r = 0.922$. Los valores críticos de $r = \pm 0.632$ se encuentran en la tabla A-6, con $n = 10$ y $\alpha = 0.05$. Observe la figura 9-6 con una gráfica que incluye el estadístico de prueba y los valores críticos.

Con el uso de cualquiera de los dos métodos encontramos que el valor absoluto del estadístico de prueba excede el valor crítico (Método 1: $6.735 > 2.306$. Método 2: $0.922 > 0.632$); 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 el número de barcos registrados y la cantidad de manatíes muertos por los barcos.

Pruebas de una cola: El ejemplo anterior y las figuras 9-5 y 9-6 ilustran una prueba de hipótesis de dos colas. Los ejemplos y los ejercicios de esta sección generalmente implicarán únicamente pruebas de dos colas, aunque podría presentar-

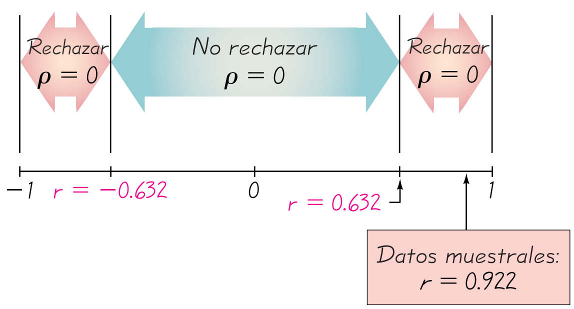


FIGURA 9-6 **Prueba de H_0: $\rho = 0$ con el método 2**

se una prueba de dos colas 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 muestran a continuación.

Aseveración de correlación *negativa* (prueba de cola izquierda)	**Aseveración de correlación *positiva* (prueba de cola derecha)**
H_0: $\rho = 0$ H_1: $\rho < 0$	H_0: $\rho = 0$ H_1: $\rho > 0$

Con estas pruebas de una cola, el método 1 puede seguirse como se hizo en capítulos previos. En el caso del método 2, hay que calcular el valor crítico como se describió en el ejercicio 31, o bien, modificar la tabla A-6 reemplazando los encabezados de columna de $\alpha = 0.05$ y $\alpha = 0.01$ por los valores de un solo lado de $\alpha = 0.025$ y $\alpha = 0.005$, respectivamente.

Fundamentos: Ya presentamos la fórmula 9-1 para el cálculo de r e ilustramos su uso; ahora lo justificaremos. La fórmula 9-1 simplifica los cálculos utilizados en esta fórmula equivalente:

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n - 1)s_x s_y}$$

De manera temporal utilizaremos esta última versión de la fórmula 9-1, ya que su forma 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 9-7.

x	1	1	2	4	7
y	4	5	8	15	23

La figura 9-7 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 suma de los productos de los momentos $(x - \bar{x})$ y $(y - \bar{y})$; es decir, en el estadístico $\Sigma(x - \bar{x})(y - \bar{y})$. 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 muestra en la figura 9-7. Si los puntos del diagrama de dispersión tienden a aproximarse a una línea ascendente (como la figura), los valores individuales del producto $(x - \bar{x})(y - \bar{y})$ tienden a ser positivos debido a que la mayoría de los puntos se encuentran en el

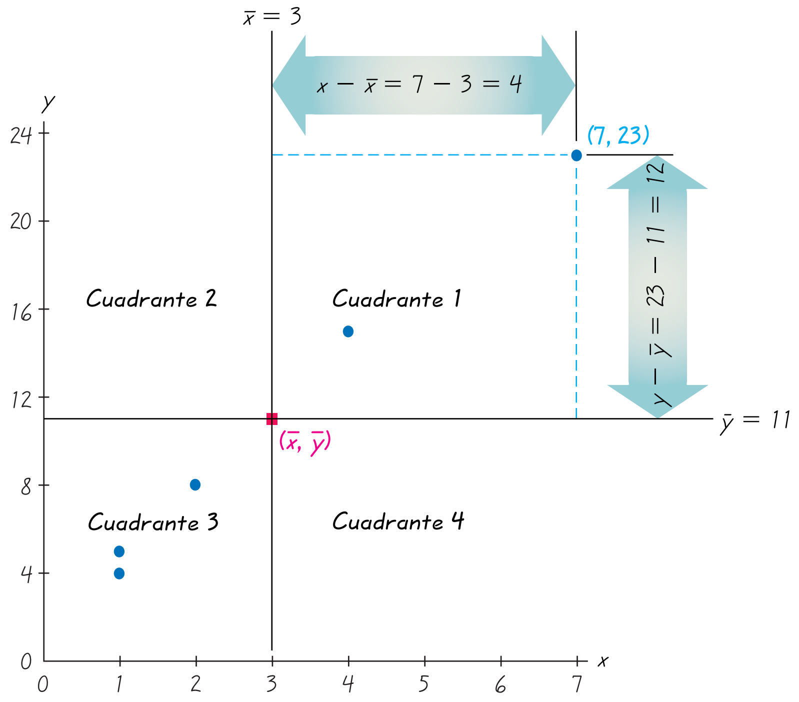


FIGURA 9-7 Diagrama de dispersión dividido cuadrantes

primero 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 signos opuestos, de modo que $\Sigma(x - \bar{x})(y - \bar{y})$ es negativo. Los puntos que no siguen un patrón lineal tienden a dispersarse en los cuatro cuadrantes, de modo que el valor de $\Sigma(x - \bar{x})(y - \bar{y})$ tiende a ser cercano 0.

La suma $\Sigma(x - \bar{x})(y - \bar{y})$ depende de la magnitud de los números utilizados. Por ejemplo, si se cambia x de pulgadas a pies, dicha suma cambiará. Para lograr que r sea independiente de la escala utilizada, incluimos la desviación estándar muestral de la siguiente manera:

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n - 1)s_x s_y}$$

Esta expresión se manipula de manera algebraica en la forma equivalente de la fórmula 9-1.

En capítulos anteriores estudiamos métodos de estadística inferencial y enfatizamos los métodos de prueba de hipótesis, así como también los métodos para construir estimados de intervalos de confianza. Es factible emplear un procedimiento similar para calcular intervalos de confianza para ρ. Sin embargo, debido a que la construcción de dichos intervalos de confianza implica transformaciones que son hasta cierto punto complicadas, ese proceso se presenta en el ejercicio 33 (Más allá de lo básico).

Podemos utilizar el coeficiente de correlación lineal para determinar si existe una relación lineal entre dos variables. Con los datos de la tabla 9-1 hemos concluido que existe una correlación lineal entre el número de barcos registrados y la cantidad de manatíes asesinados por los barcos. Una vez habiendo concluido que existe una relación, nos gustaría determinar de qué relación se trata, de modo que podamos predecir el número de muertes de manatíes para un número dado de barcos registrados. La siguiente etapa de este análisis se estudia en la siguiente sección.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, después utilice la opción **Correlation and Regression**. Introduzca los datos apareados o utilice copiar/pegar para copiar los datos. Ingrese los valores de *x* en la columna 1 y los valores *y* correspondientes en la columna 2. Ingrese 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 se obtienen gráficas, incluyendo un diagrama de dispersión, al hacer clic en los botones Plot 1 y Plot2.

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 proporcionara 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 celda de los valores para *x*, como A1:A10. También ingrese el rango de celda de los valores para *y*, como B1:B10. Para obtener un diagrama de dispersión, haga clic en el Chart Wizard del menú principal, después seleccione el tipo de gráfica identificada como **XY(Scatter)**. En el cuadro de diálogo introduzca el rango de entrada de los datos, como A1:B10. Haga clic en **Next** y proceda a utilizar los cuadros de diálogo para modificar la gráfica como lo desee.

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 icono del lápiz para la variable del eje X e introduzca el rango de valores para la variable *x*, como A1:A10. Haga clic en el icono 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.

TI-83 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 de 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. Establezca las etiquetas de la lista *X* y *Y* para L1 y L2 y presione la tecla de **ZOOM**, finalmente elija **ZoomStat** y presione la tecla **Enter.**

9-2 Destrezas y conceptos básicos

En los ejercicios 1 a 4, utilice un nivel de significancia de $\alpha = 0.05$.

1. Tamaño del pecho y peso de osos Mientras ocho osos se encontraban anestesiados, algunos 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 $r = 0.993$.

a. ¿Existe una correlación lineal significativa entre tamaño del pecho y el peso? Explique.

b. ¿Qué proporción de la variación del peso se explica por la relación lineal entre el peso y el tamaño del pecho?

2. Armas de fuego y tasa de asesinatos Con datos reunidos del FBI y del Bureau of Alcohol, Tobacco and Firearms, se obtuvo el número de armas automáticas registradas y la tasa de asesinatos (en asesinatos por 10,000 individuos) de cada uno de ocho estados de la Unión Americana seleccionados al azar. Por medio de STATDISK se calculó el valor del coeficiente de correlación lineal $r = 0.885$.

a. ¿Existe una correlación lineal significativa entre el número de armas automáticas registradas y la tasa de asesinatos? Explique.

b. ¿Qué proporción de la variación de la tasa de asesinatos puede explicarse por la relación lineal entre la tasa de asesinatos y el número de armas automáticas registradas?

3. Acciones y el Súper Bowl El conjunto de datos 25 del Apéndice B incluye pares de datos del valor elevado del Promedio Industrial Dow Jones (DJIA, por sus siglas en inglés) y el número total de puntos anotados en el Súper Bowl en 21 años diferentes. Se utilizó Excel para calcular el valor del coeficiente de correlación lineal $r = -0.133$.

a. ¿Existe una correlación lineal significativa entre el valor elevado del DJIA y los puntos en el Súper Bowl? Explique.

b. ¿Qué proporción de la variación de los puntos del Súper Bowl se explica por la variación del valor elevado del DJIA?

4. Ventas de automóviles y manchas solares El conjunto de datos 25 del Apéndice B incluye pares de datos del número de manchas solares y del número de ventas de automóviles en Estados Unidos, durante 21 años distintos. Se utilizó la calculadora TI-83 Plus para calcular que el valor del coeficiente de correlación lineal es $r = -0.284$.

a. ¿Existe una correlación lineal significativa entre el número de manchas solares y el número de ventas de automóviles en Estados Unidos? Explique.

b. ¿Qué proporción de la variación del número de ventas de automóviles en Estados Unidos se explica por la variación del número de manchas solares?

Prueba de una correlación lineal. *En los ejercicios 5 y 6, utilice un diagrama de dispersión y el coeficiente de correlación lineal r para determinar si existe una correlación entre las dos variables.*

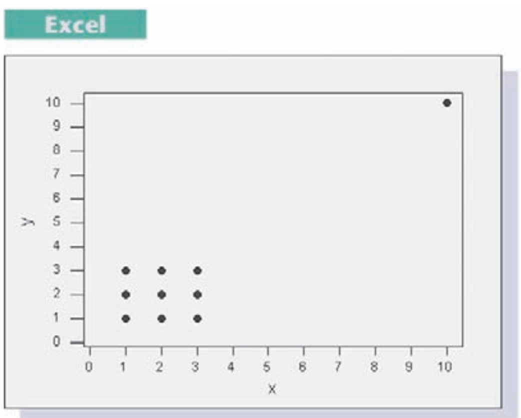
5.

x	0	1	2	3	4
y	4	1	0	1	4

6.

x	1	2	2	5	6
y	2	5	4	15	15

7. Efectos de un dato distante Remítase al siguiente diagrama de dispersión generado por Minitab.



continúa

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 significativa.
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?

Prueba de una correlación lineal. *En los ejercicios 8 a 14, construya un diagrama de dispersión, calcule el valor del coeficiente de correlación lineal r y utilice a nivel de significancia de $\alpha = 0.05$ para determinar si existe una correlación lineal significativa entre las dos variables. Guarde su trabajo, ya que utilizaremos los mismos conjuntos de datos en la siguiente sección.*

8. Incendios y acres quemados A continuación se lista el número de incendios (en miles) y los acres que resultaron quemados (en millones) en 11 estados del oeste de Estados Unidos, durante cada año de la última década (según datos de *USA Today*). ¿Existe una correlación? Los datos se listaron bajo un encabezado de “Loggers seize on fires to argue for more cutting”. ¿Sustentan los datos el argumento de que si los madereros quitan más árboles, el riesgo de incendios disminuye porque los bosques tienen menor densidad?

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

9. Compra de una audiencia televisiva El *New York Post* publicó los salarios anuales (en millones) y el número de televidentes (en millones), los cuales se presentan abajo, correspondientes a Oprah Winfrey, David Letterman, Jay Leno, Kelsey Grammer, Barbara Walters, Dan Rather, James Gandolfini y Susan Lucci, respectivamente. ¿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

10. Estaturas y pesos de supermodelos A continuación se incluyen las estaturas (en pulgadas) y los pesos (en libras) de las supermodelos Niki Taylor, Nadia Averman, Claudia Schiffer, Elle MacPherson, Christy Turlington, Bridget Hall, Kate Moss, Valerie Mazza y Kristy Hume. ¿Existe una correlación entre estatura y peso? Si es así, ¿significa esto que existe una correlación entre la estatura y el peso de todas las mujeres adultas?

Estatura (pulgadas)	71	70.5	71	72	70	70	66.5	70	71
Peso (libras)	125	119	128	128	119	127	105	123	115

11. Mediciones de la presión sanguínea Catorce estudiantes distintos del segundo año de medicina tomaron mediciones de la presión sanguínea del mismo paciente, y los resultados se presentan abajo (datos proporcionados por el doctor Marc Triola). ¿Existe una correlación entre los valores sistólicos y diastólicos? Además de la correlación, ¿habrá algún otro método que pueda utilizarse para enfatizar un aspecto importante sugerido por 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

12. Temperaturas y maratones En “The Effects of Temperature on Marathon Runner’s Performance”, de David Martin y John Buoncristiani (*Chance*, vol. 12, núm. 4), se incluyen las altas temperaturas y 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 los triunfos? ¿Parece que los tiempos ganadores 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

13. 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 qué tanto 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 significativa? 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

14. Circunferencia y altura de árboles A continuación se presentan las circunferencias (en pies) y las alturas (en pies) de árboles en Marshall, Minnesota (datos tomados de “Tree Measurements”, de Stanley Rice, *American Biology Teacher*, vol. 61, núm.9). ¿Existe una correlación? ¿Por qué debería haber una correlación?

x (circunferencia)	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
y (altura)	21.0	33.5	24.6	40.7	73.2	24.9	40.4	45.3	53.5	93.8	64.0	62.7	47.2	44.3

Prueba de una correlación lineal. *En los ejercicios 15 a 24, 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 significativa entre las dos variables. Guarde su trabajo porque utilizaremos los mismos conjuntos de datos en la siguiente sección.*

T **15.** Cereales asesinos Remítase al conjunto de datos 16 del Apéndice B y utilice las cantidades de grasa y los conteos calóricos medidos. ¿Existe una correlación?

T **16.** Tabaco y alcohol en películas infantiles Remítase al conjunto de datos 7 del Apéndice B y utilice los tiempos que las películas infantiles de dibujos animados presentan consumo de tabaco y alcohol. ¿Existe una correlación entre los tiempos para el tabaco y los tiempos para el alcohol?

T **17.** Colesterol e índice de masa corporal Remítase al conjunto de datos 1 del Apéndice B y utilice los niveles de colesterol y los valores del índice de masa corporal de las 40 mujeres. ¿Existe una correlación entre el nivel de colesterol y el índice de masa corporal?

T **18.** Niveles de lectura Remítase al conjunto de datos 14 del Apéndice B y utilice las puntuaciones de la facilidad de lectura de Flesch y los valores del grado Flesch-Kincaid para *El oso y el dragón*, de Tom Clancy. Puesto que ambas puntuaciones están diseñadas para medir la facilidad de lectura, esperaríamos una correlación entre ellas. ¿Existe correlación? ¿De qué manera se explica el valor negativo del coeficiente de correlación?

T **19.** Precios de venta de casas, precios de lista e impuestos Remítase al conjunto de datos 24 del Apéndice B.

a. Utilice los datos apareados consistentes en los precios de lista y los precios de venta de casas. Esperaríamos que estas variables estuviesen relacionadas, ¿pero existe suficiente evidencia para sustentar esta expectativa?

b. Utilice los datos apareados que consisten en el precio de venta de casas y la suma de impuestos. Se supone que el cargo de impuestos debe estar basado en el valor de la casa. ¿Es así? Explique.

T **20.** Alquitrán y nicotina Remítase al conjunto de datos 5 del Apéndice B.

a. Utilice los datos apareados de alquitrán y nicotina. Con base en el resultado, ¿parece existir una correlación lineal significativa 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?

b. Utilice los datos apareados de monóxido de carbono y nicotina. Con base en el resultado, ¿parece existir una correlación lineal significativa 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?

c. Suponga que algunos investigadores desean diseñar un método para predecir la cantidad de nicotina y sólo desean medir algún otro elemento. ¿Cuál será una mejor elección, el alquitrán o el monóxido de carbono? ¿Por qué?

T **21.** Pronóstico del tiempo Remítase al conjunto de datos 10 del Apéndice B.

a. Utilice el pronóstico de altas temperaturas para cinco días y las temperaturas altas reales. ¿Existe correlación? ¿Implica una correlación lineal significativa que las temperaturas pronosticadas a cinco días son precisas?

b. Utilice el pronóstico de altas temperaturas para un día y las temperaturas altas reales. ¿Existe una correlación? ¿Implica una correlación lineal significativa que las temperaturas pronosticadas para un día son precisas?

c. ¿Cuál esperaría que tuviera una mayor correlación con las temperaturas altas reales: el pronóstico de altas temperaturas para cinco días o el pronóstico de altas temperaturas para un día? ¿Coinciden los resultados de los incisos *a* y *b* con lo que esperaría? Si existe una correlación muy alta entre las temperaturas pronosticadas y las temperaturas reales, ¿se infiere que las temperaturas pronosticadas son precisas?

T **22.** Everglades de Florida Remítase al conjunto de datos 12 del Apéndice B.

a. Utilice las temperaturas inferiores y las mediciones de conductividad. ¿Existe una correlación?

b. Utilice las cantidades de lluvia y las mediciones de conductividad. ¿Existe una correlación?

c. Cuando se aparean los valores de conductividad con las mediciones de salinidad (contenido de sal), el coeficiente de correlación es cercano a 1. ¿Qué concluye acerca de la correlación entre la temperatura inferior y la salinidad? ¿Qué concluye acerca de la correlación entre la cantidad de lluvia y la salinidad?

T **23.** Old Faithful Remítase al conjunto de datos 13 del Apéndice B.

a. Utilice los datos apareados de las duraciones y los intervalos después de las erupciones del géiser. ¿Existe una correlación lineal significativa que sugiera que el intervalo posterior a una erupción está relacionado con la duración de la erupción?

b. Utilice los datos apareados de las alturas de las erupciones y de los intervalos después de las erupciones del géiser Old Faithful. ¿Existe una correlación lineal significativa que sugiera que el intervalo posterior a una erupción está relacionado con la altura de la erupción?

c. Suponga que usted desea crear un método para predecir el intervalo hasta la siguiente erupción. Con base en los resultados de los incisos *a* y *b*, ¿qué factores serían más relevantes: la duración de la erupción o la altura de la erupción? ¿Por qué?

T **24.** Precios, quilates y colores de diamantes Remítase al conjunto de datos 18 del Apéndice B.

a. Utilice los datos apareados de los quilates (peso) y el precio. ¿Existe una correlación lineal significativa entre el peso de un diamante, en quilates, y su precio?

b. Utilice los datos apareados color/precio. ¿Existe una correlación lineal significativa entre el color de un diamante y su precio?

c. Suponga que planea comprar un anillo de compromiso de diamantes. Al considerar el valor de un diamante, ¿qué características debería considerar más importantes: el peso en quilates o el color? ¿Por qué?

Identificación de errores de correlación. *En los ejercicios 25 a 28, describa el error en la conclusión. (Véase la lista de errores comunes incluida en esta sección).*

25. *Considere 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.

26. *Considere que:* Existe una correlación lineal significativa entre los ingresos personales y los años de educación.
Conclusión: Una mayor educación causa que se incrementen los ingresos de una persona.

27. *Considere que:* Ciertos sujetos resuelven una prueba de habilidades verbales y una prueba de destreza manual, y dichos 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.

28. *Considere que:* Existe una correlación lineal significativa entre las cargas del impuesto estatal promedio y los ingresos estatales promedio.
Conclusión: Existe una correlación lineal significativa entre las cargas de impuestos individuales y los ingresos individuales.

9-2 Más allá de lo básico

29. Uso de datos de diagramas de dispersión En ocasiones, en lugar de tener datos numéricos, tenemos únicamente datos graficados. El diagrama de dispersión adjunto de Excel es similar al que se incluyó en "The Prevalence of Nosocomial Infection in Intensive Care Units in Europe", de Vincent *et al*., *Journal of American Medical Association*, vol. 274, núm. 8. Cada punto representa un país europeo diferente. Estime el valor del coeficiente de correlación lineal y determine si existe una correlación lineal significativa entre la tasa de mortalidad y la proporción de infecciones adquiridas en las unidades de cuidados intensivos.

30. 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 tabla adjunta, 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 resulta en 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

31. Cálculo de valores críticos de r Los valores críticos de r en la tabla A-6 se calculan resolviendo

$$t = \frac{r}{\sqrt{\dfrac{1 - r^2}{n - 2}}}$$

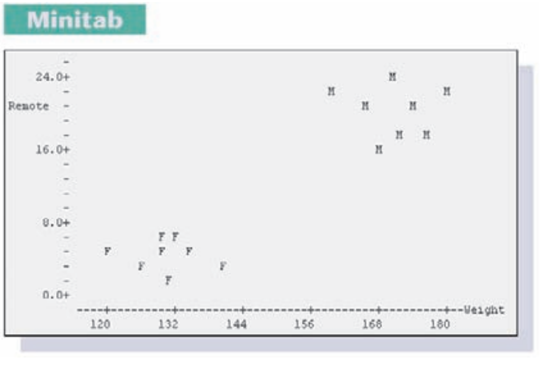
para obtener r

$$r = \frac{t}{\sqrt{t^2 + n - 2}}$$

dónde 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 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: $\rho > 0$, $n = 12$, $\alpha = 0.01$

32. Inclusión de datos categóricos en un diagrama de dispersión En ocasiones se vuelve importante incluir datos categóricos en un diagrama de dispersión. Considere los datos muestrales presentados abajo, donde el peso está representado en libras y los valores "remotos" consisten en el número de veces que el sujeto utilizó el control remoto del televisor durante un periodo de una hora. Se utiliza Minitab para generar el diagrama de dispersión, con las letras F (para las mujeres) y M (para los hombres) para identificar el género.



Género	F	F	F	F	F	F	F	F	M	M	M	M	M	M	M	M
Peso	120	126	129	130	131	132	134	140	160	166	168	170	172	174	176	180
Remoto	5	3	6	4	2	7	4	3	23	20	16	24	18	21	17	22

a. Antes de hacer cualquier cálculo, examine el diagrama de dispersión generado por Minitab. ¿Qué concluye cerca de la correlación entre el peso y el uso del control remoto?

b. Utilice los 16 pares de datos. ¿Existe una correlación entre el peso y el uso del control remoto?

c. Utilice sólo a las ocho mujeres. ¿Existe una correlación entre el peso y el uso del control remoto?

d. Utilice sólo a los ocho hombres. ¿Existe una correlación entre el peso y el uso del control remoto?

e. Con base en los resultados anteriores, ¿qué concluye?

33. Construcción de intervalos de confianza para ρ Dados n pares de datos con los que se puede calcular el coeficiente de correlación lineal r, utilice el siguiente procedimiento para construir un intervalo de confianza acerca del parámetro poblacional ρ.

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_I y w_D del intervalo:

$$w_I = \frac{1}{2}\ln\left(\frac{1 + r}{1 - r}\right) - z_{\alpha/2} \cdot \frac{1}{\sqrt{n - 3}}$$

$$w_D = \frac{1}{2}\ln\left(\frac{1 + r}{1 - r}\right) + z_{\alpha/2} \cdot \frac{1}{\sqrt{n - 3}}$$

Paso c. Ahora evalúe los límites del intervalo de confianza en la siguiente expresión.

$$\frac{e^{2w_I} - 1}{e^{2w_I} + 1} < \rho < \frac{e^{2w_D} - 1}{e^{2w_D} + 1}$$

Utilice este procedimiento para construir un intervalo de confianza del 95% para ρ, dados 50 pares de datos para los cuales $r = 0.600$.

9-3 Regresión

En la sección 9-2 analizamos pares de datos con el objetivo de determinar si existía una correlación entre dos variables. El principal objetivo 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 la relación. Esta recta se conoce como *recta de regresión* y su ecuación como *ecuación de regresión*. 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 que se presenta a continuación 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 la **variable independiente**, **variable predictora** o **variable explicativa**) y $\hat{y}$ (llamada la **variable dependiente** o **variable de respuesta**). 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 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 se utilizan 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 9-2 y 9-3.

Supuestos

1. Estamos investigando únicamente relaciones *lineales*.
2. Para cada valor de x, y es una variable aleatoria con una distribución normal (en forma de campana). Todas estas distribuciones de y tienen la misma varianza. Además, para un valor dado de x, la distribución de los valores de y tiene una media que está en la recta de regresión. (Los resultados no se ven gravemente afectados si las desviaciones de las distribuciones normales y las varianzas iguales no son demasiado extremas).

Definiciones

Dado un conjunto de datos muestrales apareados, la **ecuación de regresión**

$$\hat{y} = b_0 + b_1 x$$

describe algebraicamente la relación entre dos variables. La gráfica de la ecuación de regresión se denomina **recta de regresión** (*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	$y = \beta_0 + \beta_1 x$	$\hat{y} = b_0 + b_1 x$

Cálculo de la pendiente b_1 y del intercepto b_0 en la ecuación de regresión $\hat{y} = b_0 + b_1 x$

Fórmula 9-2 **Pendiente:** $$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Fórmula 9-3 **Intercepto *y*:** $$b_0 = \bar{y} - b_1 \bar{x}$$

El intercepto y, b_0, también se calcula usando la fórmula siguiente, pero es mucho más fácil utilizar la fórmula 9-3.

$$b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Tal vez las fórmulas 9-2 y 9-3 intimiden al estudiante, pero están consideradas en muchas calculadoras y programas de cómputo, de modo que los valores de b_0 y b_1 se calculan con facilidad. (Véase “Utilizando la tecnología” al final de esta sección). En aquellos casos en los que debemos aplicar las fórmulas en lugar de emplear una calculadora o un programa de cómputo, los cálculos requeridos serán mucho más fáciles si tenemos en mente los siguientes hechos:

1. Si se calculó el coeficiente de correlación r por medio de la fórmula 9-1, ya se tienen los valores de Σx, Σy, Σx^2 y Σxy, y pueden emplearse nuevamente en la fórmula 9-2. (Además, el numerador para r en la fórmula 9-1 es el mismo numerador que para b_1 en la fórmula 9-2; el denominador para r incluye el denominador para b_1. Si el cálculo de r se realiza con cuidado, el cálculo de b_1 requiere sólo de dividir un número conocido entre otro).
2. Si utiliza la fórmula 9-2 para calcular primero la pendiente de b_1, es fácil emplear la fórmula 9-3 para calcular b_0, el intercepto de y. [La recta de regresión siempre pasa por el centroide $(\bar{x}, \bar{y})$, de modo que $\bar{y} = b_0 + b_1\bar{x}$ siempre debe ser verdadera, y esta ecuación se expresa como la fórmula 9-3].

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 posteriormente). 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 de y

Es difícil proporcionar una regla simple universal para el redondeo de los valores de b_1 y b_0, pero generalmente tratamos de redondear cada uno de estos valores hasta *tres dígitos significativos* o utilizamos los valores proporcionados por STATDISK, Minitab, Excel o la calculadora TI-83 Plus. Puesto que estos valores son muy sensibles al redondeo durante los pasos intermedios del cálculo, trate de conservar al menos seis dígitos significativos (o utilice valores exactos) en los pasos intermedios. Dependiendo de la forma en que usted realice el redondeo, las respuestas a los ejemplos y ejercicios de este libro pueden variar ligeramente de sus respuestas.

EJEMPLO Cálculo de la ecuación de regresión En la sección 9-2 empleamos los valores listados abajo para calcular el coeficiente de correlación lineal de $r = -0.135$. (Con el uso de los métodos de la sección 9-2 no existe una correlación lineal significativa entre x y y). Use los datos muestrales dados para calcular la ecuación de regresión.

x	1	1	3	5
y	2	8	6	4

SOLUCIÓN Calculamos la ecuación de regresión por medio de las fórmulas 9-2 y 9-3. Los siguientes valores ya se obtuvieron en la tabla 9-2, en la sección 9-2.

$$n = 4 \qquad \Sigma x = 10 \qquad \Sigma y = 20$$
$$\Sigma x^2 = 36 \qquad \Sigma y^2 = 120 \qquad \Sigma xy = 48$$

continúa

LA ESTADÍSTICA EN LAS NOTICIAS

Error de pronóstico de 1° = mil millones de dólares

A pesar de que el pronóstico de las temperaturas a veces parece una ciencia inexacta, muchas compañías están trabajando con fervor para obtener estimados más precisos. El reportero de *USA Today*, Del Jones, escribió que "el costo anual de la electricidad podría disminuir por lo menos mil millones de dólares si se mejorara la precisión 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 dos 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 ubicación de la energía proveniente de generadores, plantas nucleares, plantas hidroeléctricas, carbón, gas natural y del viento. Las técnicas de pronóstico estadístico están siendo refinadas, de modo que se pueda ahorrar dinero y recursos naturales.

Teléfonos celulares y choques

Debido a que algunos países han prohibido el uso de teléfonos celulares en los automóviles y a que otros países están considerando dicha prohibición, algunos investigadores estudiaron si el empleo de teléfonos celulares al conducir incrementa la posibilidad de un choque. Se obtuvo una muestra de 699 conductores. Los miembros del grupo muestral utilizaron teléfonos celulares y sufrieron choques. Los sujetos completaron cuestionarios y se verificaron sus registros telefónicos. Se comparó el intervalo de tiempo entre el uso del teléfono celular y el choque, con un periodo comparable del día anterior. Conclusión: El riesgo de chocar es cuatro veces mayor cuando se utiliza un teléfono celular que cuando no se utiliza. (Véase "Association between Cellular-Telephone Calls and Motor Vehicle Collisions", de Redelmeier y Tibshirani, *New England Journal of Medicine,* vol. 336, núm. 7).

Primero calcule la pendiente b_1 con la fórmula 9-2:

$$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$
$$= \frac{4(48) - (10)(20)}{4(36) - (10)^2} = \frac{-8}{44} = -0.181818 = -0.182$$

Después, calcule b_0, el intercepto de y, con la fórmula 9-3 (con $\bar{y} = 20/4 = 5$ y $\bar{x} = 10/4 = 2.5$):

$$b_0 = \bar{y} - b_1\bar{x}$$
$$= 5 - (-0.181818)(2.5) = 5.45$$

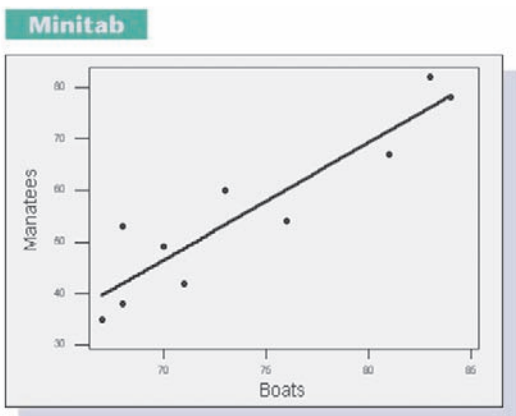
Conociendo la pendiente b_1 y b_0, el intercepto de y, ahora podemos expresar la ecuación estimada de la recta de regresión como

$$\hat{y} = 5.45 - 0.182x$$

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 probablemente produciría una ecuación ligeramente diferente.

EJEMPLO Barcos y manatíes Con los datos barco/manatí de la tabla 9-1 obtuvimos que el coeficiente de correlación lineal es $r = 0.922$. Utilice los mismos datos muestrales para calcular la ecuación de la recta de regresión.

SOLUCIÓN Si utilizamos el mismo procedimiento ilustrado en el ejemplo anterior, o empleamos herramientas tecnológicas, podemos obtener que los 10 pares de datos barco/manatí de la tabla 9-1 dan como resultado $b_0 = -113$ y $b_1 = 2.27$. Los resultados de Minitab se presentan en la página siguiente. Sustituyendo los valores calculados para b_0 y b_1, expresamos la ecuación de regresión como $\hat{y} = -113 + 2.27x$. A continuación se muestra 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

Manatees = −113 + 2.27 Boats

Predictor        Coef     SE Coef        T        P
Constant      −112.71       25.19    −4.47    0.002
Boats          2.2741      0.3388     6.71    0.000

S = 6.612        R-Sq = 84.9%    R-Sq(adj) = 83.0%
```

Uso de la ecuación de regresión para hacer predicciones

Las ecuaciones de regresión resultan útiles cuando se emplean para *predecir* el valor de una variable, a partir de algún valor particular de la otra variable. Si la recta de regresión se ajusta bastante bien a los datos, entonces es sensato utilizar esta ecuación para hacer predicciones, ya que no vamos más allá de los valores disponibles. Sin embargo, *debemos utilizar la ecuación de la recta de regresión sólo si r indica que existe una correlación lineal. En ausencia de una correlación lineal no debemos emplear la ecuación de regresión para proyectar o predecir; en su lugar, el mejor estimado de la segunda variable es sencillamente su media muestral.*

Al predecir un valor de y con base en algún valor dado de x. . .

1. **Si *no* existe una correlación lineal, el mejor valor predicho de y es $\bar{y}$.**
2. **Si existe una correlación lineal, el mejor valor predicho de y se calcula sustituyendo el valor de x en la ecuación de regresión.**

La figura 9-8 en la siguiente página resume este proceso, el cual se comprende con mayor facilidad si pensamos en r como una medida de qué tan bien se ajusta la recta de regresión a los datos muestrales. Si r se acerca a -1 o $+1$, entonces la recta de regresión se ajusta bien a los datos, pero si r es cercana a 0, entonces la recta de regresión se ajusta muy poco (y no debe emplearse para hacer predicciones).

EJEMPLO Predicción de muertes de manatíes Con los datos muestrales de la tabla 9-1, encontramos que existe una correlación lineal significativa entre el número de barcos registrados y la cantidad de manatíes asesinados por los barcos; también encontramos que la ecuación de regresión es $\hat{y} = -113 + 2.27x$. Suponga que en el año 2001 había 850,000 barcos registrados. Puesto que la tabla 9-1 lista el número de barcos registrados en decenas de miles, esto significa que para 2001 tenemos $x = 85$. Ya que $x = 85$, calcule el mejor valor predicho de y, el número de manatíes asesinados por barcos.

SOLUCIÓN Existe una fuerte tentación de brincar y sustituir x por 85 en la ecuación de regresión, pero primero debemos considerar si existe una correlación lineal que justifique el uso de dicha ecuación. En este ejemplo, tenemos

continúa

la pizza se correlaciona con la crisis

Cuando el Congreso de Estados Unidos amenazó al ex presidente Clinton con enjuiciarlo, los empleados del gobierno trabajaron hasta tarde y ordenaron un número récord de pizzas. Frank Meeks, dueño de 59 tiendas de Domino's Pizza en Washington, D.C. reportó que el sábado en el punto culminante de la crisis de la amenaza de juicio, las entregas de pizza a Capitol Hill excedieron los $10,000, mientras que las entregas de pizza a la Casa Blanca totalizaron $3,000. Meeks señaló que las ventas de pizzas también se incrementaron durante la Guerra del Golfo Pérsico y suelen aumentar anualmente durante los debates en torno al presupuesto.

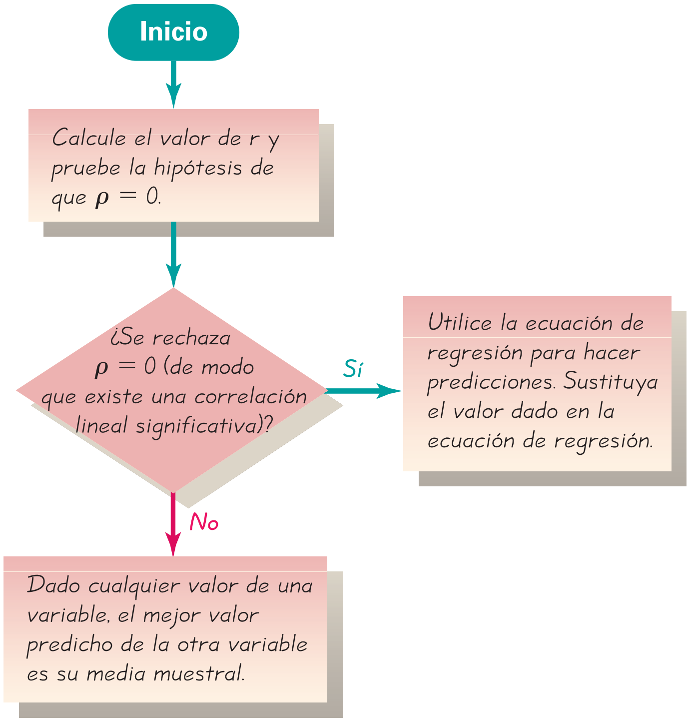


FIGURA 9-8 Procedimiento para hacer predicciones

una correlación lineal significativa (con $r = 0.922$), por lo que el valor predicho se calcula de la siguiente manera:

$$\begin{aligned}\hat{y} &= -113 + 2.27x \\ &= -113 + 2.27(85) = 80.0\end{aligned}$$

El número predicho de muertes de manatíes, a partir de 850,000 barcos registrados, es de 80.0. (Si no hubiese una correlación lineal significativa, el mejor valor predicho sería $\bar{y} = 558/10 = 55.8$). El número real de muertes de manatíes por barcos en 2001 fue de 82, de modo que el valor predicho de 80.0 es bastante cercano.

EJEMPLO Medida de sombrero y CI Evidentemente no existe una correlación lineal entre la medida de sombrero y las puntuaciones de CI de adultos. Como un individuo utiliza un sombrero de tamaño 7, calcule el mejor valor predicho de la puntuación de CI de esta persona.

SOLUCIÓN Puesto que no existe una correlación lineal, no empleamos una ecuación de regresión. No hay necesidad de reunir datos muestrales apareados consistentes de la medida de sombrero y de la puntuación de CI de una muestra de adultos seleccionados aleatoriamente. En su lugar, la mejor puntuación de CI predicho es sencillamente el CI medio 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, pero en ausencia de dicha 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 los años 1970.
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 edad y 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 barco/manatí de la tabla 9-1, la recta de regresión tiene una pendiente de 2.27, lo que demuestra que si incrementamos x (el número de barcos registrados en decenas de miles) en una unidad, el número predicho de muertes se incrementará en 2.27 manatíes. Es decir, por cada 10,000 barcos adicionales registrados, esperamos aproximadamente 2.27 muertes adicionales de manatíes debidas a los barcos.

Datos distantes y puntos de influencia

Un análisis de correlación/regresión de datos bivariados (apareados) debe incluir la investigación de *datos distantes* y *puntos de influencia*, que se definen de la siguiente manera.

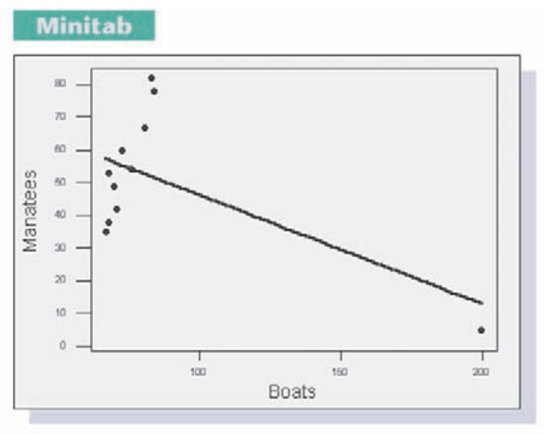
Definiciones

En un diagrama de dispersión, un **dato distante** es un punto que aparece muy lejos de los otros puntos de datos.

Los datos muestrales apareados incluyen uno o más **puntos de influencia**, que son puntos que afectan fuertemente la gráfica de la recta de regresión.

Es fácil identificar un dato distante: examine el diagrama de dispersión e identifique un punto que se aleja de los otros puntos. He aquí cómo determinar 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 con frecuencia se encuentran al identificar los datos distantes que están *horizontalmente* alejados de los demás puntos.

Por ejemplo, remítase a los resultados previos de Minitab. Suponga que incluimos el siguiente par adicional de datos: $x = 200$, $y = 5$ (en un año con 2,000,000 de barcos registrados, únicamente cinco manatíes fueron asesinados por barcos). Este punto adicional sería un punto de influencia debido a que la gráfica de la recta de regresión cambiaría considerablemente, tal como se muestra 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 añadidura de ese par de valores tiene un efecto importante en la recta de regresión.



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 la recta que 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. Dichas 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,

$$\text{residual} = y \text{ observada} - y \text{ predicha} = y - \hat{y}$$

Esta definición tal vez parezca tan clara como las instrucciones de una forma de impuestos, pero usted comprenderá fácilmente los residuales si se remite a la figura 9-9, que corresponde a los datos muestrales apareados que se listan a continuación. En la figura 9-9, 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.

x	1	2	4	5
y	4	24	8	32

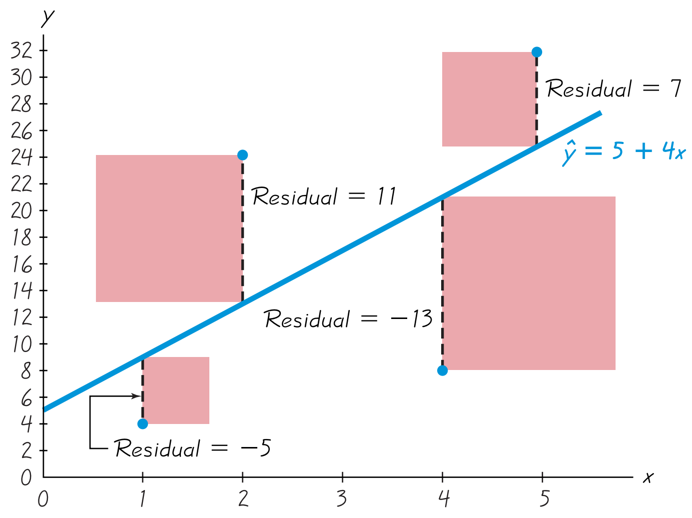


FIGURA 9-9 Residuales y cuadrados de residuales

La ecuación de regresión representa la recta que se ajusta "mejor" a los puntos, según la siguiente *propiedad de mínimos cuadrados*.

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 9-9 podemos observar que los residuales son -5, 11, -13 y 7, de modo 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 9-9, donde los cuadrados de los residuales están representados por las áreas de los cuadrados sombreados. La suma de las áreas cuadradas es 364, que es la menor suma posible. Utilice cualquier otra recta y los cuadrados se combinarán para producir una área mayor que el área combinada de 364.

Por fortuna, no necesitamos enfrentar directamente 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 9-2 y 9-3. Puesto que en las derivaciones de estas fórmulas se requiere del cálculo, no las incluimos en este libro.

Utilizando la tecnología

Debido a los cálculos complejos 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 Seleccione **Analysis** de la barra del menú principal, después utilice la opción **Correlation and Regression**. Introduzca los datos apareados o use copiar/pegar para copiar los datos. Introduzca los valores de x en la columna 1 y los valores de y correspondientes en la columna 2. Introduzca un valor para el nivel de significancia. Haga clic en el botón de **Evaluate**. Los resultados de STATDISK incluyen el valor del coeficiente de correlación lineal, junto con el valor crítico de r, la conclusión acerca de la correlación, y el intercepto y la pendiente de la ecuación de regresión, así como otros resultados. Haga clic en **Plot 1** 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 9-2 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/Regression** e 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 predictora. 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, tal como B1:B10. Introduzca el rango para los valores de x, tal 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 "coeficiente". La gráfica presentada incluirá un diagrama de dispersión de los puntos muestrales originales, junto con los puntos que serían predichos por la ecuación de regresión. Fácilmente puede obtener la recta de regresión 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 icono del lápiz para la variable de respuesta e introduzca el rango de valores para la variable y (o dependiente). Haga clic en el icono del lápiz para la variable explicatoria 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 "coeficiente".

TI-83 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 de y y la pendiente de la ecuación de regresión. La calculadora TI-83 Plus representa los valores b_0 y b_1 como a y b.

9-3 Destrezas y conceptos básicos

Haciendo predicciones. *En los ejercicios 1 a 4, utilice los datos dados para calcular el mejor valor predicho de la variable dependiente. Asegúrese de seguir el procedimiento para predicciones descrito en esta sección.*

1. En cada uno de los casos siguientes, calcule el mejor valor predicho de y, puesto que $x = 3.00$. Los estadísticos dados se resumen a partir de datos muestrales apareados.
a. $r = 0.987$, $\bar{y} = 5.00$, $n = 20$, y la recta de la ecuación de regresión es $\hat{y} = 6.00 + 4.00x$.
b. $r = 0.052$, $\bar{y} = 5.00$, $n = 20$, y la recta de la ecuación de regresión es $\hat{y} = 6.00 + 4.00x$.

2. En cada uno de los casos siguientes, calcule el mejor valor predicho de y, puesto que $x = 2.00$. Los estadísticos dados se resumen a partir de datos muestrales apareados.
a. $r = -0.123$, $\bar{y} = 8.00$, $n = 30$, y la recta de la ecuación de regresión es $\hat{y} = 7.00 - 2.00x$.
b. $r = -0.567$, $\bar{y} = 8.00$, $n = 30$, y la recta de la ecuación de regresión es $\hat{y} = 7.00 - 2.00x$.

3. Tamaño del pecho y peso de osos Cuando se anestesió a ocho osos, algunos investigadores midieron las distancias (en pulgadas) alrededor de su pecho y pesaron a los osos (en libras). Se utilizó Minitab para calcular que el valor del coeficiente de correlación lineal es $r = 0.993$ y la recta de la ecuación de regresión es $\hat{y} = -187 + 11.3x$, donde x representa el tamaño del pecho. Además, el peso medio de los ocho osos es de 234.5 libras. ¿Cuál es el mejor peso predicho de un oso con un pecho de 52 pulgadas?

4. Acciones y Súper Bowl El conjunto de datos 25 del Apéndice B incluye pares de datos del valor más alto del Promedio Industrial de Dow-Jones (DJIA) y el número total de puntos anotados en el Súper Bowl, durante 21 años distintos. 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 más alto 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 alto del DJIA de 1200?

Cálculo de la ecuación de la recta de regresión. *En los ejercicios 5 y 6, utilice los datos dados para calcular la ecuación de la recta de regresión.*

5.

x	0	1	2	3	4
y	4	1	0	1	4

6.

x	1	2	2	5	6
y	2	5	4	15	15

7. Efectos de un dato distante Remítase al diagrama de dispersión generado por Minitab que se presenta en el ejercicio 7 de la sección 9-2.
a. Utilice los pares de valores de los 10 puntos y calcule la ecuación de la recta de regresión.
b. Después 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*.

Cálculo de la ecuación de la recta de regresión y predicciones. *En los ejercicios 8 a 24, utilice los mismos conjuntos de datos que en los ejercicios de la sección 9-2. En cada caso, calcule la ecuación de regresión, permitiendo que la primera variable sea la variable independiente (x). Calcule los valores predichos indicados.* Cuidado: *Al calcular los valores predichos, asegúrese de seguir los procedimientos para predicciones descritos en esta sección.*

8. Incendios y acres quemados Calcule el mejor valor predicho para el número de acres quemados, considerando que hubo 80 incendios.

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

9. Compra de una audiencia televisiva Calcule el mejor valor predicho para el número de televidentes (en millones), considerando que el salario (en millones de dólares) de la estrella de televisión Jennifer Anniston es de $16 millones. ¿De qué forma se compara el valor predicho con el número real de televidentes, que fue de 24 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

10. Estaturas y pesos de supermodelos Calcule el mejor peso predicho de una supermodelo que mide 69 pulgadas.

Estatura (pulgadas)	71	70.5	71	72	70	70	66.5	70	71
Peso (libras)	125	119	128	128	119	127	105	123	115

11. Mediciones de la presión sanguínea Calcule la mejor presión sanguínea diastólica predicha para una persona con una lectura sistólica de 122.

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

12. Temperaturas y maratones Calcule el mejor tiempo ganador predicho para la maratón de 1990, ya que la temperatura era de 73 grados. ¿De qué forma se compara el valor predicho con 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

13. Tabaquismo y nicotina Calcule el mejor nivel de cotinina predicho para una persona que fuma 40 cigarrillos diarios.

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

14. Circunferencia y altura de árboles Calcule la mejor altura predicha de un árbol que tiene una circunferencia de 4.0 pies. ¿Cuál es una de las ventajas de poder determinar la altura de un árbol a partir de su circunferencia?

x (circunferencia)	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
y (altura)	21.0	33.5	24.6	40.7	73.2	24.9	40.4	45.3	53.5	93.8	64.0	62.7	47.2	44.3

T **15.** Cereales asesinos Remítase al conjunto de datos 16 del Apéndice B y utilice las cantidades de grasa (x) y los conteos calóricos medidos (y). Calcule el mejor conteo calórico predicho para un cereal con 0.05 gramos de grasa por gramo de cereal.

T **16.** Tabaco y alcohol en películas infantiles Remítase al conjunto de datos 7 del Apéndice B y utilice los tiempos que las películas infantiles de dibujos animados presentan consumo de tabaco (x) y alcohol (y). Calcule el mejor tiempo predicho para el consumo de alcohol, dado que una película no muestre consumo de tabaco.

T **17.** Colesterol e índice de masa corporal Remítase al conjunto de datos 1 del Apéndice B y utilice los niveles de colesterol (x) y los valores del índice de masa corporal (y) de las 40 mujeres. ¿Cuál es el mejor valor predicho del índice de masa corporal de una mujer que tiene un nivel de colesterol de 500?

T **18.** Niveles de lectura Remítase al conjunto de datos 14 del Apéndice B y utilice las puntuaciones de la facilidad de lectura de Flesch (x) y los valores del grado Flesch-Kincaid (y) para *El oso y el dragón*, de Tom Clancy. Calcule el mejor valor predicho del grado de Flesch-Kincaid para una página con una puntuación de facilidad de lectura de Flesch de 50.0.

T **19.** Precios de venta de casas, precios de lista e impuestos Remítase al conjunto de datos 24 del Apéndice B. *Cuidado*: Los valores muestrales de los precios de lista y los precios de venta están en miles de dólares, pero las cantidades de los impuestos están en dólares.

a. Utilice los datos apareados de los precios de lista (x) y de los precios de venta (y) de casas. ¿Cuál es el mejor precio de venta predicho para una casa con un precio de lista de $200,000?

b. Utilice los datos apareados que consisten en el precio de venta de casas (x) y la suma de impuestos (y). ¿Cuál es el mejor cargo de impuestos predicho para una casa que se vendió en $400,000?

T **20.** Alquitrán y nicotina Remítase al conjunto de datos 5 del Apéndice B.

a. Utilice los datos apareados de alquitrán (x) y nicotina (y). ¿Cuál es el mejor nivel de nicotina predicho para un cigarrillo que contiene 15 mg de alquitrán?

b. Utilice los datos apareados de monóxido de carbono (x) y nicotina (y). ¿Cuál es el mejor nivel de nicotina predicho para un cigarrillo que contiene 15 mg de monóxido de carbono?

T **21.** Pronóstico del tiempo Remítase al conjunto 10 de datos del Apéndice B.

a. Utilice el pronóstico de altas temperaturas para cinco días (x) y las temperaturas altas reales (y). ¿Cuál es la mejor temperatura alta real predicha, si la temperatura alta pronosticada a cinco días es de 28 grados?

b. Utilice el pronóstico de altas temperaturas para un día (x) y las temperaturas altas reales (y). ¿Cuál es la mejor temperatura alta real predicha, si la temperatura alta pronosticada para un día es de 28 grados?

c. ¿Cuál de los valores predichos es mejor: el resultado del inciso *a* o el resultado del inciso *b*? ¿Por qué?

T **22.** Everglades de Florida Remítase al conjunto de datos 12 del Apéndice B.

a. tilice las temperaturas inferiores (x) y las mediciones de conductividad (y). ¿Cuál es la mejor medición de conductividad predicha para un tiempo en el que la temperatura más baja es de 30.0°C?

b. Utilice las cantidades de lluvia (x) y las mediciones de conductividad (y). ¿Cuál es la mejor medición de conductividad predicha para un tiempo en que la cantidad de lluvia es de 0.00 pulgadas?

c. Después de identificar la mejor medición de conductividad predicha en los incisos *a* y *b*, ¿será preciso alguno de los valores predichos? ¿Por qué?

T **23.** Old Faithful Remítase al conjunto de datos 13 del Apéndice B.

a. Utilice los datos apareados de las duraciones (x) y los intervalos posteriores a las erupciones del géiser (y). ¿Cuál es el mejor tiempo predicho antes de la siguiente erupción, si la última erupción duró 210 segundos?

b. Utilice los datos apareados de las alturas de las erupciones (x) y de los intervalos después de las erupciones (y) del géiser Old Faithful. ¿Cuál es el mejor tiempo predicho antes de la siguiente erupción, si la última erupción alcanzó una altura de 275 pies?

c. ¿Cuál tiempo predicho es mejor: el resultado del inciso *a* o el resultado del inciso *b*? ¿Por qué?

T **24.** Precios, quilates y colores de diamantes Remítase al conjunto de datos 18 del Apéndice B.

a. Utilice los datos apareados del peso en quilates (x) y el precio (y). ¿Cuál es el mejor precio predicho de un diamante con un peso de 1.5 quilates?

b. Utilice los datos apareados de color (x) y precio (y). ¿Cuál es el mejor precio predicho para un diamante con un rango de color de 3?

c. ¿Cuál precio predicho es mejor: el resultado del inciso *a* o el resultado del inciso *b*? ¿Por qué?

25. Identificación de datos distantes y puntos de influencia Remítase a los datos muestrales listados en la tabla 9-1. Si incluimos otro par de valores consistente en $x = 120$ (para 1,200,000 barcos) y $y = 160$ (muertes de manatíes por barcos), ¿será el nuevo punto un dato distante? ¿Será un punto de influencia?

26. Identificación de puntos de influencia Remítase a los datos muestrales listados en la tabla 9-1. Si incluimos otro par de valores consistentes en $x = 120$ (para 1,200,000 barcos) y $y = 10$ (muertes de manatíes por barcos), ¿será el nuevo punto un dato distante? ¿Será un punto de influencia?

9-3 Más allá de lo básico

27. ¿De qué manera se ve afectada una ecuación de regresión por un cambio en la escala? Los números grandes, tales como los que se incluyen la tabla adjunta, suelen causar problemas de cálculo. Primero utilice los datos proporcionados para calcular la ecuación de la recta de regresión, después calcule la ecuación de la recta de regresión una vez que cada valor de x se ha dividido entre 1000. ¿Cómo se ven afectados los resultados por el cambio en x? ¿De qué forma se afectarían los resultados si cada valor de y se dividiera entre 1000?

x	924,736	832,985	825,664	793,427	857,366
y	142	111	109	95	119

x	1	2	4	5
y	4	24	8	32

28. 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 que $\hat{y} = 5 + 4x$ y la suma de cuadrados de los residuales es 364. Demuestre que la ecuación $\hat{y} = 8 + 3x$ da como resultado una suma de cuadrados de residuales mayor que 364.

29. Uso de logaritmos para transformar datos Si un diagrama de dispersión revela un patrón no lineal (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 ajusta mejor a los datos muestrales, y calcule la ecuación logarítmica ($y = a + b \ln x$) que se ajusta 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

30. Pruebas de hipótesis equivalentes Explique por qué una prueba de la hipótesis nula H_0: $\rho = 0$ es equivalente a una prueba de la hipótesis nula H_0: $\beta_1 = 0$, donde ρ es el coeficiente de correlación lineal de una población de datos apareados, y β_1 es la pendiente de la recta de regresión de esa misma población.

31. Gráfica residual Un diagrama de dispersión es una gráfica de los datos muestrales apareados (x,y). Una **gráfica residual** es la gráfica de los puntos con las mismas coordenadas de x, pero donde las coordenadas correspondientes de y son valores residuales. Para construir una gráfica residual, utilice el mismo eje x que en el diagrama de dispersión, pero elabore un eje vertical para los valores residuales. Dibuje una línea horizontal de referencia a lo largo del valor residual de 0, después grafique los valores apareados de (x, residual). Las gráficas residuales son útiles para identificar patrones que sugieren que la relación entre las variables es no lineal, o que el supuesto de las varianzas constantes no se satisface. Construya una gráfica residual para los datos de la tabla 9-1. ¿Existe algún patrón sobresaliente?

9-4 Variación e intervalos de predicción

Hasta ahora, hemos utilizado datos muestrales apareados para probar una correlación lineal entre x y y, y para identificar la ecuación de regresión. En esta sección continuamos analizando datos apareados (x,y), conforme procedemos a considerar distintos tipos de variación que se emplean para dos aplicaciones principales:

1. Para determinar la proporción de variación en y que se explica por la relación lineal entre x y y.
2. Para construir estimados de intervalos de valores predichos de y. Dichos intervalos se denominan *intervalos de predicción*, que se definirán de manera formal más adelante en esta sección.

Variación explicada y sin explicar

En la sección 9-2 introdujimos el concepto de correlación y empleamos el coeficiente de correlación lineal r para determinar si existe una correlación lineal significativa entre dos variables, denotadas por x y y. Además de servir como una medida de la correlación lineal entre dos variables, el valor de r también nos proporciona información adicional acerca de la variación de puntos muestrales respecto a la recta de regresión. Iniciamos con un caso muestral, que nos conduce a una definición importante (coeficiente de determinación).

Suponga que tenemos un gran conjunto de datos apareados con estos resultados:

- Existe una correlación lineal significativa.
- 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 $y = 19$.

LA ESTADÍSTICA EN LAS NOTICIAS

Diferencia en el salario por género

Aunque un reporte reciente de la revista *Working Woman* afirma que la diferencia del salario con base en el género está disminuyendo, en su mayoría, los hombres aún poseen los empleos con sueldos más altos. Los datos más recientes indican que, en promedio, las mujeres que trabajan tiempo completo ganan aproximadamente 73 centavos por cada dólar ganado por hombres que trabajan tiempo completo. Investigadores del Institute for Social Research, en la Universidad de Michigan, analizaron los efectos de diversos factores clave y encontraron que cerca de una tercera parte de la discrepancia entre los salarios de hombres y mujeres puede explicarse por diferencias en educación, antigüedad, interrupciones en el trabajo y opciones de empleo. Las restantes dos terceras partes continúan sin poder explicarse mediante este tipo de factores laborales.

- El punto (5, 13) es uno de los puntos sobre la recta de regresión, ya que al sustituir $x = 5$ en la ecuación de regresión, resulta $\hat{y} = 13$.

$$\hat{y} = 3 + 2x = 3 + 2(5) = 13$$

La figura 9-10 indica que el punto (5, 13) está sobre la recta de regresión, pero el punto (5, 19) proviene del conjunto de datos original y no está sobre la recta de regresión, debido a que no satisface la ecuación de regresión. Tome su tiempo para examinar cuidadosamente la figura 9-10 y observe las diferencias, definidas de la siguiente manera.

Desviación sin explicar, explicada y total

Definiciones

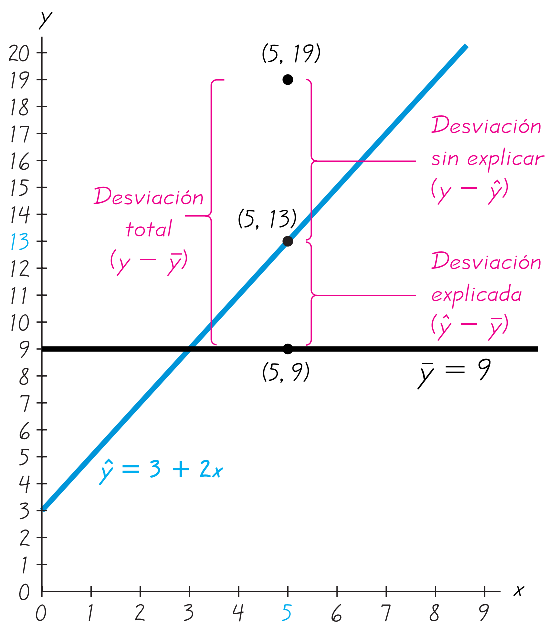
Suponga que tenemos un conjunto de datos apareados que contienen el punto muestral (x,y), que $\hat{y}$ es el valor predicho de y (obtenido por medio de la ecuación de regresión), y que la media de los valores y muestrales es $\bar{y}$.

La **desviación total** (a partir de la media) del punto particular (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 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 un *residual*, tal como se definió en la sección 9-3.)

FIGURA 9-10 Desviación sin explicar, explicada y total



Para los datos específicos bajo consideración, obtenemos los siguientes resultados:

Desviación total de (5, 19) $= y - \bar{y} = 19 - 9 = 10$

Desviación explicada de (5, 19) $= \hat{y} - \bar{y} = 13 - 9 = 4$

Desviación sin explicar de (5, 19) $= y - \hat{y} = 19 - 13 = 6$

Si fuésemos totalmente ignorantes de los conceptos de correlación y regresión, y deseáramos predecir un valor de y, dado un valor de x y un conjunto de datos apareados (x,y), nuestra mejor conjetura sería $\bar{y}$. Pero no ignoramos totalmente los conceptos de correlación y regresión: sabemos que en este caso (con una correlación lineal significativa), la forma de predecir el valor de y cuando $x = 5$ es utilizar la ecuación de regresión, que produce $\hat{y} = 13$, como se calculó antes. Podemos explicar la discrepancia entre $\bar{y} = 9$ y $\hat{y} = 13$ señalando sencillamente que existe una correlación lineal significativa que está mejor descrita por la recta de regresión. Como consecuencia, cuando $x = 5$, *y debe ser* 13, y no el valor medio de 9. Pero, aun cuando y debería ser 13, *es* 19. La discrepancia entre 13 y 19 no puede explicarse por la recta de regresión, y se denomina *desviación sin explicar* o *residual*. El caso específico ilustrado en la figura 9-10 se generaliza de la siguiente forma:

$$\begin{array}{lccccc} & \text{(desviación total} & = & \text{(desviación explicada)} & + & \text{(desviación sin explicar)} \\ \text{o} & (y - \bar{y}) & = & (\hat{y} - \bar{y}) & + & (y - \hat{y}) \end{array}$$

Esta última expresión se aplica a un punto (x,y) particular, y la misma relación se aplica a las sumas de cuadrados mostradas en la fórmula 9-4, aunque esta última expresión no es algebraicamente equivalente a la fórmula 9-4. En ésta, la **variación total** se expresa como las sumas 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 9-4

$$\begin{array}{lccccc} & \text{(variación total)} & = & \text{(variación explicada)} & + & \text{(variación sin explicar)} \\ \text{o} & \Sigma(y - \bar{y})^2 & = & \Sigma(\hat{y} - \bar{y})^2 & + & \Sigma(y - \hat{y})^2 \end{array}$$

Coeficiente de determinación

Los componentes de la fórmula 9-4 se utilizan en la siguiente definición importante.

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 9-4 o simplemente elevar al cuadrado el coeficiente de correlación lineal r, que se obtiene utilizando los métodos descritos en la sección 9-2. Por ejemplo, en la sección 9-2 señalamos que si $r = 0.922$, entonces $r^2 = 0.850$, que significa que *el 85% de la variación total de* y *puede explicarse por medio de la relación lineal entre* x *y* y *(como se describió por medio de la ecuación de regresión). Se infiere que el 15% de la variación total de* y *permanece sin explicación.*

EJEMPLO Diamantes En el ejercicio 24a, en la sección 9-2, encontramos que para los datos apareados consistentes en los pesos (en quilates) y los precios de los diamantes, el coeficiente de correlación lineal está dado por $r = 0.767$. Calcule el porcentaje de la variación en y (precio) que se explica por la relación lineal entre el peso y el precio.

SOLUCIÓN El coeficiente de determinación es $r^2 = 0.767^2 = 0.588$, lo que indica que la proporción de variación explicada en y, respecto a la variación total en y, es 0.588. Ahora podemos afirmar que el 58.8% de la variación total en y puede explicarse por la ecuación de regresión. Interpretamos que el 58.8% de la variación total de los precios de diamantes se explica por la variación en sus pesos; el restante 41.2% es atribuible a otros factores tales como el color, la claridad y el azar. Sin embargo, recuerde que estos resultados son estimados basados en datos muestrales dados. Probablemente otros datos muestrales darían como resultado estimados distintos.

Intervalos de predicción

En la sección 9-3 empleamos los datos muestrales de la tabla 9-1 para calcular la ecuación de regresión $\hat{y} = -113 + 2.27x$, donde $\hat{y}$ representa el número predicho de muertes de manatíes y x representa el número de barcos (en decenas de miles). Después utilizamos esa ecuación para predecir el valor de y, puesto que $x = 85$ (para 850,000 barcos). Encontramos que el mejor número predicho de muertes de manatíes es 80.0. Si usamos los valores sin redondear de la pendiente y del intercepto, obtenemos el resultado más preciso de 80.6 muertes de manatíes. Puesto que 80.6 es un valor único, se le conoce como *estimado del punto*. En el capítulo 6 aprendimos que los estimados del punto tienen la grave desventaja de no proporcionarnos ninguna información acerca de su precisión. Aquí, sabemos que 80.6 es el mejor valor predicho, pero no sabemos qué tan preciso es este valor. En el capítulo 6 elaboramos estimados del intervalo de confianza para superar esa desventaja, y en esta sección seguimos este precedente. Utilizaremos un **intervalo de predicción**, que es un estimado del intervalo de un valor predicho de y.

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 9-10. 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 de y observados y los valores predichos $\hat{y}$ que se obtienen por medio de la ecuación de regresión. Está dado por

$$s_e = \sqrt{\frac{\Sigma(y - \hat{y})^2}{n - 2}} \qquad \text{(donde } \hat{y} \text{ es el valor predicho de } y\text{)}$$

o por la siguiente fórmula equivalente:

Fórmula 9-5
$$s_e = \sqrt{\frac{\Sigma y^2 - b_0 \Sigma y - b_1 \Sigma xy}{n - 2}}$$

STATDISK, Minitab, Excel y la calculadora TI-83 Plus están diseñados para calcular de manera automática el valor de s_e. Véase el apartado “Utilizando la tecnología” al final de esta sección.

La elaboración del error estándar del estimado s_e se asemeja mucho a la de la desviación estándar ordinaria introducida en el capítulo 2. Así como la desviación estándar es una medida de qué tanto los valores se desvían de su media, el error estándar del estimado s_e es una medida de qué tanto los puntos de los datos muestrales se desvían de su recta de regresión. La lógica que subyace a 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 valores relativamente grandes ocurren con puntos que se alejan de la recta de regresión.

La fórmula 9-5 es algebraicamente equivalente a la otra expresión en la definición, pero la fórmula 9-5 suele ser más fácil debido a que no requiere que calculemos cada uno de los valores predichos $\hat{y}$ por medio de su sustitución en la ecuación de regresión. Sin embargo, la fórmula 9-5 sí requiere que calculemos b_0, el intercepto y y la pendiente b_1 de la recta de regresión estimada.

PROBLEMA DEL CAPÍTULO 9

EJEMPLO Utilice la fórmula 9-5 para calcular el error estándar del estimado s_e para los datos muestrales barco/manatí listados en la tabla 9-1.

SOLUCIÓN Con los datos muestrales de la tabla 9-1, calculamos estos valores:

$$n = 10 \qquad \Sigma y^2 = 33{,}456 \qquad \Sigma y = 558 \qquad \Sigma xy = 42{,}214$$

En la sección 9-3 empleamos los datos muestrales de la tabla 9-1 para obtener el intercepto y y la pendiente de la recta de regresión. Dichos valores se presentan aquí con más decimales para una mayor precisión.

$$b_0 = -112.7098976 \qquad b_1 = 2.274087687$$

continúa

Ahora podemos usar estos valores en la fórmula 9-5 para calcular el error estándar del estimado s_e.

$$s_e = \sqrt{\frac{\Sigma y^2 - b_0\Sigma y - b_1\Sigma xy}{n - 2}}$$

$$= \sqrt{\frac{33{,}456 - (-112.7098976)(558) - (2.274087687)(42{,}214)}{10 - 2}}$$

$$= 6.6123487 = 6.61 \qquad \text{(redondeado)}$$

Medimos la dispersión de los puntos muestrales alrededor de la recta de regresión, con el error estándar del estimado $s_e = 6.61$.

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 del punto 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 24 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 que representa el valor dado de x, $t_{\alpha/2}$ tiene $n - 2$ grados de libertad, y s_e se calcula partir de la fórmula 9-5.

EJEMPLO **Barcos y manatíes** Para los datos apareados barco/manatí de la tabla 9-1, encontramos que cuando $x = 85$ (para 850,000 barcos), el mejor número predicho de muertes de manatíes es 80.0, pero obtenemos un valor predicho de 80.6 cuando empleamos valores más precisos de b_0, el intercepto y y la pendiente b_1. Construya un intervalo de predicción del 95% para el número de manatíes asesinados por barcos, considerando que el número de barcos es de 850,000 (de modo que $x = 85$). Esto nos proporcionará una idea de cuán preciso es el valor predicho de 80.6.

SOLUCIÓN En secciones anteriores hemos demostrado que existe una correlación lineal significativa (a nivel 0.05 de significancia), y la ecuación de regresión es $\hat{y} = -113 + 2.27x$. En el ejemplo anterior encontramos que $s_e = 6.6123487$, y los siguientes estadísticos se obtienen a partir de los datos muestrales de la tabla 9-1:

$$n = 10 \qquad \bar{x} = 74.1 \qquad \Sigma x = 741 \qquad \Sigma x^2 = 55{,}289$$

En la tabla A-3 encontramos que $t_{\alpha/2} = 2.306$. (Utilizamos $10 - 2 = 8$ grados de libertad con $\alpha = 0.05$ en dos colas). Primero calculamos el margen de error E permitiendo que $x_0 = 85$, ya que buscamos el intervalo de predicción del número de muertes de manatíes, puesto que $x = 85$ (para 850,000 barcos).

$$\begin{aligned}
E &= t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}} \\
&= (2.306)(6.6123487)\sqrt{1 + \frac{1}{10} + \frac{10(85 - 74.1)^2}{10(55{,}289) - (741)^2}} \\
&= (2.306)(6.6123487)(1.1882420) = 18.1
\end{aligned}$$

Con $\hat{y} = 80.6$ y $E = 18.1$, obtenemos el intervalo de predicción de la siguiente forma:

$$\begin{gathered}
\hat{y} - E < y < \hat{y} + E \\
80.6 - 18.1 < y < 80.6 + 18.1 \\
62.5 < y < 98.7
\end{gathered}$$

Es decir, para $x = 85$ (para 850,000 barcos) tenemos una certeza del 95% de que el número de muertes de manatíes está entre 62.5 y 98.7. Se trata de un rango relativamente grande. (Un factor que contribuye a lo grande del rango es que el tamaño de la muestra es muy pequeño, debido a que estamos utilizando únicamente 10 pares de datos muestrales).

Minitab es útil para calcular los límites del intervalo de predicción. Si utilizamos Minitab, nos dará el resultado de (62.5, 98.7) bajo el encabezado "95.0% P.I.". Éste corresponde al mismo intervalo de predicción calculado antes.

Además de saber que para $x = 85$, el número predicho de muertes de manatíes es 80.6, ahora tenemos una idea de qué tan confiable es en realidad el estimado. El intervalo de predicción del 95% calculado en este ejemplo indica que el valor real de y llega a variar sustancialmente del valor predicho de 80.6.

Utilizando la tecnología

STATDISK STATDISK es útil 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. Seleccione **Analysis** de la barra del menú principal, después utilice la opción **Correlation and Regression.** Introduzca los datos apareados o utilice las funciones copiar/pegar para copiar los datos. Ingrese los valores de x en la columna 1 y los valores correspondientes de y en la columna 2. Introduzca un valor para el nivel de significancia. 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 se utiliza para calcular la ecuación de regresión, el error estándar del estimado s_e (denotado por s), el

continúa

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 **Opciones** 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, tal como B1:B10. Ingrese el rango para los valores de x, tal como A1:A10. 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 icono del lápiz para la variable de respuesta e introduzca el rango de valores para la variable y (o dependiente). Haga clic en el icono del lápiz para la variable explicativa e introduzca el rango de valores para la variable x (o independiente). Haga clic en **OK**.

TI-83 Plus La calculadora TI-83 Plus permite 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 r^2). Ingrese los datos apareados en las listas L1 y L2, después presione **STAT** y seleccione **TESTS**, luego elija la opción **LinRegTTest.**

9-4 Destrezas y conceptos básicos

Interpretación del coeficiente de determinación*. En los ejercicios 1 a 4, utilice el valor del coeficiente de correlación lineal r para calcular el coeficiente de determinación y el porcentaje de variación total que explica la relación lineal entre las dos variables.*

1. $r = 0.8$

2. $r = -0.6$

3. $r = -0.503$

4. $r = 0.636$

Interpretaciones de resultados de un programa de cómputo. *En los ejercicios 5 a 8, 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 5 del Apéndice B. Junto con los datos muestrales apareados, se le indicó a Minitab una cantidad de alquitrán de 17 mg a utilizar para predecir la cantidad de nicotina.*

Minitab

```
The regression equation is
Nicotine = 0.154 + 0.0651 Tar

Predictor        Coef      SE Coef        T        P
Constant      0.15403      0.04635     3.32    0.003
Tar          0.065052     0.003585    18.15    0.000

S = 0.08785   R-Sq = 92.4%   R-Sq(adj) = 92.1%

Predicted Values for New Observations

New Obs    Fit SE Fit      95.0% CI           95.0% PI
1       1.2599 0.0240 ( 1.2107, 1.3091)( 1.0731, 1.4468)
```

5. Prueba de correlación Con la información proporcionada en los resultados, determine el valor del coeficiente de correlación lineal. Como hay 29 pares de datos, ¿existe una correlación lineal significativa entre la cantidad de alquitrán y la cantidad de nicotina en un cigarrillo?

6. Identificación de la variación total ¿Qué porcentaje de la variación total de nicotina se explica por la relación lineal entre alquitrán y nicotina?

7. 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).

8. Cálculo del intervalo de predicción Para una cantidad de alquitrán dada de 17 mg, identifique el estimado del intervalo de predicción del 95% de la cantidad de nicotina y redacte una afirmación que interprete ese intervalo.

Cálculo de medidas de variación. *En los ejercicios 9 a 12, 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 modo que es razonable utilizar la ecuación de regresión para hacer predicciones.*

9. Estatura y peso de supermodelos A continuación se listan las estaturas (en pulgadas) y los pesos (en libras) de las supermodelos Niki Taylor, Nadia Auerman, Claudia Schiffer, Elle MacPherson, Christy Turlington, Bridget Hall, Kate Moss, Valerie Mazza y Kristy Hume.

Estatura (pulg.)	71	70.5	71	72	70	70	66.5	70	71
Peso (libras)	125	119	128	128	119	127	105	123	115

10. Mediciones de presión sanguínea Catorce estudiantes diferentes del segundo año de medicina tomaron mediciones de la presión sanguínea del mismo paciente, y los resultados se presentan abajo (datos proporcionados por el 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

11 Circunferencia y altura de árboles A continuación se listan las circunferencias (en pies) y las alturas (en pies) de árboles en Marshall, Minnesota (datos tomados de "Tree Measurements", de Stanley Rice, *American Biology Teacher*, vol. 61, núm. 9).

x (circ.)	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
y (ht)	21.0	33.5	24.6	40.7	73.2	24.9	40.4	45.3	53.5	93.8	64.0	62.7	47.2	44.3

12. Niveles de facilidad de lectura Remítase al conjunto de datos 14 del Apéndice B y utilice los caracteres por palabra (x) y la puntuaciones de facilidad de lectura de Flesch (y) para *Harry Potter y la piedra filosofal*, de J. K. Rowling.

13. Efecto de la variación en el intervalo de predicción Remítase a los datos dados en el ejercicio 9 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.

a. Calcule el peso predicho de una supermodelo que mide 69 pulgadas.

b. Calcule un estimado del intervalo de predicción del 95% del peso de una supermodelo que mide 69 pulgadas.

14. Cálculo del valor predicho e intervalo de predicción Remítase al ejercicio 10 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
a. Calcule la lectura diastólica predicha, dado que la lectura sistólica es de 120.
b. Calcule un estimado del intervalo de predicción del 95% de la lectura diastólica, dado que la lectura sistólica es de 120.

15. Cálculo del valor predicho e intervalo de predicción Remítase a los datos del ejercicio 11 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
a. Calcule la altura predicha de un árbol que tiene una circunferencia de 4.0 pies.
b. Calcule un estimado del intervalo de predicción del 99% de la altura de un árbol que tiene una circunferencia de 4.0 pies.

16. Cálculo del valor predicho e intervalo de predicción Remítase a los datos descritos en el ejercicio 12 y suponga que se cumplen las condiciones necesarias de normalidad y varianza.
a. Calcule la puntuación de facilidad de lectura de Flesch predicha para una página que tiene un promedio de 4.0 caracteres por palabra.
b. Calcule un estimado del intervalo de predicción del 99% de la puntuación de facilidad de lectura de Flesch, predicha para una página que tiene un promedio de 4.0 caracteres por palabra.
c. ¿De qué manera se comparan los resultados de los incisos *a* y *b* con el par de datos observados, consistentes en 4.0 caracteres por palabra y una calificación de facilidad de lectura de Flesch de 86.2?

Cálculo de un intervalo de predicción. *En los ejercicios 17 a 20, remítase a los datos muestrales de la tabla 9-1. Permita que x represente el número de barcos registrados (en decenas de miles) y permita que y represente el número de manatíes muertos por barcos. Utilice el número de barcos registrados (en decenas de miles) y el nivel de confianza indicado para construir un estimado del intervalo de predicción del número de manatíes muertos por barcos. (Véase el ejemplo en esta sección).*

17. $x = 85$ (para 850,000 barcos); 99% de confianza

18. $x = 85$ (para 850,000 barcos); 90% de confianza

19. $x = 90$ (para 900,000 barcos); 95% de confianza

20. $x = 90$ (para 900,000 barcos); 99% de confianza

9-4 Más allá de lo básico

21. 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 ($y = \beta_0 + \beta_1 x$) se obtienen evaluando los límites en los intervalos que siguen.

$$b_0 - E < \beta_0 < b_0 + E$$

donde

$$E = t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}}$$

$$b_1 - E < \beta_1 < b_1 + E$$

donde

$$E = t_{\alpha/2} \cdot \frac{s_e}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}}$$

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 barco/manatí en la tabla 9-1, calcule los estimados del intervalo de confianza del 95% de β_0 y β_1.

22. Comprensión de la variación

a. Si un conjunto de datos apareados incluye al menos tres pares de valores, ¿qué sabe usted acerca del coeficiente de correlación lineal si $s_e = 0$?

b. Si un conjunto de datos apareados es tal que la variación explicada total es 0, ¿qué sabe acerca de la pendiente de la recta de regresión?

23. Comprensión de la variación

a. Encuentre una expresión para la variación no explicada en términos del tamaño de muestra n y el error estándar del estimado s_e.

b. Encuentre una expresión para la variación explicada en términos del coeficiente de determinación r^2 y la variación sin explicar.

c. Suponga que tenemos un conjunto de datos apareados para los que $r^2 = 0.900$ y la ecuación de regresión es $\hat{y} = 3 - 2x$. Calcule el coeficiente de correlación lineal.

24. Cálculo del intervalo de confianza para un valor predicho de la media A partir de la expresión que se dio en esta sección para el margen de error correspondiente a un intervalo de predicción para y, podemos obtener la expresión

$$s_{\hat{y}} = s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

que es el *error estándar de la predicción* cuando se predice para una *sola y*, ya que $x = x_0$. Cuando se predice la *media* de todos los valores de y para los que $x = x_0$, el estimado del punto $\hat{y}$ es el mismo, pero $s_{\hat{y}}$ es de la siguiente manera:

$$s_{\hat{y}} = s_e \sqrt{\frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

Utilice los datos de la tabla 9-1 y extienda el último ejemplo de esta sección para calcular un estimado del punto y un estimado del intervalo de confianza del 95% del número medio de muertes de manatíes por barcos, ya que el número de barcos registrados es de 850,000 (de manera que $x = 85$).

9-5 Regresión múltiple

Hasta ahora hemos utilizado métodos de correlación y regresión para investigar relaciones entre exactamente *dos* variables, pero algunas circunstancias requieren más de dos variables. Al predecir el precio de un diamante, por ejemplo, podríamos considerar variables tales como el peso (en quilates), el color y la claridad, de manera que interviene un total de cuatro variables. Esta sección presenta un método para analizar relaciones que incluyen *más de dos* variables. Nos enfocaremos en tres elementos clave: 1. la ecuación de regresión múltiple, 2. el valor de R^2 ajustada y 3. el valor P. Igual que en las secciones anteriores de este capítulo, trabajaremos únicamente con relaciones *lineales*. Iniciamos con la *ecuación de regresión múltiple*.

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 los puntos anotados 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) fuesen 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 precisas debido a otras variables que no se consideraron, tales como la popularidad del jugador.

Ecuación de regresión múltiple

Definición

Una **ecuación de regresión múltiple** expresa una relación lineal entre una variable dependiente y y dos o más variables independientes ($x_1, x_2, \ldots, x_k$). La forma general de una ecuación de regresión múltiple es
$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$.

Emplearemos la siguiente notación, que surge de manera natural de la notación utilizada en la sección 9-3.

Notación

$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$ (forma general de la ecuación de regresión múltiple estimada)

n = tamaño de la muestra

k = número de variables *independientes*. (Las variables independientes también se conocen como **variables predictoras** o variables x).

$\hat{y}$ = valor predicho de la variable dependiente y (que se calcula por medio de la ecuación de regresión múltiple).

$x_1, x_2, \ldots, x_k$ son las variables independientes.

β_0 = intercepto y, o el valor de y cuando todas las variables predictoras 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).

$\beta_1, \beta_2, \ldots, \beta_k$ son los coeficientes de las variables independientes $x_1, x_2, \ldots, x_k$

$b_1, b_2, \ldots, b_k$ son estimados muestrales de los coeficientes $\beta_1, \beta_2, \ldots, \beta_k$

Los cálculos que se requieren para la regresión múltiple son tan complicados que *debe* utilizarse un programa de cómputo de estadística, por lo 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 del STATDISK, Minitab, Excel y la calculadora TI-83 Plus.

EJEMPLO Osos Por cuestiones de seguridad, un estudio de osos implicó la recolección de diversas mediciones una vez que los osos estaban anestesiados. Cuando se obtienen mediciones de un oso salvaje anestesiado, es relativamente fácil utilizar una cinta métrica para obtener valores como el tamaño del pecho, el tamaño del cuello y la altura total, pero es difícil calcular el peso debido a que se debe alzar al oso. En lugar de pesar realmente a un oso, ¿podemos predecir su peso con base en otras mediciones que son más fáciles de obtener? El conjunto de datos 9 del Apéndice B incluye mediciones tomadas de 54 osos, pero nosotros consideraremos los datos de únicamente ocho osos, que se listan en la tabla 9-3. Utilice los datos de la tabla 9-3 y calcule la ecuación

Tabla 9-3 Datos de osos machos anestesiados

Variable	Columna de Minitab	Nombre	Datos muestrales							
y	C1	WEIGHT(PESO)	80	344	416	348	262	360	332	34
x_2	C2	AGE(EDAD)	19	55	81	115	56	51	68	8
x_3	C3	HEADLEN(LONGI-TUD DE CABEZA)	11.0	16.5	15.5	17.0	15.0	13.5	16.0	9.0
x_4	C4	HEADWDTH(ANCHO DE CABEZA)	5.5	9.0	8.0	10.0	7.5	8.0	9.0	4.5
x_5	C5	NECK(CUELLO)	16.0	28.0	31.0	31.5	26.5	27.0	29.0	13.0
x_6	C6	LENGTH(ALTURA)	53.0	67.5	72.0	72.0	73.5	68.5	73.0	37.0
x_7	C7	CHEST(PECHO)	26	45	54	49	41	49	44	19

de regresión múltiple en la que la variable dependiente (y) es el peso y las variables independientes son la longitud de la cabeza (HEADLEN) y la altura total (LENGTH).

SOLUCIÓN Con el uso de Minitab, obtenemos los resultados mostrados en la imagen de abajo. La ecuación de regresión múltiple se presenta como

WEIGHT = −374 + 18.8 HEADLEN + 5.87 LENGTH

Si empleamos la notación presentada anteriormente en esta sección, podríamos escribir esta ecuación como

$$\hat{y} = -374 + 18.8x_3 + 5.87x_6$$

Minitab

```
The regression equation is
WEIGHT = -374 + 18.8 HEADLEN + 5.87 LENGTH

Predictor          Coef       Stdev     t-ratio        p
Constant         -374.3       134.1       -2.79    0.038
HEADLEN           18.82       23.15        0.81    0.453
LENGTH            5.875       5.065        1.16    0.299

s = 68.56   R-sq = 82.8% R-sq(adj) = 75.9%
```

Ecuación ① de regresión múltiple

```
Analysis of Variance
```

$R^2 = 0.828$ ② Ajustada $R^2 = 0.759$

```
SOURCE        DF        SS       MS        F        p
Regresión      2    113142    56571    12.03    0.012
Error          5     23506     4701
Total          7    136648
```

③ Significancia general de la ecuación de regresión múltiple

Fabricación de música 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 modo 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 dio como resultado la disminución de discos defectuosos en favor de una mayor calidad. El uso de métodos de regresión múltiple condujo a costos más bajos de producción y a un mejor control del proceso de fabricación.

Si una ecuación de regresión múltiple se ajusta bien a los datos muestrales, resulta útil para hacer predicciones. Por ejemplo, si determinamos que la ecuación es adecuada para predecir, y tenemos un oso con una longitud de cabeza de 14.0 pulgadas y una altura total de 71.0 pulgadas, podemos predecir su peso sustituyendo esos valores en la ecuación de regresión para obtener un peso predicho de 306 libras. Además, los coeficientes $b_3 = 18.8$ y $b_6 = 5.87$ se emplean para determinar el cambio marginal, como se describió en la sección 9-3. Por ejemplo, el coeficiente $b_3 = 18.8$ indica que cuando la altura total de un oso permanece constante, el peso predicho se incrementa en 18.8 libras por cada pulgada de aumento en la longitud de la cabeza.

R^2 ajustada

R^2 denota el **coeficiente múltiple de determinación**, que es una medida de qué tan bien 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 da como resultado un valor cercano a 1. Un ajuste muy pobre se relaciona con un valor de R^2 cercano a 0. El valor de $R^2 = 0.828$ en los resultados de Minitab indica que el 82.8% de la variación del peso de los osos puede explicarse por la longitud de la cabeza x_3 y la altura total x_6. 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 sencillo hecho de incluir todas las variables disponibles, pero la mejor ecuación de regresión múltiple no necesariamente utiliza todas las variables de que se dispone. Debido a 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 9-6.

Fórmula 9-6
$$R^2 \text{ ajustada} = 1 - \frac{(n-1)}{[n-(k+1)]}(1-R^2)$$

donde n = tamaño de la muestra

k = número de variables independientes (x)

Los resultados de Minitab para los datos de la tabla 9-3 indican que el coeficiente ajustado de determinación es `R-sq(adj) = 75.9%`. Si utilizamos la fórmula 9-6 con el valor de $R^2 = 0.828$, $n = 8$ y $k = 2$, encontramos que el valor ajustado de R^2 es 0.759, lo que confirma el valor de 75.9% de los resultados de Minitab. Para los datos de la tabla 9-3 relativos al peso, la longitud de la cabeza y la altura, el valor de R^2 de 82.8% indica que el 82.8% de la variación del peso puede explicarse por la longitud en la cabeza x_3 y la altura total x_6, pero cuando comparamos esta ecuación de regresión múltiple con otras, es mejor utilizar la R^2 ajustada de 75.9% (o 0.759).

Valor *P*

El valor *P* es una medida de la significancia general de la ecuación de regresión múltiple. El valor *P* de 0.012 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, es lógico predecir pesos de osos con base en la longitud de su cabeza y de su altura total. 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 0.012 resulta de una prueba de la hipótesis nula de que $\beta_3 = \beta_6 = 0$. El rechazo de $\beta_3 = \beta_6 = 0$ implica que al menos uno de β_3 y β_6 no es 0, lo que indica que esta ecuación de regresión es eficaz para determinar los pesos de osos. Un análisis completo de los resultados de Minitab podría llevarnos a concluir otros elementos importantes, tales 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

La tabla 9-3 incluye siete variables distintas de medición de ocho osos diferentes. El resultado de Minitab se basa en la selección del peso como variable dependiente y en la selección de la longitud de la cabeza y de la altura total como variables independientes. Pero si deseamos predecir el peso de un oso, ¿existe alguna otra combinación de variables que podría ser mejor que la longitud de la cabeza y la altura total? La tabla 9-4 lista algunas de las combinaciones de variables, y ahora nos confrontamos con el objetivo importante de calcular la *mejor* ecuación de regresión múltiple. Puesto que la determinación de la mejor 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 va más allá de los objetivos de este libro,* pero los siguientes lineamientos proporcionan cierta ayuda.

Lineamientos para el cálculo de la mejor ecuación de regresión múltiple

1. *Utilice el sentido común y consideraciones prácticas para incluir o excluir variables.* Por ejemplo, podríamos excluir la variable de edad debido a que investigadores inexpertos tal vez no sepan cómo determinar la edad de un oso y, al preguntarles, los osos se rehúsan a revelar su edad. A diferencia de otras variables independientes, la edad de un oso no se obtiene fácilmente con una cinta de medición. Por lo tanto, parece lógico excluir la edad como variable independiente.

Tabla 9-4 Búsqueda de la mejor ecuación de regresión múltiple

	ALTURA	PECHO	LONGITUD DE CABEZA/ ALTURA	EDAD/ CUELLO/ ALTURA/PECHO	EDAD/LONGITUD DE CABEZA/ ANCHO DE CABEZA/CUELLO/ ALTURA/PECHO
R^2	0.805	0.983	0.828	0.999	0.999
R^2 ajustada	0.773	0.980	0.759	0.997	0.996
Significancia general	0.002	0.000	0.012	0.000	0.046

Predictores de éxito

Cuando una universidad acepta a un nuevo estudiante, es conveniente tener algunos indicadores positivos de que el estudiante tendrá éxito en sus estudios. Los decanos universitarios de admisiones toman en cuenta las calificaciones del SAT, las pruebas estándar de aprovechamiento, el lugar que ocupa el estudiante en la clase, la dificultad de los cursos de preparatoria, las calificaciones de preparatoria y actividades extracurriculares. En un estudio de las características que suelen ser buenos predictores 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 predictores que las calificaciones del SAT. Una ecuación de regresión múltiple con el promedio general en la universidad predicho por el lugar que ocupa el estudiante en la clase y la puntuación en pruebas de aprovechamiento no mejoró al incluir la calificación del SAT como otra variable. Este estudio en particular sugiere que las calificaciones del 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.

2. *Considere el valor P.* Seleccione una ecuación que tiene significancia general, tal como determina el valor P indicado en los resultados del programa de cómputo. Por ejemplo, observe los valores de la significancia general en la tabla 9-4. El uso de las seis variables independientes da como resultado una significancia general de 0.046, que es apenas significativa a nivel $\alpha = 0.05$; la variable sola `PECHO` es mejor, ya que tiene una significancia general de 0.000.

3. *Considere ecuaciones con valores altos de R^2 y trate de incluir sólo unas cuantas variables.* En lugar de incluir casi todas las variables disponibles, trate de incluir relativamente pocas variables independientes (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. Por ejemplo, la tabla 9-4 muestra que si empleamos únicamente la variable independiente `PECHO`, la R^2 ajustada es 0.980, pero cuando incluimos las seis variables, la R^2 ajustada se incrementa a 0.996. Incluir las seis variables, en lugar de sólo una, es un precio demasiado alto para un incremento tan pequeño en la R^2 ajustada. Es mejor que utilicemos únicamente la variable independiente `PECHO`, que las seis variables independientes.
- Para un número dado de variables independientes (x), seleccione la ecuación con el valor más grande de la R^2 ajustada.
- Para suprimir las variables independientes que no tienen mucho efecto sobre la variable dependiente, sería útil calcular el coeficiente de correlación lineal r para cada par de variables en consideración. Por ejemplo, con los datos de la tabla 9-3, encontraremos que existe una correlación lineal de 0.955 para los datos apareados `CUELLO/LONGITUD DE CABEZA`. Debido a que existe una correlación tan alta entre el tamaño del cuello y la longitud de la cabeza, no hay necesidad de incluir ambas variables. Para elegir entre `CUELLO` y `LONGITUD DE CABEZA`, debemos inclinarnos por el `CUELLO` por la siguiente razón: el `CUELLO` es un mejor predictor del `PESO`, ya que los datos apareados `CUELLO/PESO` tienen un coeficiente de correlación lineal de $r = 0.971$, que es más alto que $r = 0.884$ de los datos apareados `LONGITUD DE CABEZA/PESO`.

Si seguimos estos lineamientos al intentar calcular la mejor ecuación para predecir los pesos de osos, encontramos que, para los datos de la tabla 9-3, la mejor ecuación de regresión utiliza la variable independiente del tamaño del pecho (`PECHO`). Parece que la mejor ecuación de regresión es

$$\text{PESO} = -195 + 11.4\ \text{PECHO}$$

o

$$\hat{y} = -195 + 11.4x_7$$

Algunos programas estadísticos de cómputo incluyen un programa para realizar la **regresión por pasos**, de modo que los cálculos se realizan con distintas combinaciones de variables independientes, pero este procedimiento implica graves problemas, incluyendo los siguientes: la regresión por pasos no necesariamente produce el mejor modelo si algunas variables predictoras tienen una alta correlación; produce valores inflados de R^2; utiliza demasiado papel y no nos permite *pensar* en el problema. Como siempre, debemos ser cuidadosos al emplear los resultados de las computadoras como una herramienta que nos ayude a tomar decisiones inteligentes; no debemos permitir que la computadora sea quien tome las decisiones. En lugar 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 eliminamos la variable EDAD (como en el lineamiento 1) y después corremos el programa de regresión por pasos de Minitab, obtendremos una pantalla de resultados que sugieren que la mejor ecuación de regresión es aquella en la que PECHO es la única variable independiente. (Si incluimos las seis variables independientes, Minitab selecciona una ecuación de regresión con las variables independientes EDAD, CUELLO, ALTURA y PECHO, con un valor ajustado de R^2 de 0.997 y una significancia general de 0.000). Parece que podemos estimar el peso de un oso con base en el tamaño de su pecho, y la ecuación de regresión nos conduce a esta regla: se estima que el peso de un oso (en libras) es 11.4 veces el tamaño de su pecho (en pulgadas) menos 195.

Cuando estudiamos la regresión en la sección 9-3, 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.

Utilizando la tecnología

STATDISK Seleccione **Analysis**, luego **Multiple Regression**. Ingrese los datos en las diferentes columnas o utilice las funciones copiar/pegar para obtener las columnas de datos deseadas. Introduzca los datos de la variable dependiente en la columna 1. Haga clic en **Evaluate** y aparecerá un cuadro de diálogo. Identifique las columnas que desea incluir. STATDISK proporcionará 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 **Statistics** 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 el coeficiente múltiple de determinación 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 las funciones copiar/pegar para mover las columnas como desee). Los resultados incluirán el coeficiente múltiple de determinación R^2, la R^2 ajustada y una lista de los valores del intercepto y coeficiente utilizados para la ecuación de regresión múltiple.

TI-83 Plus El programa A2MULREG de la calculadora TI-83 Plus puede bajarse del CD-ROM incluido en este libro. Seleccione el archivo de *software*, después *TI83PlusPRGMS*. Debe bajar el programa a su calculadora y después introducir los datos muestrales como una matriz D, en donde la primera columna contenga los valores de la variable dependiente (y). Presione **2nd** y la tecla $\boldsymbol{x^{-1}}$, gire hacia la derecha hasta **EDIT**, gire hacia abajo hasta **[D]**, después presione **ENTER** y proceda a introducir el número de valores listados por cada variable, seguido por el número total de variables (incluyendo la variable dependiente). Ahora presione **PRGM,** seleccione **A2MULREG** y luego **ENTER**. Cuando se le solicite, ingrese el número de variables independientes, después introduzca los números de las columnas de las variables independientes que desea incluir. La pantalla proporcionará 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** nuevamente 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 aparece, vaya a la tabla A-3 y busque el valor t crítico correspondiente, introdúzcalo y proceda a ingresar los valores que se emplearán para las variables independientes. Presione **ENTER** para seleccionar la opción **QUIT**.

9-5 Destrezas y conceptos básicos

Interpretación de resultados de programas de cómputo*. En los ejercicios 1 a 4, remítase a los resultados de Minitab que se presentan aquí y responda las preguntas o identifique los elementos indicados. Los resultados de Minitab están basados en la muestra de 54 osos incluida en el conjunto de datos 9 del Apéndice B.*

1. Mediciones de osos Identifique la ecuación de regresión múltiple que expresa el peso en términos de la longitud de la cabeza, la altura y el tamaño del pecho.

2. 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

3. 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 altura y el tamaño del pecho? ¿Por qué?

4. Mediciones de osos Se encuentra que un oso tiene una longitud de cabeza de 14.0 pulgadas, una altura de 70.0 pulgadas y un tamaño del pecho de 50.0 pulgadas.
a. Calcule el peso predicho del oso.
b. El oso en cuestión en realidad pesaba 320 libras. ¿Qué tan preciso es el peso predicho en el inciso *a*?

Minitab

```
The regression equation is
WEIGHT = −272 − 0.87 HEADLEN + 0.55 LENGTH + 12.2 CHEST
Predictor          Coef    SE Coef        T        P
Constant        −271.71      31.62    −8.59    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 = 33.66          R-Sq = 92.8%      R-Sq(adj) = 92.4%

Analysis of Variance

Source           DF      SS        MS        F        P
Regresión         3  729645    243215   214.71    0.000
Residual Error   50   56638      1133
Total            53  786283
```

Datos de automóviles: cálculo de la mejor ecuación de regresión múltiple. *En los ejercicios 5 a 8, remítase a la tabla adjunta, que se obtuvo utilizando el conjunto de datos 22 del Apéndice B. La variable dependiente es el consumo de combustible en ciudad (en millas/galón), y las variables independientes están listadas en la tabla. CRT denota el consumo de combustible en carretera, PS denota el peso del automóvil y DSPZ el desplazamiento del motor del automóvil.*

Variables independientes	Valor *P*	R^2	R^2 ajustada	Ecuación de regresión
CRT, PS, DSPZ	0.000	0.882	0.860	$\hat{y} = 5.9 + 0.742x_1 - 0.00162x_2 - 0.441x_3$
CRT, PS	0.000	0.876	0.861	$\hat{y} = 4.6 + 0.794x_1 - 0.00209x_2$
CRT, DSPZ	0.000	0.873	0.859	$\hat{y} = -3.23 + 0.892x_1 - 0.626x_2$
PS, DSPZ	0.000	0.788	0.763	$\hat{y} = 41.5 - 0.00535x_1 - 0.950x_2$
CRT	0.000	0.860	0.853	$\hat{y} = -9.73 + 1.05x$
PS	0.000	0.759	0.746	$\hat{y} = 44.2 - 0.00708x$
DSPZ	0.000	0.620	0.599	$\hat{y} = 29.5 - 2.74x$

5. Si se utiliza únicamente una variable independiente para predecir la cantidad de consumo de combustible en la ciudad (en mi/gal), ¿cuál variable es mejor? ¿Por qué?

6. Si se van a utilizar exactamente dos variables independientes para predecir la cantidad del consumo de combustible en la ciudad, ¿cuáles dos variables deben elegirse? ¿Por qué?

7. ¿Cuál ecuación de regresión es mejor para predecir la cantidad de consumo de combustible en la ciudad? ¿Por qué?

8. Si un automóvil tiene una tasa de consumo de combustible en carretera de 35 mi/gal, un peso de 2675 libras y un desplazamiento de motor de 3.8 L, ¿cuál es el mejor valor predicho de la tasa de consumo de combustible en la ciudad? ¿Es posible que ese valor predicho constituya un buen estimado? ¿Es posible que el valor predicho sea muy preciso?

T **9.** Estaturas de padres e hijos Remítase al conjunto de datos de 2 del Apéndice B.

a. Calcule la ecuación de regresión que expresa la variable dependiente de la estatura de un hijo en términos de la variable independiente de la estatura de la madre.

b. Calcule la ecuación de regresión que expresa la variable dependiente de la estatura de un hijo en términos de la variable independiente de la estatura del padre.

c. Calcule la ecuación de regresión que expresa la variable dependiente de la estatura de un hijo en términos de las variables independientes de la estatura de la madre y la estatura del padre.

d. Respecto a las ecuaciones de regresión obtenidas en los incisos *a*, *b* y *c*, ¿cuál es la mejor ecuación para predecir la estatura de un hijo? ¿Por qué?

e. ¿Será la mejor ecuación de regresión, identificada en el inciso *d*, una *buena* ecuación para predecir la estatura de un hijo? ¿Por qué?

T **10.** Facilidad de lectura de Harry Potter Remítase al conjunto de datos 14 del Apéndice B y utilice los valores de *Harry Potter y la piedra filosofal*, de J. K. Rowling.

a. Calcule la ecuación de regresión que expresa la variable dependiente de la puntuación de facilidad de lectura de Flesch en términos de la variable independiente de las palabras por oración.

b. Calcule la ecuación de regresión que expresa la variable dependiente de la puntuación de facilidad de lectura de Flesch en términos de la variable independiente de los caracteres por palabra.

c. Calcule la ecuación de regresión que expresa la variable dependiente de la puntuación de facilidad de lectura de Flesch en términos de las variables independientes de las palabras por oración y los caracteres por palabra.

d. Respecto a las ecuaciones de regresión obtenidas en los incisos *a*, *b* y *c*, ¿cuál es la mejor ecuación para predecir una puntuación de la facilidad de lectura de Flesch? ¿Por qué?

e. ¿Será la mejor ecuación de regresión, identificada en el inciso *d*, una *buena* ecuación para predecir una puntuación de facilidad de lectura de Flesch? ¿Por qué?

T **11.** Cereales y calorías Remítase al conjunto de datos 16 del Apéndice B.

a. Calcule la ecuación de regresión que expresa la variable dependiente de calorías en términos de la variable independiente de la cantidad de grasa.

b. Calcule la ecuación de regresión que expresa la variable dependiente de calorías en términos de la variable independiente de la cantidad de azúcar.
c. Calcule la ecuación de regresión que expresa la variable dependiente de calorías en términos de las variables independientes de la cantidad de grasa y de la cantidad de azúcar.
d. Respecto a las ecuaciones de regresión obtenidas en los incisos *a*, *b* y *c*, ¿cuál es la mejor ecuación para predecir el número de calorías? ¿Por qué?
e. ¿Será la mejor ecuación de regresión, identificada en el inciso *d*, una *buena* ecuación para predecir el número de calorías? ¿Por qué?

T **12.** Uso de la basura para predecir el tamaño poblacional Remítase al conjunto de datos 23 del Apéndice B.
a. Calcule la ecuación de regresión que expresa la variable dependiente del tamaño de los hogares en términos de la variable independiente del peso de los desechos de comida.
b. Calcule la ecuación de regresión que expresa la variable dependiente del tamaño de los hogares en términos de la variable independiente del peso de los desechos plásticos.
c. Calcule la ecuación de regresión que expresa la variable dependiente del tamaño de los hogares en términos de las variables independientes del peso de los desechos de comida y el peso de los deshechos plásticos.
d. Respecto a las ecuaciones de regresión obtenidas en los incisos *a*, *b* y *c*, ¿cuál es la mejor ecuación para predecir el tamaño de los hogares? ¿Por qué?
e. ¿Será la mejor ecuación de regresión, identificada en el inciso *d*, una *buena* ecuación para predecir el tamaño de los hogares? ¿Por qué?

9-5 Más allá de lo básico

T **13.** Nicotina de cigarrillos: cálculo de la mejor ecuación de regresión múltiple Remítase al conjunto de datos 5 del Apéndice B y calcule la mejor ecuación de regresión múltiple, con la nicotina como variable dependiente. ¿Será esta "mejor" ecuación buena para predecir la cantidad de nicotina en un cigarrillo con base en la cantidad de alquitrán y monóxido de carbono?

T **14.** Precio de un diamante: cálculo de la mejor ecuación de regresión múltiple Remítase al conjunto de datos 18 del Apéndice B.
a. Utilice únicamente los tres factores tradicionales del quilate, color y claridad para calcular la mejor ecuación de regresión múltiple que podría emplearse para predecir el precio de un diamante.
b. Las variables profundidad y meseta describen el corte de un diamante que, se supone, afecta su color. ¿Existe una relación lineal significativa entre la variable dependiente del color y las variables dependientes de profundidad y meseta? Si no existe una relación lineal significativa, ¿quiere esto decir que el color no se ve afectado por la profundidad y la meseta?

T **15.** Precio de venta de casas: cálculo de la mejor ecuación de regresión múltiple Remítase al conjunto de datos 24 del Apéndice B y calcule la mejor ecuación de regresión múltiple con el precio de venta como variable dependiente. ¿Será esta "mejor" ecuación buena para predecir el precio de venta de una casa?

x	1	3	4	7	5
y	5	14	19	42	26

T **16.** 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 un ecuación como ésta es una parábola. Utilice el conjunto de datos listado al margen, permita que $x_1 = x$, permita que $x_2 = x^2$, y calcule la ecuación de regresión múltiple para la parábola, que se ajusta mejor a los datos. Con base en el valor del coeficiente múltiple de determinación, ¿qué tan bien se ajusta esta ecuación a los datos?

9-6 Elaboración de modelos

No, no ese tipo de modelos. Esta sección introduce algunos conceptos básicos del desarrollo de un **modelo matemático**, que 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. Esto es muy parecido a los métodos de regresión de la sección 9-3, excepto que ya no estamos restringidos a un modelo que deba ser lineal. Además, en lugar 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 se emplea la tecnología para obtener un buen modelo matemático.

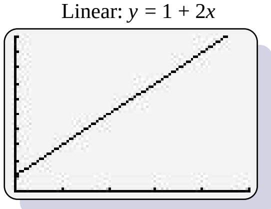
A continuación se presentan algunos modelos genéricos como aparecen en un menú de la calculadora TI-83 Plus (presione **STAT** y luego seleccione **CALC**):

Lineal:	$y = a + bx$	Cuadrático:	$y = ax^2 + bx + c$
Logarítmico:	$y = a + b \ln x$	Exponencial:	$y = ab^x$
Potencia:	$y = ax^b$	Logístico:	$y = \dfrac{c}{1 + ae^{-bx}}$

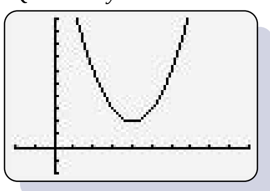
El modelo particular que usted seleccione depende de la naturaleza de los datos muestrales, y un diagrama de dispersión resulta muy útil para tomar esta determinación. Las ilustraciones a continuación son gráficas de algunos modelos comunes elaborados en una calculadora TI-83 Plus.

TI-83 Plus

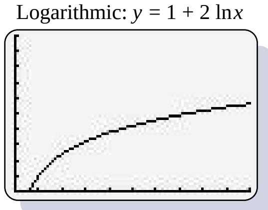
Linear: $y = 1 + 2x$



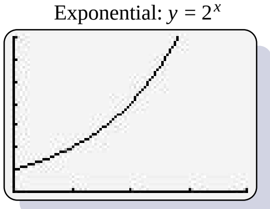
Quadratic: $y = x^2 - 8x + 18$



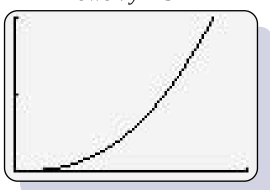
Logarithmic: $y = 1 + 2 \ln x$



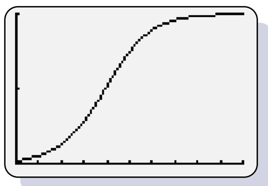
Exponential: $y = 2^x$



Power: $y = 3x^{2.5}$



Logistic: $y = \dfrac{2}{1 + 50e^{-x}}$



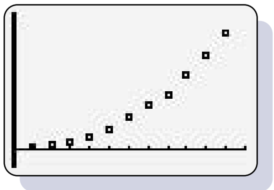
He aquí las reglas básicas para la creación de un buen modelo matemático:

1. *Busque un patrón en la gráfica*. Examine la gráfica con los puntos y compare el patrón básico con las gráficas genéricas conocidas de una función lineal, una función cuadrática, una función exponencial, una función potencial, etcétera. (Remítase a las gráficas que se presentan en los ejemplos de los resultados de la calculadora TI-83 Plus). Cuando trate de seleccionar un modelo, considere únicamente 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 Plus para obtener el valor del coeficiente de determinación R^2. Los valores de R^2 se interpretan aquí de la misma forma que se interpretaron en la sección 9-5. Al delimitar sus posibles modelos, seleccione funciones que dan como resultado valores más grandes de R^2, porque valores más grandes corresponden a funciones que se ajustan mejor a los puntos observados. Sin embargo, no dé demasiada importancia a las diferencias pequeñas, tales 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 10).
3. *Piense*. Aplique el sentido común. No utilice un modelo que conduzca a valores predichos que son poco realistas. Utilice el modelo para calcular valores futuros, valores pasados y valores de años perdidos; luego determine si los resultados son realistas.

Tabla 9-5 Población de Estados Unidos (en millones)

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

TI-83 Plus

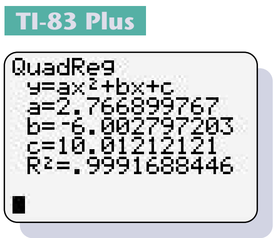


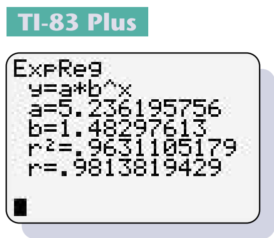
EJEMPLO La tabla 9-5 lista la población de Estados Unidos en diferentes años. Encuentre un buen modelo matemático para el tamaño poblacional, después haga una predicción del tamaño de la población de Estados Unidos para el año 2020.

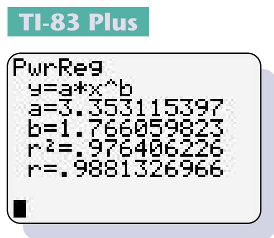
SOLUCIÓN Primero "codificamos" los valores del año utilizando 1, 2, 3. . . en lugar de 1800, 1820, 1840. . . La razón de esta codificación es que de esta manera se utilizan valores de x más pequeños y que tienen muchas menos posibilidades de causar problemas de cálculo, como los que podrían ocurrir al emplear 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 Plus (mostrados al margen) y compare el patrón con los modelos genéricos presentados anteriormente en esta sección. El patrón de estos puntos no es una recta, por lo que descartamos un modelo lineal. Tampoco consideramos un modelo logístico, porque los puntos no presentan el patrón de "S" de esa gráfica, ya que existe un aplanamiento de la gráfica en la zona derecha. 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 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, pero 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 modo que $x = 1$ representa 1800, $x = 2$ representa 1820, y así sucesivamente) y la población y (en millones).







Para predecir la población de Estados Unidos para el año 2020, primero observe que el año 2020 esta codificado como $x = 12$ (véase la tabla 9-5). Sustituyendo $x = 12$ en el modelo cuadrático de $y = 2.77x^2 - 6.00x + 10.01$, obtenemos el resultado $y = 337$, que indica una estimación de que en el año 2020 la población de Estados Unidos será de 337 millones.

Piense. El resultado predicho de 337 millones en 2020 parece razonable. (Una proyección del Bureau of the Census de Estados Unidos 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. Para estimados futuros, únicamente el modelo logístico presenta el comportamiento típico de poblaciones crecientes: la población empieza a estabilizarse cuando alcanza la *capacidad de soporte del ambiente*, la máxima población que puede sostenerse con los recursos limitados. 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 de tiempo.

En el artículo "Modeling the U.S. Population" (*AMATYC Review*, vol. 20, núm. 2), Sheldon Gordon emplea más datos que los de la tabla 9-5 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 analizan y requiere de ejercitar el juicio, más allá de los cálculos".

Utilizando la tecnología

Cualquier sistema capaz de realizar regresión múltiple resulta útil 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 se emplea con los datos de la tabla 9-5 para generar el modelo cuadrático de la siguiente manera: seleccione **Analysis**, después **Multiple Regression**, luego proceda a introducir los valores poblacionales en la columna 1. Introduzca 1, 2, 3, . . . , 11 en la columna 2 e introduzca 1, 4, 9, . . . , 121 en la columna 3. 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 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 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,** y 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.

9-6 Destrezas y conceptos básicos

T ***Obtención del mejor modelo.*** *En los ejercicios 1 a 8, construya un diagrama de dispersión e identifique el modelo matemático que se ajusta mejor a los datos dados. 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.*

1.

x	1	2	3	4	5	6
y	8	2	0	2	8	18

2.

x	1	2	3	4	5	6
y	3	8	13	18	23	28

3.

x	1	2	3	4	5	6
y	3	9	27	80	245	725

4.

x	1	2	3	4	5	6
y	2.000	2.828	3.464	4.000	4.472	4.899

T **5.** Muertes de manatíes por barcos La tabla adjunta lista el número de muertes de manatíes en Florida, relacionadas con encuentros con embarcaciones (datos tomados de el *The New York Times*). ¿Cuál es el mejor valor predicho para 2001? En 2001 hubo 82 muertes de manatíes relacionadas con embarcaciones. ¿De qué manera se compara el valor predicho con el valor real?

Año	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Muertes	16	24	20	15	34	33	33	39	43	50	47

Año	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Muertes	53	38	35	49	42	60	54	67	82	78

6. Mercado bursátil Remítase a los valores altos anuales del Promedio Industrial Dow-Jones, listados en el conjunto de datos 25 del Apéndice B. ¿Cuál es el mejor valor predicho para el año 2001? Considerando que el alto valor real en 2001 fue de 11,350, ¿qué tan bueno fue el valor predicho? ¿Qué sugiere el patrón acerca del mercado bursátil para propósitos de inversión? (Actos de terrorismo y condiciones económicas negativas causaron grandes pérdidas en el mercado bursátil en 2002).

7. Tiendas “Target” La siguiente tabla lista el número de tiendas departamentales “Target” en Estados Unidos (de acuerdo con datos de “Target”). ¿Cuál es el mejor valor predicho para el número de tiendas “Target” en 2005?

Año	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Tiendas	420	463	506	554	611	670	736	796	851	914	984

8. Recuperación de inversión Kendra Korbin, propietaria y operadora de Cyber Video Game Store, registra los costos y los ingresos de su negocio durante varios años. Los resultados se presentan abajo.

Cantidad invertida (en miles de dólares)	1	2	5	11	20	31	41	46	48
Ingresos (en dólares)	2001	2639	3807	5219	6629	7899	8834	9250	9409

9-6 Más allá de lo básico

9. 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: 1971: 2.3; 1978: 31; 1982: 110; 1985: 280; 1989: 1200; 1993: 3100; 1995: 5500; 1999: 14,000. Permita que 1971 sea el año base representado por $x = 1$.

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, logístico? ¿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?

10. Uso del criterio de suma de cuadrados Se señaló que, 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 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 $\Sigma(y - \hat{y})^2$, la suma de cuadrados de los residuales que resultan del modelo lineal.

b. Calcule la suma de cuadrados de los residuales que resulta del modelo cuadrático.

c. Verifique que, según el criterio de la suma de cuadrados, el modelo cuadrático es mejor que el modelo lineal.

11. Cálculo de suma de cuadrados y R^2 Si utilizamos los datos de la tabla 9-5, el modelo logístico es

$$y = \frac{465.9305}{1 + 72.5260e^{-0.425483x}}$$

a. Calcule $\Sigma(y - \hat{y})^2$, la suma de los cuadrados de los residuales.
b. Calcule

$$R^2 = 1 - \frac{\Sigma(y - \hat{y})^2}{\Sigma(y - \bar{y})^2}$$

c. Después de comparar los valores de R^2 y las sumas de cuadrados de los residuales, determine si el modelo logístico es mejor que el modelo cuadrático.

Repaso

Este capítulo presentó métodos básicos para investigar relaciones o correlaciones entre dos o más variables.

- La sección 9-2 empleó diagramas de dispersión y el coeficiente de correlación lineal para decidir si existe una correlación lineal entre dos variables.
- La sección 9-3 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 permite predecir el valor de una variable a partir de un valor de la otra variable.
- La sección 9-4 introdujo el concepto de variación total, con componentes de variación explicada y sin explicar. Definimos el coeficiente de determinación r^2 como el cociente obtenido al dividir la variación explicada entre la variación total. También desarrollamos métodos para construir intervalos de predicción, los cuales sirven para juzgar la precisión de valores predichos.
- En la sección 9-5 consideramos la regresión múltiple, que nos permite investigar relaciones entre diversas variables. 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 9-6 exploramos conceptos básicos para el desarrollo de un modelo matemático, que es una función que se emplea para describir una relación entre dos variables. A diferencia de las secciones anteriores de este capítulo, la sección 9-6 incluyó varias funciones no lineales.

Ejercicios de repaso

1. CIS y prisión Se realizó un estudio para investigar la relación entre la edad (en años) y la CAS (concentración de alcohol en sangre) medida cuando presos CIS (detenidos por conducir bajo el influjo de sustancias tóxicas) fueron arrestados por primera vez. A continuación se presentan datos muestrales de sujetos seleccionados aleatoriamente (según datos de Dutchess County STOP-DWI Program). Con base en el resultado, ¿parece que el nivel de CAS está relacionado con la edad de las personas sometidas a prueba?

Edad	17.2	43.5	30.7	53.1	37.2	21.0	27.6	46.3
CAS	0.19	0.20	0.26	0.16	0.24	0.20	0.18	0.23

2. Propinas Muchos de nosotros hemos escuchado que la propina debe corresponder al 15% de la cuenta. La lista adjunta incluye algunos datos muestrales reunidos por los alumnos del autor. Utilice los datos muestrales para lo siguiente.

a. ¿Existe suficiente evidencia para concluir que hay una relación entre el monto de la cuenta y el monto de la propina?

b. Si existe una relación, ¿cómo la utilizamos para determinar la cantidad de propina que debemos dejar?

Cuenta ($)	33.46	50.68	87.92	98.84	63.60	107.34
Propina ($)	5.50	5.00	8.08	17.00	12.00	16.00

Datos de helados: comprensión de la correlación y la regresión. *En los ejercicios 3 a 6, utilice los datos de la tabla adjunta (tomados de Kadiyala,* Econometrica*, vol. 38). Los datos provienen de un estudio del consumo de helado que abarcó las primaveras y veranos de tres años. El consumo de helado se midió en pintas* per cápita *por semana, el precio del helado en dólares, el ingreso familiar de los consumidores en dólares por semana y la temperatura en grados Fahrenheit.*

Consumo	0.386	0.374	0.393	0.425	0.406	0.344	0.327	0.288	0.269	0.256
Precio	1.35	1.41	1.39	1.40	1.36	1.31	1.38	1.34	1.33	1.39
Ingreso	351	356	365	360	342	351	369	356	342	356
Temperatura	41	56	63	68	69	65	61	47	32	24

3. a. Utilice un nivel de significancia de 0.05 para probar una correlación lineal entre consumo y precio.

b. ¿Qué porcentaje de la variación del precio se explica por la relación lineal entre precio y consumo?

c. Calcule la ecuación de la recta de regresión que expresa el consumo (y) en términos del precio (x).

d. ¿Cuál es la mejor cantidad de consumo predicho si el precio es de $1.38?

4. a. Utilice un nivel de significancia de 0.05 para probar una correlación lineal entre consumo e ingreso.

b. ¿Qué porcentaje de la variación en el consumo se explica por la relación lineal entre consumo e ingreso?

c. Calcule la ecuación de la recta de regresión que expresa el consumo (y) en términos del ingreso (x).

d. ¿Cuál es la mejor cantidad de consumo predicho si el ingreso es de $365?

5. a. Utilice un nivel de significancia de 0.05 para probar una correlación lineal entre consumo y temperatura.

b. ¿Qué porcentaje de la variación del consumo se explica por la relación lineal entre consumo y temperatura?

c. Calcule la ecuación de la recta de regresión que expresa el consumo (y) en términos de la temperatura (x).

d. ¿Cuál es la mejor cantidad de consumo predicha si la temperatura es de 32°F?

6. Utilice programas de cómputo como STATDISK, Minitab o Excel para calcular la ecuación de regresión múltiple de la forma $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3 x_3$, donde la variable dependiente y representa el consumo, x_1 representa el precio, x_2 representa el ingreso y x_3 representa la temperatura. También 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. ¿Sirve la ecuación de regresión para predecir el consumo de helado? ¿Serán mejores algunas de las ecuaciones de los ejercicios 3 a 5?

Ejercicios de repaso acumulativo

1. *La guerra y la paz* de León Tolstoi Remítase a los datos muestrales de 12 páginas seleccionadas al azar de la obra *La guerra y la paz*, de León Tolstoi, tal como se listan en el conjunto de datos 14 del Apéndice B.
   a. Las puntuaciones de facilidad de lectura de Flesch y las puntuaciones de nivel de Flesch-Kincaid fueron diseñadas para medir la facilidad de lectura. Pruebe si existe una correlación entre esas dos variables.
   b. Calcule la ecuación de regresión en la que la puntuación de facilidad de lectura de Flesch es la variable dependiente y la puntuación del nivel de lectura de Flesch-Kincaid es la variable independiente.
   c. ¿Será posible probar la aseveración de que para la población de todas las páginas de *La guerra y la paz*, la puntuación media de facilidad de lectura de Flesch es igual a la puntuación media de nivel de lectura de Flesch-Kincaid? ¿Tendría sentido hacer una prueba como ésta?
   d. Construya un estimado del intervalo de confianza del 95% para la puntuación media de facilidad de lectura de Flesch para la población de todas las páginas de *La guerra y la paz.*

2. Efectos de la herencia y del ambiente sobre el CI Al estudiar los efectos de la herencia y del ambiente sobre la inteligencia, ha sido de gran utilidad analizar los CI de gemelos idénticos que fueron separados inmediatamente después de su nacimiento. Los gemelos idénticos comparten genes idénticos heredados del mismo huevo fertilizado. Al estudiar gemelos idénticos criados de manera separada, podemos eliminar la variable de la herencia y aislar mejor los efectos del ambiente. La siguiente tabla incluye los CI de padres de gemelos idénticos (los gemelos más grandes son x) criados de forma separada (datos tomados de "IQs of Identical Twins Reared Apart", de Arthur Jensen, *Behavioral Genetics*). Los datos muestrales son típicos de los obtenidos en otros estudios.
   a. Calcule la media y la desviación estándar de la muestra de los gemelos más grandes.
   b. Calcule la media y la desviación estándar de la muestra de los gemelos más jóvenes.
   c. Con base en los resultados de los incisos a y b, ¿parece existir una diferencia entre las medias de las dos poblaciones? Al explorar la relación entre el CI de gemelos, ¿será la comparación de las dos medias muestrales el mejor método? ¿Por qué?
   d. Combine todas las puntuaciones de CI muestrales y después utilice un nivel de significancia de 0.05 para probar la aseveración de que la puntuación media del CI de gemelos criados por separado es diferente del CI medio de 100.
   e. ¿Existe una relación entre los CI de gemelos que fueron separados inmediatamente después de su nacimiento? ¿Qué método utilizó? Escriba resúmenes acerca del efecto de la herencia y del ambiente sobre la inteligencia, y señale que sus conclusiones se basan en esta muestra relativamente pequeña de 12 pares de gemelos idénticos.

x	107	96	103	90	96	113	86	99	109	105	96	89
y	111	97	116	107	99	111	85	108	102	105	100	93

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 estatura y estatura umbilical? Si es así, calculen la ecuación de regresión con la estatura expresada en términos de la estatura umbilical. Según una vieja teoría, la proporción de la estatura respecto a la estatura umbilical de la persona promedio es el cociente de oro: $(1 + \sqrt{5})/2 \approx 1.6$. ¿Esta teoría parece razonablemente precisa?

2. Actividad en clase Formen grupos de 8 a 12 personas. Para cada miembro de cada grupo, *midan* la estatura y el largo de brazo. Para el largo de brazo el sujeto debe estar de pie con los brazos extendidos, como las alas de un avión. Es fácil marcar la estatura y el largo de brazo en un pizarrón y después medir las distancias desde ahí. Con los datos muestrales apareados, ¿existe correlación entre la estatura y el largo de brazo? Si es así, calculen la ecuación de regresión con la estatura expresada en términos del largo de brazo. ¿Puede emplearse el largo de brazo como un predictor suficientemente bueno de la estatura?

3. Actividad en clase Formen grupos de 8 a 12 personas. Para cada miembro, 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 Organicen grupos de tres a cuatro personas. El Apéndice B incluye muchos conjuntos de datos que no se han analizado con los métodos de este capítulo. Por ejemplo, con el conjunto de datos 25, podemos investigar la correlación entre los valores altos del Promedio Industrial Dow Jones y el número de ventas de automóviles en Estados Unidos. Busquen en el apéndice B un par de variables de interés, después investiguen la correlación y la regresión. Enuncien sus conclusiones y traten de identificar aplicaciones prácticas.

5. 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? 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 beisbol (basquetbol o futbol) y sus logros por temporada?
   - Tasas *versus* pesos: ¿existe relación entre las tasas de consumo de combustible de los automóviles y el peso de los automóviles? Si es así, ¿cuál es?
   - ¿Existe una relación entre un 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?

6. Actividad en clase Dividan la clase en grupos de tres o cuatro personas. El problema del capítulo y el problema de la sección “De los datos a la decisión” tratan sobre el tema de las muertes de manatíes por los barcos. Identifiquen otras dos variables ecológicas o ambientales que pueden estar relacionadas, tales como la población de ballenas y el número de barcos pesqueros. Lleven a cabo una investigación, reúnan y analicen datos, y enuncien sus conclusiones. Con base en los resultados, planteen recomendaciones para mejorar nuestro mundo.

Proyecto tecnológico

En el ejercicio 2 de los ejercicios de repaso acumulativo señalamos que, al estudiar los efectos de la herencia y el ambiente sobre la inteligencia, ha sido útil analizar los CI de gemelos idénticos que fueron separados inmediatamente después de su nacimiento. 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. Utilice el mismo procedimiento descrito en el Proyecto tecnológico al final del capítulo 5, genere una lista de 100 puntuaciones de CI simuladas, seleccionadas aleatoriamente de una población distribuida normalmente, con una media de 100 y una desviación estándar de 15. Ahora utilice el mismo procedimiento para generar una segunda lista de 100 puntuaciones de CI simuladas, que también se seleccionan aleatoriamente 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, trátelas como datos apareados, de modo 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 9-2 con un nivel de significancia de 0.05 para probar una correlación lineal significativa y enuncie sus resultados.

Considere que el procedimiento anterior es un ensayo. Dada la forma en que se generaron los datos muestrales, ¿qué proporción de dichos ensayos conduce 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; en este caso, significa que concluimos que existe una correlación lineal significativa, cuando en realidad no existe dicha correlación.

de los DATOS a la DECISIÓN

Pensamiento crítico: ¿deben imponerse mayores restricciones para salvar a los manatíes?

Con los datos muestrales de la tabla 9-1, encontramos que existe una correlación lineal significativa entre el número de registros de barcos en Florida y el número de muertes de manatíes por encuentros con barcos de Florida. Como consecuencia, se han creado refugios para manatíes, se han impuesto límites de velocidad para los barcos en ciertas zonas y se ha pospuesto la construcción de nuevos muelles. Puesto que un mayor número de barcos corresponde a un mayor número de muertes de manatíes, ¿deberán imponerse más restricciones?

Este tema parece ser muy claro para los ambientalistas. Después de todo, las autopsias prueban claramente que los manatíes mueren a causa de los barcos. Sin embargo, otros argumentan que, además de considerar el número de registros de barcos y el número de muertes de manatíes por los barcos, también es importante tomar en cuenta los cambios en el tamaño de la población de manatíes. Es difícil identificar los valores de la población de manatíes, ya que con frecuencia residen en aguas con poca visibilidad y suelen descansar en el fondo de aguas profundas. Los estudios aéreos se han criticado por considerarse poco confiables. En la sección 9-2 señalamos que el investigador Thomas Fraser sugirió en un reporte que "el estado debe implementar un programa vigoroso de captura-etiquetación y recaptura para obtener mayor información acerca del tamaño y los cambios de la población". Cathy Beck, una bióloga que trabaja para el Geological Survey de Estados Unidos, afirma que el uso de métodos estadísticos con los patrones de las cicatrices en la espalda de los manatíes podría ser útil para comprender a la población de manatíes.

Analice los resultados

La siguiente lista incluye el número de registros de barcos (en decenas de miles), las muertes de manatíes por los barcos y los conteos aéreos de manatíes de varios años. Utilice los datos para investigar y analizar aspectos relevantes. Determine si se deben imponer mayores restricciones para salvar a los manatíes.

Año	Barcos	Muertes de manatíes	Población de manatíes
1976	41	10	738
1977	42	13	
1978	43	21	
1979	45	24	
1980	46	16	
1981	48	24	
1982	47	20	
1983	50	15	
1984	52	34	
1985	55	33	1200
1986	64	33	
1987	62	39	
1988	64	43	
1989	66	50	
1990	67	47	
1991	68	53	1267
1992	68	38	1856
1993	67	35	
1994	70	49	
1995	71	42	1443
1996	73	60	2639
1997	76	54	2229
1998	81	67	2022
1999	83	82	1873
2000	84	78	2223
2001		81	3276

PROYECTO DE INTERNET Regresión lineal

El coeficiente de correlación lineal es una herramienta que se utiliza para medir la potencia de una relación lineal entre dos conjuntos de mediciones. Estrictamente desde el punto de vista de los cálculos, el coeficiente de correlación puede obtenerse para cualesquiera dos conjuntos de datos de valores apareados, sin importar lo que esos datos representen. Por esta razón es necesario plantear ciertas preguntas cuando se investiga una correlación. ¿Es razonable esperar una correlación lineal? ¿Podría una correlación obtenida ser causada por una tercera cantidad relacionada con cada una de las variables estudiadas? Localice la página Web de este libro de texto:

http://www.aw.com/triola

El proyecto de Internet para este capítulo lo guiará hasta varios conjuntos de datos apareados en las áreas de deportes, medicina y economía. En el proyecto 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.

La estadística @ en el trabajo

"Quienes solicitan empleo deben tener conocimientos fundamentales de estadística y de sus implicaciones en el mundo de los negocios".

Angela Gillespie

Analista de tráfico, Lycos.com

Como analista de tráfico para Lycos, Inc., Angela realiza reportes sobre mediciones amplias y menores de tráfico. Ella verifica los cambios en las tendencias y los patrones de comportamiento del uso del sitio de Internet, mejorándolo para incrementar su alcance y magnitud (la cantidad de tiempo que las personas están conectadas a cualquier sitio Web en particular).

¿Cuál es su trabajo en Lycos?

Realizo reportes de tráfico de las actividades de nuestro sitio cada semana. Los reportes pasan después a revisión con nuestros equipos de trabajo de producción y gerentes. Ellos ven qué aumenta, qué disminuye y toman decisiones relativas al gasto de los recursos.

Mis reportes analizan básicamente las tendencias en los sitios y proyectan dónde estaremos en un año o en cualquier periodo de tiempo.

¿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 de siguiente año fiscal. Con el uso de datos de visitantes de los dos años anteriores, hice una proyección que mostró en dónde 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 hubiese sabido en dónde enfocar su energía y sus recursos. Puesto que forman un equipo eficiente, hemos alcanzado nuestra meta inalcanzable.

¿Estará aumentando, disminuyendo o permanecerá estable el uso de la probabilidad y la estadística?

Conforme Lycos se vuelve más sofisticado, ellos (la gerencia) esperan reportes cada vez más sofisticados. Está en aumento.

¿Considera que los solicitantes de empleo que tienen algunos estudios de estadística son evaluados de forma más favorable?

Por completo, y no sólo en lo que respecta a los reportes de Lycos, sino también en mercadotecnia y finanzas.

10

Experimentos multinomiales y tablas de contingencia

PROBLEMA DEL CAPÍTULO

PROBLEMA 10 DEL CAPÍTULO

Uso de la estadística para detectar fraudes

En el artículo del *New York Times* “Following Benford’s Law, or Looking Out for No. 1”, Malcolm Browne escribe que “las agencias de recaudación de impuestos de varias naciones y varios estados, entre ellos California, al igual que diversas compañías grandes y negocios contables, están utilizando programas de cómputo de detección, que se basan en la ley de Benford”. De acuerdo con la ley de Benford, una variedad de conjuntos diferentes de datos incluyen números con dígitos líderes (los primeros) que siguen la distribución que se muestra en los primeros dos renglones de la tabla 10-1. 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 en las devoluciones de impuestos, longitudes de ríos y montos de cheques.

Cuando trabajaba para el Brooklyn District Attorney, 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 que se espera de cheques con montos con un dígito líder de 1 es 235.984 (puesto que el 30.1% de 784 es 235.984). Las otras frecuencias que se esperaban se listan en el tercer renglón de la tabla 10-1. El último renglón de la tabla 10-1 lista las frecuencias de los dígitos líderes de los montos de 784 cheques que expidieron 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 tal discordancia? ¿Son *significativas* tales discrepancias? ¿Hay bastante evidencia para justificar la conclusión de que se cometió un fraude? ¿La evidencia va más allá de una “duda razonable”? En este capítulo abordaremos estas preguntas.

Tabla 10-1 Ley de Benford: Distribución de dígitos líderes

Dígito líder	1	2	3	4	5	6	7	8	9
Frecuencia de acuerdo con la ley de Benford	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%
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 de 784 cheques que se analizarón por fraude	0	15	0	76	479	183	8	23	0

10-1 Panorama general

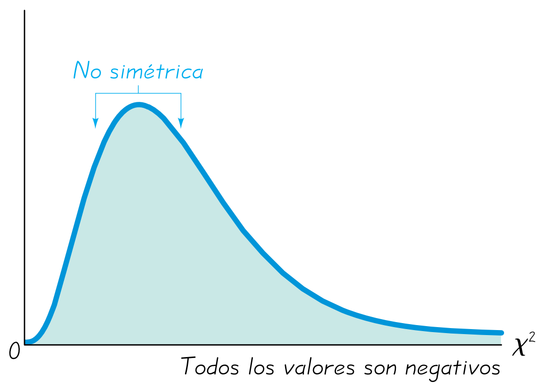
En este capítulo continuamos con la aplicación de métodos inferenciales a diferentes configuraciones de datos. Recuerde que en el capítulo 1 vimos que los datos categóricos (o cualitativos o de atributo) son aquellos datos que pueden separarse en categorías diferentes (que suelen llamarse **celdas**) y se distinguen por alguna característica no numérica. Por ejemplo, es posible separar una muestra de dulces M&M en las categorías de colores rojo, anaranjado, amarillo, café, azul y verde. Después de calcular el conteo de frecuencia para cada categoría, procederíamos a probar la aseveración de que las frecuencias se ajustan (o concuerdan) con la distribución de color que asevera 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 10-2 consideraremos experimentos multinomiales, que consisten en conteos de frecuencias que se observan en un solo renglón o en una columna (que se conoce como tabla de frecuencias de un factor), y probaremos la aseveración de que los conteos de frecuencias que se observan concuerdan con alguna distribución aseverada. En la sección 10-3 consideraremos tablas de contingencia (o tablas de frecuencias de dos factores), consistentes en conteos de frecuencias ordenados en una tabla con al menos dos renglones y dos columnas. Utilizaremos las tablas de contingencia para dos tipos de pruebas muy similares: 1. pruebas de independencia, que prueban la aseveración de que las variables del renglón y la columna son independientes; y 2. pruebas de homogeneidad, que prueban la aseveración de que poblaciones diferentes tienen la misma proporción de alguna característica especificada.

Veremos que los métodos de este capítulo, utilizan la misma distribución χ^2 (chi cuadrada) que se introdujo en la sección 6-5. A continuación se listan las propiedades importantes de la distribución chi cuadrada:

1. A diferencia de las distribuciones normal y t de Student, la distribución chi cuadrada no es simétrica. (Véase la figura 10-1).
2. Los valores de la distribución chi cuadrada pueden ser 0 o positivos, pero no negativos. (Véase la figura 10-1).
3. La distribución chi cuadrada es diferente para cada número de grados de libertad. (Véase la figura 10-2).

Los valores críticos de la distribución chi cuadrada se encuentran en la tabla A-4.

FIGURA 10-1 La distribución chi cuadrada



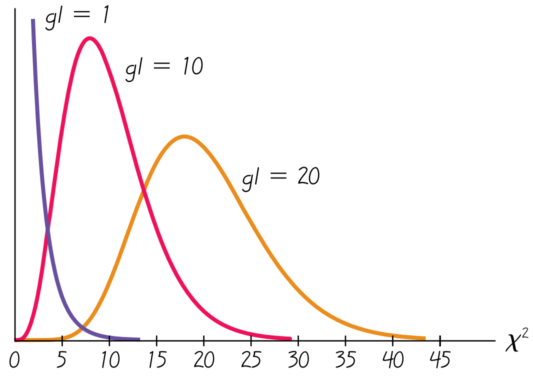


FIGURA 10-2 **Distribución chi cuadrada para 1, 10 y 20 grados de libertad**

10-2 Experimentos multinomiales: bondad de ajuste

Cada conjunto de datos en esta sección consta de datos que se separaron en categorías diferentes. El objetivo principal es determinar si la distribución concuerda o "se ajusta" con alguna distribución que se asevera. Definimos un *experimento multinomial* de la misma manera que un experimento binomial (sección 4-3), excepto que en un experimento multinomial hay más de dos categorías (de manera diferente al 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 **Análisis del último dígito de distancias de jonrón**
En 2001, Barry Bonds anotó 73 jonrones y se convirtió en el nuevo poseedor del récord como el jugador de beisbol que conectó el mayor número de jonrones en una temporada. El conjunto de datos 30 en el Apéndice B lista las distancias que se registraron de estos jonrones, en tanto que la tabla 10-2 resume los *últimos dígitos* de tales distancias. Si se miden realmente las distancias, por lo regular esperaríamos que los últimos dígitos ocurrieran con frecuencias relativas (o probabilidades) que son aproximadamente las mismas. En contraste, los valores estimados tienden a tener 0 o 5, con una ocurrencia mucho más frecuente como últimos dígitos. En la tabla 10-2 parece que hay muchos más ceros de los que obtendríamos con las mediciones reales. Más tarde analizaremos los datos; por ahora, simplemente verificamos que se satisfagan las cuatro condiciones de un experimento multinomial.

continúa

Tabla 10-2
Últimos dígitos de las distancias de jonrones de Barry Bonds

Último dígito	Frecuencia
0	47
1	3
2	1
3	0
4	3
5	11
6	3
7	3
8	1
9	1

SOLUCIÓN Aquí la verificación de que se satisfacen las cuatro condiciones del experimento multinomial:

1. El número de ensayos (últimos dígitos) es el número fijo 73.
2. Los ensayos son independientes, puesto que el último dígito de la longitud de un jonrón no afecta al último dígito de la longitud de cualquier otro jonrón.
3. Cada resultado (último dígito) se clasifica exactamente en una de 10 categorías diferentes. Las categorías se identifican como 0, 1, 2, . . . , 9.
4. Finalmente, si suponemos que se miden las distancias de jonrones, los últimos dígitos deberían ser igualmente probables, para que cada posible dígito tenga una probabilidad de 1/10.

En esta sección presentamos un método para probar la aseveración de que, en un experimento multinomial, las frecuencias que se observan 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 que se observó a alguna distribución teórica que se especifica, 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 concuerda con) alguna distribución que se asevera.

Por ejemplo, utilizando los datos de la tabla 10-2, probaremos la hipótesis de que los datos se ajustan a una distribución uniforme, en la que todos los dígitos son igualmente probables. Nuestras pruebas de bondad de ajuste incorporarán la siguiente notación.

Notación

O representa la *frecuencia* que se observa de un resultado.
E representa la *frecuencia* que se espera 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 10-2 vemos que las frecuencias O que se observan se denotan por 47, 3, 1, 0, 3, 11, 3, 3, 1 y 1. La suma de las frecuencias que se observa es 73, entonces $n = 73$. Si suponemos que los 73 dígitos se obtuvieron de una población en la que todos los dígitos son igualmente probables, entonces *esperamos* que cada dígito ocurra en 1/10 de los 73 ensayos, de manera que cada una de las 10 frecuencias esperadas están dadas por $E = 7.3$. Si generalizamos el resultado, obtendremos un procedimiento sencillo para calcular las frecuencias que se esperan, siempre y cuando supongamos que todas las frecuencias que se esperan son iguales: simplemente hay que dividir 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, suele ser posible calcular las frecuencias que se esperan para cada categoría multiplicando la suma de todas las frecuencias que se observaron por la probabilidad p de la categoría, entonces $E = np$. Aquí resumimos estos dos procedimientos.

- **Si todas las frecuencias esperadas son iguales, entonces cada frecuencia que se espera es la suma de todas las frecuencias observadas, que se divide entre el número de categorías, de manera que $E = n/k$.**
- **Si las frecuencias que se esperan no son todas iguales, entonces cada frecuencia esperada se calcula multiplicando la suma de todas las frecuencias que se observan por la probabilidad para la categoría, entonces $E = np$ para cada categoría.**

Aun cuando estas dos fórmulas para E pueden ser muy buenas, sería mejor utilizar un método informal que se base en la comprensión de las circunstancias. Sólo pregúntese: "¿Cómo es posible repartir las frecuencias que se observan entre las diferentes categorías, de manera que haya un acuerdo perfecto con la distribución que se asevera?". Además, reconozca que las frecuencias que se *observan* deben ser todas números enteros, puesto que representan conteos reales, pero las frecuencias que se *esperan* 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 ocurra con una frecuencia de 5.5, aunque es imposible obtener el resultado de que el 3 ocurra exactamente 5.5 veces.

Sabemos que las frecuencias muestrales por lo regular se desvían un poco de los valores que esperamos teóricamente, así que presentamos ahora la pregunta clave: ¿Son estadísticamente significativas las diferencias entre los valores O reales que se observan y los valores E teóricos que se esperan? Necesitamos una medida de la discrepancia entre los valores O y E, entonces utilizamos el estadístico de prueba dado con los supuestos y los valores críticos. (Más adelante explicaremos cómo se desarrolló este estadístico de prueba, aunque note que incluye diferencias de $O - E$ como componente clave).

Supuestos

1. Los datos se seleccionaron aleatoriamente.
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 hay ningún requisito de que la frecuencia que se *observa* 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 forma del estadístico de prueba χ^2 es de manera tal que una *concordancia cercana* entre los valores que se observan y esperan llevará a un valor de χ^2 *pequeño* y a un valor *P grande*. Una discrepancia grande entre los valores observados y esperados llevará a un valor de χ^2 *grande* y a un valor *P pequeño*. Por lo tanto, 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. Tales relaciones se resumen e ilustran en la figura 10-3.

Una vez que conocemos cómo calcular el valor del estadístico de prueba y el valor crítico, probaremos hipótesis utilizando el procedimiento que se introdujo en el capítulo 7, resumido en la figura 7-8.

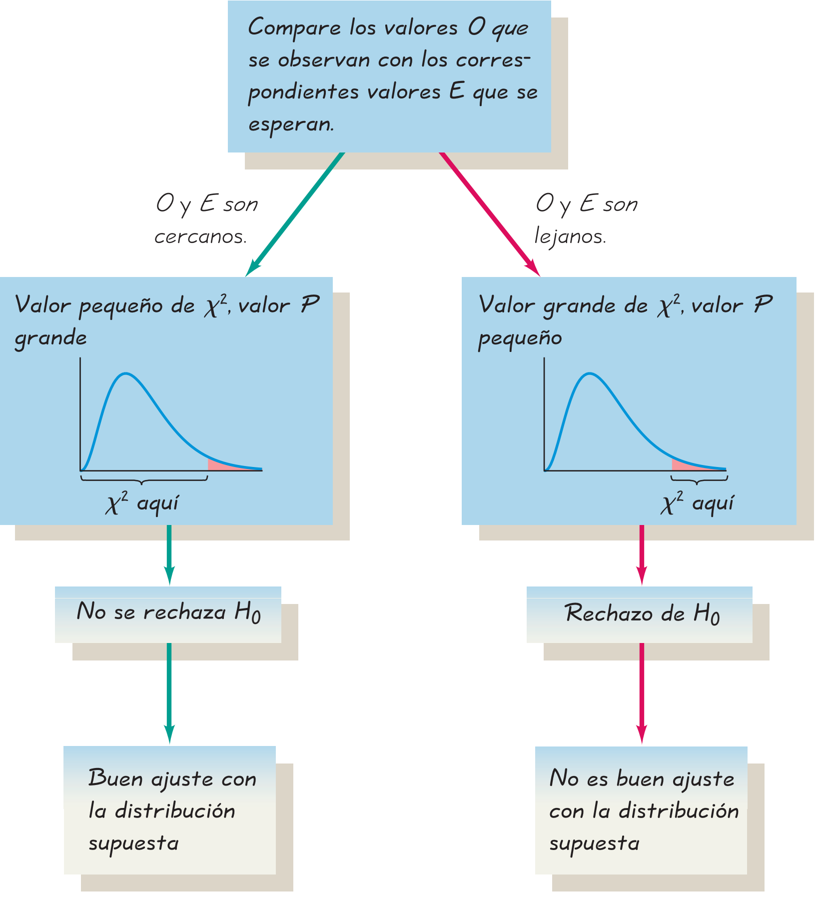


FIGURA 10-3 Relaciones entre el estadístico de prueba χ^2, el valor *P* y la bondad de ajuste

EJEMPLO Análisis del último dígito de jonrones: frecuencias iguales que se esperan Remitámonos otra vez en la tabla 10-2 a los últimos dígitos de las distancias de jonrones de Barry Bonds. El valor 0 parece ocurrir con mayor frecuencia, pero ¿es esto en verdad significativo? Pruebe la aseveración de que los dígitos no ocurren con la misma frecuencia.

SOLUCIÓN 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 ($p_0, p_1, \ldots, p_9$) no son todas iguales. Aplicaremos nuestro procedimiento estándar para la prueba de hipótesis.

Paso 1: La aseveración original es que los dígitos no ocurren con la misma frecuencia. Es decir, al menos una de las probabilidades $p_0, p_1, \ldots, p_9$ es diferente de las otras.

Paso 2: Si la aseveración original es falsa, entonces todas las probabilidades son las mismas. Esto es, $p_0 = p_1 = \ldots = p_9$.

Paso 3: La hipótesis nula debe contener la condición de igualdad, entonces 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: Al menos una de las probabilidades es diferente de las otras.

Paso 4: No se especificó un nivel de significancia, entonces seleccionamos $\alpha = 0.05$, una elección muy común.

Paso 5: Ya que probamos una aseveración acerca de que la distribución de los últimos dígitos es una distribución uniforme, utilizamos la prueba de bondad de ajuste descrita en esta sección. Se emplea la distribución χ^2 con el estadístico de prueba que se dio al principio.

Paso 6: Las frecuencias O que se observan se listan en la tabla 10-2; cada frecuencia E correspondiente que se espera es igual a 7.3 (si los 73 dígitos se distribuyeran uniformemente a través de las 10 categorías). La tabla 10-3, en la siguiente página, muestra el cálculo del estadístico de prueba χ^2. El estadístico de prueba es $\chi^2 = 251.521$ (que se redondea). El valor crítico es de $\chi^2 = 16.919$ (que se encontró en 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 10-4, en la siguiente página.

Paso 7: Puesto que el estadístico de prueba cae dentro de la región crítica, hay evidencia suficiente para rechazar la hipótesis nula.

Paso 8: Existe suficiente evidencia para sustentar la aseveración de que los últimos dígitos no ocurren con la misma frecuencia relativa. Ahora tenemos evidencia muy fuerte que sugiere que las distancias de jonrones realmente no se midieron. Es razonable especular que las distancias son estimados en lugar de mediciones reales.

Las técnicas de esta sección resultan útiles para probar si una distribución de frecuencias que se observan tiene un buen ajuste con alguna distribución de frecuencias de carácter teórico. El ejemplo anterior probó la bondad de ajuste con una distribución uniforme. Puesto que muchos análisis estadísticos requieren de una población que se distribuye normalmente, es posible utilizar la prueba chi cuadrada de esta sección para ayudar a determinar si las muestras dadas se obtienen a partir de poblaciones que se distribuyen normalmente (véase el ejercicio 25).

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 se ilustra en el siguiente ejemplo.

LA ESTADÍSTICA EN LAS NOTICIAS

Los asientos de avión más seguros

Muchos de nosotros creemos que, en un choque aéreo, los asientos de atrás son los más seguros. Los expertos en seguridad no están de acuerdo con 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 de la 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 que sufrieron accidentes se dividirían en las secciones frontal, media y trasera. Entonces, las frecuencias que se observan de decesos se compararían 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 que se esperan, a la vez que revelaría si algunas secciones son más seguras que las otras.

Tabla 10-3 Cálculo del estadístico de prueba χ^2 para los últimos dígitos de distancias de jonrones

Último dígito	Frecuencia O que se observa	Frecuencia E que se espera	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
0	47	7.3	39.7	1576.09	215.9027
1	3	7.3	−4.3	18.49	2.5329
2	1	7.3	−6.3	39.69	5.4370
3	0	7.3	−7.3	53.29	7.3000
4	3	7.3	−4.3	18.49	2.5329
5	11	7.3	3.7	13.69	1.8753
6	3	7.3	−4.3	18.49	2.5329
7	3	7.3	−4.3	18.49	2.5329
8	1	7.3	−6.3	39.69	5.4370
9	1	7.3	−6.3	39.69	5.4370
	73	73			

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} = 251.5206$$

(Estos dos totales deben coincidir).

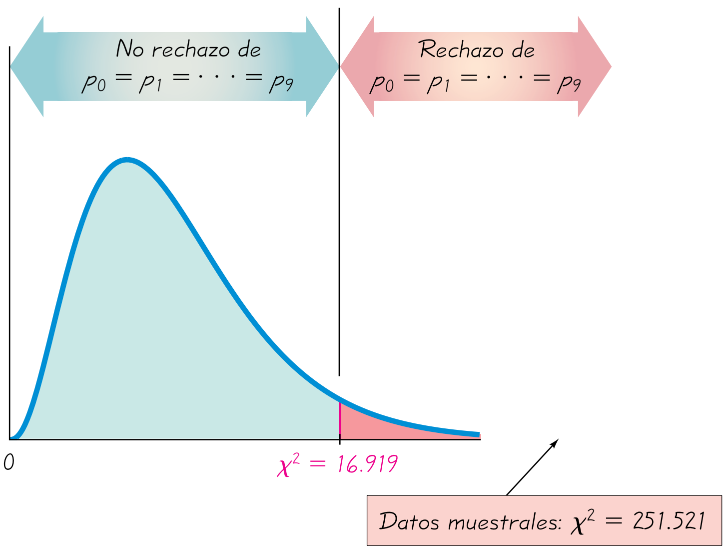


FIGURA 10-4 **Prueba de $p_0 = p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9$**

EJEMPLO Detección de fraude En el problema del capítulo se señaló que algunas veces se utiliza la estadística para detectar fraudes. El segundo renglón de la tabla 10-1 lista porcentajes para dígitos líderes, tal como se esperarían según la ley de Benford, y el tercer renglón lista las frecuencias que se esperan cuando los porcentajes de la ley de Benford se aplican a 784 dígitos líderes. El último renglón de la tabla 10-1 lista las frecuencias que se observan de los dígitos líderes que se esperan de los montos de 784 cheques que expidieron siete compañías diferentes. Pruebe la aseveración de que hay una discrepancia significativa entre los dígitos líderes que se esperaba por la ley de Benford y los dígitos líderes que se observó en los 784 cheques. Utilice un nivel de significancia de 0.01.

SOLUCIÓN En la prueba de la aseveración dada, los pasos 1, 2 y 3 dan como resultado las siguientes hipótesis:

H_0: La distribución de dígitos líderes es la distribución descrita por la ley de Benford. Es decir, $p_1 = 0.301$ 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 que se listan para la ley de Benford en la tabla 10-1).

H_1: Al menos una de las proporciones de arriba es diferente del valor que se asevera.

Los pasos 4, 5 y 6 nos llevan a utilizar la prueba de bondad de ajuste con un nivel de significancia de 0.01 y un estadístico de prueba que se calcula a partir de la tabla 10-4.

continúa

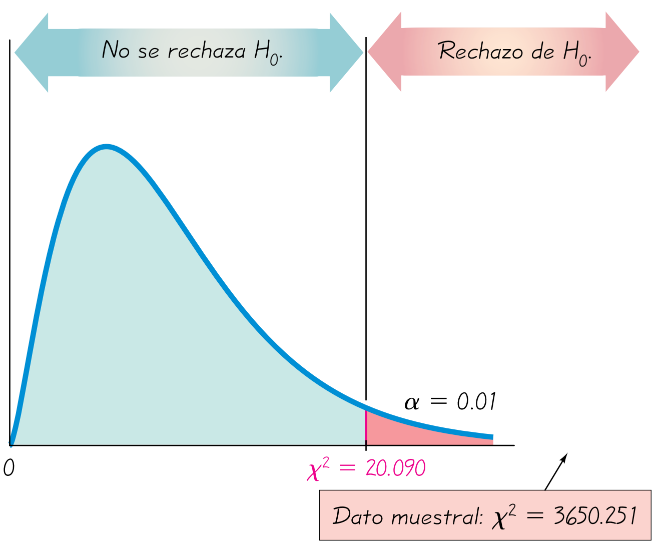
Tabla 10-4 Frecuencias que se observan y frecuencias que se esperan con la ley de Benford

Dígito	Frecuencia que se observa	Frecuencia que se espera	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
1	0	235.984	−235.984	55688.4483	235.9840
2	15	137.984	−122.984	15125.0643	109.6146
3	0	98.000	−98.000	9604.0000	98.0000
4	76	76.048	−0.048	0.0023	0.0000
5	479	61.936	417.064	173942.3801	2808.4213
6	183	52.528	130.472	17022.9428	324.0737
7	8	45.472	−37.472	1404.1508	30.8795
8	23	39.984	−16.984	288.4563	7.2143
9	0	36.064	−36.064	1300.6121	36.0640
				Total: $\chi^2 = \sum \frac{(O - E)^2}{E} =$	3650.2514

El estadístico de prueba es $\chi^2 = 3650.251$. El valor crítico de χ^2 es 20.090, que se encuentra en la tabla A-4 (utilizando $\alpha = 0.01$ en la cola derecha con $k - 1 = 8$ grados de libertad). El estadístico de prueba y el valor crítico se muestran en la figura 10-5. Puesto que el estadístico de prueba cae dentro de la región crítica, hay suficiente evidencia para justificar el rechazo de la hipótesis nula. Existe suficiente evidencia para sustentar la aseveración de que hay una discrepancia significativa entre los dígitos líderes esperados según la ley de Benford y los dígitos líderes que se observaron en los 784 cheques.

En la figura 10-6*a*, graficamos las proporciones que se aseveran de 0.301, 0.176, 0.125, 0.097, 0.079, 0.067, 0.058, 0.051 y 0.046, junto con las proporciones que se observan de 0.000, 0.019, 0.000, 0.097, 0.611, 0.233, 0.010, 0.029 y 0.000, para visualizar la discrepancia entre la distribución de la ley de Benford, que se aseveró y las frecuencias que se observaron. Como se aprecia, se trazó una línea que une los puntos que representan las proporciones que se esperan y otra que corresponde a los puntos de las proporciones que se observan. Los pares de puntos correspondientes están muy separados; esto indica que las frecuencias que se esperan son muy diferentes de las frecuencias correspondientes que se observan. La gran disparidad entre la línea de las frecuencias que se observan y la línea de las frecuencias que se esperan sugiere que los montos de los cheques no son el resultado de transacciones típicas; parece que hay un fraude. De hecho, el Brooklyn District Attorney levantó cargos por fraude utilizando esta línea de razonamiento. Para hacer una comparación, véase la figura 10-6*b*, que se basa en los dígitos líderes de las cantidades de los últimos 200 cheques firmados por el autor. Note cómo las proporciones que se observan de los cheques del autor concuerdan bastante bien con las proporciones que se esperan con la ley de Benford. Los cheques del autor parecen ser típicos en lugar de mostrar un patrón que sugeriría un fraude. En general, las gráficas como la de la figura 10-6 son muy útiles para comparar visualmente las frecuencias esperadas y las frecuencias observadas, tanto como para sugerir cuáles categorías resultan en las discrepancias principales.

FIGURA 10-5 Prueba de concordancia entre frecuencias que se observan y frecuencias que se espera con la ley de Benford



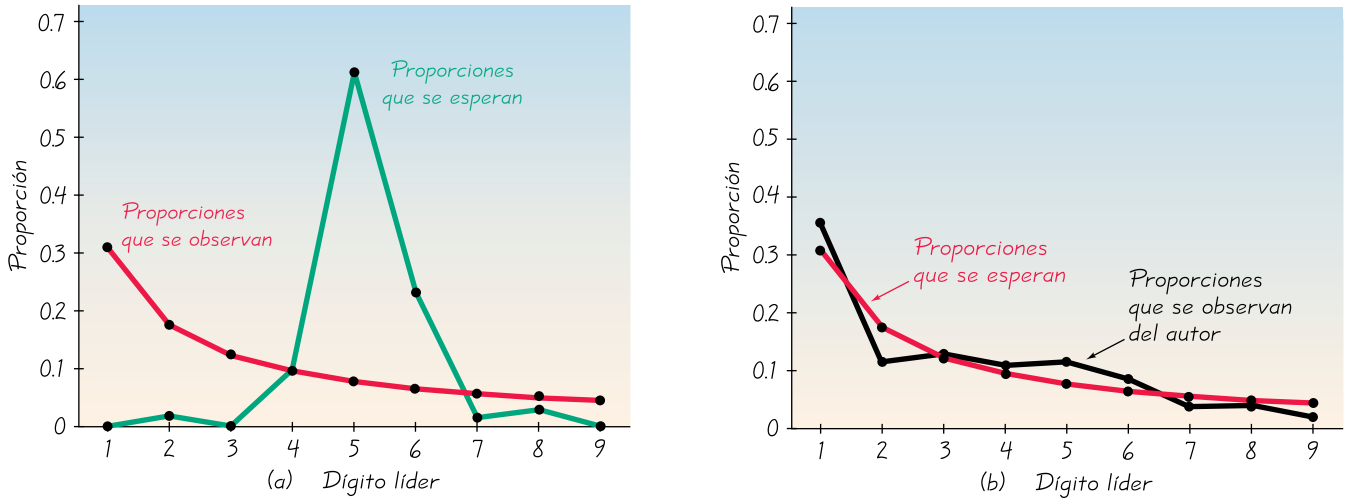


FIGURA 10-6 Comparación de las frecuencias que se observan y las frecuencias que se esperan con la ley de Benford

Fundamentos del estadístico de prueba: Los ejemplos anteriores resultan útiles para tener una idea de la función del estadístico de prueba χ^2. Es claro que queremos medir la cantidad de discordancia entre las frecuencias observadas y esperadas. Sumar simplemente las diferencias entre los valores que se observan y se esperan no resulta una medida eficaz, puesto que esa suma siempre es 0, como se indica abajo.

$$\Sigma(O - E) = \Sigma O - \Sigma E = n - n = 0$$

Se obtiene un mejor estadístico al elevar al cuadrado los valores $O - E$, lo que refleja las diferencias entre las frecuencias que se observan y las que se esperan. (Las razones para elevar al cuadrado los valores $O - E$ son esencialmente las mismas que aquellas para elevar al cuadrado los valores $x - \bar{x}$ en la fórmula de la desviación estándar). El valor de $\Sigma(O - E)^2$ sólo mide la magnitud de las diferencias, pero necesitamos calcular la magnitud de las diferencias en relación con lo que se esperaba. Dicha magnitud relativa se calcula a través de la división entre las frecuencias esperadas, como en el estadístico de prueba.

La distribución teórica de $\Sigma(O - E)^2/E$ es una distribución discreta, puesto que el número de valores posibles se limita a un número finito. La distribución puede aproximarse por una distribución chi cuadrada, que es continua. Tal aproximación por lo regular se considera aceptable, siempre y cuando todos los valores E que se esperan sean al menos 5. Incluimos este requisito con los supuestos que se aplican en esta sección. En la sección 5-6 vimos que la distribución de probabilidad normal continua puede aproximarse razonablemente a la distribución de probabilidad binomial discreta, siempre y cuando np y nq sean ambas de al menos 5. Ahora vemos que la distribución continua chi cuadrada puede aproximar razonablemente la distribución discreta de $\Sigma(O - E)^2/E$, siempre y cuando todos los valores de E sean de al menos 5. (Hay formas para evitar el problema de una frecuencia que se espera menor que 5, como combinar categorías de manera que todas las frecuencias que se esperan sean de al menos 5).

El número de grados de libertad refleja el hecho de que es posible asignar libremente frecuencias a $k - 1$ categorías, antes de que se determine la frecuencia para cada categoría. (Aun cuando decimos que asignamos “con libertad” frecuencias a $k - 1$ categorías, no podemos tener frecuencias negativas, ni frecuencias tan grandes que su suma exceda el total de las frecuencias que se observan de todas las categorías combinadas).

Valores *P*

Los ejemplos de esta sección utilizaron el método tradicional de prueba de hipótesis, pero también es posible utilizar el método del valor *P*. Los valores *P* se obtienen automáticamente con el STATDISK o la calculadora TI-83 Plus, o bien, con los métodos descritos en el capítulo 7. Así, el ejemplo anterior 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, veremos 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 el STATDISK, la pantalla 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 lleva 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.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego la opción **Multinomial Experiments**. Escoja entre “equal expected frequencies” y “unequal expected frequencies”, e ingrese los datos en el cuadro de diálogo. Si elige “unequal expected frequencies”, ingrese los valores esperados en la segunda columna como “conteos” (con las frecuencias reales que se esperan) o como “proporciones” (ingresando las *probabilidades*).

TI-83 Plus Los métodos de esta sección no están disponibles como procedimiento directo en la calculadora TI-83 Plus, pero se puede utilizar un sencillo truco (gracias a Rich Stephens, de la Universidad de Alaska). Primero identifique las frecuencias que se observan y que se esperan, entonces ingréselas como una *matriz*. Oprima **2nd** $\mathbf{x^{-1}}$ para obtener el menú **MATRIX** (o la tecla **MATRIX** de la TI-83). Seleccione **EDIT** y oprima **ENTER**. Ingrese las dimensiones de la matriz (dos renglones por el número de columnas) y proceda a ingresar las frecuencias que se observan en el renglón superior. Para el renglón inferior, introduzca las frecuencias que se esperan y se multiplicaron por un número sumamente grande; por ejemplo, 10^{30}. (Para frecuencias que se esperan de 25, 15 y 50, ingrese 25E30, 15E30 y 50E30). Cuando termine, oprima **STAT**, seleccione **TESTS**, luego la opción $\boldsymbol{\chi^2}$**-Test**. Asegúrese de que la matriz que se observa sea la que ingresa como la matriz A. Lleve el cursor hacia abajo hasta **Calculate** y oprima **ENTER** para obtener el estadístico de prueba, el valor *P* y el número de grados de libertad.

Minitab **Excel** Los métodos de esta sección no están disponibles como procedimientos preestablecidos.

10-2 Destrezas y conceptos básicos

1. Prueba para categorías igualmente probables Las siguientes son las frecuencias que se observaron de cuatro categorías: 5, 6, 8 y 13. Suponga que queremos utilizar un nivel de significancia de 0.05 para probar la aseveración de que las cuatro categorías son todas igualmente probables.
a. ¿Cuál es la hipótesis nula?
b. ¿Cuál es la frecuencia que se espera para cada una de las cuatro categorías?
c. ¿Cual es el valor del estadístico de prueba?
d. ¿Cuál es el valor crítico?
e. ¿Qué concluye acerca de la aseveración que se plantea?

2. Prueba para categorías con proporciones diferentes Las siguientes son las frecuencias que se observan para cinco categorías: 9, 8, 13, 14 y 6. Suponga que queremos utilizar un nivel de significancia de 0.05 para probar la aseveración de que las cinco categorías tienen proporciones de 0.2, 0.2, 0.2, 0.3 y 0.1, respectivamente.
a. ¿Cuál es la hipótesis nula?
b. ¿Cuál es la frecuencia que se espera para cada una de las cinco categorías?
c. ¿Cual es el valor del estadístico de prueba?
d. ¿Cuál es el valor crítico?
e. ¿Qué concluye acerca de la aseveración que se plantea?

3. Prueba de balance de rueda de ruleta El autor observó 500 giros de una rueda de ruleta en el Mirage Resort and Casino. (Para la 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 las 38 ranuras diferentes que se supone son igualmente probables. Cuando se utilizó el STATDISK para probar la aseveración de que las ranuras son de hecho igualmente probables, se obtuvo el estadístico de prueba $\chi^2 = 38.232$.
a. Calcule el valor crítico suponiendo que el nivel de significancia es 0.10.
b. El 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 las 38 ranuras?
c. Escriba una conclusión acerca de la aseveración de que los 38 resultados son igualmente probables.

4. Prueba de una máquina tragamonedas El autor compró una máquina tragamonedas (Bally modelo 809), que probó jugando 1197 veces. En la prueba de la aseveración de que los resultados que se observaron concuerdan con las frecuencias que se esperan, se obtuvo el estadístico de prueba de $\chi^2 = 8.185$. Hay 10 categorías de resultados diferentes, incluyendo no ganar, ganar el premio mayor, ganar con tres campanas, etcétera.
a. Calcule el valor crítico suponiendo que el nivel de significancia es de 0.05.
b. ¿Qué concluye acerca del valor *P* de la tabla A-4, si sabe que el estadístico de prueba es $\chi^2 = 8.185$ y que hay 10 categorías?
c. Establezca una conclusión acerca de la aseveración de que los resultados que se observan concuerdan con las frecuencias que se esperan. ¿Parece que la máquina tragamonedas del autor funciona correctamente?

5. Dado cargado El autor taladró un hoyo en un dado y lo rellenó con plomo, luego procedió a lanzarlo 200 veces. Las siguientes son las frecuencias que se observaron 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 a un dado balanceado?

6. El neumático desinflado y la clase perdida Un cuento clásico incluye a cuatro estudiantes que comparten un automóvil. Ellos perdieron un examen y dieron como excusa un neumático desinflado. En el examen de recuperación, el instructor pidió a los estudiantes que identificaran el neumático en particular que se desinfló. Si ellos en realidad no tuvieron un neumático desinflado, ¿serían capaces de identificar el mismo neumático? El autor pidió a otros 41 estudiantes que identificaran la llanta que ellos seleccionarían. Los resultados se listan 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 ellos realmente no tuvieron un neumático desinflado?

Neumático	Frontal izquierdo	Frontal derecho	Trasero izquierdo	Trasero derecho
Número que se seleccionó	11	15	8	6

7. ¿Los choques de automóviles ocurren con la misma frecuencia en diferentes días? Es una creencia común que los choques fatales de automóviles ocurren más en ciertos días de la semana, como viernes o sábado. Se selecciona aleatoriamente una muestra de muertes en vehículos de motor en Montana en un año reciente. El número de decesos para los diferentes días de la semana se lista en la tabla adjunta. Con un nivel de significancia de 0.05, pruebe la aseveración de que los accidentes ocurren con igual frecuencia en los diferentes días.

Día	Dom	Lun	Mar	Mié	Jue	Vie	Sáb
Número de muertes	31	20	20	22	22	29	36

Datos del Insurance Institute for Highway Safety.

8. ¿Las muertes CIS son resultado de beber el fin de semana? Muchas personas creen que los choques fatales CIS (en los que intervienen conductores bajo la influencia de sustancias tóxicas) ocurren a causa de los bebedores casuales que tienden a emborracharse las noches de viernes y sábado, mientras que otros creen que los choques fatales CIS los causan personas que beben todos los días de la semana. En un estudio de choques automovilísticos fatales se seleccionaron aleatoriamente 216 casos del grupo donde se encontró que el conductor tenía un contenido de alcohol en la sangre que estaba por encima de 0.10. Tales casos se separaron de acuerdo con el día de la semana, con los resultados que se listan en la tabla adjunta. Con un nivel de significancia de 0.05, pruebe la aseveración de que esta clase de choques fatales ocurren en los diferentes días de la semana con igual frecuencia. ¿Sustenta la evidencia la teoría de que los choques fatales CIS se deben a bebedores casuales o la teoría de que los causan quienes beben diariamente?

Día	Dom	Lun	Mar	Mié	Jue	Vie	Sáb
Número	40	24	25	28	29	32	38

Datos del Dutchess County STOP-DWI Program.

9. Prueba para accidentes industriales que se distribuyen de manera uniforme Se realizó un estudio de 147 accidentes industriales que requirieron atención médica. De tales accidentes, 31 ocurrieron en lunes, 42 en martes, 18 en miércoles, 25 en jueves y 31 en viernes (según resultados de "Counted Data CUSUM's", de Lucas, *Technometrics*, vol. 27, núm. 2). Pruebe la aseveración de que los accidentes ocurren con proporciones iguales en los cinco días de trabajo. Si las proporciones no son las mismas, ¿qué factores explicarían las diferencias?

10. Calificación y lugar para sentarse ¿Tienden los estudiantes con calificación "A" a sentarse en una parte 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, nueve se sentaron en medio y cinco se sentaron atrás. ¿Hay suficiente evidencia para sustentar la aseveración de que los estudiantes de calificación "A" no se distribuyen en un patrón regular 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?

11. Posición asignada y carreras de caballos ganadores Muchas personas creen que, cuando corre un caballo, tiene mayores oportunidades de ganar si su posición en la línea de arranque es más cercana al carril interior de la pista. La posición de arranque de 1 es la más cercana al carril interior, a la que le sigue la posición 2, etcétera. La tabla adjunta lista el número de triunfos de caballos en las diferentes posiciones de arranque. Pruebe la aseveración de que las probabilidades de ganar en las diferentes posiciones que se asignaron no son todas las mismas.

Posición de arranque	1	2	3	4	5	6	7	8
Número de triunfos	29	19	18	25	17	10	15	11

Datos del *New York Post*.

12. 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 distribuirse uniformemente, pero si son estimados o reportados, los últimos dígitos tienden a tener, con mucha desproporción, más ceros o cincos. Remítase al conjunto de datos 1 en el 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 and Examination Survey. Pruebe la aseveración de que los últimos dígitos de 0, 1, 2, . . . , 9 ocurren con la misma frecuencia. Con base en los dígitos que se observan, ¿qué se infiere acerca del procedimiento que se utilizó para obtener los pulsos?

13. ¿Los cuatro números ganadores son aleatorios? Remítase al conjunto de datos 26 en el Apéndice B y tenga en cuenta los 160 dígitos que se seleccionaron en el juego de lotería Win 4 del estado de Nueva York. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que esos dígitos se seleccionan de manera tal que los 10 dígitos posibles son todos igualmente probables. ¿Cambia la conclusión si se utiliza un nivel de significancia de 0.01 en lugar de 0.05? ¿Cuál sería una implicación de la conclusión de que los dígitos no son igualmente probables?

14. ¿Los crímenes violentos se distribuyen uniformemente? Con base en datos del Federal Bureau of Investigation (FBI), los crímenes violentos en un año reciente ocurrieron con la distribución que se presenta en la tabla adjunta. Los crímenes violentos incluyen el asesinato, el homicidio sin premeditación, la violación, el robo y el asalto con agravantes. Los porcentajes que se listan se basan en un total de 1,424,287 casos de crímenes violentos. Maneje un nivel de significancia de 0.01 para probar la aseveración de que los crímenes violentos se distribuyen por igual entre los 12 meses. ¿Cómo explica la conclusión, considerando que los porcentajes que se listan no parecen ser muy diferentes? ¿Hay una explicación razonable de por qué los crímenes violentos no pueden estar distribuidos por igual en los 12 meses?

Mes	Ene.	Feb.	Marzo	Abril	Mayo	Junio	Julio	Ago.	Sept.	Oct.	Nov.	Dic.
Porcentaje	7.7	7.4	8.4	8.3	9.2	8.6	9.0	8. 9	8.6	8.7	7.6	7.7

15. Dulces M&M Mars, Inc. asevera que sus dulces M&M clásicos se distribuyen con los siguientes porcentajes de color: 30% café, 20% amarillo, 20% rojo, 10% anaranjado, 10% verde y 10% azul. Remítase al conjunto de datos 19 del Apéndice B y utilice los datos muestrales para probar la aseveración de que la distribución de color es como asevera Mars, Inc. Utilice un nivel de significancia de 0.05.

16. Choques de automóvil y rangos de edad Entre conductores que tuvieron un choque de automóvil en el último año, se selecciona aleatoriamente a 88 y se ordenan por categorías de edad, con los resultados que se listan en la tabla adjunta. Si todas las edades tienen la misma tasa de choques, esperaríamos (por la distribución de la edad de los conductores con licencia) que las categorías dadas incluyan el 16%, 44%, 27% y 13% de los sujetos, en ese orden. Con un nivel de significancia de 0.05, pruebe la aseveración de que la distribución de choques concuerda con la distribución de edades. ¿Parece que algún grupo de edades sufre un número desproporcionado de choques?

Edad	Menores de 25	25–44	45–64	Mayores de 64
Conductores	36	21	12	19

Datos del Insurance Information Institute.

17. Distribución de dígitos en el número irracional Pi El número π es un número irracional con la propiedad de que, cuando tratamos de expresarlo en una forma decimal, requiere un número infinito de lugares decimales y no hay un patrón de repetición. En la representación decimal de π, los primeros 100 dígitos ocurren con las frecuencias que se describen en la tabla adjunta. Con un nivel de significancia de 0.05, pruebe la aseveración de que los dígitos se distribuyen de manera uniforme.

Dígito	0	1	2	3	4	5	6	7	8	9
Frecuencia	8	8	12	11	10	8	9	8	12	14

18. Distribución de dígitos en el número racional $22/7$ El número $22/7$ es similar a π en el sentido de que ambos requieren un número infinito de lugares decimales. Sin embargo, $22/7$ es un número racional porque es posible expresarlo como la proporción de dos enteros, mientras que con π esto no es así. Cuando los números racionales como $22/7$ se expresan en forma decimal, hay un patrón de repetición. En la representación decimal de $22/7$, los primeros 100 dígitos ocurren con las frecuencias descritas en la tabla adjunta. Con un nivel de significancia de 0.05, pruebe la aseveración de que los dígitos se distribuyen de manera uniforme. ¿Cómo difiere el resultado de aquel que se calculó en el ejercicio 17?

Dígito	0	1	2	3	4	5	6	7	8	9
Frecuencia	0	17	17	1	17	16	0	16	16	0

19. Montos de cheques del autor y la ley de Benford La figura 10-6*b* ilustra las frecuencias que se observan de los dígitos líderes de las cantidades de los últimos 200 cheques que firmó el autor. Las frecuencias que se observan de estos dígitos líderes se listan abajo. Utilizando un nivel de significancia de 0.05, pruebe la aseveración de que provienen de una población de dígitos líderes que cumplen con la ley de Benford. (Véanse los primeros dos renglones de la tabla 10-1, que se incluyen 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

20. ¿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 zumbadoras V-1 en la Segunda Guerra Mundial, el sur de Londres se subdividió en regiones, cada una con un área de 0.25 km^2. En la sección 4-5 presentamos un ejemplo e incluimos una tabla de frecuencias reales de impactos, así como las frecuencias que se esperan 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. Además, 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 que se espera de regiones (de la distribución de Poisson)	227.5	211.4	97.9	30.5	8.7

10-2 Más allá de lo básico

21. Prueba de efectos de datos distantes Al realizar una prueba para la bondad de ajuste, como se describe en esta sección, ¿tiene un dato distante un gran efecto sobre el valor del estadístico de prueba χ^2? Pruebe el efecto de un dato distante repitiendo el ejercicio 6, después de cambiar las frecuencias para el neumático trasero derecho de 6 a 60. Describa el efecto general de un dato distante.

22. Detección de datos experimentales alterados Cuando Gregorio Mendel realizó sus famosos experimentos de hibridación con chícharos, aparentemente su asistente de jardinería conocía los resultados que Mendel esperaba, por lo cual alteró los resultados para ajustarlos a las expectativas de Mendel. Un análisis subsiguiente de los resultados condujo a la conclusión de que hay una probabilidad de sólo 0.00004 de que los resultados esperados y los resultados reportados coincidieran tanto. ¿Cómo utilizar los métodos de la presente sección para detectar resultados como éste, que son muy perfectos como para ser realistas?

23. 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 7-3). Suponga que un experimento multinomial en particular tiene sólo dos resultados posibles, A y B, con frecuencias que se observan de f_1 y f_2, respectivamente.

a. Calcule una expresión para el estadístico de prueba χ^2 y 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, $(f_1 + f_2)/2$.

b. Se utiliza el estadístico de prueba $z = (\hat{p} - p)/\sqrt{pq/n}$ 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*.

24. Prueba de bondad de ajuste con una distribución binomial La distribución de una frecuencia que se observa es como sigue:

Número de éxitos	0	1	2	3
Frecuencia	89	133	52	26

a. 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

b. Utilizando las probabilidades que se calcularon en el inciso *a*, determine la frecuencia que se espera para cada categoría.

c. Utilice un nivel de significancia de 0.05 para probar la aseveración de que las frecuencias que se observan ajustan con una distribución binomial para la que $n = 3$ y $p = 1/3$.

25. Prueba de bondad de ajuste con una distribución normal La distribución de frecuencias que se observan de una muestra de puntuaciones de CI es como sigue:

Puntuación de CI	Menor que 80	80–95	96–110	111–120	Mayor que 120
Frecuencia	20	20	80	40	40

a. Suponga una distribución normal con $\mu = 100$ y $\sigma = 15$ y utilice los métodos que se describen en el capítulo 5 para calcular la probabilidad de que un sujeto seleccionado aleatoriamente pertenezca a cada clase. (Utilice fronteras de clase de 79.5, 95.5, 110.5 y 120.5).

b. Utilice las probabilidades calculadas del inciso *a* y calcule la frecuencia que se espera para cada categoría.

c. Utilice un nivel de significancia de 0.01 para probar la aseveración de que las puntuaciones de CI fueron seleccionadas aleatoriamente de una población que se distribuye normalmente con $\mu = 100$ y $\sigma = 15$.

10-3 Tablas de contingencia: independencia y homogeneidad

En la sección 10-2 consideramos datos categóricos que se resumen con conteos de frecuencias que se listan en un solo renglón o una columna. Puesto que las celdas de un solo renglón o una columna corresponden a las categorías de una sola variable (como sería el color), las tablas en la sección 10-2 suelen llamarse *tablas de frecuencias de un factor*. En esta sección consideraremos otra vez datos categóricos que se resumen con conteos de frecuencias, pero las celdas corresponden a dos variables diferentes. Las tablas que consideramos en la sección se llaman *tablas de contingencia* o *tablas de frecuencias de dos factores*.

Definiciones

Una **tabla de contingencia** (o **tabla de frecuencias de dos factores**) es una tabla en donde las frecuencias corresponden a dos variables. (Una variable se utiliza para categorizar renglones y una segunda variable para categorizar columnas).

La tabla 10-5, que resume el destino de los pasajeros y la tripulación cuando se hundió el *Titanic* el lunes 15 de abril de 1912, tiene dos variables: una variable de renglón, que indica si la persona sobrevivió o murió; y una variable de columna, que lista las categorías demográficas —hombres, mujeres, niños y niñas.

Las tablas de contingencia tienen especial importancia, puesto que suelen utilizarse para analizar resultados de encuesta. Por ejemplo, podemos hacer una pregunta a los sujetos en la que identifiquen su género (masculino/femenino), y otra en la que describan la frecuencia de su uso del control remoto del televisor (frecuentemente/algunas veces/nunca). Entonces los métodos de esta sección resultan útiles para determinar si el uso del control remoto del televisor es independiente del género. (Probablemente ya sabemos la respuesta a esto). Las aplicaciones de este tipo son muy numerosas, así que los métodos que se presentan en esta sección se utilizan con mucha frecuencia.

Esta sección presenta dos tipos de prueba de hipótesis que se basan en tablas de contingencia. Primero consideramos las pruebas de independencia, que se usan 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 sirven para determinar si situaciones diferentes tienen las mismas proporciones de alguna característica. Buenas noticias: ambos tipos de prueba de hipótesis utilizan los mismos métodos básicos. Comenzamos con las pruebas de independencia.

Tabla 10-5 Mortalidad del *Titanic*

	Hombres	Mujeres	Niños	Niñas	**Total**
Sobrevivientes	332	318	29	27	**706**
Muertos	1360	104	35	18	**1517**
Total	**1692**	**422**	**64**	**45**	**2223**

Prueba de independencia

Una de las dos pruebas que se incluyen en esta sección es la *prueba de independencia* entre la variable de renglón y la variable de columna.

Definición

Prueba de independencia: prueba la hipótesis nula de que no hay 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 esto sólo es una dependencia estadística, por lo cual no es posible utilizarla para establecer una cadena directa de causa-efecto entre las dos variables en cuestión. Por ejemplo, después de analizar los datos de la tabla 10-5, concluiríamos que si una persona sobrevivió al hundimiento del *Titanic*, depende de si era un hombre, una mujer, un niño o una niña, lo cual no significa que la categoría género/edad tenga algún efecto causal directo para sobrevivir.

Cuando se prueba la hipótesis nula de independencia entre las variables de renglón y de columna, en una tabla de contingencia, los supuestos, el estadístico de prueba y los valores críticos son como se describe en el siguiente cuadro.

Supuestos

1. Los datos muestrales se seleccionan aleatoriamente.
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 E que se espera es de al menos 5. (No existe el requisito de que cada frecuencia *observada* deba ser de al menos 5. Además, no existe el requisito de que la población deba tener una distribución normal o cualquiera 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

   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*.

Muerte pospuesta

El sociólogo David Phillips, de la Universidad de California, estudia la capacidad de las personas para posponer su muerte hasta después de algún acontecimiento importante. Al analizar las tasas de decesos de hombres judíos que murieron cerca de la Pascua, encontró que la tasa de decesos disminuía drásticamente la semana anterior a la Pascua, pero se elevaba la semana posterior. Él encontró un fenómeno similar entre mujeres chino-estadounidenses; su tasa de decesos disminuía la semana anterior a su importante *Harvest Moon Festival* y se elevaba la semana posterior.

El estadístico de prueba nos permite medir el grado de discordancia entre las frecuencias que se observan en la realidad y aquellas que se esperarían teóricamente cuando las dos variables son independientes. Los valores pequeños del estadístico de prueba χ^2 resultan de una gran concordancia entre las frecuencias que se observan y las frecuencias que se esperan, con variables de renglón y de columna independientes. Los valores grandes del estadístico de prueba χ^2 están en la región de la extrema derecha de la distribución chi cuadrada; reflejan diferencias significativas entre las frecuencias que se observan y las que se esperan. En muestreos grandes repetidos, la distribución del estadístico de prueba χ^2 puede aproximarse por la distribución chi cuadrada, tomando en cuenta que todas las frecuencias esperadas sean de al menos 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 que se observan para ese renglón (o columna)].

En la sección anterior vimos las probabilidades correspondientes y logramos determinar con facilidad los valores esperados; sin embargo, la tabla de contingencia típica no incluye las probabilidades relevantes. Para cada celda en la tabla de frecuencia, la frecuencia que se espera E se calcula aplicando la regla de la multiplicación de probabilidad para sucesos independientes. Suponiendo que las variables de renglón y de columna son independientes (lo que se asume en la hipótesis nula), la probabilidad de que un valor esté en una celda en particular es la probabilidad de que esté en el renglón que contiene la celda (a saber, el total del renglón que se divide entre la suma de todas las frecuencias), multiplicado por la probabilidad de estar en la columna que contiene la celda (a saber, el total de la columna que se dividió entre la suma de todas las frecuencias) multiplicado por la suma de todas las frecuencias. ¿Parece muy complicado? La frecuencia que se espera para una celda queda simplificada en la siguiente ecuación.

Frecuencia esperada para una tabla de contingencia

$$\text{frecuencia esperada} = \frac{(\text{total de renglón})(\text{total de columna})}{(\text{gran total})}$$

El *gran total* se refiere al total de todas las frecuencias que se observan en la tabla. Por ejemplo, la frecuencia que se espera para la celda superior izquierda de la tabla 10-6 (un duplicado de la tabla 10-5 con las frecuencias que se esperan y se incluyen entre paréntesis) es de 537.360, que se calcula observando que el total de todas las frecuencias para el primer renglón es de 706, el total de las frecuencias de la columna es 1692, y la suma de todas las frecuencias en la tabla es 2223. Así obtenemos una frecuencia esperada de

$$E = \frac{(\text{total de renglón})(\text{total de columna})}{(\text{gran total})} = \frac{(706)(1692)}{2223} = 537.360$$

Tabla 10-6 Frecuencias que se observan (y frecuencias que se esperan)

	Género/Categoría de edad				
	Hombres	Mujeres	Niños	Niñas	**Totales de renglón**
Sobrevivientes	332 (537.360)	318 (134.022)	29 (20.326)	27 (14.291)	**706**
Muertos	1360 (1154.640)	104 (287.978)	35 (43.674)	18 (30.709)	**1517**
Total de columnas:	**1692**	**422**	**64**	**45**	**Gran total: 2223**

EJEMPLO Cálculo de la frecuencia que se espera La frecuencia que se espera para la celda superior izquierda de la tabla 10-6 es de 537.360. Calcule la frecuencia que se espera para la celda inferior izquierda, suponiendo independencia entre las variables de renglón (si la persona sobrevivió) y las variables de columna (si la persona es un hombre, una mujer, un niño o una niña).

SOLUCIÓN La celda inferior izquierda está en el segundo renglón (con un total de 1517) y en la primera columna (con un total de 1692). La frecuencia esperada es

$$E = \frac{(\text{total de renglón})(\text{total de columna})}{(\text{gran total})} = \frac{(1517)(1692)}{2223} = 1154.640$$

INTERPRETACIÓN Al interpretar este resultado para la celda inferior izquierda, afirmamos que, aunque realmente murieron 1360 hombres, podríamos esperar que murieran 1154.640 hombres si la supervivencia es independiente de que la persona sea un hombre, una mujer, un niño o una niña. Hay una discrepancia entre $O = 1360$ y $E = 1154.640$; este tipo de discrepancias son componentes clave del estadístico de prueba.

Para entender mejor los fundamentos del cálculo de frecuencias que se esperan con este procedimiento, pretendamos conocer sólo los totales del renglón y de la columna, así como que debemos llenar la celda de las frecuencias que se esperan suponiendo independencia (o no relación) entre las dos variables que se implican; esto es, vamos a pretender que sólo conocemos los totales del renglón y de la columna que se muestra en la tabla 10-6. Comencemos con la celda en la esquina superior izquierda. Puesto que sobrevivieron 706 de las 2223 personas, tenemos $P(\text{sobrevivientes}) = 706/2223$. De manera similar, 1692 de esas personas eran hombres, entonces $P(\text{hombre}) = 1692/2223$. Puesto que suponemos independencia entre la supervivencia y la columna de categoría de género/edad, utilizamos la regla de la multiplicación de la probabilidad para obtener

$$P(\text{sobreviviente y hombre}) = P(\text{sobreviviente}) \cdot P(\text{hombre}) = \frac{706}{2223} \cdot \frac{1692}{2223}$$

Ventaja del equipo local

En el artículo de la revista *Chance* "Predicting Professional Sports Game Outcomes from Intermediate Game Scores", 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 basquetbol. Con 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 basquetbol que van al frente después de tres cuartos ganan alrededor de cuatro de cada cinco ocasiones, pero los equipos de beisbol que van ganando después de siete entradas ganan alrededor de 19 de cada 20 ocasiones. Los métodos de análisis estadístico incluyeron la distribución chi cuadrada que se aplicó a una tabla de contingencia.

Esta ecuación es una aplicación de la regla de la multiplicación para sucesos independientes, que se expresa en general como sigue: $P(A \text{ y } B) = P(A) \cdot P(B)$. Conociendo la probabilidad de estar en la celda superior izquierda, ahora calcularemos el *valor que se espera* para esa celda, el cual se obtiene multiplicando la probabilidad para esa celda por el número total de personas, como se muestra en la ecuación siguiente:

$$E = n \cdot p = 2223\left[\frac{706}{2223} \cdot \frac{1692}{2223}\right] = 537.360$$

La forma de este producto sugiere una forma general para obtener la frecuencia que se espera 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 así

$$E = \frac{(\text{total de renglón}) \cdot (\text{total de columna})}{(\text{gran total})}$$

Ahora, procedamos a utilizar los datos de la tabla de contingencia para probar las hipótesis, como en el siguiente ejemplo.

EJEMPLO Hundimiento del *Titanic* Remítase a los datos de mortalidad del *Titanic* en la tabla 10-5. Trataremos a las 2223 personas a bordo del *Titanic* como una muestra. Podríamos tomar la postura de que los datos del *Titanic* constituyen una *población* y, por lo tanto, no deberían tratarse como una muestra; por consiguiente, no se aplican los métodos de la estadística inferencial. Estipularemos que los datos son datos muestrales que se seleccionaron en forma aleatoria de una población de personas que teóricamente se encontrarían en las mismas condiciones. En la realidad ninguna persona se encontrará realmente en las mismas condiciones, pero supondremos esto para nuestro tema y su análisis. Entonces determinaremos si las diferencias que se observan tienen significancia estadística. (Véase también el programa de cómputo *ActivStats* de Paul Velleman para el ejemplo sobre el *Titanic*).

Utilice un nivel de significancia de 0.05 para probar la aseveración de que, cuando se hundió el *Titanic*, el hecho de sobrevivir o morir era independiente de si la persona era un hombre, una mujer, un niño o una niña.

SOLUCIÓN La hipótesis nula y la hipótesis alternativa son las siguientes:

H_0: El hecho de sobrevivir o morir es independiente de si la persona es un hombre, una mujer, un niño o una niña.

H_1: El hecho de sobrevivir al hundimiento del *Titanic* y ser un hombre, una mujer, un niño o una niña son dependientes.

El nivel de significancia es $\alpha = 0.05$.

Puesto que los datos se presentan en una tabla de contingencia, utilizamos la distribución χ^2 con este estadístico de prueba:

$$
\begin{aligned}
\chi^2 &= \sum \frac{(O - E)^2}{E} \\
&= \frac{(332 - 537.360)^2}{537.360} + \frac{(318 - 134.022)^2}{134.022} + \frac{(29 - 20.326)^2}{20.326} \\
&\quad + \frac{(27 - 14.291)^2}{14.291} + \frac{(1360 - 1154.640)^2}{1154.640} + \frac{(104 - 287.978)^2}{287.978} \\
&\quad + \frac{(35 - 43.674)^2}{43.674} + \frac{(18 - 30.709)^2}{30.709} \\
&= 78.481 + 252.555 + 3.702 + 11.302 \\
&\quad + 36.525 + 117.536 + 1.723 + 5.260 \\
&= 507.084
\end{aligned}
$$

(El estadístico de prueba más exacto de 507.080 se obtiene con más lugares decimales en los cálculos intermedios. El STATDISK, el Minitab, y la calculadora TI-83 Plus coinciden en que 507.080 es un mejor resultado). El valor crítico es $\chi^2 = 7.815$; éste se encuentra en la tabla A-4, observando que $\alpha = 0.05$ en la cola derecha y que el número de grados de libertad se da por $(r - 1)(c - 1) = (2 - 1)(4 - 1) = 3$. El estadístico de prueba y el valor crítico se muestran en la figura 10-7. Puesto que el estadístico de prueba está dentro de la región crítica, rechazamos la hipótesis nula de que el hecho de sobrevivir es independiente de si la persona es un hombre, una mujer, un niño o una niña. Parece que el hecho de sobrevivir al hundimiento del *Titanic* y el de ser un hombre, una mujer, un niño o una niña son variables dependientes.

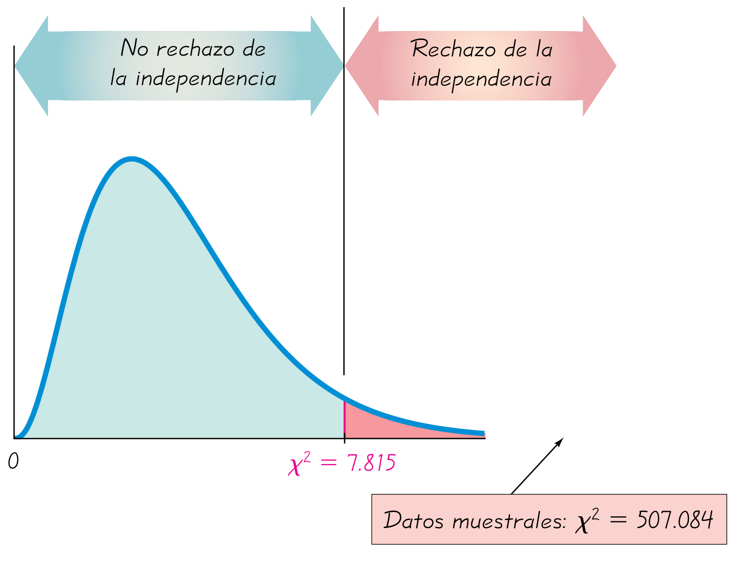


FIGURA 10-7 Prueba de independencia para los datos de mortalidad del *Titanic*

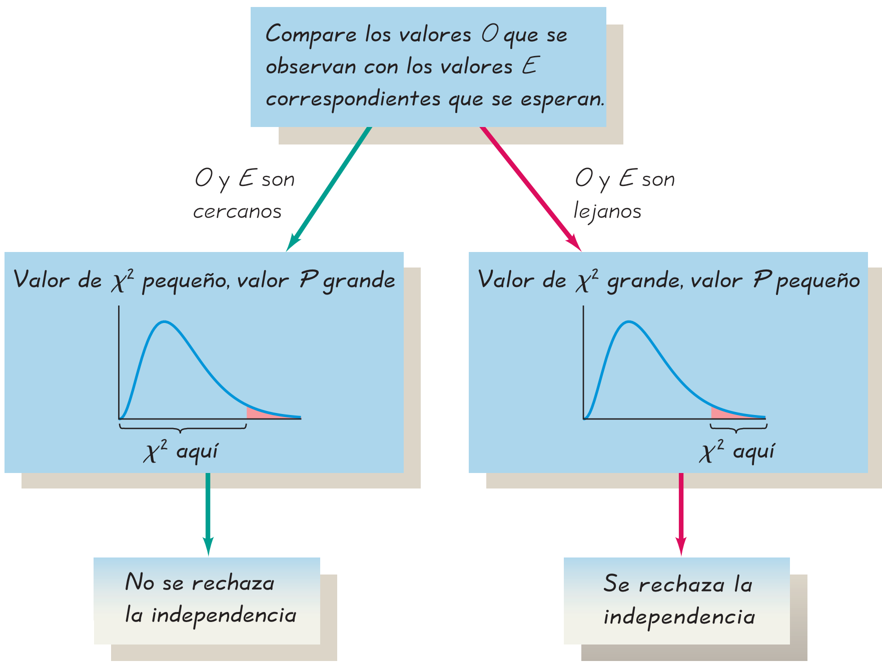
Valores *P*

En el ejemplo anterior se utilizó el método tradicional de prueba de hipótesis, pero es posible manejar con facilidad el método del valor *P*.

STATDISK, Minitab, Excel y la calculadora TI-83 Plus proporcionan valores *P* para pruebas de independencia de tablas de contingencia. Si no tiene una calculadora o un programa de cómputo adecuados, estime los valores *P* con la tabla A-4, en el Apéndice A. Localice el número apropiado de grados de libertad para ubicar un renglón particular de la tabla. Encuentre dónde cae el estadístico de prueba en ese renglón y logrará identificar un rango de valores *P* posibles remitiéndose a las áreas dadas en la parte superior de cada columna. En el ejemplo anterior hay tres grados de libertad, entonces vaya al tercer renglón de la tabla A-4. Ahora utilice el estadístico de prueba $\chi^2 = 507.084$, para ver que el estadístico de prueba es mayor que (y se ubica más a la derecha de) cada valor crítico de χ^2 que se encuentra en el tercer renglón, entonces el valor *P* es menor que 0.005. Con base en este pequeño valor *P*, otra vez rechazamos la hipótesis nula, por lo que concluimos que existe suficiente evidencia muestral para justificar el rechazo de la hipótesis nula de independencia.

Igual que en la sección 10-2, si la frecuencia que se observa y la que se espera son cercanas, el estadístico de prueba χ^2 será pequeño y el valor *P* será grande. Si la frecuencia que se observa y la que se espera se alejan mucho, el estadístico de prueba χ^2 será grande y el valor *P* será pequeño. Dichas relaciones se resumen e ilustran en la figura 10-8.

FIGURA 10-8 Relaciones entre componentes clave en la prueba de independencia



Prueba de homogeneidad

En el ejemplo anterior ilustramos una prueba de independencia utilizando una muestra de 2223 personas que estaban a bordo del *Titanic*. Tratamos a las 2223 personas como una muestra aleatoria obtenida a partir de *una* población hipotética en la que todas las personas se encontraban en circunstancias similares. Sin embargo, algunas otras muestras se obtienen de poblaciones *diferentes* y queremos 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"; 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 realizar una prueba de homogeneidad, podemos utilizar los procedimientos que ya presentamos en esta sección, como se ilustra en el ejemplo siguiente.

EJEMPLO Influencia del género ¿Produce un efecto el género del encuestador en las respuestas de encuesta de hombres? Un artículo del *U.S. News & World Report* acerca de encuestas afirmó: "En temas sensibles, las personas tienden a dar respuestas 'aceptables' en lugar de respuestas honestas; sus respuestas pueden depender del género o la raza del entrevistador". Para sustentar dicha 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 hombres que se encuestaron y la tabla 10-7 se basa en tales datos. Suponga que la encuesta se diseñó de manera que los entrevistadores hombres recibieron instrucciones para obtener 800 respuestas de sujetos hombres, en tanto que las entrevistadoras mujeres recibieron instrucciones para obtener 400 respuestas de sujetos hombres. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las proporciones de las respuestas de acuerdo/desacuerdo son las mismas para los sujetos que entrevistaron hombres y los sujetos que entrevistaron mujeres.

SOLUCIÓN Puesto que tenemos dos poblaciones separadas (sujetos que entrevistaron hombres y sujetos que entrevistaron mujeres), probamos la homogeneidad con estas hipótesis:

H_0: Las proporciones de las respuestas acuerdo/desacuerdo son iguales para los sujetos que entrevistaron hombres y los sujetos que entrevistaron mujeres.

H_1: Las proporciones son diferentes.

Tabla 10-7 Género y respuestas de encuesta

	Género del entrevistador	
	Hombre	Mujer
Hombres que están de acuerdo	560	308
Hombres que están en desacuerdo	240	92

continúa

El nivel de significancia es $\alpha = 0.05$. Utilizamos el mismo estadístico de prueba χ^2 ya descrito, que se calcula por medio del mismo procedimiento. En lugar de hacer una lista de los detalles de este cálculo, presentamos la pantalla del Minitab que resulta de los datos de la tabla 10-7.

Minitab

```
Expected counts are printed below observed counts

                   C1         C2      Total
    1             560        308        868
               578.67     289.33

    2             240         92        332
               221.33     110.67

    Total         800        400       1200

    Chi-Sq = 0.602 + 1.204 + 1.574
            + 3.149 = 6.529

    DF = 1, P-Value = 0.011
```

La pantalla del Minitab muestra las frecuencias que se esperan de 578.67, 289.33, 221.33 y 110.67. Los resultados incluyen también el estadístico de prueba $\chi^2 = 6.529$ y el valor P de 0.011. Con el uso del método del valor P para la prueba de hipótesis, rechazamos la hipótesis nula de proporciones iguales (homogéneas) (puesto que el valor P de 0.011 es menor que 0.05). Hay suficiente evidencia para sustentar el rechazo de la aseveración de que las proporciones son las mismas. Parece que la respuesta y el género del entrevistador son dependientes. Aunque tal análisis estadístico no puede utilizarse para justificar ninguna afirmación acerca de la causalidad, quizás a los hombres los influyó el género del entrevistador.

Usando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **Contingency Tables**, ahora proceda a ingresar las frecuencias como aparecen en la tabla de contingencia. Haga clic en **Evaluate**. El resultado del STATDISK incluye el estadístico de prueba, el valor crítico, el valor P y la conclusión.

Minitab Primero ingrese las frecuencias que se observan en las columnas, luego seleccione **Stat** de la barra del menú principal, después la opción **Tables**, luego **Chi Square Test**. Ahora 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 Plus Primero ingrese la tabla de contingencia como una *matriz* presionando **2nd** $\mathbf{x^{-1}}$ para obtener el menú **MATRIX** (o el botón **MATRIX** del teclado de la TI-83). Seleccione **EDIT** y presione **ENTER**. Ingrese las dimensiones de la matriz (renglones por columnas) y proceda a hacer lo mismo con las frecuencias individuales. Cuando termine, oprima **STAT**, seleccione **TESTS**, luego la opción $\boldsymbol{\chi^2}$**-Test**. Asegúrese de que ingresa la matriz observada como la matriz A. Las frecuencias esperadas se calcularán automáticamente y se guardarán en la matriz que se separó e identificó como "Esperada". 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 Debe ingresar las frecuencias que se observan, así como determinar e ingresar las frecuencias que se esperan. Cuando termine, haga clic en el icono ***fx*** en la barra del menú, seleccione la categoría de función **Statistical**, luego el nombre de la función **CHITEST**. Debe ingresar el rango de valores para las frecuencias observadas y el rango de valores para las frecuencias esperadas. Sólo se da el valor P.

10-3 Destrezas y conceptos básicos

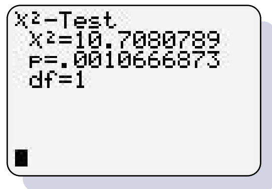
1. ¿Existe discriminación racial? La *discriminación racial* es la práctica controversial de señalar que alguien tiene una conducta criminal con base en su raza, nación de origen o grupo étnico. La tabla adjunta resume resultados de conductores que se seleccionaron al azar y que detuvo la policía en 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 que se detenga a alguien es independiente de la raza y del grupo étnico. Con base en la evidencia disponible, ¿concluiríamos que hay discriminación racial?

	Raza y grupo étnico	
	Negros y no hispanos	Blancos y no hispanos
Detenidos por la policía	24	147
No detenidos por la policía	176	1253

Minitab

```
Chi-Sq = 0.322 + 0.046 +
         0.039 + 0.006 = 0.413
DF = 1, P-Value = 0.521
```

TI-83 Plus



2. Prueba de la eficacia de cascos de ciclista Se realizó un estudio de 531 personas heridas en choques de bicicleta; los resultados de una muestra que se seleccionó al azar se resumen en la tabla adjunta. También se presentan los resultados de la calculadora TI-83 Plus. Con un nivel de significancia de 0.05, pruebe la aseveración de que usar un casco no tiene efecto en el sufrimiento de heridas faciales. Con base en dichos resultados, ¿parece que el casco es eficaz en ayudar a prevenir heridas faciales en un choque?

	Uso de casco	Sin casco
Heridas faciales recibidas	30	182
Todas las heridas no faciales	83	236

Datos tomados de “A Case-Control Study of the Effectiveness of Bicycle Safety Helmets in Preventing Facial Injury”, de Thompson, Thompson, Rivara y Wolf, *American Journal of Public Health*, vol. 80, núm. 12.

3. Correo electrónico y privacidad Se preguntó a trabajadores y a jefes de alto nivel si era poco ético vigilar el correo electrónico de los empleados; los resultados se resumen en la tabla (según 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 trabajador o un jefe de alto nivel. ¿Cambia la conclusión si se emplea un nivel de significancia de 0.01 en lugar de 0.05? ¿Parecen estar de acuerdo los trabajadores y los jefes en este tema?

	Sí	No
Trabajadores	192	244
Jefes	40	81

4. Exactitud de pruebas de polígrafo Los datos en la tabla adjunta resumen resultados de pruebas de exactitud de polígrafos (de acuerdo con datos de la Office of Technology Assessment). Utilice un nivel de significancia de 0.05 para probar la aseveración de que el hecho de que el sujeto mienta 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

5. Prueba de la influencia del género La tabla 10-7 resume datos de sujetos hombres que se encuestaron y la tabla adjunta resume datos de una muestra de mujeres. Utilice un nivel de significancia de 0.01 y suponga que los tamaños muestrales de 800 hombres y 400 mujeres están predeterminados, y pruebe la aseveración de que las proporciones de las respuestas de acuerdo/desacuerdo son las mismas para los sujetos que entrevistaron hombres y los sujetos que entrevistaron mujeres.

	Género del entrevistador	
	Hombre	Mujer
Mujeres que están de acuerdo	512	336
Mujeres en desacuerdo	288	64

Datos del Eagleton Institute.

6. Prueba de discriminación En el caso judicial de *Estados Unidos* versus *la ciudad de Chicago*, se pusieron en tela de juicio prácticas de empleo injustas. Un grupo minoritario (grupo A) y un grupo mayoritario (grupo B) realizaron el examen para ser capitán de bomberos. Suponga que el estudio comenzó con tamaños muestrales que se predeterminaron de 24 candidatos minoritarios (grupo A) y 562 candidatos mayoritarios (grupo B), con los resultados que se muestran en la tabla. Con un nivel de significancia de 0.05, pruebe la aseveración de que la proporción de candidatos minoritarios que aprobaron es la misma que la proporción de candidatos mayoritarios que aprobaron. Con base en los resultados, ¿parece que el examen discrimina?

	Aprobados	Reprobados
Grupo A	10	14
Grupo B	417	145

7. Diferencia de género en el temor a volar El Marist Institute for Public Opinion realizó una encuesta entre 1014 adultos, 48% de los cuales eran hombres. Los resultados de encuesta muestran que el 12% de los hombres y el 33% de las mujeres temen volar. Después de construir una tabla de contingencia que resuma los datos en forma de conteos de frecuencias, utilice un nivel de significancia de 0.05 para probar la aseveración de que el género es independiente del temor a volar.

8. 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 se basan en resultados de los Centers for Disease Control and Prevention. Maneje un nivel de significancia de 0.05 para probar la aseveración de que el éxito es independiente del método que se utilice. Si alguien quiere dejar de fumar, ¿la elección del método provocará una diferencia?

	Goma de mascar de nicotina	Parche de nicotina
Fuma	191	263
No fuma	59	57

9. No fumar Repita el ejercicio 8 después de incluir los datos adicionales que se muestran en la tabla.

	Goma de mascar de nicotina	Parche de nicotina	Inhalador de nicotina
Fuma	191	263	95
No fuma	59	57	27

10. Tabaquismo en China La tabla de abajo resume los resultados de una encuesta que se realizó a hombres de 15 años de edad o mayores que viven en el distrito Minhang en China (datos que se tomaron de "Cigarrette Smoking in China", de Gong, Koplan, Feng, *et al*., *Journal of the American Medical Association*, vol. 274, núm. 15). Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que el hecho de fumar es independiente del nivel educativo. ¿Qué concluye acerca de la relación entre el consumo de tabaco y la educación en China?

	Educación primaria	Educación media	Universidad
Fumador	606	1234	100
Nunca ha fumado	205	505	137

11. 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 se basa en datos del Departamento del Trabajo de Estados Unidos, Bureau of Labor Statistics. ¿Parece que alguna ocupación en particular sea más propensa a los homicidios? Si así fuera, ¿cuál es?

	Policía	Cajero	Taxista	Guardia
Homicidio	82	107	70	59
Otra causa de muerte diferente de homicidio	92	9	29	42

12. ¿Es la precisión del escáner 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 que se etiquetaron. La tabla adjunta resume resultados de una muestra de 819 artículos. Cuando los almacenes manejan escáner para cobrar los artículos, ¿son las tasas de error las mismas para los artículos con precio normal que para los artículos en oferta? ¿Cómo cambiaría la conducta de los consumidores si creen que ocurre desproporcionadamente una cantidad mayor de cobros de más 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

Datos tomados de "UPC Scanner Pricing Systems: Are They Accurate?", de Ronald Goodstein, *Journal of Marketing*, vol. 58.

13. Rechazos de encuestas y rango de edad Un estudio de personas que se rehusaron a responder preguntas de encuesta proporcionó los datos muestrales, que se seleccionaron aleatoriamente, los cuales se muestran en la tabla. Con un nivel de significancia de 0.01, pruebe la aseveración de que la cooperación del sujeto (responder o rehusarse) es independiente de la categoría de edad. ¿Parece que alguno de los grupos de edad es poco cooperativo en particular?

	Edad					
	18–21	22–29	30–39	40–49	50–59	60 o mayores
Respondieron	73	255	245	136	138	202
Se rehusaron	11	20	33	16	27	49

Datos tomados de "I Hear You Knocking But You Can't Come In", de Fitzgerald y Fuller, *Sociological Methods and Research*, vol. 11, núm. 1.

14. Curso de armas de fuego y seguridad ¿Un curso de entrenamiento en armas de fuego da como resultado prácticas más seguras por parte de los propietarios de armas? En un estudio se encuestó a sujetos que se seleccionaron al azar, con los resultados que se muestran en la tabla adjunta. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el entrenamiento formal en armas de fuego es independiente de cómo se guardan las armas. ¿El entrenamiento formal parece ejercer un efecto positivo?

	¿Las armas se guardan cargadas y sin seguro?	
	Sí	No
Recibieron un curso formal en armas de fuego	122	329
No recibieron un curso formal en armas de fuego	49	299

Datos tomados de “Firearm Training and Storage”, de Hemenway, Solnick y Azrael, *Journal of American Medical Association*, vol. 273, núm. 1.

15. El crimen y los extraños La tabla adjunta lista resultados de encuesta que se obtuvieron de una muestra aleatoria de víctimas de diferentes crímenes. 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 afectarían 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 familiar o un conocido	39	106	642

Datos del Departamento de Justicia de Estados Unidos.

16. ¿El uso del cinturón de seguridad depende de la cantidad de cigarrillos que se fuma? Un estudio de usuarios y no usuarios de cinturón de seguridad produjo los datos que se seleccionaron al azar, los cuales se resumen en la siguiente tabla. Pruebe la aseveración de que la cantidad de fumadores es independiente del uso del cinturón de seguridad. Una teoría plausible es que la gente que fuma mucho se preocupa menos por su salud y su seguridad; por lo tanto, se inclina menos a utilizar el cinturón de seguridad. ¿Sustentan los datos muestrales dicha teoría?

	Número de cigarrillos que se fuman por día			
	0	1–14	15–34	35 o más
Utilizan cinturón de seguridad	175	20	42	6
No utilizan cinturón de seguridad	149	17	41	9

Datos tomados de “What Kinds of People Do Not Use Seat Belts?”, de Helsing y Comstock, *American Journal of Public Health*, vol. 67, núm. 11.

17. ¿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, que se seleccionaron al azar, de casos de acusados de robo en San Francisco. A todos los sujetos los sentenciaron a prisión. 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 inocencia. Si usted fuera el abogado defensor de un acusado culpable, ¿sugieren estos resultados que debe fomentar una declaración de culpabilidad?

	Enviados a prisión	No enviados a prisión
Declaración de culpabilidad	392	58
Declaración de inocencia	564	14

Datos tomados 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.

18. ¿La ventaja en los deportes depende de ser un equipo local? Se reunieron datos del equipo ganador para equipos de 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/visitantes son independientes del tipo de deporte. Ya que de los cuatro deportes incluidos, el beisbol es el único en el cual el equipo de casa llega a modificar las dimensiones del campo a favor de sus jugadores, ¿parece que los equipos de beisbol son eficientes al utilizar dicha ventaja?

	Basquetbol	Beisbol	Jockey	Futbol
Triunfos del equipo de casa	127	53	50	57
Triunfos del equipo visitante	71	47	43	42

Datos tomados de "Predicting Professional Sports Game Outcomes from Intermediate Game Scores", de Cooper, DeNeve y Mosteller, *Chance*, vol. 5, núm. 3-4.

19. Pruebas clínicas de Lipitor El fármaco Lipitor, reductor del colesterol, contiene atorvastatin de calcio. En la tabla se incluyen resultados de ensayos clínicos que consideran los dolores de cabeza como una reacción adversa (según datos de Parke-Davis). Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que un dolor de cabeza es independiente de la cantidad de atorvastatin que se administre como tratamiento. (*Sugerencia:* Puesto que no todos los valores que se esperan son de 5 o mayores, combine los resultados de los tratamientos que consisten en dosis de 20 y de 40 miligramos de atorvastatin).

	Placebo	10 mg de atorvastatin	20 mg de atorvastatin	40 mg de atorvastatin	80 mg de atorvastatin
Dolor de cabeza	19	47	6	2	6
Sin dolor de cabeza	251	816	30	77	88

20. Ejercicio y tabaquismo Un estudio de los efectos del ejercicio en mujeres incluyó los resultados que se resumen en la tabla (de acuerdo con datos de "Physical Activity and Coronary Hearth Disease in Women", de Lee, Rexrode, Cook, Manson y Buring, *Journal of the American Medical Association*, vol. 285, núm. 11). Los valores del ejercicio están en kilocalorías de actividad física por semana. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el nivel de tabaquismo es independiente del nivel de ejercicio.

	Por debajo de	200–599	600–1499	1500 o más
Nunca han fumado	4997	5205	5784	4155
Fuman menos de 15 cigarrillos al día	604	484	447	359
Fuman 15 o más cigarrillos al día	1403	830	644	350

10-3 Más allá de lo básico

21. 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 que se espera 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

$$\sum \frac{(O-E)^2}{E} \quad \text{con} \quad \sum \frac{(|O-E|-0.5)^2}{E}$$

Dada la tabla de contingencia del ejercicio 1, encuentre el valor del estadístico de prueba χ^2 con y sin la corrección de Yates. ¿Qué efecto tiene la corrección de Yates?

22. 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, en tanto que las frecuencias de c y d están en el segundo renglón.

a. Verifique que el estadístico de prueba se exprese 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 que $\hat{p}_2 = b/(b + d)$. Demuestre que el estadístico de prueba

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\dfrac{\bar{p}\,\bar{q}}{n_1} + \dfrac{\bar{p}\,\bar{q}}{n_2}}}$$

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 8-2.

Repaso

En este capítulo trabajamos con datos que se resumen como conteos de frecuencias para diferentes categorías. En la sección 10-2 describimos métodos para probar la bondad de ajuste en un experimento multinomial, que es similar a un experimento binomial, sólo que hay más de dos categorías de resultados. Los experimentos multinomiales resultan en conteos de frecuencias que se acomodan en un solo renglón o una columna; realizamos pruebas para determinar si las frecuencias muestrales observadas concuerdan (o se "ajustan") con alguna distribución que se asevera.

En la sección 10-3 describimos métodos para probar aseveraciones que incluyen 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 otra variable sirve para determinar la columna que describe un valor muestral. La sección 10-3 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 cuentan con las mismas proporciones de algunas características. He aquí algunos componentes clave de los métodos que se analizan en este capítulo.

- *Sección 10-2 (prueba de bondad de ajuste):*

  El estadístico de prueba es $\chi^2 = \sum \dfrac{(O - E)^2}{E}$

  La prueba es de cola derecha con $k - 1$ grados de libertad. Todas las frecuencias que se esperan deben ser de al menos 5.

- *Sección 10-3 (prueba de independencia u homogeneidad de la tabla de contingencia):*

  El estadístico de prueba es $\chi^2 = \sum \dfrac{(O - E)^2}{E}$

  La prueba es de cola derecha con $(r - 1)(c - 1)$ grados de libertad. Todas las frecuencias que se esperan deben ser de al menos 5.

Ejercicios de repaso

1. Datos de central de llamadas La tabla lista las llamadas que recibió una central telefónica durante una semana en un año reciente. (Los datos son de un gran productor estadounidense de electrónica que desea permanecer anónimo). Utilice un nivel de significancia de 0.05 para probar la aseveración de que las llamadas se distribuyen de manera uniforme durante los días laborales de la semana. ¿Qué sugiere el resultado acerca de los requisitos de personal en esta central de llamadas?

	Lun	Mar	Mié	Jue	Vie
Llamadas	98	68	89	64	56

2. ¿Ocurren con mayor frecuencia las muertes por arma de fuego durante los fines de semana? Cuando la revista *Time* investigó las muertes por arma de fuego en Estados Unidos durante una semana, se obtuvieron los resultados que se presentan en la tabla adjunta. Con un nivel de significancia 0.05, pruebe la aseveración de que las tasas de muerte por arma de fuego son las mismas durante los diferentes días de la semana. ¿Hay algún sustento para la teoría de que ocurren más muertes por arma de fuego los fines de semana, cuando más personas están en casa?

Día de la semana	Lun	Mar	Mié	Jue	Vie	Sáb	Dom
Número de muertes por arma de fuego	74	60	66	71	51	66	76

3. ¿Los tipos de crimenes cometidos dependen del consumo de alcohol? La tabla adjunta lista datos muestrales que el estadístico Karl Pearson utilizó en 1909. ¿Parece que el tipo de crimen se relaciona con el hecho de que el criminal fuera bebedor o abstemio?

	Incendiario	Violación	Violencia	Robo	Falsificación de moneda	Fraude
Bebedor	50	88	155	379	18	63
Abstemio	43	62	110	300	14	144

4. Prueba de dependencia entre la alta temprana del hospital y el reingreso de recién nacidos al hospital ¿Es seguro dar pronto de alta del hospital a los recién nacidos después de su nacimiento? La tabla adjunta muestra los resultados de un estudio sobre este tema. Utilice un nivel de significancia 0.05 para probar la aseveración de que el hecho de que se dé de alta pronto o más tarde a un recién nacido es independiente de su reingreso al hospital durante la semana posterior a su salida. ¿Se altera la conclusión si el nivel de significancia se cambia a 0.01?

	¿Reingreso durante la semana posterior a la alta?	
	Sí	No
Alta temprana (menos de 30 horas)	622	3997
Alta tardía (30 a 78 horas)	631	4660

Datos tomados de "The Safety of Newborn Early Discharge", de Liu *et al.*, *Journal of the American Medical Association*, vol. 278, núm. 4.

Ejercicios de repaso acumulativo

Tabla 10-8

	A	B	C	D
x	66	80	82	75
y	77	89	94	84

1. Cálculo de estadísticos Suponga que en la tabla 10-8, los títulos del renglón y la columna carecen de significado, de manera que la tabla contiene calificaciones de pruebas de ocho prisioneros que se seleccionaron 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 cinco números.

2. Cálculo de probabilidad Suponga que en la tabla 10-8 las letras A, B, C y D representan las opciones de la primera pregunta de un examen de opción múltiple. También que x representa a hombres y y representa a mujeres, así como que los números de la tabla son conteos de frecuencias, de forma tal que 66 hombres escogen la respuesta A, 77 mujeres la respuesta A, 80 hombres la respuesta B, etcétera.
a. Si se selecciona al azar una respuesta, calcule la probabilidad de que sea la respuesta C.
b. Si se selecciona al azar una respuesta, calcule la probabilidad de que la escoja un hombre.
c. Si se selecciona al azar una respuesta, calcule la probabilidad de que sea C o que la elija un hombre.
d. Si se seleccionan al azar dos respuestas diferentes, calcule la probabilidad de que ambas las elija una mujer.

3. Prueba para proporciones iguales Utilice los mismos supuestos que en el ejercicio 2, pruebe la aseveración de que los hombres y las mujeres escogen las diferentes respuestas en las mismas proporciones.

4. Prueba para una relación Suponga que la tabla 10-8 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 hay una relación entre las puntuaciones x y y.

5. Prueba de eficacia de entrenamiento Suponga que la tabla 10-8 lista puntuaciones de prueba para cuatro personas, donde la puntuación x corresponde a una prueba previa que se aplica antes de una sesión de entrenamiento para desarrollo de la memoria, en tanto que la puntuación y corresponde a una prueba que se aplica después del entrenamiento. Pruebe la aseveración de que la sesión de entrenamiento es eficaz para elevar las puntuaciones.

6. Prueba para igualdad de medias Suponga que en la tabla 10-8, las letras A, B, C y D representan diferentes versiones de la misma prueba de razonamiento. Las puntuaciones x se obtuvieron de cuatro hombres que se seleccionaron al azar y las puntuaciones y se obtuvieron de cuatro mujeres que se seleccionaron al azar. Pruebe la aseveración de que los hombres y las mujeres tienen la misma puntuación media.

Actividades de cooperación en equipo

1. Actividad fuera de clase Forme grupos de cuatro o cinco estudiantes. Vea los primeros dos renglones de la tabla 10-1, en el problema del capítulo, de la distribución de dígitos líderes que se esperan de acuerdo con la ley de Benford. Reúna datos y utilice los métodos de la sección 10-2 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 firmado
- Los precios de las acciones
- Las poblaciones de los condados en Estados Unidos

2. Actividad fuera de clase Divida la clase en grupos de cuatro o cinco estudiantes y reúna resultados anteriores 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 10-2 para probar que los números se seleccionan de manera que todos los posibles resultados son igualmente probables.

3. Actividad fuera de clase Forme grupos de cuatro o cinco estudiantes. Cada miembro del grupo debe encuestar al menos a 15 estudiantes hombres y a 15 estudiantes mujeres en la misma universidad, haciéndoles dos preguntas: 1. ¿Cuál partido político favorece más el sujeto? 2. Si a quien se encueste fuese a inventar la excusa de un neumático que se desinfló para justificar su ausencia, ¿cuál neumático diría él o ella que se desinfló, si el instructor preguntara? (Véase el ejercicio 6 en la sección 10-2). 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 que se reunieron. Incluya estas pruebas:

- La elección de un partido político es independiente del género del sujeto.
- El neumático que se identificó como el que se desinfló es independiente del género del sujeto.
- La elección de un partido político es independiente de si el sujeto es diestro o zurdo.
- El neumático que se identificó como el que se desinfló es independiente de la mano dominante del sujeto.
- El género es independiente de la mano dominante del sujeto.
- La elección de un partido político es independiente del neumático que se identificó como el que se desinfló.

4. Actividad fuera de clase Forme grupos de cuatro o cinco estudiantes. Cada miembro del grupo debe elegir alrededor de 15 estudiantes y pedirles primero a cada uno que seleccionen "aleatoriamente" cuatro dígitos. Después de registrar los cuatro dígitos, pedirá 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. Con el conjunto de datos "aleatorios", pruebe la aseveración de que los estudiantes seleccionan dígitos aleatoriamente. 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 aleatoriamente? ¿Seleccionan quizás algunos dígitos con más frecuencia que otros? ¿Parece que los últimos dígitos de los números del seguro social se seleccionaron aleatoriamente?

5. Actividad en clase Forme grupos de tres o cuatro estudiantes. A cada grupo se le entrega un dado junto con la instrucción de que debe probar su "balance". ¿El dado está balanceado o está cargado? Describa el análisis y los resultados.

6. Actividad fuera de clase Forme 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. (Véase los ejemplos de Barry Bonds en la sección 10-2 y el ejercicio 12 en la misma sección). Se señaló que el análisis de los últimos dígitos en ocasiones revela si los valores son resultado de mediciones reales o si son estimados que se reportan. Remítase a un almanaque y encuentre las longitudes de los ríos en el mundo, luego analice los últimos dígitos para determinar si dichas longitudes parecen ser mediciones reales o si parece el reporte de estimados. (En lugar 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 Plus o cualquier otro programa de cómputo o una 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 (para la que todos los dígitos son igualmente probables). ¿El generador de números aleatorios parece funcionar como debería?

Dígito	0	1	2	3	4	5	6	7	8	9
Frecuencia										

de los DATOS a la DECISIÓN

Pensamiento crítico: ¿Es el acusado culpable de fraude?

En el juicio del *estado de Arizona* versus *Wayne James Nelson*, al sujeto se le acusó de expedir cheques a un vendedor que no existía. Las cantidades de los cheques se listan abajo y se ordenaron por renglón.

Análisis de los resultados

¿Cumplen los dígitos líderes la ley de Benford descrita en el problema del capítulo? Cuando se prueba la bondad de ajuste con las proporciones que se esperan por la ley de Benford, es necesario combinar categorías, puesto que no todos los valores que se esperan son de al menos 5. Utilice una categoría con dígitos líderes de 1, una segunda categoría con dígitos líderes de 2, 3, 4 y 5, así como una tercera categoría con dígitos líderes de 6, 7, 8 y 9. ¿Son todos los valores que se esperan para estas tres categorías de al menos 5? ¿Hay evidencia suficiente para concluir que los dígitos líderes en los cheques no cumplen con 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 lugar de resultar de transacciones típicas y reales? Con base en la evidencia, si fuera parte de un jurado, ¿concluiría que los montos de los cheques son el resultado de un fraude? ¿Cuál sería un argumento que presentaría si usted fuera el abogado defensor?

$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	

PROYECTO DE INTERNET

Tablas de contingencia

Una característica importante de las pruebas de independencia con tablas de contingencia es que los datos reunidos no necesitan ser de naturaleza cuantitativa. Una tabla de contingencia resume observaciones por medio de las categorías o etiquetas de los renglones y las columnas. Como resultado, características como el género, la raza y el partido político se convierten en información susceptible de someterse a los procedimientos formales de prueba de hipótesis. El proyecto de Internet para este capítulo se encuentra en el sitio Web de *Estadística*:

http://www.pearsoneducacion.net/triola

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 sacará conclusiones que se relacionan 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 leen más de un millón de inversionistas.

Por favor describa su ocupación.

Trabajo para *Published Image*, donde utilizo la estadística para generar gráficos y datos que utilizamos en nuestras publicaciones financieras; hago uso de muchos estadísticos y sus aplicaciones. Escribimos notas informativas para bancos y sociedades de inversión.

¿Qué conceptos de estadística utiliza?

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 brutos. Se trata por lo regular de rendimientos mensuales, diarios o anuales. Luego uso Excel para graficar los datos y así obtener una imagen de lo que estoy tratando. A partir de esto, procedo a realizar un análisis. Algunas veces los resultados no respaldan un punto 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 por lo bajo, 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 señaló con claridad la cuestión.

En términos de estadística, ¿que le recomendaría a quienes buscan un empleo?

Es una herramienta lógica que, cuando se utiliza informativamente, puede convencerle a uno y a su audiencia del punto que usted está tratando de señalar con mucha más eficacia que las palabras. 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 quienes 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 arquitectura como licenciatura y un postgrado en negocios.

11

Análisis de varianza

PROBLEMA DEL CAPÍTULO

PROBLEMA 11 DEL CAPÍTULO

Clancy, Rowling y Tolstoi: ¿Hay diferencias en el nivel de lectura?

El conjunto de datos 14 del Apéndice B incluye valores que se obtuvieron de 12 páginas seleccionadas aleatoriamente de cada uno de los tres libros siguientes: *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. Se obtuvo la puntuación de la facilidad de lectura de Flesch para cada una de estas páginas y los resultados se presentan en el conjunto de datos 14. En el sistema de puntuación de facilidad de lectura de Flesch, las puntuaciones más altas de un texto implican una mayor facilidad de lectura. Las puntuaciones más bajas corresponden a textos que son difíciles de leer.

Con la intención de explorar datos para investigar el centro, la variación, la distribución, los datos distantes y los patrones de cambio a través del tiempo (CVDDT), obtendremos los estadísticos muestrales que se incluyen en la tabla 11-1. Además, los histogramas de los tres conjuntos de datos sugieren que las muestras provienen de poblaciones con distribuciones que son aproximadamente normales. Al investigar datos distantes, los únicos candidatos son la puntuación más baja de Clancy, que está a 2.37 desviaciones estándar por debajo de la media, y la puntuación más baja de Rowling, que está a 2.10 desviaciones estándar por debajo de la media. Si se juzgan en el contexto de las demás puntuaciones, dichos valores no parecen estar muy lejos del resto, por lo que asumiremos que no hay datos distantes. Si deseamos ser muy cuidadosos, analizaremos los datos con y sin esos dos valores que se incluyen para ver si la conclusión final es afectada. (De hecho, los resultados no se ven afectados de forma importante por estos valores). Puesto que los libros nunca cambiarán, el patrón de cambio a través del tiempo no es relevante aquí.

Cuando pensamos en una comparación de la facilidad de lectura de los libros de Tom Clancy, J. K. Rowling y León Tolstoi, esperaríamos que el libro de Rowling fuese el de lectura más fácil, porque fue escrito para niños. También que el libro de Tolstoi fuese el más difícil, ya que es una traducción de un clásico ruso. Ahora observe las puntuaciones medias de facilidad de lectura y vea que parecen sustentar dichas expectativas, puesto que el libro de Rowling tiene la puntuación de facilidad de lectura más alta y el de Tolstoi la más baja. ¿Pero concluiríamos realmente que las medias son diferentes? He aquí un tema importante que abordaremos en este capítulo: ¿proporcionan los datos muestrales de las puntuaciones de facilidad de lectura de Flesch, del conjunto de datos 14, evidencia suficiente que sustente la aseveración de que los libros de Clancy, Rowling y Tolstoi tienen diferentes medias?

Tabla 11-1 Puntuaciones de facilidad de lectura de los tres libros

	Puntuación de facilidad de lectura de Flesch		
	Clancy	Rowling	Tolstoi
n	12	12	12
$\bar{x}$	70.73	80.75	66.15
s	11.33	4.68	7.86

11-1 Panorama general

En lugar de "Análisis de varianza", un mejor título para este capítulo sería "Prueba de igualdad de tres o más medias poblacionales". Aun cuando no es muy práctico, el último título describe mejor el objetivo del presente capítulo. Deseamos introducir un procedimiento para probar hipótesis que establecen que tres o más medias poblacionales son iguales, de manera que una hipótesis nula típica sería H_0: $\mu_1 = \mu_2 = \mu_3$, en tanto que la hipótesis alternativa sería la afirmación de que al menos una media es diferente de las otras. En la sección 8-3 presentamos 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 lugar de referirnos al objetivo principal de probar medias iguales, el término *análisis de varianza* se refiere al *método* que empleamos, el cual se basa 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.

El ANOVA se utiliza en aplicaciones tales como las siguientes:

- Si tratamos un grupo con dos tabletas de aspirina diariamente, un segundo grupo con una tableta de aspirina diariamente y un tercer grupo con un placebo diariamente, es posible hacer una prueba para determinar si hay evidencia suficiente para sustentar la aseveración de que los tres grupos cuentan con distintos niveles medios de presión sanguínea.
- Se asevera que los supermercados colocan los cereales con alto contenido de azúcar en estantes que están a la altura de los ojos de los niños, de manera que eso nos permite probar la aseveración de que los cereales en los estantes tienen el mismo contenido de azúcar.

¿Por qué no probar sencillamente dos muestras al mismo tiempo? ¿Por qué necesitamos un nuevo procedimiento, cuando la igualdad de dos medias se demuestra utilizando los métodos que se presentan en el capítulo 8? Por ejemplo, si deseamos utilizar los datos muestrales de la tabla 11-1, 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: $\mu_2 = \mu_3$; entonces, H_0: $\mu_1 = \mu_3$? Este método (probar dos a la vez) requiere de tres pruebas de hipótesis diferentes, de forma que el grado de confianza sería tan bajo como 0.95^3 (o 0.857). En general, conforme incrementamos el número de pruebas de significancia individuales, incrementamos la posibilidad de obtener una diferencia únicamente por el azar (en lugar de una diferencia real en las medias). El riesgo de cometer un error tipo I (es decir, de encontrar una diferencia en uno de los pares cuando en realidad tal diferencia no existe) es demasiado alto. El método del análisis de varianza nos sirve para evitar este problema en particular (rechazar una hipótesis nula verdadera), si utilizamos una prueba de igualdad de varias medias.

Distribución *F*

Los métodos del ANOVA de este capítulo requieren de la distribución *F*, que se explicó en la sección 8-5. En esta sección señalamos que la distribución *F* tiene las siguientes propiedades importantes (véase la figura 11-1):

1. La distribución *F* es no simétrica; se sesga hacia la derecha.
2. Los valores de *F* pueden ser 0 o positivos, pero no negativos.
3. Hay 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) se basa 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 11-2. El término *un factor* se utiliza porque los datos muestrales se separan en grupos según una característica o un factor. Por ejemplo, las puntuaciones de facilidad de lectura, que se resumen en la tabla 11-1, se separaron en tres grupos diferentes, de acuerdo con la característica (o el factor) del autor (Clancy, Rowling, Tolstoi). En la sección 11-3 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, separaríamos la estatura de las personas utilizando los siguientes dos factores: 1. género (hombre o mujer) y 2. mano dominante derecha o izquierda.

Estrategia de estudio sugerida: Puesto que los procedimientos que se emplean en este capítulo requieren de cálculos complicados, pondremos énfasis en el uso y la interpretación de programas de cómputo, tales como STATDISK, Minitab y Excel, o de una calculadora TI-83 Plus. Sugerimos que inicie la sección 11-2 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 se complican, nuestro procedimiento será fácil porque se basa 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.

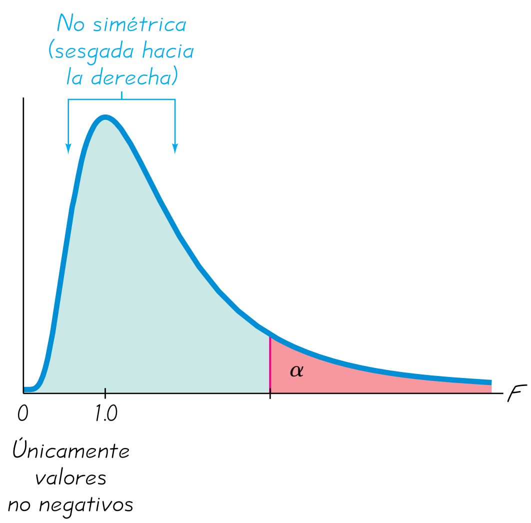


FIGURA 11-1
Distribución *F*

Existe una distribución *F* distinta para cada par de grados de libertad diferente para el numerador y el denominador.

Después de comprender este procedimiento básico y sencillo, proceda a la comprensión de los fundamentos subyacentes.

11-2 ANOVA de un factor

En esta sección consideramos las pruebas de hipótesis de que tres o más medias poblacionales son iguales, como en H_0: $\mu_1 = \mu_2 = \mu_3$. Los cálculos se complican mucho, de manera que recomendamos el siguiente método:

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. Logre comprender el fundamento subyacente estudiando los ejemplos de esta sección.
3. Familiarícese con la naturaleza de los valores de la SC (suma de cuadrados) y los CM (cuadrados medios), así como con el papel que tienen en la determinación del estadístico de prueba F, pero utilice programas estadísticos de cómputo o una calculadora para obtener dichos 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

Tratamiento (o **factor**): es una propiedad o característica que nos permite distinguir entre sí a las distintas poblaciones.

Por ejemplo, las puntuaciones de facilidad de lectura que se resumen en la tabla 11-1 se distinguen de acuerdo con el tratamiento (o factor) del autor (Clancy, Rowling, Tolstoi). Se utiliza el término *tratamiento* ya que las primeras aplicaciones del análisis de varianza implicaron experimentos de agricultura en los cuales distintas porciones de tierra se trataban con diferentes fertilizantes, tipos de semillas, insecticidas, etcétera. El siguiente recuadro incluye los supuestos requeridos y los procedimientos que utilizaremos.

Supuestos

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 12-5).
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 estadístico de la

Universidad de Wisconsin, George E. P. Box demostró que, siempre y cuando los tamaños muestrales sean iguales [o casi iguales], las varianzas llegan a diferir de tal forma que la más grande es nueve veces el tamaño de la más pequeña, en tanto que los resultados del ANOVA seguirán 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. (Las muestras no están aparejadas ni asociadas de ninguna forma).
5. Las diferentes muestras provienen de poblaciones que se categorizaron de una sola forma. (De ahí el nombre del método: análisis de varianza de *un factor*).

Procedimiento de prueba de H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots$

1. **Utilice STATDISK, Minitab, Excel o una calculadora TI-83 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 $P \le \alpha$, 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 hay 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 permiten identificar las medias específicas, que son diferentes. Dichas pruebas se conocen como procedimientos de comparación múltiple. La comparación de intervalos de confianza, la prueba de Scheffé, la prueba de Tukey extendida y la prueba de Bonferroni son procedimientos de comparación múltiple comunes).

EJEMPLO Facilidad de lectura de Clancy, Rowling y Tolstoi A partir de las puntuaciones de facilidad de lectura que se resumen en la tabla 11-1, y con un nivel de significancia de $\alpha = 0.05$, utilice STATDISK, Minitab, Excel o una calculadora TI-83 Plus para probar la aseveración de que las tres muestras provienen de poblaciones con medias que no son iguales.

SOLUCIÓN La hipótesis nula es H_0: $\mu_1 = \mu_2 = \mu_3$, en tanto que la hipótesis alternativa es la aseveración de que al menos una de las medias es diferente de las otras.

Paso 1: Al final de esta sección describiremos procedimientos específicos para obtener resultados por computadora o calculadora, ahora consideraremos los resultados que se despliegan en la página siguiente.

Paso 2: Todas las pantallas de resultados muestran que el valor *P* es 0.000562 o 0.001 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.

continúa

Resistencia a las encuestas

Las encuestas que se basan 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. La organización Council of American Survey Research reportó que, en un año reciente, el 38% de los consumidores se rehusaron a responder encuestas. El director de una compañía de investigación de mercado 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 encargada de hacer investigación de mercados, multimillonaria en dólares, afectan los productos que compramos, los programas de televisión que vemos y muchas otras facetas de nuestras vidas.

STATDISK

Equal Length Samples	
Total Num Values	36
Upper Deg Free	2
Lower Deg Free	33
SS(treatment)	1338.0
SS(error)	2331.4
SS(total)	3669.4
MS(treatment)	669.00
MS(error)	70.648
MS(total)	104.84
Test Statistic, F	9.4695
Critical F	3.2849
P-Value	0.000562

Reject the Null Hypothesis

Data provides evidence that the sample means are unequal

TI-83 Plus

```
One-way ANOVA
 F=9.469487401
 p=5.6213335E-4
 Factor
  df=2
  SS=1338.00222
↓ MS=669.001111
```

```
One-way ANOVA
↑ MS=669.001111
 Error
  df=33
  SS=2331.38667
  MS=70.6480808
 Sxp=8.40524127
```

Excel

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Column 1	12	848.8	70.73333333	128.2806061
Column 2	12	969	80.75	21.91545455
Column 3	12	793.8	66.15	61.74818182

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1338.002222	2	669.0011111	**9.469487401**	**0.000562133**	3.284924333
Within Groups	2331.386667	33	70.64808081			
Total	3669.388889	35				

Minitab

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	2	1338.0	669.0	9.47	0.001
Error	33	2331.4	70.6		
Total	35	3669.4			

INTERPRETACIÓN Hay evidencia suficiente para sustentar la aseveración de que las tres medias poblacionales no son iguales. Con base en páginas que se seleccionaron al azar de las obras *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, concluimos que estos libros tienen niveles de facilidad de lectura diferentes. Con base en la prueba ANOVA, no concluimos que alguna media en particular sea distinta de las otras.

Fundamentos

El método de análisis de varianza se basa en el siguiente concepto fundamental: bajo el supuesto de que las poblaciones tienen la misma varianza σ^2, estimamos el valor común de σ^2 por medio de dos métodos diferentes. El estadístico de prueba F es la proporción de dichos estimados, de forma que un estadístico de prueba F significativamente *grande* (que se ubica a la extrema derecha de la gráfica de distribución F) constituye evidencia en contra de que las medias poblacionales son iguales. La figura 11-2 muestra la relación entre el estadístico de prueba F y el valor P.

Los dos métodos para estimar el valor común de σ^2 son los siguientes:

1. La **varianza entre muestras** (también se le llama **variación debida al tratamiento**) es un estimado de la varianza poblacional común σ^2, que se basa en la variación entre las *medias* muestrales.
2. La **varianza dentro de las muestras** (también se le llama **variación debida al error**) es un estimado de la varianza poblacional común σ^2, que se basa en las *varianzas* muestrales.

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 afecta por las diferencias entre las medias muestrales. Como consecuencia, las medias muestrales con valores cercanos dan como

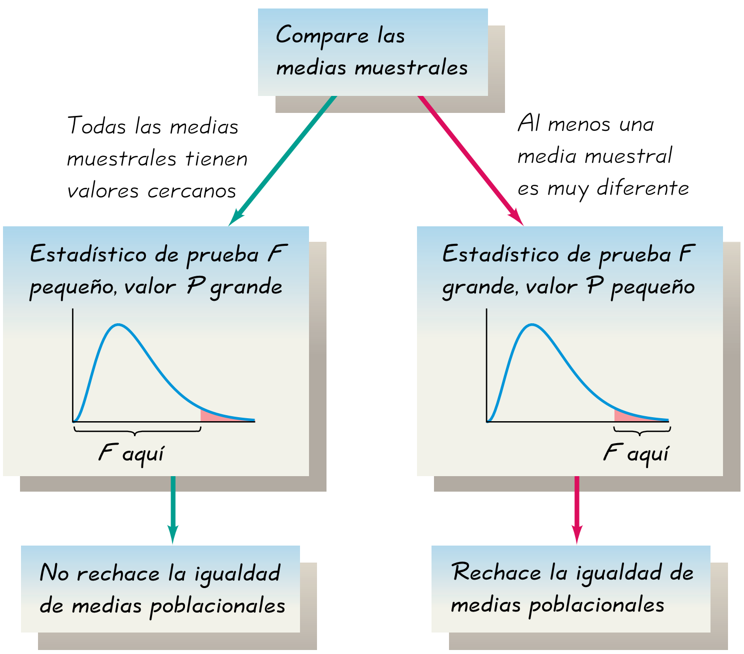


FIGURA 11-2 Relación entre el estadístico de prueba *F* y el valor *P*

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 términos vagos “pequeño” y “excesivamente grande” se vuelven objetivos por medio del valor P correspondiente, que 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.

Cálculos con tamaños muestrales *n* iguales

Remítase al conjunto de datos A en la tabla 11-2. Si todos los conjuntos de datos tienen el mismo tamaño de muestra (como en $n = 4$ para el conjunto de datos A de la tabla 11-2), los cálculos que se requieren no son muy difíciles. Primero, calcule la varianza entre muestras al evaluar $ns_{\bar{x}}^2$, donde $s_{\bar{x}}^2$ 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. (A partir del teorema de límite central, en $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ se despeja σ para obtener $\sigma = \sqrt{n} \cdot \sigma_{\bar{x}}$, de forma que estimamos σ^2 con $ns_{\bar{x}}^2$). Por ejemplo, las medias

Tabla 11-2 Efecto de una media sobre el estadístico de prueba F

A añadir 10 **B**

	A: Muestra 1	A: Muestra 2	A: Muestra 3	B: Muestra 1	B: Muestra 2	B: 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
	$n_1 = 4$	$n_2 = 4$	$n_3 = 4$	$n_1 = 4$	$n_2 = 4$	$n_3 = 4$
	$\bar{x}_1 = 5.5$	$\bar{x}_2 = 6.0$	$\bar{x}_3 = 6.0$	$\bar{x}_1 = 15.5$	$\bar{x}_2 = 6.0$	$\bar{x}_3 = 6.0$
	$s_1^2 = 3.0$	$s_2^2 = 2.0$	$s_3^2 = 2.0$	$s_1^2 = 3.0$	$s_2^2 = 2.0$	$s_3^2 = 2.0$

	A	B
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	$s_p^2 = \dfrac{3.0 + 2.0 + 2.0}{3} = 2.3333$	$s_p^2 = \dfrac{3.0 + 2.0 + 2.0}{3} = 2.3333$
Estadístico de prueba F	$F = \dfrac{ns_{\bar{x}}^2}{s_p^2} = \dfrac{0.3332}{2.3333} = 0.1428$	$F = \dfrac{ns_{\bar{x}}^2}{s_p^2} = \dfrac{120.3332}{2.3333} = 51.5721$
Valor P (obtenido con Excel)	Valor $P = 0.8688$	Valor $P = 0.0000118$

muestrales del conjunto de datos A de la tabla 11-2 son 5.5, 6.0 y 6.0. Estos tres valores tienen una varianza de $s_{\bar{x}}^2 = 0.0833$, de forma 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 que se agrupa que se obtiene al calcular la media de las varianzas muestrales. Las varianzas muestrales en la tabla 11-2 son 3.0, 2.0 y 2.0, de forma que

$$\begin{aligned}\text{varianza dentro de las muestras} &= s_p^2 \\ &= \frac{3.0 + 2.0 + 2.0}{3} = 2.3333\end{aligned}$$

Finalmente, evalúe el estadístico de prueba F de la siguiente manera:

$$F = \frac{\text{varianza entre muestras}}{\text{varianza dentro de muestras}} = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{0.3332}{2.3333} = 0.1428$$

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 de la siguiente manera.

Grados de libertad:
(k = número de muestras y *n* = tamaño de la muestra)

$$\text{numerador de grados de libertad} = k - 1$$
$$\text{denominador de grados de libertad} = k(n - 1)$$

Para el conjunto de datos A de la tabla 11-2, $k = 3$ y $n = 4$, entonces 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 nueve 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 11-2, 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; 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 de la tabla 11-2. 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 hay 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 11-2 produce un efecto importante 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 surte un efecto drástico en el valor P, que 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 se separan más). También note que las varianzas dentro de las muestras son de 2.3333 en ambas partes, puesto que la varianza dentro de una muestra no se afecta 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* $\bar{x}_1$. Lo anterior 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 11-2: 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 provoca 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 que se requieren para los casos con tamaños muestrales iguales son razonables, las cosas se complican bastante cuando los tamaños muestrales son desiguales. 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 muestra a continuación.

$$F = \frac{\text{varianza entre muestras}}{\text{varianza dentro de muestras}} = \frac{\left[\dfrac{\Sigma n_i(\bar{x}_i - \bar{\bar{x}})^2}{k-1}\right]}{\left[\dfrac{\Sigma(n_i-1)s_i^2}{\Sigma(n_i-1)}\right]}$$

donde $\bar{\bar{x}}$ = media de todos los valores muestrales que se combinan

k = número de medias poblacionales que se comparan

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 se incluye, 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.

Ya que el cálculo de este estadístico de prueba a veces conduce 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 cuadrados) y los CM (cuadrados medios). A pesar de que la siguiente notación y los componentes son complicados y tediosos, la idea básica es la misma: el estadístico de prueba F es un cociente con un numerador que refleja la variación *entre* las medias de las muestras, en tanto que un denominador 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 de cuadrados total, es una medida de la variación total (alrededor de $\bar{\bar{x}}$) en todos los datos muestrales que se combinan.

Fórmula 11-1 $\text{SC(total)} = \Sigma(x - \bar{\bar{x}})^2$

La SC(total) se puede separar en los componentes de la SC (tratamiento) y la SC (error), descritas a continuación.

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 11-2

$$\begin{aligned}\text{SC(tratamiento)} &= n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \cdots + n_k(\bar{x}_k - \bar{\bar{x}})^2 \\ &= \Sigma n_i(\bar{x}_i - \bar{\bar{x}})^2\end{aligned}$$

Si las medias poblacionales ($\mu_1, \mu_2, \ldots, \mu_k$) son iguales, entonces todas las medias muestrales $\bar{x}_1, \bar{x}_2 \ldots, \bar{x}_k$ tenderán a acercarse entre sí y también a acercarse a $\bar{\bar{x}}$. El resultado será un valor de SC (tratamiento) relativamente pequeño. Sin embargo, si las medias poblacionales no son todas iguales, entonces al menos una de $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ tenderá a estar lejos de las demás y también de $\bar{\bar{x}}$. El resultado será un valor relativamente grande de SC (tratamiento).

La **SC (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 que se consideran.

Fórmula 11-3

$$\begin{aligned}\text{SC(error)} &= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2 \\ &= \Sigma(n_i - 1)s_i^2\end{aligned}$$

Dadas las expresiones anteriores para SC (total), SC (tratamiento) y SC (error), siempre deben mantenerse las siguientes relaciones.

Fórmula 11-4 SC (total) = SC (tratamiento) + SC (error)

SC (tratamiento) y SC (error) son ambas sumas de cuadrados, por lo que si dividimos cada una de ellas entre su número correspondiente de grados de libertad, obtendremos los cuadrados *medios*. Algunas de las siguientes expresiones, para los cuadrados medios, incluyen la notación N:

N = número total de valores en todas las muestras que se combinan

CM (tratamiento) es un cuadrado medio de tratamiento, que se obtiene como sigue:

Fórmula 11-5 $$\text{CM(tratamiento)} = \frac{\text{SC(tratamiento)}}{k - 1}$$

CM (del error) es un cuadrado medio del error, que se obtiene como sigue:

Fórmula 11-6 $$\text{CM(error)} = \frac{\text{SC(error)}}{N - k}$$

CM (total) es un cuadrado medio de la variación total, que se obtiene como sigue:

Fórmula 11-7 $$\text{CM(total)} = \frac{\text{SC(total)}}{N - 1}$$

Estadístico de prueba para ANOVA con tamaños muestrales desiguales

Al probar la hipótesis nula H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ en contra de la hipótesis alternativa, de que todas estas medias no son iguales, el estadístico de prueba

Fórmula 11-8 $$F = \frac{\text{CM(tratamiento)}}{\text{CM(error)}}$$

tiene una distribución F (cuando la hipótesis nula H_0 es verdadera) con grados de libertad dados por

grados de libertad del numerador $= k - 1$
grados de libertad del denominador $= N - k$

Este estadístico de prueba es esencialmente el mismo que se introdujo antes y su interpretación también es igual a la ya descrita. El denominador sólo depende de las varianzas muestrales que miden la variación dentro de los tratamientos y no se afecta por las diferencias entre las medias muestrales. En contraste, el numerador se afecta 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; por lo tanto, la prueba ANOVA es de cola derecha.

Las tablas implican un formato conveniente para resumir los resultados más importantes en los cálculos del ANOVA, en tanto que la tabla 11-3 tiene un formato que suele utilizarse en el despliegue de resultados de las computadoras. (Véanse los resultados anteriores de Minitab y Excel). Las cifras de la tabla 11-3 resultan de los datos de facilidad de lectura de la tabla 11-1.

Tabla 11-3 Tabla de ANOVA para los datos de facilidad de lectura

Fuente de variación	Suma de cuadrados (SC)	Grados de libertad	Cuadrado medio (CM)	Estadístico de prueba *F*
Tratamientos	1338.00	2	669.000	9.4695
Error	2331.39	33	70.648	
Total	3669.39	35		

Diseño del experimento: Cuando utilizamos un análisis de varianza de un factor (o de una entrada) y concluimos que las diferencias entre las medias son significativas, no estaremos completamente seguros de que el factor dado es el responsable de las diferencias. Es posible que la variación de algún otro factor desconocido sea el responsable. Una manera de reducir el efecto de factores extraños es diseñar el experimento de forma que sea un **diseño completamente aleatorizado**, en el cual se da a cada elemento la misma posibilidad de pertenecer a las diferentes categorías o tratamientos. Por ejemplo, podría asignar sujetos a un grupo de tratamiento, a un grupo placebo y a un grupo control por medio de un proceso de selección aleatoria equivalente a sacar papeles de un tazón. Otra manera 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. Por ejemplo, tratar a una niña saludable de siete años de edad de Texas, mientras que le da un placebo a otra niña saludable de siete años de edad de Texas y coloca a una tercera niña saludable de siete años de edad de Texas en un grupo control que no recibe nada. Además de la salud, la edad, el género y el estado de residencia, identificaría otros factores relevantes que convendría tomar en cuenta. En general, los buenos resultados requieren que el experimento se diseñe y ejecute de manera cuidadosa.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **One-Way Analysis of Variance** y proceda a ingresar los datos muestrales. Haga clic en **Evaluate** al finalizar.

Minitab Primero ingrese los datos muestrales en las columnas C1, C2, C3, . . . Después, elija **Stat, ANOVA ONEWAY (UNSTACKED)** e introduzca C1, C2, C3, . . . , en el recuadro que se identifica como Responses (en columnas separadas).

Excel Primero introduzca los datos en las columnas A, B, C, . . . Después 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 se ubica en el renglón 12 de la columna C).

TI-83 Plus Primero introduzca los datos en listas en L1, L2, L3, . . . , después presione **STAT**, seleccione **TESTS** y elija la opción **ANOVA**. Ingrese las etiquetas 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**.

11-2 Destrezas y conceptos básicos

1. **Facilidad de lectura de autores** El problema del capítulo utiliza las puntuaciones de facilidad de lectura de Flesch para páginas que se seleccionaron aleatoriamente de libros de Tom Clancy, J. K. Rowling y León Tolstoi. Si en su lugar se utilizan las puntuaciones del nivel de lectura de Flesch-Kincaid (véase el conjunto de datos 14 del Apéndice B), los resultados del análisis de varianza, contenidos de Minitab, son los que se incluyen en la siguiente tabla. Suponga que deseamos utilizar un nivel de significancia de 0.05 para probar la hipótesis nula de que los tres autores tienen puntuaciones del nivel de lectura de Flesch-Kincaid con la misma media.
   a. ¿Cuál es la hipótesis nula?
   b. ¿Cuál es la hipótesis alternativa?
   c. Identifique el valor del estadístico de prueba.
   d. Calcule el valor crítico para un nivel de significancia de 0.05.
   e. Identifique el valor *P*.
   f. Con base en los resultados anteriores, ¿qué concluye acerca de la igualdad de las medias poblacionales?

Minitab

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	2	68.19	34.09	8.98	0.001
Error	33	125.31	3.80		
Total	35	193.50			

2. **Prueba de inflamabilidad de tela en diferentes laboratorios** Se realizaron pruebas de inflamabilidad en ropa de dormir infantil. Se utilizó la prueba Vertical Semirestrained, en la que se incendiaron pedazos de tela en condiciones controladas. Una vez que se detuvo el incendio, se midió y registró la longitud de la porción quemada. Las mismas muestras de tela se probaron en cinco laboratorios diferentes. Abajo se presentan los resultados del análisis de varianza realizado con Excel.
   a. ¿Cuál es la hipótesis nula?
   b. ¿Cuál es la hipótesis alternativa?
   c. Identifique el valor del estadístico de prueba.
   d. Calcule el valor crítico para un nivel de significancia de 0.05.
   e. Identifique el valor *P*.
   f. ¿Existe evidencia suficiente para sustentar la aseveración de que las medias de los distintos laboratorios no son iguales?

Excel

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2.087194264	4	0.521798566	2.949333035	0.030665893	2.588834036
Within Groups	7.607597403	43	0.17692087			
Total	9.694791667	47				

3. **Tiempos de maratón** Una muestra aleatoria de hombres que completaron la carrera de maratón de Nueva York se divide en tres categorías, con las edades 21-29, 31-39 y 40 o más. Del conjunto de datos 8 del Apéndice B se obtienen los tiempos (en segundos). Los resultados del análisis de varianza, que se obtuvieron por medio de Excel se presentan a continuación.

a. ¿Cuál es la hipótesis nula?
b. ¿Cuál es la hipótesis alternativa?
c. Identifique el valor del estadístico de prueba.
d. Calcule el valor crítico para un nivel de significancia de 0.05.
e. Identifique el valor *P*.
f. ¿Existe evidencia suficiente para sustentar la aseveración de que los hombres de las distintas categorías tienen tiempos medios diferentes?

Excel

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	3532063.284	2	1766031.642	0.188679406	0.828324293	3.080387501
Within Groups	1010875649	108	9359959.71			
Total	1014407712	110				

4. Presión sanguínea sistólica en distintos grupos de edad A una muestra aleatoria de 40 mujeres se le dividió en tres categorías con edades por debajo de 20, de 20 a 40 y mayores de 40. Los niveles de presión sanguínea sistólica se obtienen del conjunto de datos 1 del Apéndice B. Los resultados del análisis de varianza, que se obtuvieron por medio de Excel, se presentan a continuación.
a. ¿Cuál es la hipótesis nula?
b. ¿Cuál es la hipótesis alternativa?
c. Identifique el valor del estadístico de prueba.
d. Identifique el valor *P*.
e. ¿Hay evidencia suficiente para sustentar la aseveración de que las mujeres en las distintas categorías tienen niveles medios de presión sanguínea diferentes?

Minitab

Source	DF	SS	MS	F	P
Factor	2	938	469	1.65	0.205
Error	37	10484	283		
Total	39	11422			

En los ejercicios 5 y 6, utilice los datos muestrales que se listan relativos a experimentos de choques de automóviles, que realizó la National Transportation Safety Administration. Se adquirieron coches nuevos y se chocaron contra una barrera fija a 35 millas/hora; las mediciones se registraron con respecto al maniquí colocado 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 Suburban 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.

T **5.** Traumatismo craneal en un choque de automóvil A continuación se presentan los datos de traumatismo craneal. 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

T **6.** Deceleración del pecho en un choque de automóvil A continuación se presentan datos de deceleración 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

T **7.** Arqueología: anchura de cráneos de distintas épocas Los valores en la tabla corresponden a las anchuras máximas medidas de cráneos de hombres egipcios de distintas épocas (datos que se tomaron 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.

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

T **8.** Energía solar en diferentes climas Un alumno del autor vive en una casa con un sistema eléctrico solar. A la misma hora, cada día, reunió las lecturas de voltaje de un medidor que conectó al sistema, cuyos resultados se listan en la tabla adjunta. 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. ¿Hay evidencia suficiente para sustentar una aseveración de medias poblacionales diferentes? Esperaríamos que un sistema solar proporcione más energía eléctrica los días soleados que los días nublados o lluviosos. ¿Concluiríamos que los días soleados dan como resultado mayores cantidades de energía eléctrica?

Días soleados	Días nublados	Días lluviosos
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

T **9.** Pesos medios de dulces M&M Remítase al conjunto de datos 19 del Apéndice B. Con un nivel de significancia de 0.05, pruebe la aseveración de que el peso medio de los dulces M&M es el mismo para cada una de las seis poblaciones de diferente color. Si la intención de Mars, Inc., es fabricar los dulces de modo que las poblaciones de diferente color tengan el mismo peso medio, ¿sugieren los resultados que la compañía enfrenta un problema que amerita corregirse?

T **10.** Distancias de jonrones Remítase al conjunto de datos 30 del Apéndice B. Utilice un nivel de significancia de 0.05 para probar la aseveración de que los jonrones que anotaron Barry Bonds, Mark McGwire y Sammy Sosa tienen distancias medias que no son iguales. ¿Explican las distancias de los jonrones el hecho de que, hasta ahora, Barry Bonds conectó el mayor número de jonrones en una temporada, mientras que Mark McGwire posee el segundo número más grande de jonrones?

11. Azúcar en el cereal Remítase al conjunto de datos 16 del Apéndice B y combine las cantidades de azúcar de los estantes 3 y 4, que son los dos más altos. Utilice un nivel de significancia de 0.05 para probar la hipótesis nula de que las cantidades medias de azúcar en los diferentes estantes son iguales. ¿Qué sugieren los resultados acerca de la creencia común de que los supermercados colocan los cereales con alto contenido de azúcar en estantes que están al nivel de los ojos de los niños?

12. Tabaquismo pasivo en distintos grupos Remítase al conjunto de datos 6 del Apéndice B. Utilice un nivel de significancia de 0.05 para probar la aseveración de que el nivel medio de nicotina es diferente en los siguientes tres grupos: 1. individuos que no fuman y que no están expuestos al humo del tabaco; 2. personas que no fuman y que están expuestas al humo del tabaco, y 3. personas que fuman. ¿Qué sugieren los resultados acerca del tabaquismo pasivo?

11-2 Más allá de lo básico

13. Uso de la prueba *t* Se seleccionan al azar cinco muestras independientes, de 50 valores cada una, provenientes de poblaciones que se distribuyen normalmente con varianzas iguales. Deseamos probar la aseveración de que $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$.

a. Si utilizáramos únicamente los métodos que se presentan en la sección 8-3, probaríamos las aseveraciones individuales $\mu_2 = \mu_2$, $\mu_1 = \mu_3$, etcétera. ¿De cuántas formas distintas es posible hacer pares con las cinco medias?

b. Suponga que para cada prueba de igualdad entre dos medias hay una probabilidad de 0.95 de no cometer un error tipo I. Si se prueba la igualdad de todos los pares posibles de medias, ¿cuál es la probabilidad de no cometer errores tipo I? (Aun cuando las pruebas no son independientes en realidad, suponga que sí lo son).

c. Si utilizamos el análisis de varianza para probar la aseveración de que $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$, con un nivel de significancia de 0.05, ¿cuál es la probabilidad de no cometer un error tipo I?

d. Compare los resultados de los incisos *b* y *c*. ¿Cuál método es mejor porque nos ofrece una mayor oportunidad de no cometer un error tipo I?

14. Pruebas equivalentes En este ejercicio comprobará que cuando tiene dos conjuntos de datos muestrales, la prueba *t* para muestras independientes y el método ANOVA de esta sección son equivalentes. Remítase a las mediciones de facilidad de lectura de la tabla 11-1, pero utilice únicamente los datos de Clancy y Rowling. Los datos originales se encuentran en el conjunto de datos 14 del Apéndice B.

a. Utilice un nivel de significancia de 0.05 y el método de la sección 8-3 para probar la aseveración de que las dos muestras provienen de poblaciones con la misma media. (Suponga que ambas poblaciones tienen la misma varianza).

b. Utilice un nivel de significancia de 0.05 y el método ANOVA de esta sección para probar la aseveración que se planteó en el inciso *a*.

c. Verifique que los cuadrados del estadístico de prueba *t* y el valor crítico del inciso *a* son iguales al estadístico de prueba *F* y el valor crítico del inciso *b*.

11-3 ANOVA de dos factores

En la sección 11-2 empleamos el análisis de varianza para decidir si tres o más poblaciones tienen la misma media. Esa sección utiliza procedimientos que se conocen como análisis de varianza de *un* factor (o análisis de varianza de una entrada), ya que los datos se categorizan en grupos de acuerdo con un *solo* factor (o tratamiento). Recuerde que un factor o tratamiento es una propiedad que es la base para categorizar los diferentes grupos de datos. Véase la tabla 11-4 , que incluye los tiempos (en segundos) de corredores que finalizaron una maratón reciente en la ciudad de Nueva York. Los tiempos que se listan se seleccionaron al azar del conjunto de datos 8 del Apéndice B y se dividen en seis categorías de acuerdo con dos variables: 1. la variable de renglón género, y 2. la variable de columna edad. El **análisis de varianza de dos factores** considera *dos* factores, como el género y la edad en la tabla 11-4. Las seis subcategorías de la tabla 11-4 se conocen como *celdas*, de modo que la tabla 11-4 tiene seis celdas con cinco valores cada una.

Tabla 11-4 Tiempo (en segundos) de corredores de la maratón de Nueva York

	Edad		
	21-29	30-39	40 o más
Hombre	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
Mujer	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

Al analizar los datos muestrales de la tabla 11-4, ya estudiamos el análisis de varianza de un factor, por lo que sería razonable proceder sencillamente con el ANOVA de un factor para el factor del género y otra para el factor de edad. Por desgracia, el hecho de realizar dos pruebas ANOVA de un factor separadas desperdicia información e ignora por completo un aspecto muy importante: el efecto de una interacción entre los dos factores.

Definición

Hay 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 del alimento y el vino en un restaurante de calidad. Se sabe que ciertos alimentos y vinos interactúan bien al producir un sabor agradable, mientras que otros interactúan de tal manera que producen un sabor desagradable. Hay una buena interacción entre el vino Chablis y las ostras; la piedra caliza que se entierra en la tierra donde se hace el Chablis deja un residuo en el vino que interactúa muy bien con las ostras. La mantequilla de cacahuate y la mermelada también interactúan bien. En contraste, el jarabe de chocolate y los *hot dogs* interactúan de una forma que provoca un sabor desagradable. Si utilizamos el ANOVA de dos factores para los datos de la tabla 11-4, consideraríamos tres efectos posibles en los tiempos de la carrera de maratón: **1.** los efectos de una interacción entre género y edad; **2.** los efectos del género; **3.** los efectos de la edad. Los cálculos son bastante complejos, por lo que *suponemos que se utiliza un programa de cómputo o una calculadora TI-83 Plus*. (Al final de esta sección se describen procedimientos para el uso de herramientas tecnológicas). La siguiente es la pantalla de los resultados de Minitab para los datos de la tabla 11-4.

Minitab

```
Analysis of Variance for TIME
Source         DF        SS         MS        F        P
GENDER          1  15225413   15225413     1.69    0.206
AGE             2  92086979   46043490     5.10    0.014
Interaction     2  21042069   10521034     1.17    0.329
Error          24 216683456    9028477
Total          29 345037917
```

Los resultados de Minitab incluyen componentes de SC (suma de cuadrados), similares a los descritos en la sección 11-2. Como las circunstancias de la sección 11-2 incluyen un solo factor, utilizamos SC (tratamiento) como una medida de la variación consecuencia de las diferentes categorías de tratamiento y SC (error) como una medida de variación por el error de muestreo. Aquí manejamos SC (género) como una medida de variación entre las medias de género. Empleamos SC (edad) como una medida de la variación entre las medias de edades. Continuamos empleando SC (error) como una medida de la variación que ocasiona el error de muestreo. De manera similar, trabajamos con CM (género) y CM (edad) para los dos cuadrados medios distintos, y continuamos con CM (error) como antes. Además, utilizamos gl (género) y gl (edad) para los dos distintos grados de libertad.

A continuación se listan los supuestos que se requieren y el procedimiento básico del análisis de varianza de dos factores (ANOVA). El procedimiento también se resume en la figura 11-3.

Supuestos

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 s).
3. Las muestras son muestras aleatorias simples. (Es decir, las muestras del mismo tamaño tienen la misma probabilidad de que se seleccionen).
4. Las muestras son independientes entre sí. (Las muestras no están apareadas ni asociadas de ninguna manera).
5. Los valores muestrales se categorizan en dos factores. (De ahí el nombre del método: análisis de varianza de *dos factores*).
6. Todas las celdas tienen el mismo número de valores muestrales. (A éste se le denomina un diseño *balanceado*).

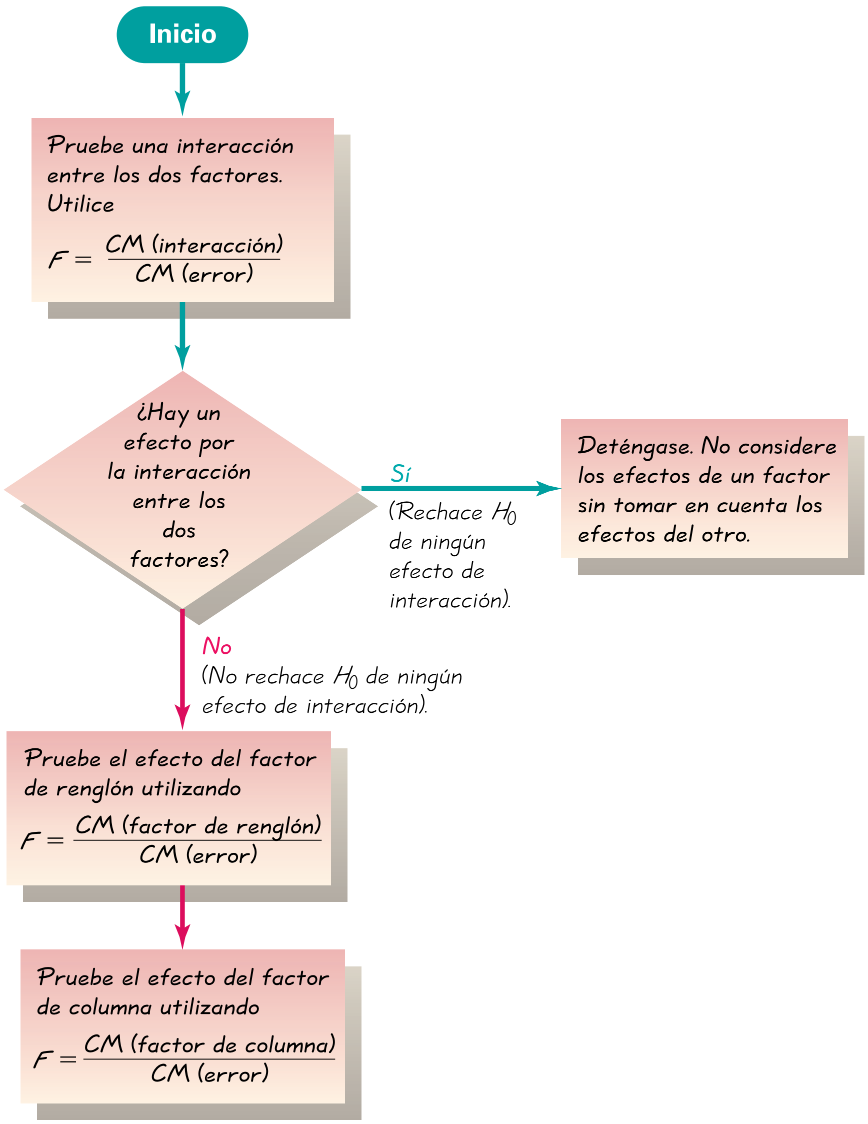


FIGURA 11-3 **Procedimiento del ANOVA de dos factores**

Procedimiento del ANOVA de dos factores (véase la figura 11-3)

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 11-4, obtenemos el siguiente estadístico de prueba:

$$F = \frac{\text{CM(interacción)}}{\text{CM(error)}} = \frac{10{,}521{,}034}{9{,}028{,}477} = 1.17$$

Interpretación: El valor P correspondiente aparece en los resultados de Minitab como 0.329, por lo que no rechazamos la hipótesis nula de ninguna interacción entre los dos factores. No parece que los tiempos de la carrera de maratón se afecten por una interacción entre el género y la edad.

Paso 2: *Efectos renglón/columna*: Si rechazamos la hipótesis nula de ninguna interacción entre factores, entonces debemos detenernos aquí; no habremos de proceder con las dos pruebas adicionales. (Si existe una interacción entre los factores, no consideremos los efectos de alguno de los factores sin tomar en cuenta los del otro).

Si no rechazamos la hipótesis nula de ninguna interacción entre los factores, entonces tenemos que proceder a probar las siguientes dos hipótesis:

H_0: No hay efectos del factor de renglón (es decir, las medias de renglón son iguales).

H_0: No hay 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 que se identificaron en el paso 2.

Para el factor de renglón del género obtenemos

$$F = \frac{\text{CM(género)}}{\text{CM(error)}} = \frac{15{,}225{,}413}{9{,}028{,}477} = 1.69$$

Interpretación: Este valor no es significativo, ya que el valor P correspondiente aparece en los resultados de Minitab como 0.206. No rechazamos la hipótesis nula de que no hay efectos por el género. Es decir, el género del corredor no parece tener un efecto sobre el tiempo. Como el ganador de este tipo de maratones suele ser casi siempre un hombre, esperaríamos encontrar un efecto por género, pero no fue así. Quizá no hay valores muestrales suficientes para que los efectos se consideren significativos.

Para el factor de columna de la edad obtenemos

$$F = \frac{\text{CM(edad)}}{\text{CM(error)}} = \frac{46{,}043{,}490}{9{,}028{,}477} = 5.10$$

Interpretación: Este valor es significativo, ya que el valor P correspondiente se indica como 0.014. (Con un valor P de 0.014 tenemos una significancia al nivel de 0.05, pero no al nivel de 0.01). Por lo tanto, rechazamos la hipótesis nula de ningún efecto de la edad. Parece que la edad del corredor produce un efecto sobre el tiempo. Con base en los datos muestrales de la tabla 11-4, concluimos que los tiempos parecen tener medias desiguales en las diferentes categorías de edad, pero los tiempos parecen con medias iguales en ambos géneros.

Caso especial: una observación por celda y ninguna interacción La tabla 11-4 contiene cinco observaciones por celda. Si nuestros datos muestrales consisten únicamente en una observación por celda, perderemos CM (interacción), SC (interacción) y gl (interacción), ya que dichos valores se basan en varianzas muestrales que se calculan para cada celda individual. Si existe sólo una observación por celda, no hay variación dentro de las celdas individuales, por lo que dichas varianzas muestrales no pueden calcularse. Cuando tenemos una observación por celda, procedemos de la siguiente manera: *si parece razonable suponer (con base en el conocimiento sobre 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 11-4. Si usamos esos primeros valores, las dos medias por renglón son 14,273.3 y 16,339.3. ¿Es esta diferencia significativa como para sugerir que hay un efecto por el género? De nuevo, si utilizamos únicamente el primer valor de cada celda, las medias de las tres columnas son 15,008.0, 15,017.0 y 15,894.0. ¿Son significativas tales diferencias como para sugerir que hay un efecto por la edad? Es razonable creer que los tiempos de los corredores de la maratón no se ven afectados por la interacción entre el género y la edad. (Si creemos que existe una interacción, el método descrito aquí no se aplica). A continuación se presenta la pantalla de resultados de Minitab para los datos de la tabla 11-4, únicamente con el primer valor de cada celda.

Minitab

Analysis of Variance for TIME

Source	DF	SS	MS	F	P
GENDER	1	6402534	6402534	8.88	0.097
AGE	2	1036137	518069	0.72	0.582
Error	2	1441476	720738		
Total	5	8880147			

Primero empleamos los resultados de la pantalla de Minitab para probar la hipótesis nula de ningún efecto del factor de renglón del género.

$$F = \frac{\text{CM}(\text{género})}{\text{CM}(\text{error})} = \frac{6{,}402{,}534}{720{,}738} = 8.88$$

Este estadístico de prueba no es significativo, porque el valor P correspondiente en la pantalla de Minitab es 0.0972. No rechazamos la hipótesis nula; parece que los tiempos de la maratón no se ven afectados por el género del corredor.

Ahora utilizamos la pantalla de Minitab para probar la hipótesis nula de ningún efecto del factor de columna de la categoría de edad. El estadístico de prueba es

$$F = \frac{\text{CM}(\text{edad})}{\text{CM}(\text{error})} = \frac{518{,}069}{720{,}738} = 0.72$$

Este estadístico de prueba no es significativo, ya que el valor P correspondiente que se dio en Minitab es 0.582. No rechazamos la hipótesis nula, de forma que parece que el tiempo de los corredores no es afectada por la edad del corredor. Con el uso del primer valor de cada celda, concluimos que los tiempos de los corredores no parecen verse afectados por el género ni por la edad, pero cuando tomamos cinco valores de cada celda, concluimos que los tiempos parecen ser afectados por la categoría de edad. Tal 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, a la vez que omitimos los cálculos y las fórmulas manuales, que son bastante engorrosos.

Utilizando la tecnología

STATDISK Hasta el momento en que se escribe esto, no se incluía un módulo para el análisis de varianza de dos factores en STATDISK. Verifique la última versión en el sitio www.pearsoneducacion.net/triola para determinar si ya existe.

Minitab Primero ingrese todos los valores muestrales en la columna C1. Introduzca los números por renglón correspondientes en la columna C2. Introduzca 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, ponga etiquetas 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 las etiquetas correspondientes al conjunto de datos en la columna A y el renglón 1, como en este ejemplo, que corresponde a la tabla 11-4:

	A	B	C	D
1		21-29	30-39	40 o más
2	Hombre	13,615	14,677	14,528
3	Hombre	18,784	16,090	17,034
⋮	⋮	⋮	⋮	⋮

Después de ingresar los datos muestrales y las etiquetas, 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 11-4, ingrese A1:D11. Para "rows per sample", introduzca el número de valores en cada celda; ingrese 5 para los datos de la tabla 11-4. Haga clic en **OK.**

Para tablas de dos factores con exactamente un dato por celda, no se requieren las etiquetas. 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 etiquetas en el rango de entrada. Haga clic en **OK.**

TI-83 Plus El programa A1ANOVA de la calculadora TI-83 Plus puede bajarse del CD-ROM que se incluye con este libro. Seleccione el archivo del *software*. El programa debe bajarse a la calculadora y los datos muestrales ingresarse como una matriz D con tres columnas. Presione **2nd** y la tecla $\boldsymbol{x^{-1}}$. Muévase a la derecha hasta **Edit**, luego hacia abajo hasta [**D**], ahora presione **ENTER** y proceda a ingresar el número total de valores de datos, que será 3 (para las tres columnas). La primera columna de D lista todos los datos muestrales, la segunda lista el número de renglón correspondiente y la tercera 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, que seguirá el valor *P* correspondiente. (Es necesario presionar **ENTER** para ver el resto de los resultados). F (AB) es el estadístico de prueba *F* para el efecto interacción, al cual le sigue por el valor *P* correspondiente.

11-3 Destrezas y conceptos básicos

Interpretación de una pantalla de resultados de computadora. *Algunos de los ejercicios 1 a 7 requieren la pantalla de resultados de Minitab, acerca de las cantidades de pesticida DDT que se detectan en halcones en tres categorías diferentes (joven, de mediana edad, viejo) en tres lugares distintos (Estados Unidos, Canadá, Región Ártica). El conjunto de datos se incluye con el programa de Minitab, en el archivo FALCON.MTW.*

Minitab

```
Analysis of Variance for DDT
Source        DF        SS        MS         F        P
Site           2  17785.41   8892.70   2581.75    0.000
Age            2   1721.19    860.59    249.85    0.000
Interaction    4     17.70      4.43      1.28    0.313
Error         18     62.00      3.44
Total         26  19586.30
```

1. **Significado del ANOVA de dos factores** El método de esta sección se conoce como análisis de varianza de dos factores o ANOVA de dos factores. ¿Por qué se emplea el término *dos factores*? ¿Por qué se utiliza el término *análisis de varianza*?

2. **¿Por qué no ANOVA de un factor?** La pantalla de Minitab es el resultado de las cantidades de DDT que se detectan en halcones, las cuales se dividen en nueve celdas de acuerdo con un factor de lugar y otro factor de la edad del halcón. Cada celda incluye tres mediciones de DDT. ¿Por qué no es posible realizar un análisis exhaustivo de los datos ejecutando sencillamente dos pruebas separadas con un ANOVA de un factor (descrito en la sección 11-2), de manera que una prueba incluya las diferencias en los lugares y la otra prueba las diferencias en la edad? Es decir, ¿por qué se requiere de un ANOVA de dos factores, en lugar de dos aplicaciones separadas de un ANOVA de un factor?

3. **Efecto de interacción** Suponga que un análisis de varianza de dos factores revela que hay un efecto significativo de una interacción entre dos factores. ¿Por qué *no* debemos proceder a probar el efecto a partir del factor de renglón?

4. **¿Por qué no utilizar un ANOVA de dos factores?** ¿Por qué no es posible utilizar el método del análisis de varianza de dos factores con las tablas de dos factores tal como se describió en la sección 10-3?

5. **Efecto de interacción** Remítase a los resultados de Minitab y pruebe la hipótesis nula de que las cantidades de DDT no se ven afectadas por una interacción entre el lugar y la edad. ¿Qué concluye?

6. **Efecto del lugar** Remítase a los resultados de Minitab y suponga que las cantidades de DDT que se detectan en los halcones no se ven afectadas por una interacción entre el lugar y la edad. ¿Existe suficiente evidencia para sustentar la aseveración de que el lugar tiene un efecto en la cantidad de DDT?

7. **Efecto de la edad** Remítase a los resultados de Minitab y suponga que las cantidades de DDT no se ven afectadas por una interacción entre el lugar y la edad. ¿Hay suficiente evidencia para sustentar la aseveración de que la edad produce un efecto en la cantidad de DDT?

Interpretación de una pantalla de resultados de computadora. *En los ejercicios 8 a 10, utilice los resultados de Minitab que se obtuvieron a partir de las puntuaciones que se listan en la siguiente tabla. Los datos muestrales son calificaciones del SAT en las secciones verbal y matemática del SAT-I, que se basan en estadísticos que reportó 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

Analysis of Variance for SAT

Source	DF	SS	MS	F	P
Gender	1	52635	52635	5.03	0.031
Ver/Math	1	6027	6027	0.58	0.453
Interaction	1	31528	31528	3.01	0.091
Error	36	376748	10465		
Total	39	466938			

8. Efecto de interacción Pruebe la hipótesis nula de que las calificaciones del SAT no se ven afectadas por una interacción entre el género y la prueba (verbal/matemáticas). ¿Qué concluye?

9. Efecto del género Suponga que las calificaciones del SAT no se ven afectadas por una interacción entre el género y el tipo de prueba (verbal/matemáticas). ¿Hay suficiente evidencia para sustentar la aseveración de que el género tiene un efecto sobre las calificaciones del SAT?

3. Efecto del tipo de prueba del SAT Suponga que las calificaciones del SAT no se ven afectadas por una interacción entre el género y el tipo de prueba (verbal/matemáticas). ¿Hay evidencia suficiente para sustentar la aseveración de que el tipo de prueba (verbal/matemáticas) produce un efecto sobre las calificaciones del SAT?

Interpretación de una pantalla de resultados de computadora. *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 en la cual se utilizaron 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 son igualmente difíciles de comprender. En la tabla original de las puntuaciones de la prueba de audición, cada celda contiene un dato. Los datos originales provienen de* 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 efecto de una interacción entre el sujeto y las listas en las puntuaciones de las pruebas de audición, ¿hay evidencia suficiente para sustentar la aseveración de que la selección del sujeto produce 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 lista de palabras Suponiendo que no existe efecto de una interacción entre el sujeto y las listas en las puntuaciones de las pruebas de audición, ¿hay 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?

T **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?

continúa

	Edad		
	Menor de 20	Entre 20 y 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. Consumo de combustible de automóviles La siguiente tabla lista el consumo de combustible en carretera (en millas/galón) del conjunto de datos 22 del Apéndice B. Suponga que el consumo de combustible no es afectado por la interacción entre el tipo de transmisión (manual o automática) y el número de cilindros. ¿Es afectado el consumo de combustible por el tipo de transmisión? ¿Es afectado el consumo de combustible por el número de cilindros?

	Cilindros		
	4	6	8
Manual	33	30	28
Automática	31	27	24

11-3 Más allá de lo básico

15. Transformaciones de datos Suponga que se utiliza un ANOVA de dos factores para analizar datos muestrales que constan de más de un dato por celda. ¿De qué manera se ven afectados los resultados del ANOVA en cada uno de los siguientes casos?
a. Se añade la misma constante a cada valor muestral.
b. Cada valor muestral se multiplica por la misma constante distinta de cero.
c. Se transpone el formato de la tabla, de manera que se intercambien los factores de renglón y de columna.
d. Se cambia el primer valor muestral de la primera celda, de forma que se convierte en un dato distante.

Repaso

En la sección 8-3 presentamos un procedimiento para probar la igualdad entre *dos* medias poblacionales, pero en la sección 11-2 utilizamos el análisis de varianza (o ANOVA) 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 a partir de poblaciones que se caracterizan según un solo factor. Las siguientes son características clave del análisis de varianza de un factor:

- El estadístico de prueba F se basa en el cociente de dos estimados diferentes de la varianza poblacional común σ^2, como se muestra a continuación.

$$F = \frac{\text{varianza entre muestras}}{\text{varianza dentro de muestras}} = \frac{CM(tratamiento)}{CM(error)}$$

- Los valores críticos de F se encuentran en la tabla A-5, pero nos enfocamos en la interpretación de los valores P que se incluyen como parte de un resultado por computadora.

En la sección 11-3 consideramos el análisis de varianza de dos factores con los datos que se categorizaron 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 de análisis de varianza de dos factores se resume en la figura 11-3 y requiere que primero probemos si hay una interacción entre los dos factores. Si no existe una interacción significativa, entonces procederemos 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 hay una observación por celda.

Por la naturaleza de los cálculos que se requieren a lo largo de este capítulo, ponemos énfasis en la interpretación de resultados por computadora.

Ejercicios de repaso

1. Beber y conducir El Associated Insurance Institute financia estudios de los efectos de las bebidas alcohólicas en los conductores. En uno de estos estudios se seleccionaron aleatoriamente tres grupos de hombres adultos para un experimento que pretendía medir los niveles de alcohol en la sangre después de consumir 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 tales 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

```
Analysis of Variance
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. Lugares, lugares, lugares La lista adjunta presenta precios de venta (en miles de dólares) de casas ubicadas en Long Beach Island, en Nueva Jersey. Se espera encontrar precios medios de venta diferentes en los distintos lugares. ¿Sustentan estos datos muestrales la aseveración de distintos precios medios de venta? Utilice un nivel de significancia de 0.05.

A un costado del mar:	235	395	547	469	369	279
Frente al mar:	538	446	435	639	499	399
A un costado de la bahía:	199	219	239	309	399	190
Frente a la bahía:	695	389	489	489	599	549

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 que se listan en la tabla adjunta. Los datos muestrales son estimados de los estudiantes de la longitud (en pies) 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

```
Analysis of Variance for LENGTH
Source        DF        SS        MS       F       P
GENDER         1      29.4      29.4    0.78   0.395
MAJOR          2      10.1       5.1    0.13   0.876
Interaction    2      14.1       7.1    0.19   0.832
Error         12     453.3      37.8
Total         17     506.9
```

3. Efecto de interacción Pruebe la hipótesis nula de que las longitudes que se estiman no son afectados por una interacción entre el género y el área.

4. Efecto del género Suponga que las longitudes estimadas no se ven afectadas por una interacción entre el género y el área. ¿Hay evidencia suficiente para sustentar la aseveración de que la longitud que se estima se ve afectada por el género?

5. Efecto del área Suponga que las longitudes que se estiman no se ven afectadas por una interacción entre el género y el área. ¿Hay evidencia suficiente para sustentar la aseveración de que la longitud que se estima se ve afectada por el área?

6. Contaminación de automóviles La tabla adjunta lista las cantidades de gases invernadero que emitieron diferentes automóviles en un año. (Véase el conjunto de datos 22 del Apéndice B). La pantalla de Minitab resulta de esta tabla.

a. Suponiendo que no existe un efecto de interacción, ¿hay evidencia suficiente para sustentar la aseveración de que las cantidades de gases invernadero que se emiten se ven afectadas por el tipo de transmisión (automática/manual)?

b. Suponiendo que no existe un efecto de interacción, ¿hay evidencia suficiente para sustentar la aseveración de que las cantidades de gases invernadero que se emiten se ven afectadas por el número de cilindros?

c. Con base en los resultados de los incisos *a* y *b*, ¿podemos concluir que la emisión de gases invernadero no se ve afectada por el tipo de transmisión o el número de cilindros? ¿Por qué?

Emisión de gases invernadero (toneladas/año)

	4 cilindros	6 cilindros	8 cilindros
Automática	10	12	14
Manual	10	12	12

Minitab

```
Analysis of Variance for GASES
Source      DF        SS        MS       F       P
TRANS        1     0.667     0.667    1.00   0.423
CYL          2     9.333     4.667    7.00   0.125
Error        2     1.333     0.667
Total        5    11.333
```

Ejercicios de repaso acumulativo

1. Estadísticas de lluvia en Boston Remítase a las cantidades de lluvia en Boston de los lunes, tal como se listan en el conjunto de datos 11 del Apéndice B.

a. Calcule la media.
b. Calcule la desviación estándar.
c. Calcule el resumen de los cinco números.
d. Identifique si existen datos distantes.
e. Construya un histograma.
f. Suponga que desea probar la hipótesis nula de que la cantidad media de lluvia es la misma los siete días la semana. ¿Puede utilizar un ANOVA de un factor? ¿Por qué sí o por qué no?
g. Con base en los datos muestrales, estime la probabilidad de que llueva en Boston un lunes seleccionado al azar.

2. Tratamiento M&M La tabla de abajo incluye 60 calificaciones del SAT, separadas en categorías de acuerdo con el color de los dulces M&M que se utilizaron como tratamiento. Las calificaciones del SAT se basan en datos del Consejo Universitario, en tanto que el elemento del color de los dulces M&M se basa en un capricho del autor.

a. Calcule la media de las 20 calificaciones del SAT en cada una de las tres categorías. ¿Parecen las tres medias ser aproximadamente iguales?
b. Calcule la mediana de las 20 calificaciones del SAT en cada una de las tres categorías. ¿Parecen las tres medianas ser aproximadamente iguales?
c. Calcule la desviación estándar de las 20 calificaciones del SAT en cada una de las tres categorías. ¿Parecen las tres desviaciones estándar ser aproximadamente iguales?
d. Pruebe la hipótesis nula de que no hay una diferencia entre la calificación media del SAT de los sujetos que se trataron con M&M rojos y la calificación media del SAT de sujetos tratados con M&M verdes.
e. Construya un estimado del intervalo de confianza del 95% de la puntuación media del SAT, para la población de sujetos que recibe el tratamiento con M&M rojos.
f. Pruebe la hipótesis nula de que las tres poblaciones (tratamientos con M&M rojos, verdes y azules) obtuvieron la misma calificación media del 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	1611	939	1004	821	915
	866	848	1408	793	1097	1244	996	1131	1039	1159

3. 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 libras y una desviación estándar de 1.09 libras (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 libras?
b. Si se seleccionan al azar a 16 bebés recién nacidos, ¿cuál es la probabilidad de que su peso medio sea mayor que 8.0 libras?
c. ¿Cuál es la probabilidad de que cada uno de los siguientes tres bebés tenga un peso al nacer mayor que 7.54 libras?

Actividades de cooperación en equipo

1. Actividad fuera de clase El *World Almanac and Book of Facts* incluye una sección que se llama "personalidades notables", con apartados específicos que se dedican a arquitectos, artistas, líderes de negocios, dibujantes, científicos sociales, líderes militares, filósofos, líderes políticos, científicos, escritores, compositores, animadores y otros. Diseñe y realice un estudio observacional que inicie con la selección de muestras de grupos selectos, seguido por una comparación de los periodos de vida de personas de distintas categorías. ¿Algunos grupos en particular parecen tener periodos de vida diferentes de los otros grupos? ¿Puede explicar tales 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, forme grupos de tres o cuatro miembros y utilice los datos de toda la clase para plantear las siguientes preguntas:
   - ¿Hay una diferencia significativa entre el estimado medio de hombres y el estimado medio de mujeres?
   - ¿Existe evidencia suficiente para rechazar la igualdad de los estimados medios en las diferentes áreas de estudio? Describa cómo se categorizaron las áreas de estudio.
   - ¿Tiene la interacción entre el género y el área de estudio un efecto sobre la longitud que se estima?
   - ¿Parece que el género tiene un efecto sobre la longitud que se estima?
   - ¿Parece que el área de estudio tiene un efecto sobre la longitud que se estima?

3. Actividad fuera de clase Forme grupos de tres o cuatro estudiantes. Cada grupo debe encuestar a otros estudiantes de la misma universidad y pedirles que identifiquen su área de estudio y su género. También podría incluir factores tales como el empleo (ninguno, de medio tiempo, de tiempo completo) y edad (menos de 21, 21-30, más de 30). Para cada sujeto que se encueste, determine la precisión de la hora de su reloj de pulso. Primero ponga su reloj a la hora correcta por medio de una fuente precisa y confiable ("Cuando escuche el tono, la hora es..."). Registre una hora positiva para los relojes que se adelantaron y una hora negativa para los relojes que se atrasaron. Utilice los datos muestrales para plantear preguntas como éstas:
   - ¿Parece que el género tiene un efecto sobre la precisión de los relojes de pulso?
   - ¿Ejerce el área de estudio algún efecto en la precisión de los relojes de pulso?
   - ¿Una interacción entre el género y el área de estudio tiene un efecto en la precisión de los relojes de pulso?

Proyecto tecnológico

Con respecto a los presidentes de Estados Unidos, los pontífices y los monarcas británicos a partir de 1690, la lista que se presenta a continuación incluye el número de años que vivieron después de que tomaron posesión, de ser elegidos o de su coronación. Utilice una gráfica de cuadro y el análisis de varianza para determinar si el tiempo de supervivencia de los grupos difiere. Realice el análisis de varianza por medio de STATDISK, Minitab, Excel o una calculadora TI-83 Plus, o por medio de otro programa estadístico de cómputo. Imprima copias de los resultados de computadora y anote sus observaciones y conclusiones.

Presidentes		Papas		Reyes y reinas	
Washington	10	Alejandro VIII	2	Jaime II	17
J. Adams	29	Inocencio XII	9	María II	6
Jefferson	26	Clemente XI	21	Guillermo III	13
Madison	28	Inocencio XIII	3	Ana	12
Monroe	15	Benedicto XIII	6	Jorge I	13
J. Q. Adams	23	Clemente XII	10	Jorge II	33
Jackson	17	Benedicto XIV	18	Jorge III	59
Van Buren	25	Clemente XIII	11	Jorge IV	10
Harrison	0	Clemente XIV	6	Guillermo IV	7
Tyler	20	Pío VI	25	Victoria	63
Polk	4	Pío VII	23	Eduardo VII	9
Taylor	1	León XII	6	Jorge V	25
Fillmore	24	Pío VIII	2	Eduardo VIII	36
Pierce	16	Gregorio XVI	15	Jorge VI	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		
Cleveland	24	Pablo VI	15		
Harrison	12	Juan Pablo I	0		
McKinley	4				
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				

Fuente: *Computer-Interactive Data Analysis*, de Lunn y McNeil, John Wiley & Sons.

de los DATOS a la DECISIÓN

Pensamiento crítico: ¿Debe usted aprobar este fármaco?

Los fármacos deben someterse a pruebas exhaustivas antes de aprobarse para su uso general. Además de probar sus reacciones adversas, también hay que probar su eficacia, por lo que el análisis de este tipo de resultados de pruebas suele incluir métodos estadísticos. Considere la creación del xinamine, un nuevo fármaco que se diseñó para disminuir el pulso. 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 que se trataron con xinamine en dosis de 10 miligramos y de un grupo de hombres que se trataron con xinamine en dosis de 20 miligramos. El gerente de producción del fármaco realiza 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 reporte afirma que el fármaco es eficaz, de acuerdo con 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 miligramos de xinamine tiene una media más baja de 66.2 en tanto que el grupo que se trató con dosis de 20 miligramos de xinamine tiene la media más baja de 65.2.

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

Análisis de resultados

Analice los datos utilizando los métodos de este capítulo. Con base en los resultados, ¿parece que hay evidencia suficiente para sustentar la aseveración de que el fármaco reduce el pulso? ¿Existen algunos problemas graves con el diseño de este experimento? Puesto que únicamente se incluyeron hombres en el experimento, ¿se aplican los resultados también a las mujeres? El gerente del proyecto comparó los pulsos postratamiento con el pulso medio de hombres adultos. ¿Existe una mejor forma 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.

PROYECTO DE INTERNET Análisis de varianza

Entre al sitio Web de este libro de texto en

http://www.pearsoneducacion.net/triola

Siga el vínculo del "Proyecto de Internet" de este capítulo. El proyecto describe antecedentes para experimentos en áreas tan variadas como el desempeño atlético, el proceso de etiquetar productos destinados al 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 que se estudiaron en este capítulo. Usted formulará las hipótesis apropiadas, después realizará y resumirá pruebas ANOVA.

La estadística @ en el trabajo

Joseph Marvin

Gerente de cartera y director de manejo de cartera y comercialización del grupo de bonos en State Street Global Advisors (SSGA)

Como gerente de cartera, Joseph se especializa en la comercialización de derivados de ingresos fijos. En su trabajo utiliza métodos estadísticos para evaluar valores relativos entre diversos instrumentos financieros. SSGA es una de las empresas de manejo de dinero más grandes de Estados Unidos, ya que maneja más de 580 mil millones de dólares en acciones.

"El conocimiento básico de la estadística es fundamental en finanzas".

¿En qué consiste su trabajo?

Soy gerente de cartera y director de manejo de cartera en State Street Global Advisors (SSGA). Ésta es una de las compañías de manejo de inversiones más grandes de Estados Unidos y es filial de State Street Corporation. Con más de 580 mil millones de dólares en acciones, la tarea principal de SSGA es el manejo de acciones de cuentas públicas y privadas de pensiones y jubilaciones.

¿Recomienda el estudio de la estadística a los universitarios de hoy? ¿Por qué?

La recomiendo por completo a TODOS los estudiantes universitarios. La estadística proporciona a los estudiantes fundamentos excelentes para tomar mejores decisiones. En puestos que se relacionan con economía y finanzas, se valora más a los individuos con conocimientos de estadística.

¿Qué conceptos de estadística utiliza?

Empleamos activamente el análisis de probabilidad y la prueba de hipótesis. También utilizamos análisis de regresión lineal y no lineal. Para la creación de cartera, utilizamos optimización de la varianza media. Empleamos estos estadísticos para evaluar el valor de bonos. Sin una buena comprensión de las bases y los fundamentos de la estadística, no sería capaz de cumplir con eficacia mis responsabilidades. El conocimiento básico de la estadística es fundamental en finanzas.

Por favor, describa un ejemplo específico que ilustre cómo el uso de la estadística contribuyó a mejorar un producto o servicio.

Para determinar qué bono se debe comprar, utilizamos pruebas simples de hipótesis. La diferencia en el rendimiento entre el beneficio de un bono corporativo con la madurez y la tasa libre de riesgos, que representa un bono del tesoro de Estados Unidos, representa una prima de riesgo de bono o una gama. La gama de un bono es lo que los gerentes utilizamos para comparar el valor de un bono con otro. Las gamas de los bonos tienden a que la "media se revierta" a lo largo del tiempo, lo que da como resultado una distribución casi normal. Si asumimos una reversión de la media y una distribución normal, podemos utilizar la prueba simple de hipótesis para buscar significancia estadística. En otras palabras, si la gama de un bono tiene una amplitud estadísticamente significativa en comparación con la gama de otro bono, cuando todo lo demás permanece igual, consideraríamos que el bono es barato. Una gama más amplia implica un rendimiento más alto. Cuando las gamas se acortan, el precio o valor del bono sube.

Los gerentes del mercado de bonos incrementan sus ganancias al comprar bonos con gamas que se espera se acorten. Buscamos comprar bonos con gamas que estén a 1.0 o 2.0 desviaciones estándar más amplias que el promedio y vender bonos con gamas que estén a 1.0 o 2.0 desviaciones estándar (más cortas) que el promedio. ¡Simples pruebas de hipótesis!

12

Estadística no paramétrica

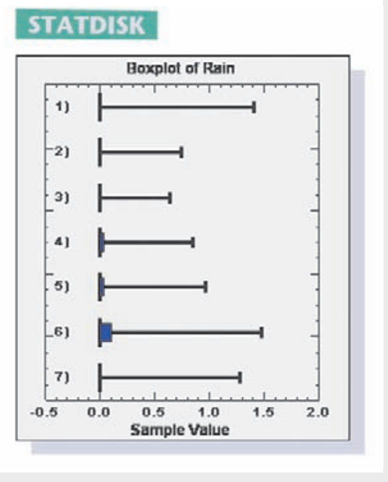
¿¡Llueve más durante los fines de semana!?

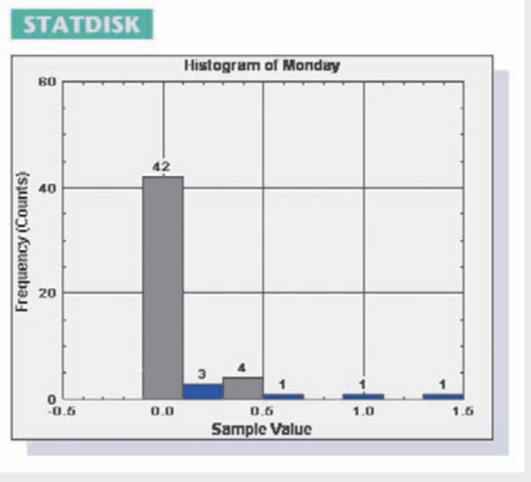
En un artículo para el Knight Ridder News Service, Usha Lee McFarling escribió que "sus peores temores por el clima son ciertos. Llueve más durante los fines de semana. Científicos que a través de muchos años reunieron gran cantidad de datos sobre la lluvia han descubierto un patrón claro y desalentador. Viernes, sábados y domingos son los días más lluviosos de la semana a todo lo largo de la costa este de Estados Unidos, desde Maine hasta Florida". La nota se refiere a un estudio que realizaron los científicos Randall S. Ceverny y Robert C. Balling de Arizona State University. Pero, ¿son correctas sus conclusiones? ¿Los datos se reportaron e interpretaron correctamente?

El conjunto de datos 11 del Apéndice B incluye cantidades de lluvia de un año reciente en Boston. Al seleccionar una ciudad para verificar el fenómeno de la lluvia durante el fin de semana, Boston debería ser una buena opción, puesto que se localiza en la costa este. Si utilizamos el STATDISK para calcular la cantidad de lluvia para cada día de la semana, obtendremos las gráficas de cuadro adjuntas. De la parte superior a la inferior, las gráficas de cuadro representan el lunes, el martes, . . . y el domingo. La gráfica de cuadro del lunes, en la parte superior, es peculiar, ya que parece estar desapareciendo del cuadro, pero las cantidades de lluvia del lunes tienen tantos ceros que el mínimo, el primer cuartil y la mediana son todos ceros, ocasionando que el cuadro se comprima hacia la izquierda. Mayores cantidades de lluvia en viernes, sábado y domingo serían visibles con las distribuciones situadas más a la derecha. ¿Es éste el caso en realidad? ¿Son realmente significativas tales diferencias?

Consideremos el uso de los métodos que se presentan en capítulos anteriores para investigar este tema. El análisis de varianza (sección 11-2) sería una buena opción, pero ese método requiere que las muestras provengan de poblaciones con una distribución normal. El histograma generado por el STATDISK para las cantidades de lluvia del lunes muestra con claridad que esos valores no provienen de una población distribuida normalmente. Los otros días muestran distribuciones similares que, a todas luces, no son normales. Una ventaja importante de los métodos que se analizan en este capítulo es que no requieren de una distribución normal o de cualquiera otra distribución en particular.

¿Es posible utilizar los datos de Boston para sustentar la aseveración de más lluvias durante los fines de semana? ¿Son significativamente diferentes las cantidades de lluvia para los diferentes días? Abordaremos tales preguntas más tarde en el capítulo.





12-1 Panorama general

A los métodos de estadística diferencial que se presentan en los capítulos 6, 7, 8, 9 y 11 se les llama *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. Estos métodos paramétricos por lo regular deben cumplir con algunas condiciones bastante estrictas, como el requisito de que los datos muestrales provengan de una población que se distribuya normalmente. Este capítulo introduce métodos no paramétricos, que están libres de tan estrictos requisitos.

Definiciones

Las **pruebas paramétricas** requieren supuestos acerca de la naturaleza o forma de las poblaciones involucradas; las **pruebas no paramétricas** no requieren supuestos acerca de las distribuciones poblacionales. En consecuencia, las pruebas de hipótesis no paramétricas suelen llamarse **pruebas de distribución libre**.

Aunque el término *no paramétrica* sugiere que la prueba no se basa en un parámetro, hay algunas pruebas no paramétricas que sí dependen de un parámetro, como la mediana. Sin embargo, las pruebas no paramétricas no requieren 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*. Las siguientes son las ventajas y desventajas principales de los métodos no paramétricos.

Ventajas de los métodos no paramétricos

1. Los métodos no paramétricos se aplican 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 poblaciones distribuidas normalmente.
2. A diferencia de los métodos paramétricos, los métodos no paramétricos con frecuencia se aplican 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; por lo tanto, son más fáciles de entender y aplicar.

Desventajas de los métodos no paramétricos

1. Los métodos no paramétricos tienden a desperdiciar información, pues los datos numéricos exactos suelen reducirse a una forma cualitativa. Por ejemplo, en la prueba del signo no paramétrica (descrita en la sección 12-2), las pérdidas de peso de las personas que se someten a una dieta se registran simplemente como signos negativos; las magnitudes reales de las pérdidas de peso se ignoran.
2. Las pruebas no paramétricas no son tan eficientes como las pruebas paramétricas, por lo que para una prueba no paramétrica generalmente necesitaremos

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, aunque la reducción en la eficacia puede compensarse con un tamaño muestral más grande. Por ejemplo, la sección 12-6 presentará un concepto que se llama *correlación de rangos*, con una tasa de eficacia de 0.91, cuando se compara con la correlación lineal que se presenta en el capítulo 9. Esto significa que si todas las demás cosas son iguales, la correlación de rangos no paramétrica requiere 100 observaciones muestrales, para obtener los mismos resultados que 91 observaciones muestrales que se analicen con la correlación lineal paramétrica, suponiendo que se satisfacen los requisitos más estrictos para la aplicación del método paramétrico. La tabla 12-1 lista los métodos no paramétricos que se presentan en este capítulo, junto con el método paramétrico correspondiente y la tasa de **eficiencia**. La tabla 12-1 muestra que varias pruebas no paramétricas tienen tasas de eficiencia por encima de 0.90, por lo que la eficiencia más baja no sería un factor crítico para elegir entre los métodos paramétricos y no paramétricos. Sin embargo, como las pruebas paramétricas no tienen tasas de eficiencia más altas que sus contrapartes no paramétricas, generalmente es mejor utilizar las pruebas paramétricas cuando los supuestos que se requieren se satisfacen.

Rangos

Las secciones 12-3 a 12-6 utilizan métodos que se basan en rangos, que ahora describiremos.

Definición

Los datos se *ordenan* cuando se acomodan de acuerdo con algún criterio, como de más pequeño a más grande, o de mejor a peor. Un **rango** es un número que se asigna a un elemento muestral individual de acuerdo con su orden en la lista ordenada. Al primer elemento se le asigna un rango de 1, al segundo elemento se le asigna un rango de 2, etcétera.

Tabla 12-1 Eficiencia: comparación de pruebas paramétricas y no paramétricas

Aplicación	Prueba paramétrica	Prueba no paramétrica	Tasa de eficiencia de prueba no paramétrica con población normal
Datos apareados de los datos muestrales	Prueba t o prueba z	Prueba del signo	0.63
		Prueba de rangos con signo de Wilcoxon	0.95
Dos muestras independientes	Prueba t o prueba z	Prueba de la suma de rangos de Wilcoxon	0.95
Varias muestras independientes	Análisis de varianza (prueba F)	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 comparación

EJEMPLO Los números 5, 3, 40, 10 y 12 pueden ordenarse (si se acomodan de menor a mayor) como 3, 5, 10, 12 y 40, con rangos de 1, 2, 3, 4 y 5, respectivamente:

5	3	40	10	12	Valores originales
3	5	10	12	40	Valores ordenados (que se acomodaron en orden)
↑	↑	↑	↑	↑	
1	2	3	4	5	Rangos

Manejo de empates en rangos: Si ocurre un empate en los rangos, el procedimiento habitual es calcular la media de los rangos que intervienen y luego asignar este rango medio a cada uno de los elementos empatados, como en el ejemplo siguiente.

EJEMPLO Los números 3, 5, 5, 10 y 12 tienen rangos dados de 1, 2.5, 2.5, 4 y 5, respectivamente. En este caso, los rangos 2 y 3 empataron; por lo tanto, calculamos la media de 2 y 3 (que es 2.5) y la asignamos a los valores que crearon el empate:

3	5	5	10	12	Valores originales
↑	↑	↑	↑	↑	
1	2.5	2.5	4	5	Rangos
	↑	↑			

2 y 3 están empatados

12-2 Prueba del signo

El objetivo principal de esta sección es entender el procedimiento de la *prueba del signo*, que es uno de los más sencillos de las pruebas no paramétricas.

Definición

Prueba del signo: una prueba no paramétrica (de distribución libre) que utiliza signos positivos y negativos para probar diferentes aseveraciones, incluyendo:

1. Aseveraciones que incluyen datos apareados de datos muestrales
2. Aseveraciones que incluyen datos nominales
3. Aseveraciones acerca de la mediana de una sola población

Concepto básico de la prueba del signo La idea básica que está detrás de 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 que se diseñó para disminuir la presión sanguínea. Si se trata a 100 sujetos y 51 de ellos experimentan una presión sanguínea más baja, mientras que los otros 49 tienen una presión sanguínea incrementada, el sentido

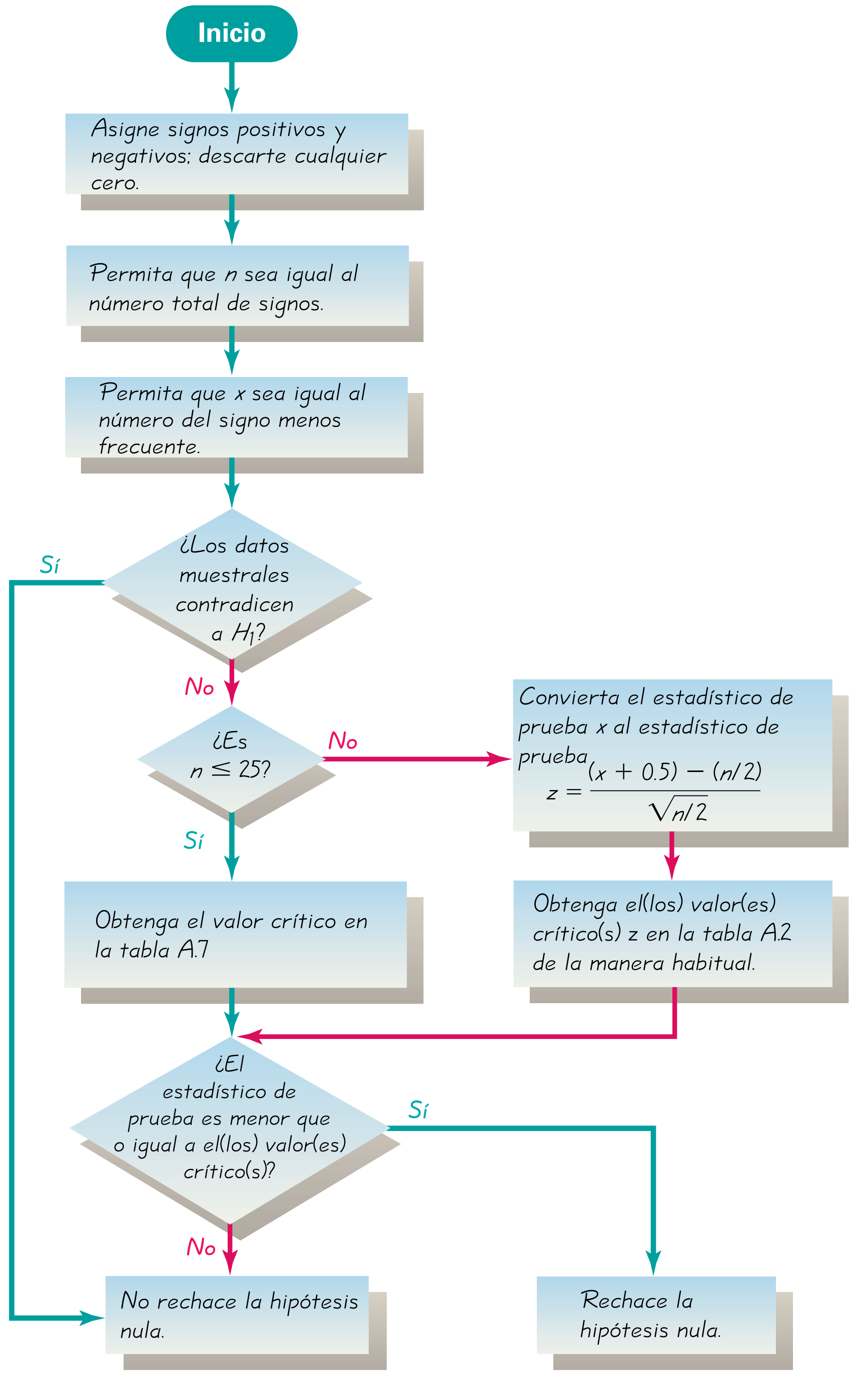


FIGURA 12-1
Procedimiento de la prueba del signo

común sugiere que no hay evidencia suficiente para decir que el fármaco es eficaz, puesto que 51 disminuciones en 100 casos no son significativas. Pero ¿qué sucede con 52 disminuciones y 48 incrementos? ¿O con 90 disminuciones y 10 incrementos? La prueba del signo nos permite determinar cuándo son significativos este tipo de resultados.

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 12-1 resume el procedimiento de la prueba del signo, que se ilustrará con los ejemplos que siguen.

Prueba del signo

Supuestos

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)

Para $n > 25$: $$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

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 no sacar la conclusión incorrecta cuando un signo ocurre significativamente con más frecuencia que el otro, aunque los datos muestrales contradicen la hipótesis alternativa. Por ejemplo, suponga que estamos probando la aseveración de que una técnica de selección de género favorece a los niños, pero tenemos 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 de inmediato no rechazamos la hipótesis nula y no procedemos con la prueba del signo. La figura 12-1 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 pensar acerca de los datos y evitar la confianza a ciegas en cálculos o resultados de computadora.*

Aseveraciones que incluyen datos apareados

Cuando se utiliza la prueba del signo con datos que se ordenan en pares, convertimos los datos brutos en 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 que se encontró 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 tras 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 Medición de inteligencia en niños Las mediciones mentales de niños pequeños se hacen 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) listados en la tabla 12-2 (datos tomados de "Tower Building", de Johnson y Courtney, *Child Development*, vol. 3). Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que no hay diferencia entre los tiempos de la primera y segunda pruebas.

SOLUCIÓN La siguiente es la idea básica: si no hay diferencia entre los tiempos de la primera prueba y los tiempos de la segunda prueba, los números de signos positivos y negativos deben ser aproximadamente iguales. En la tabla 12-2 tenemos 12 signos positivos y 2 signos negativos. ¿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 perfilaron en la figura 12-1.

Pasos 1, 2 y 3: La hipótesis nula es la aseveración de no diferencia entre los tiempos de la primera y la segunda pruebas, en tanto que la hipótesis alternativa es la aseveración de que hay 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 12-2 Tiempos de construcción de torres con cubos

Niño	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Primera prueba	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
Segunda prueba	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15
Signo de la diferencia	0	+	+	+	+	+	+	−	−	+	+	+	+	+	+

Paso 6: El estadístico de prueba x es el número de veces que ocurre el signo menos frecuente. La tabla 12-2 incluye diferencias con 12 signos positivos y dos signos negativos; descartamos el único caso con una diferencia de cero. Permitimos que x sea igual al menor entre 12 y 2; por lo tanto, $x = 2$. Además, $n = 14$ (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 2 para $n = 14$ y $\alpha = 0.05$ en dos colas. (Véase la figura 12-1).

Paso 7: Con un estadístico de prueba de $x = 2$ y un valor crítico de 2, rechazamos la hipótesis nula de no diferencia. [Consulte la nota 2 de 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 $x = 2$ es menor que o igual al valor crítico de 2, rechazamos la hipótesis nula].

Paso 8: Hay suficiente evidencia para sustentar el rechazo de la aseveración de que la mediana de las diferencias es igual a 0; esto es, existe suficiente evidencia para sustentar el rechazo de la aseveración de que no hay una diferencia entre los tiempos de la primera prueba y los tiempos de la segunda prueba. Es la misma conclusión que se alcanzaría utilizando la prueba paramétrica t con los datos apareados de la sección 8-4, aunque los resultados de la prueba del signo no siempre coinciden con los resultados de la prueba paramétrica.

Asistencia a clases y calificaciones

En un estudio entre 424 estudiantes universitarios 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 recibieron, en general, calificaciones de B o superiores. El estudio también mostró que los estudiantes que se sientan al frente en el salón de clases significativamente obtienen mejores calificaciones.

Aseveraciones que incluyen datos nominales

Recuerde que los datos nominales consisten sólo en nombres, etiquetas o categorías. Aunque dichos conjuntos de datos nominales limitan los cálculos posibles, se identifica la *proporción* de datos muestrales que pertenece a una categoría en particular y se prueban aseveraciones acerca de la proporción poblacional p correspondiente. El siguiente ejemplo utiliza datos nominales que consisten en el género (hombre/mujer). La prueba del signo se utiliza representando a los hombres con signos positivos (+) y a las mujeres con signos negativos (−). (Créanme, los signos se eligieron arbitrariamente). También observe el procedimiento para manejar casos en los que $n > 25$.

EJEMPLO Discriminación por género La cadena de restaurantes Hatters recibió acusaciones de discriminación por género porque sólo contrató a 30 hombres junto a 70 mujeres solicitantes. Una representante de la compañía aceptó que los solicitantes calificados son aproximadamente la mitad hombres y la mitad mujeres, pero además asevera que "Hatters no discrimina y el hecho de que 30 de los últimos 100 empleados nuevos sean hombres es sólo una casualidad". Utilice la prueba del signo con un nivel de significancia de 0.05 y pruebe la hipótesis nula de que esta compañía contrata a hombres y a mujeres por igual.

SOLUCIÓN Permita que p denote la proporción poblacional de hombres contratados. La aseveración de no discriminación implica que las proporciones de hombres y mujeres contratados son iguales a 0.5; entonces, $p = 0.5$. Por lo tanto, las hipótesis nula y la alternativa pueden establecerse de la siguiente manera:

H_0: $p = 0.5$ (la proporción de hombres contratados es igual a 0.5)

H_1: $p \neq 0.5$

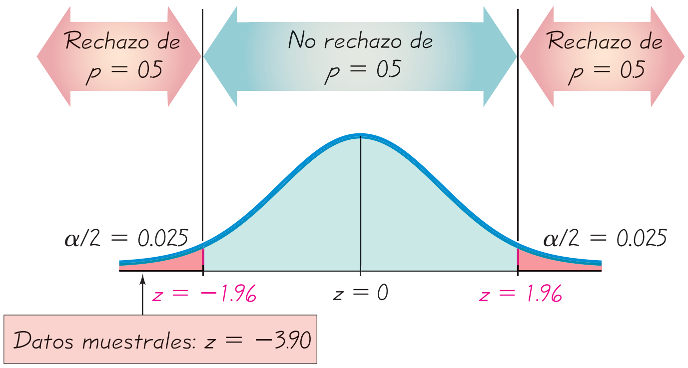


FIGURA 12-2 **Prueba de la aseveración de prácticas de contratación injustas**

Si denotamos a los hombres contratados con $+$ y a las mujeres contratadas con $-$, tenemos 30 signos positivos y 70 signos negativos. Ahora remítase al procedimiento de prueba del signo que se resume en la figura 12-1. El estadístico de prueba x es el menor entre 30 y 70; entonces, $x = 30$. Esta prueba es de dos colas, puesto que un número desproporcionadamente bajo de cualquier género nos causará el rechazo de la aseveración de igualdad. Los datos muestrales no contradicen la hipótesis alternativa, ya que 30 y 70 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 12-1, notamos que el valor de $n = 100$ es superior a 25, por lo cual el estadístico de prueba x se convierte (utilizando una corrección por continuidad) al estadístico de prueba z como sigue:

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

$$= \frac{(30 + 0.5) - \left(\frac{100}{2}\right)}{\frac{\sqrt{100}}{2}} = -3.90$$

Con $\alpha = 0.05$ en una prueba de dos colas, los valores críticos son $z = \pm 1.96$. El estadístico de prueba $z = -3.90$ es menor que -1.96 (véase la figura 12-2), por lo que rechazamos la hipótesis nula de que la proporción de hombres contratados es igual a 0.5. Hay suficiente evidencia de muestra para justificar el rechazo de la aseveración de que las prácticas de contratación son justas, con proporciones de hombres contratados y mujeres contratadas igual a 0.5. Parece que esta compañía discrimina por no contratar proporciones iguales de hombres y de mujeres.

Aseveraciones acerca de la mediana de una sola población

El siguiente ejemplo ilustra el procedimiento para utilizar la prueba del signo en la prueba de una aseveración acerca de la mediana de una sola población. Observe cómo los signos positivos y negativos se basan en el valor que se asevera para la mediana.

EJEMPLO Temperaturas corporales El conjunto de datos 4 del Apéndice B incluye temperaturas corporales medidas en adultos. Utilice las 106 temperaturas listadas para las 12:00 AM 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 encima de 98.6°F y 15 sujetos con temperaturas iguales a 98.6°F.

SOLUCIÓN La aseveración de que la mediana es menor que 98.6°F es la hipótesis alternativa, mientras la hipótesis nula es la aseveración de que la mediana es igual a 98.6°F.

H_0: La mediana es igual a 98.6°F. (mediana $=$ 98.6°F)
H_1: La mediana es menor que 98.6°F. (mediana $<$ 98.6°F)

Siguiendo el procedimiento que se perfiló en la figura 12-1, descartamos los 15 ceros, utilizamos el signo negativo ($-$) para denotar cada temperatura por debajo de 98.6°F y utilizamos el signo positivo ($+$) para denotar cada temperatura por encima de 98.6°F. Así, 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 presentan un conflicto con la hipótesis alternativa, terminaríamos 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:

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

$$= \frac{(23 + 0.5) - \left(\frac{91}{2}\right)}{\frac{\sqrt{91}}{2}} = -4.61$$

En la 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 12-3 vemos 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.

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 da como resultado un estadístico de prueba de $t = -6.611$ con un valor P de 0.000000000813. Puesto que el valor P de la prueba del signo no es tan bajo como

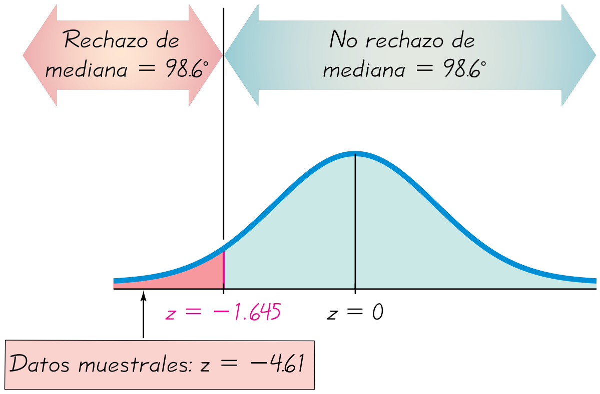


FIGURA 12-3 Prueba de la aseveración de que la mediana es menor que 98.6°F

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, parcialmente porque la prueba del signo utiliza sólo información acerca de la *dirección* de los datos, ignorando las *magnitudes* de los valores de los datos. La siguiente sección introduce la prueba de rangos con signo de Wilcoxon, que supera con creces tal 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 5-6 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 4-4 vimos que $\mu = np$ y $\sigma = \sqrt{npq}$ para distribuciones de probabilidad binomial. Puesto que la prueba del signo supone que $p = q = 1/2$, satisfacemos los prerrequisitos 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 la distribución de probabilidad continua, un valor discreto como 10 se representa en realidad por el intervalo de 9.5 a 10.5. Ya que x representa el signo menos frecuente, actuamos conservadoramente interesándonos sólo por $x + 0.5$. Así obtenemos el estadístico de prueba z, como aparece en la ecuación y en la figura 12-1.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **Sign Test**. Elija la opción **Given Number of Signs** si conoce el número de signos positivos y negativos, o **Given Pairs of Values** si prefiere ingresar pares de datos asociados. Después de realizar las entradas que se requieren 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 pares de datos asociados o el número de signos positivos y negativos. (Para detalles véase el manual *Minitab Student Laboratory Manual and Workbook*). Seleccione **Stat**, y luego **Nonparametrics** y **1-Sample Sign**. Haga clic en el botón para **Test Median**. Ingrese el valor de la mediana y elija el tipo de prueba, luego haga clic en **OK**. Minitab producirá el valor *P*; por lo tanto, rechace la hipótesis nula si el valor *P* es menor que o igual al nivel de significancia. De otra forma, no rechace la hipótesis nula.

Excel Excel no tiene una función predeterminada a la prueba del signo, pero usted puede utilizar la función BINOMDIST del programa 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 entonces **BINOMDIST**. En el cuadro de diálogo, ingrese primero *x*, luego el número de ensayos *n* y una probabilidad de 0.5. Teclee **TRUE** en el cuadro para *cumulative*. El valor resultante es la probabilidad de obtener *x* o menos éxitos entre *n* ensayos. *Duplique 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 que o igual al nivel de significancia. De no ser así, no rechace la hipótesis nula.

TI-83 Plus La calculadora TI-83 Plus no tiene una función predeterminada para la prueba del signo, pero usted 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. *Duplique 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 que o igual al nivel de significancia. De otra forma, no rechace la hipótesis nula.

12-2 Destrezas y conceptos básicos

En los ejercicios 1 a 4, suponga que los datos apareados dan como 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.

1. Signos positivos: 10; signos negativos: 5; empates: 3

2. Signos positivos: 6; signos negativos: 16; empates: 2

3. Signos positivos: 50; signos negativos: 40; empates: 5

4. Signos positivos: 10; signos negativos: 30; empates: 3

En los ejercicios 5 a 16 utilice la prueba del signo.

5. Prueba para la diferencia entre estaturas de hombres reportadas y medidas Como parte de la National Health and Nutrition Examination Survey, realizada por el Department of Health and Human Services de Estados Unidos, se obtuvieron las estaturas y medidas de hombres de 12 a 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 hombres de 12 a 16 años de edad? Utilice un nivel de significancia de 0.05.

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

6. Prueba para la diferencia entre estaturas reportadas y medidas La tabla de abajo lista datos apareados de estaturas que se midieron de 12 hombres estudiantes de estadística. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que no hay diferencia entre las estaturas reportadas y las estaturas medidas.

Estatura reportada	68	74	82.25	66.5	69	68	71	70	70	67	68	70
Estatura medida	66.8	73.9	74.3	66.1	67.2	67.9	69.4	69.9	68.6	67.9	67.6	68.8

7. Prueba para temperatura corporal media de 98.6°F En una clase de estadística, se le pide a una estudiante del curso propedéutico de medicina que realice un proyecto de clase. Inspirada por las temperaturas corporales del conjunto de datos 4 en el 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, encuentra que sólo alcanzará a reunir datos de 12 personas. Después de planear con cuidado un procedimiento para obtener una muestra aleatoria de 12 adultos saludables, 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 dichas 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

8. Prueba para mediana de peso bajo La Prince County Bottling Company surte botellas de limonada que se etiquetan de 12 onzas. Cuando el Prince County Department of Weights and Measures prueba una muestra aleatoria de botellas, obtiene las cantidades que se listan abajo. Si utilizamos un nivel de significancia de 0.05, ¿existe suficiente evidencia para documentar la acusación de que la compañía embotelladora está engañando a los consumidores dando cantidades con una mediana menor que 12 onzas?

11.4 11.8 11.7 11.0 11.9 11.9 11.5 12.0 12.1 11.9 10.9 11.3 11.5 11.5 11.6

9. Datos nominales: encuesta de votantes En una encuesta que se aplicó a 1002 personas, 701 dijeron que votaron en la elección presidencial reciente (según datos del ICR Research Group). ¿Hay suficiente evidencia para sustentar la aseveración de que la mayoría de las personas votaron en la elección?

10. Datos nominales: tabaquismo y parches de nicotina En un estudio de 71 fumadores que intentaban dejar de fumar con la terapia de parches de nicotina, 41 siguieron fumando un año después del tratamiento (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 0.05 y pruebe la aseveración de que, de los fumadores que tratan de dejar el hábito con la terapia de parches de nicotina, la mayoría siguieron fumando un año después del tratamiento.

11. Prueba de la mediana del volumen de latas de Coca Cola Remítase al conjunto de datos 17 en el Apéndice B y utilice los volúmenes de Coca Cola clásica. Pruebe la aseveración de que las latas de Coca Cola clásica tienen volúmenes con una mediana mayor que 12 onzas. ¿Parece que las latas de Coca Cola se llenan correctamente?

12. Prueba de la mediana de la cantidad de azúcar Dominó en sobres Remítase al conjunto de datos 28 del Apéndice B y utilice los datos muestrales para probar la aseveración de que la mediana de la cantidad de azúcar en los sobres es igual a 3.5 onzas. ¿Parece que los sobres de azúcar se llenan correctamente?

13. Prueba de la mediana del intervalo del tiempo del géiser Old Faithful Remítase al conjunto de datos 13 en el Apéndice B. Pruebe la aseveración de que los intervalos entre las erupciones del géiser Old Faithful tienen una mediana mayor a 77 minutos, que es la mediana que se registró desde hace 30 años.

14. Prueba de la diferencia entre temperaturas pronosticadas y reales Remítase al conjunto de datos 10 en el Apéndice B y utilice las temperaturas máximas reales y el pronóstico de temperaturas máximas de tres días. ¿Parece haber una diferencia?

12-2 Más allá de lo básico

15. Procedimientos para manejar empates En el procedimiento de la prueba del signo descrito en esta sección, excluimos los empates (representados por 0 en lugar de un signo de $+$ o de $-$). Un segundo método consiste en tratar a la mitad de los ceros como signos positivos y a la mitad como negativos. (Si el número de ceros es impar, excluya uno para que se dividan 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, pues cualquier signo sustenta la hipótesis nula. Suponga que en el uso de la prueba del signo para una aseveración de que el valor de la mediana es menor que 100, tenemos 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.

16. Cálculo de valores críticos La tabla A-7 lista valores críticos para alternativas limitadas de α. Utilice la tabla A-1 para añadir una nueva columna en la tabla A-7 (bajando hasta $n = 15$) que representa un nivel de significancia de 0.03 en una cola o de 0.06 en dos colas. Para cualquier n particular, utilice $p = 0.5$, ya que la prueba del signo requiere el supuesto de que P (signo positivo) $= P$ (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.

17. Error de aproximación normal Entre sus últimos 54 empleados nuevos, la compañía Compulife.com contrató a 18 mujeres. De los solicitantes, alrededor de la mitad son mujeres y la otra mitad hombres, y todos son calificados. Utilizando un nivel de significancia de 0.01 con la prueba del signo, ¿hay suficiente evidencia para acusar de favoritismo? ¿La conclusión cambia si se utiliza la distribución binomial en lugar de la distribución normal aproximada?

12-3 Prueba de rangos con signo de Wilcoxon para datos apareados

En la sección 12-2 utilizamos la prueba del signo para analizar tres tipos diferentes de datos, incluyendo datos muestrales consistentes en datos apareados. La prueba del signo incluyó sólo los signos de las diferencias y no sus magnitudes reales (qué tan grandes son los números). Esta sección introduce la *prueba de rangos con signo de Wilcoxon*, que también se usa con datos muestrales apareados. Mediantel uso de rangos, esta prueba toma en cuenta las magnitudes de las diferencias. (Véase la sección 12-1 para una descripción de los rangos). Puesto que la prueba de rangos con signo de Wilcoxon incorpora y utiliza más información que la prueba del signo, tiende a proporcionar conclusiones que reflejan mejor la verdadera naturaleza de los datos.

Definición

Prueba de rangos con signo de Wilcoxon: Una prueba no paramétrica que utiliza rangos ordenados de datos muestrales consistentes en datos apareados. Se usa para probar las diferencias en las distribuciones poblacionales, por lo que las hipótesis nula y alternativa son las siguientes:

H_0: Las dos muestras provienen de poblaciones con la misma distribución.
H_1: Las dos muestras provienen de poblaciones con distribuciones diferentes.

(La prueba de rangos con signo de Wilcoxon también resulta útil para probar la aseveración de que una muestra proviene de una población con una mediana específica. Véase el ejercicio 9 para dicha aplicación).

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. Guarde 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 más baja a la más alta y reemplace las diferencias por el valor del rango correspondiente (como se describe en la sección 12-1). Cuando las diferencias tengan el mismo valor numérico, asígneles la media de los rangos implicados en el empate.

Paso 3: Adjunte 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, la suma de los rangos positivos.

Paso 5: Permita que T sea la *más pequeña* de las dos sumas que se calcularon en el paso 4. Es posible utilizar cualquier suma, aunque para simplificar el procedimiento seleccionamos arbitrariamente la más pequeña de las dos sumas. (Véase la notación para T en el cuadro adjunto).

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 cuadro adjunto.

Paso 8: Cuando plantee la conclusión, rechace la hipótesis nula si los datos muestrales le llevan a un estadístico de prueba que está en la región crítica, esto es, cuando el estadístico de prueba es menor que o igual al(los) valor(es) crítico(s). De otra forma, no rechace la hipótesis nula.

Prueba de rangos con signo de Wilcoxon

Supuestos

1. Los datos consisten en datos apareados que se seleccionaron aleatoriamente.
2. La población de las diferencias (calculadas 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).

continúa

Notación

Véanse los pasos del procedimiento adjunto para calcular la suma de rangos T.

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.

Si $n > 30$, el estadístico de prueba es $z = \dfrac{T - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}}$

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.

EJEMPLO **Medición de la inteligencia en niños** Los datos en la tabla 12-3 son datos apareados de tiempos (en segundos) que se obtuvieron a partir de una muestra aleatoria de niños a quienes se les pidió que construyeran, usando cubos, una torre tan alta como fuera posible (según datos de "Tower Building", de Johnson y Courtney, *Child Development*, vol. 3). Este procedimiento se utiliza para medir la inteligencia de los niños. Utilice la prueba de rangos con signo de Wilcoxon con un nivel de significancia de 0.05 y pruebe la aseveración de que no hay diferencia entre los tiempos de la primera y de la segunda pruebas.

SOLUCIÓN Las hipótesis nula y alternativa son como sigue:

H_0: No hay diferencia entre los tiempos de la primera y de la segunda pruebas.

H_1: Existe una diferencia entre los tiempos de la primera y de la segunda pruebas.

Tabla 12-3 Tiempos para la construcción de torres con cubos

Niño	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Primera prueba	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
Segunda prueba	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15
Diferencias d	0	13	5	15	15	126	28	−2	−5	31	3	51	3	14	37
Rangos de |diferencias|		6	4.5	8.5	8.5	14	10	1	4.5	11	2.5	13	2.5	7	12
Rangos con signo		6	4.5	8.5	8.5	14	10	−1	−4.5	11	2.5	13	2.5	7	12

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 que ya se presentó en esta sección.

Paso 1: En la tabla 12-3, el renglón de diferencias se obtiene calculando esta diferencia para cada par de datos:

$$d = \text{tiempo de la primera prueba} - \text{tiempo de la segunda prueba}$$

Paso 2: Ignorando sus signos, ordenamos los rangos de las diferencias absolutas de la más baja a la más alta. Note que los empates en los rangos se manejan asignando la media de los rangos implicados a cada uno de los valores empatados y las diferencias de 0 se descartan.

Paso 3: El renglón inferior de la tabla 12-3 se crea insertando a cada rango el signo de la diferencia correspondiente. Si en realidad no hay diferencia entre los tiempos de la primera prueba y los tiempos de la segunda prueba (como en la hipótesis nula), esperamos que el número de rangos positivos sea aproximadamente igual al número de 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.

Suma de los valores absolutos de los rangos negativos: 5.5
Suma de los rangos positivos: 99.5

Paso 5: Permitiendo que T sea la menor de las dos sumas calculadas en el paso 4, encontramos que $T = 5.5$.

Paso 6: Permitiendo que n sea el número de pares de datos para los que la diferencia d no es 0, tenemos $n = 14$.

Paso 7: Puesto que $n = 14$, tenemos que $n \leq 30$, por lo cual utilizamos un estadístico de prueba de $T = 5.5$ (y no calculamos un estadístico de prueba z). Además, puesto que n ≤ 30, utilizamos la tabla A-8 para encontrar el valor crítico de 21.

Paso 8: El estadístico de prueba $T = 5.5$ es menor que o igual al valor crítico de 21, por lo que rechazamos la hipótesis nula. Parece que hay una diferencia entre los tiempos de la primera prueba y los tiempos de la segunda prueba.

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, hay otros casos en los que no concuerdan.

Fundamentos: En este ejemplo los rangos sin signo de 1 hasta 14 tienen un total de 105, por lo que, si no hay diferencias significativas, cada uno de los dos totales de rangos con signo debe ser de alrededor de $105 \div 2$, o 52.5. Esto es, los rangos negativos y los rangos positivos deberían repartirse como 52.5-52.5 o algo cercano, tal como 51-54. La tabla de valores críticos muestra que a un nivel de significancia de 0.05, con 14 pares de datos, un reparto de 21-84 representa una desviación significativa de la hipótesis nula y cualquier reparto que se separe más (como 20-85 o 5.5-99.5) también representará una desviación significativa de la hipótesis nula. Por el contrario, repartos como 22-83 no representan desviaciones significativas de un reparto de 52.5-52.5, por lo que no justificarían el rechazo de la hipótesis nula. La prueba de rangos con signo de Wilcoxon se basa en el total del rango más bajo, por lo que en lugar de analizar los dos números que constituyen el reparto, consideramos sólo el número más bajo.

La suma $1 + 2 + 3 + \cdots + n$ 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

$$1^2 + 2^2 + 3^3 + \cdots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$$

La prueba de rangos con signo de Wilcoxon se utiliza sólo con datos apareados. La siguiente sección describirá una prueba de suma de rangos que puede aplicarse a dos conjuntos de datos independientes que no se asocian en pares.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **Wilcoxon Tests**, ahora **Signed-Ranks Test.** Proceda a ingresar los datos muestrales apareados. Haga clic en **Evaluate**. La pantalla del STATDISK incluirá el estadístico de prueba, el valor crítico y la conclusión.

Minitab Ingrese los datos apareados en las columnas C1 y C2. Haga clic en **Editor** y luego en **Enable Command Editor**. Ingrese el comando **LET C3 = C1 − C2**. Oprima la tecla **Enter**. Elija las opciones **Stat**, **Nonparametrics** y **1-Sample Wilcoxon**. Ingrese C3 para la variable y haga clic en el botón para **Test Median**. La pantalla del Minitab incluirá el valor **P**. Rechace la hipótesis nula de distribuciones iguales si el valor *P* es menor que o igual al nivel de significancia. No rechace la hipótesis nula si el valor *P* es mayor que el nivel de significancia.

Excel Excel no está programado para la prueba de rangos con signo de Wilcoxon.

TI-83 Plus La calculadora TI-83 Plus no está programada para la prueba de rangos con signo de Wilcoxon.

12-3 Destrezas y conceptos básicos

Uso de la prueba de rangos con signo de Wilcoxon. *En los ejercicios 1 y 2, remítase a los datos muestrales apareados que se presentan; utilice la prueba de rangos con signo de Wilcoxon para probar la aseveración de que ambas muestras provienen de poblaciones que tienen la misma distribución. Utilice un nivel de significancia de 0.05.*

1.

x	12	14	17	19	20	27	29	30
y	12	15	15	14	12	18	19	20

2.

x	8	6	9	12	22	31	34	35	37
y	8	8	12	17	29	39	24	47	49

Uso de la prueba de rangos con signo de Wilcoxon. *En los ejercicios 3 y 4, remítase a los datos muestrales para los ejercicios de la sección 12-2. En lugar de la prueba del signo, utilice la prueba de rangos con signo de Wilcoxon para probar la aseveración de que ambas muestras provienen de poblaciones con la misma distribución.*

3. Ejercicio 5 **4.** Ejercicio 6

En los ejercicios 5 a 8, utilice la prueba de rangos con signo de Wilcoxon.

5. Prueba de la diferencia entre mediciones en posición de sentado y acostado En un estudio de técnicas que se utilizan para medir los volúmenes de los pulmones, se reunieron datos fisiológicos de 10 sujetos. Los valores dados en la tabla vienen en litros y representan las capacidades residuales funcionales medidas de los 10 sujetos en posición de sentado y en posición supina (acostado). Con un nivel de significancia de 0.5, pruebe la aseveración de que no hay diferencias significativas entre las mediciones que se tomaron en las dos posiciones.

Sentado	2.96	4.65	3.27	2.50	2.59	5.97	1.74	3.51	4.37	4.02
Supino	1.97	3.05	2.29	1.68	1.58	4.43	1.53	2.81	2.70	2.70

Datos tomados de "Validation of Esophageal Balloon Technique at Different Lung Volumes and Postures", de Baydur, Cha y Sassoon, *Journal of Applied Physiology*, vol. 62, núm. 1.

6. Prueba de eficacia de fármacos El captopril es un fármaco que se diseñó para disminuir la presión sanguínea sistólica. Cuando se probó este fármaco en sujetos, sus lecturas de presión sanguínea sistólica (en milímetros de mercurio) se midieron antes y después de que se tomara el fármaco, con los resultados que se presentan en la tabla adjunta. ¿Hay suficiente evidencia para sustentar la aseveración de que el fármaco no surte efecto? ¿Parece que el captopril disminuye la presión sanguínea sistólica?

Sujeto	A	B	C	D	E	F	G	H	I	J	K	L
Antes	200	174	198	170	179	182	193	209	185	155	169	210
Después	191	170	177	167	159	151	176	183	159	145	146	177

Datos tomados de "Essential Hypertension: Effect of an Oral Inhibitor of Angiotensin-Converting Enzyme", de MacGregor *et al*., *British Medical Journal*, vol. 2.

7. Prueba de la diferencia entre temperaturas pronosticadas y reales Remítase al conjunto de datos 10 del Apéndice B. Utilice las temperaturas máximas reales y el pronóstico de temperaturas máximas de tres días. ¿Parece haber una diferencia?

8. Prueba para la diferencia entre los tiempos en que se muestra consumo de alcohol y de tabaco Remítase al conjunto de datos 7 del Apéndice B. Utilice sólo aquellas películas que muestran 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 haber una diferencia?

12-3 Más allá de lo básico

9. Uso de la prueba de rangos con signo de Wilcoxon para aseveraciones acerca de una mediana La prueba de rangos con signo de Wilcoxon se utiliza para probar la aseveración de que una muestra proviene de una población con una mediana específica. El procedimiento que se emplea es el mismo que el descrito 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 consistentes en las 106 temperaturas corporales listadas para las 12 AM del día 2 en el conjunto de datos 4 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.

12-4 Prueba de la suma de rangos de Wilcoxon para dos muestras independientes

Aquí se introduce la *prueba de la suma de rangos de Wilcoxon*, que es la prueba no paramétrica de que dos conjuntos independientes de datos muestrales provienen de poblaciones con la misma distribución. Dos muestras son independientes, si los valores muestrales seleccionados de una población no se relacionan, asocian o se aparean de ninguna forma con los valores muestrales de la otra población. (Para evitar confusiones acerca de la 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 correspondientes al impuesto sobre la renta como técnica mnémica para recordarnos "**i**ndependiente: **s**uma de **r**angos").

Definición

Prueba de la suma de rangos de Wilcoxon: Una prueba no paramétrica que utiliza rangos de datos muestrales a partir de dos poblaciones independientes. Se emplea para probar la hipótesis nula de que dos muestras independientes provienen de poblaciones con la misma distribución (es decir, las dos poblaciones son idénticas). La hipótesis alternativa es la aseveración de que las dos distribuciones poblacionales son diferentes en alguna forma.

H_0: Las dos muestras provienen de poblaciones con la misma distribución (esto es, las dos poblaciones son idénticas).

H_1: Las dos muestras provienen de poblaciones con distribuciones diferentes (esto es, las dos poblaciones son diferentes en alguna forma).

Concepto básico: La prueba de la suma de rangos de Wilcoxon es equivalente a la **prueba *U* de Mann-Whitney** (véase el ejercicio 11), que se incluye en algunos otros libros de texto y programas de cómputo (como el Minitab). Ésta es la idea clave que sustenta la prueba de la suma de rangos de Wilcoxon: si dos muestras se obtienen de poblaciones idénticas y los valores individuales se acomodan en rangos como un conjunto de valores que se combina, entonces por lo regular los rangos alto y bajo deberían caer 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 no son idénticas. Esta idea clave se refleja en el siguiente procedimiento para calcular el valor del estadístico de prueba.

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 empatados, asígneles la media de los rangos que implica el empate. Véase la sección 12-1 para obtener una descripción de los rangos y el procedimiento para manejar empates).

2. Calcule la suma de los rangos para cualquiera de las dos muestras.
3. Calcule el valor del estadístico de prueba z como se indica en el siguiente recuadro, donde cualquier muestra puede utilizarse como la "muestra 1". (Si se está probando la hipótesis nula de poblaciones idénticas y si ambos tamaños muestrales son mayores que 10, la distribución muestral de R es aproximadamente normal, con media μ_R y desviación estándar σ_R, en tanto que el estadístico de prueba es como se indica en el recuadro siguiente).

Prueba de la suma de rangos de Wilcoxon

Supuestos

1. Hay dos muestras independientes de datos seleccionados aleatoriamente.
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 cualquiera 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

R_2 = suma de rangos de la muestra 2

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 son idénticas

σ_R = desviación estándar de los valores muestrales R que se espera, cuando las dos poblaciones son idénticas

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 están en la tabla-A.2 (puesto que el estadístico de prueba se basa en la distribución normal).

Note que, a diferencia de las pruebas de hipótesis correspondientes en la sección 8-3, la prueba de la suma de rangos de Wilcoxon no requiere poblaciones que se distribuyen normalmente. Además, la prueba de la suma de rangos de Wilcoxon se utiliza con datos en el nivel de medición ordinal, como los datos consistentes en rangos. En contraste, los métodos paramétricos de la sección 8-3 no se utilizan con datos en el nivel de medición ordinal. En la tabla 12-1 notamos que la prueba de la suma de rangos de Wilcoxon tuvo una tasa de eficiencia de 0.95 cuando se compara con la prueba paramétrica t o con la prueba z. Puesto que dicha prueba tiene una alta tasa de eficiencia y supone cálculos más fáciles, suele preferirse sobre las pruebas paramétricas que se presentan en la sección 8-3, aun cuando se satisfaga el requisito de normalidad.

La expresión de μ_R se basa en el resultado de inducción matemática siguiente: la suma de los primeros n enteros positivos está dada por $1 + 2 + 3 + \cdots + n = n(n + 1)/2$. La expresión de σ_R se basa en un resultado que establece que los enteros $1, 2, 3, \ldots, n$ tienen una desviación estándar $\sqrt{(n^2 - 1)/12}$.

Tabla 12-4
Calificaciones de lectura

Rowling	Tolstoi
85.3 (24)	69.4 (7)
84.3 (22)	64.2 (4)
79.5 (18)	71.4 (9.5)
82.5 (20)	71.6 (11)
80.2 (19)	68.5 (6)
84.6 (23)	51.9 (1)
79.2 (17)	72.2 (12)
70.9 (8)	74.4 (15)
78.6 (16)	52.8 (2)
86.2 (25)	58.4 (3)
74.0 (14)	65.4 (5)
83.7 (21)	73.6 (13)
71.4 (9.5)	
$n_1 = 13$	$n_2 = 12$
$R_1 = 236.5$	$R_2 = 88.5$

EJEMPLO Rowling y Tolstoi El conjunto de datos 14 del Apéndice B incluye las calificaciones de facilidad de lectura de Flesch para páginas que se seleccionaron aleatoriamente de cada uno de dos libros: *Harry Potter y la piedra filosofal*, de J. K. Rowling, y *La guerra y la paz*, de León Tolstoi. La tabla 12-4 incluye valores del conjunto de datos 14 junto con un valor adicional que se diseñó para ilustrar mejor el procedimiento de la suma de rangos de Wilcoxon. (El valor de 71.4 se añadió al final de la lista de Rowling para que hubiera un empate y los conjuntos de datos tuvieran números de valores diferentes). Utilice los dos conjuntos de datos muestrales independientes de la tabla 12-4, con un nivel de significancia de 0.05, y pruebe la aseveración de que las puntuaciones de facilidad de lectura para páginas de los dos libros tienen la misma distribución.

SOLUCIÓN Las hipótesis nula y alternativa son las siguientes:

H_0: Los libros de Rowling y Tolstoi tienen calificaciones de facilidad de lectura de Flesch, con la misma distribución.

H_1: Las dos poblaciones tienen distribuciones de calificaciones de facilidad de lectura de Flesch, que son diferentes en alguna forma.

Acomode en rangos las 25 calificaciones de lectura que se combinaron, comenzando con un rango de 1 (que se asignó al valor más bajo de 51.9). Los empates en los rangos se manejan como se describe en la sección 12-1: calcule la media de los rangos implicados y asigne este rango medio a cada uno de los valores empatados. Los valores 9o. y 10o. son ambos de 71.4, por lo que se asigna el rango de 9.5 a cada uno de dichos valores. En la tabla 12-4, los rangos correspondientes a los valores muestrales individuales se presentan entre paréntesis. R denota la suma de los rangos para la muestra que escogimos como muestra 1. Si elegimos las calificaciones de Rowling, tenemos

$$R = 24 + 22 + 18 + \cdots + 9.5 = 236.5$$

Puesto que hay 13 valores de Rowling, tenemos $n_1 = 13$. Además, $n_2 = 12$, ya que existen 12 valores para Tolstoi. Ahora determinemos los valores de μ_R, σ_R, y el estadístico de prueba z.

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{13(13 + 12 + 1)}{2} = 169$$

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}} = \sqrt{\frac{(13)(12)(13 + 12 + 1)}{12}} = 18.385$$

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{236.5 - 169}{18.385} = 3.67$$

La prueba es de dos colas, puesto que un valor positivo grande de z indicaría que los rangos más altos se encuentran desproporcionadamente en la primera muestra, en tanto que un valor negativo grande de z indicaría que la primera muestra tuvo una porción 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 la misma distribución.

La significancia del estadístico de prueba z puede tratarse de la misma manera que en los capítulos anteriores. Ahora estamos probando (con $\alpha = 0.05$) la hipótesis de que las dos poblaciones tienen la misma distribución; entonces, tenemos una prueba de dos colas con valores críticos z de 1.96 y -1.96. El estadístico de prueba de $z = 3.67$ cae dentro de la región crítica, por lo que rechazamos la hipótesis nula de que los libros de Rowling y de Tolstoi tienen las mismas calificaciones de lectura. Parece que las páginas de Rowling y de Tolstoi provienen de poblaciones con distribuciones diferentes. Puesto que los rangos más bajos parecen ocurrir, en su mayoría, en los valores de Tolstoi, parece que Tolstoi obtuvo calificaciones significativamente más bajas de facilidad de lectura. Esto sugiere que *La guerra y la paz*, de Tolstoi, suele ser más difícil de leer que *Harry Potter y la piedra filosofal*, de Rowling.

Verifiquemos que si intercambiamos los dos conjuntos de valores muestrales y consideramos que la muestra de Tolstoi es la primera, $R = 88.5$, $\mu_R = 156$, $\sigma_R = 18.385$ y $z = -3.67$; por lo tanto, la conclusión es exactamente la misma.

PROBLEMA 12 DEL CAPÍTULO

EJEMPLO Lluvia en miércoles y sábado El problema del capítulo se refirió a las cantidades de lluvia en Boston, listadas en el conjunto de datos 11 en el Apéndice B. El problema del capítulo incluye gráficas de cuadro de las cantidades de lluvia para los siete días de la semana, iniciando con el lunes en la parte superior. Las comparaciones de tales gráficas de cuadro indican que el miércoles y el sábado parecen ser los dos días que más difieren. Pero ¿son significativas dichas diferencias? Utilice la prueba de la suma de rangos de Wilcoxon para probar la aseveración de que las cantidades de precipitación pluvial de los miércoles y los sábados provienen de la misma distribución.

SOLUCIÓN Las hipótesis nula y alternativa son las siguientes:

H_0: Las cantidades de lluvia del miércoles y del sábado provienen de poblaciones con la misma distribución.

H_1: Las dos distribuciones son diferentes en alguna forma.

En lugar de calcular manualmente las sumas de los rangos, nos remitimos a la pantalla de Minitab que se muestra aquí. En esta pantalla de Minitab,

continúa

Brecha de género en las pruebas de fármacos

Un estudio de la relación entre los ataques cardiacos y las dosis de aspirina incluyó a 22,000 médicos hombres. Este estudio, como muchos otros, excluyó a las mujeres. La General Accounting Office criticó hace poco a los National Institutes of Health 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 se sacan 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.

"ETA1" y "ETA2" denotan la mediana de la primera muestra y la mediana de la segunda muestra, respectivamente. La pantalla sugiere que estamos probando la hipótesis nula de medianas iguales, pero la prueba de la suma de rangos de Wilcoxon se basa en las distribuciones completas, no sólo en las medianas. He aquí los componentes clave de la pantalla del Minitab: la suma de rangos para el miércoles es M = 2639.0, el valor *P* es 0.2773 (o 0.1992 después de un ajuste por los empates) y la conclusión es que no podemos rechazar (la hipótesis nula) con un nivel de significancia de 0.05. Conclusión final: las diferencias entre el miércoles y el sábado no son significativas. Esto parece contradecir los reportes de los medios de comunicación de que llueve más durante los fines de semana, pero consideraremos de nuevo este tema en la siguiente sección.

Minitab

```
WED        N =  53     Median =      0.0000
SAT        N =  52     Median =      0.0000
Point estimate for ETA1-ETA2 is      0.0000
95.1 Percent CI for ETA1-ETA2 is (0.0000,0.0000)
W = 2639.0
Test of ETA1 = ETA2  vs  ETA1 not = ETA2 is significant at 0.2773
The test is significant at 0.1992 (adjusted for ties)

Cannot reject at alpha = 0.05
```

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **Wilcoxon Tests**, seguida por la opción **Rank-Sum Test**. Ingrese los datos muestrales en el cuadro de diálogo 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**; ahora 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 $\alpha = 0.05$, en tanto que el cuadro "alternate: not equal" se refiere a la hipótesis alternativa, donde "not equal" corresponde a una prueba de hipótesis de dos colas. Minitab produce el valor *P* y la conclusión. Véase la pantalla de muestra de Minitab que se incluye en el ejemplo anterior.

Excel Excel no está programado para la prueba de la suma de rangos de Wilcoxon.

TI-83 Plus La calculadora TI-83 Plus no está programada para la prueba de la suma de rangos de Wilcoxon.

12-4 Destrezas y conceptos básicos

***Identificación de las sumas de rangos**. En los ejercicios 1 y 2, 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 y los valores críticos z; luego establezca la conclusión.*

1. Valores de la muestra 1: 1 3 4 6 8 12 15 16 17 22 26
Valores de la muestra 2: 2 5 7 9 11 13 14 18 19 20 25 26

2.

Valores de la muestra 1:	1	3	4	6	8	12	15	16	17	22	26
Valores de la muestra 2:	22	25	28	33	34	35	37	39	41	43	45

Uso de la prueba de la suma de rangos de Wilcoxon. *En los ejercicios 3 a 10, utilice la prueba de la suma de rangos de Wilcoxon.*

3. ¿Se relacionan los trastornos psiquiátricos severos con factores biológicos? Un estudio utilizó tomografía computadorizada (TC) por rayos X para reunir datos de volúmenes cerebrales de un grupo de pacientes con trastorno obsesivo compulsivo y un grupo control de personas saludables. La lista adjunta presenta los resultados muestrales (en mililitros) para volúmenes del hemisferio derecho (datos que se tomaron de "Neuroanatomical Abnormalities in Obsesive-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 los mismos volúmenes cerebrales. Con base en este resultado, ¿concluiríamos que el trastorno obsesivo compulsivo tiene una base biológica?

Pacientes obsesivo compulsivos

0.308	0.210	0.304	0.344
0.407	0.455	0.287	0.288
0.463	0.334	0.340	0.305

Grupo control

0.519	0.476	0.413	0.429
0.501	0.402	0.349	0.594
0.334	0.483	0.460	0.445

4. Prueba del efecto de anclaje Se pidió a estudiantes de estadística, seleccionados aleatoriamente, que en cinco segundos estimaran el valor de un producto de números con los resultados que vienen en la tabla adjunta. (Véanse las "Actividades de cooperación en equipo", al final del capítulo 2). ¿Hay suficiente evidencia para sustentar la aseveración de que las dos muestras provienen de poblaciones con distribuciones diferentes?

Estimados de estudiantes a los que se les 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 los que se les 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

5. ¿Afecta el orden de las preguntas de examen la calificación? Se estudió el orden de preguntas de examen para ver su efecto en la ansiedad. Los resultados muestrales se listan abajo. Utilizando un nivel de significancia de 0.05, pruebe la aseveración de que las dos muestras provienen de poblaciones con las mismas calificaciones. (Los datos se basan en "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4.)

Fácil a difícil

24.64	39.29	16.32	32.83
28.02	33.31	20.60	21.13
26.69	28.90	26.43	24.23
7.10	32.86	21.06	28.89
28.71	31.73	30.02	21.96
25.49	38.81	27.85	30.29
30.72			

Difícil a fácil

33.62	34.02	26.63	30.26
35.91	26.68	29.49	35.32
27.24	32.34	29.34	33.53
27.62	42.91	30.20	32.54

6. Prueba de poblaciones idénticas de dulces M&M rojos y cafés Abajo se listan los pesos (en gramos) de dulces M&M que se tomaron del conjunto de datos 19 del Apéndice B. Utilice un nivel de significancia de 0.05, y pruebe la aseveración de que los dulces M&M clásicos rojos y cafés tienen pesos con la misma distribución. Esto es, pruebe la aseveración de que las poblaciones de dulces M&M clásicos rojos y cafés son idénticas.

Rojos:	0.870	0.933	0.952	0.908	0.911	0.908	0.913	0.983	0.920
	0.936	0.891	0.924	0.874	0.908	0.924	0.897	0.912	0.888
	0.872	0.898	0.882						
Cafés:	0.932	0.860	0.919	0.914	0.914	0.904	0.930	0.871	1.033
	0.955	0.876	0.856	0.866	0.858	0.988	0.936	0.930	0.923
	0.867	0.965	0.902	0.928	0.900	0.889	0.875	0.909	0.976
	0.921	0.898	0.897	0.902	0.920	0.909			

T **7.** Facilidad de lectura de Rowling y Tolstoi Un ejemplo de esta sección utilizó las puntuaciones de facilidad de lectura de Flesch para páginas que se seleccionaron aleatoriamente de *Harry Potter y la piedra filosofal*, de J. K. Rowling, y *La guerra y la paz*, de León Tolstoi. (Ese ejemplo incluyó un valor muestral adicional que no se lista en el Apéndice B). Remítase al conjunto de datos 14 del Apéndice B y utilice las puntuaciones del grado de Flesch-Kincaid para las páginas de Rowling y Tolstoi. Con un nivel de significancia de 0.05, pruebe la aseveración de que las dos muestras provienen de poblaciones con la misma distribución.

T **8.** Récord de jonrones Remítase a las distancias de los jonrones que anotaron Barry Bonds y Mark McGwire en el conjunto de datos 30 del Apéndice B. Considerando que estas distancias son datos muestrales, utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las muestras de Bonds y McGwire provienen de poblaciones con la misma distribución.

T **9.** Polizones del *Queen Mary* Remítase al conjunto de datos 15 del Apéndice B, utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las edades de los polizones de la costa oeste de Estados Unidos y las edades de los polizones de la costa este provienen de poblaciones con la misma distribución.

T **10.** Índice de masa corporal (IMC) Remítase al conjunto de datos 1 del Apéndice B de los valores de índice de masa corporal para hombres y mujeres. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las dos muestras de valores de IMC provienen de poblaciones con la misma distribución.

12-4 Más allá de lo básico

11. 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

$$z = \frac{U - \dfrac{n_1 n_2}{2}}{\sqrt{\dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

donde

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R$$

Utilice las mediciones de facilidad de lectura de Rowling y Tolstoi que se listan en la tabla 12-4 de esta sección y calcule el estadístico de prueba z para la prueba U de Mann-Whitney y compárelo con el estadístico de prueba z de 3.67 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 tenemos sólo dos observaciones del tratamiento A y dos observaciones del tratamiento B. No podemos utilizar el estadístico de prueba dado en esta sección, ya que los tamaños muestrales no exceden a 10.

Rango				Suma de rangos
1	2	3	4	del tratamiento A
A	A	B	B	3

a. 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.

b. 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].

c. ¿Es posible, con un nivel de significancia de 0.10, rechazar la hipótesis nula de que no hay diferencia entre los tratamientos A y B? Explique.

12-5 Prueba de Kruskal-Wallis

Esta sección introduce la *prueba de Kruskal-Wallis*, que se utiliza para probar la hipótesis nula de que tres o más muestras independientes provienen de poblaciones idénticas. En la sección 11-2 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

Prueba de Kruskal-Wallis (también se le llama **prueba *H***): Es una prueba no paramétrica que utiliza rangos de datos muestrales de tres o más poblaciones independientes. Se emplea para probar la hipótesis nula de que las muestras independientes provienen de poblaciones con la misma distribución; la hipótesis alternativa es la aseveración de que las distribuciones poblacionales son diferentes en alguna forma.

H_0: Las muestras provienen de poblaciones con la misma distribución.

H_1: Las muestras provienen de poblaciones con distribuciones diferentes.

En la aplicación de la prueba de Kruskal-Wallis, calculamos el *estadístico de prueba H, el cual tiene una distribución que puede aproximarse por 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 6-5).

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. (Acomode los valores del más bajo al más alto; 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 el tamaño muestral.
3. Calcule H utilizando los resultados del paso 2, con la notación y el estadístico de prueba que vienen en el siguiente recuadro.

Prueba de Kruskal-Wallis

Supuestos

1. Tenemos al menos tres muestras independientes, que se seleccionan aleatoriamente.
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 cualquiera otra distribución particular.

Notación

N = número total de observaciones en todas las muestras que se combinaron

k = número de muestras

R_1 = suma de los rangos de la muestra 1

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; 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 una distribución chi cuadrada, utilice la tabla A-4 con $k - 1$ grado de libertad, donde k es el número de muestras diferentes).

El estadístico de prueba H es básicamente una medida de la varianza de las sumas de rangos $R_1, R_2, \ldots, R_k$. Si los rangos se distribuyen de manera equitativa entre los grupos muestrales, entonces H debe ser un número relativamente pequeño. Si las muestras son muy diferentes, luego 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 **Clancy, Rowling y Tolstoi** El conjunto de datos 14 del Apéndice B incluye datos obtenidos de 12 páginas que se seleccionaron aleatoriamente de tres libros diferentes: *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. Se obtuvo la puntuación de facilidad de lectura de Flesch para cada una de estas obras; los resultados se listan en la tabla 12-5. El sistema de puntuación de facilidad de lectura de Flesch da como resultado calificaciones más altas para el texto que es más fácil de leer. Las calificaciones bajas resultan de trabajos que son más difíciles de leer. En la sección 11-2 utilizamos el análisis de varianza para probar la hipótesis nula de que tres muestras de puntuaciones de lectura provienen de poblaciones con la misma media. Ahora usaremos la prueba de Kruskal-Wallis, con la hipótesis nula de que tres muestras provienen de poblaciones con la misma distribución.

Tabla 12-5 Calificaciones de facilidad de lectura

Clancy	Rowling	Tolstoi
58.2 (4)	85.3 (34)	69.4 (10.5)
73.4 (19)	84.3 (32)	64.2 (6)
73.1 (18)	79.5 (28)	71.4 (13)
64.4 (7)	82.5 (30)	71.6 (14)
72.7 (16)	80.2 (29)	68.5 (9)
89.2 (36)	84.6 (33)	51.9 (2)
43.9 (1)	79.2 (27)	72.2 (15)
76.3 (23)	70.9 (12)	74.4 (22)
76.4 (24)	78.6 (25)	52.8 (3)
78.9 (26)	86.2 (35)	58.4 (5)
69.4 (10.5)	74.0 (21)	65.4 (8)
72.9 (17)	83.7 (31)	73.6 (20)
$n_1 = 12$	$n_2 = 12$	$n_3 = 12$
$R_1 = 201.5$	$R_2 = 337$	$R_3 = 127.5$

SOLUCIÓN Las hipótesis nula y alternativa son las siguientes:

H_0: Las poblaciones de calificaciones de facilidad de lectura para las páginas de los tres libros son idénticas.

H_1: Las tres poblaciones no son idénticas.

Para determinar el valor del estadístico de prueba H, primero tenemos que ordenar en rangos todos los datos. Comenzamos con el valor más bajo de 43.9, al cual se le asigna un rango de 1. En la tabla 12-5 los rangos se muestran entre paréntesis con las calificaciones de facilidad de lectura originales. Después calculamos el tamaño muestral, n, y la suma de rangos, R, para cada muestra. Estos valores se listan al final de la tabla 12-5. Puesto que el número total de

continúa

observaciones es 36, tenemos $N = 36$. Ahora evaluamos el estadístico de prueba como sigue:

$$
\begin{aligned}
H &= \frac{12}{N(N+1)}\left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right) - 3(N+1) \\
&= \frac{12}{36(36+1)}\left(\frac{201.5^2}{12} + \frac{337^2}{12} + \frac{127.5^2}{12}\right) - 3(36+1) \\
&= 16.949
\end{aligned}
$$

Ya que cada muestra tuvo 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 $k = 3$; entonces, tenemos $3 - 1$ grados de libertad. Remítase a la tabla A-4 para encontrar el valor crítico de 5.991, que corresponde a 2 grados de libertad y a un nivel de significancia de 0.05 (con un área de 0.05 en la cola derecha).

El estadístico de prueba $H = 16.949$ está en la región crítica acotada por 5.991; por lo tanto, rechazamos la hipótesis nula de poblaciones idénticas. (En la sección 11-2, rechazamos la hipótesis nula de medias iguales).

INTERPRETACIÓN Hay suficiente evidencia para sustentar la conclusión de que las poblaciones de calificaciones de facilidad de lectura para las páginas de los tres libros no son idénticas. Los libros parecen tener calificaciones de facilidad de lectura diferentes. Examinando las sumas de rangos, vemos que Tolstoi tuvo la suma de rangos más baja, lo que sugiere que su libro es el más difícil de leer. Rowling tiene la suma de rangos más alta, lo que sugiere que su libro es el más fácil de leer de los tres.

PROBLEMA DEL CAPÍTULO 12

EJEMPLO ¿Llueve más durante los fines de semana? En el problema del capítulo señalamos que los medios de comunicación reportaron que llueve más durante los fines de semana, a lo largo de la costa este de Estados Unidos, desde Maine hasta Florida. El conjunto de datos 11 del Apéndice B incluye cantidades de lluvia en un año reciente en Boston. Utilizando ese conjunto de datos, pruebe la aseveración de que los siete días de la semana tienen distribuciones que no son las mismas.

SOLUCIÓN Parecería que es conveniente someter el conjunto de datos 11 a los métodos de análisis de varianza que se introducen en la sección 11-2, pero esos métodos requieren que los valores muestrales provengan de poblaciones con distribuciones que sean aproximadamente normales. El problema del capítulo incluye un histograma para las cantidades de lluvia del lunes, por lo que es evidente que no hay una distribución normal. Los histogramas de los otros días de la semana tienen la misma forma básica que la del lunes. Puesto que los datos no indican distribuciones normales, no es posible utilizar el análisis de varianza, por lo cual la prueba de Kruskal-Wallis resulta una alternativa ideal. El conjunto de datos 11 incluye datos para cada uno de los 365 días, de forma que estamos tratando con conjuntos de datos grandes; en consecuencia, los cálculos manuales serían muy engorrosos. En lugar de ello utilizaremos un programa de cómputo. Abajo se muestran las últimas dos líneas de la pantalla de Minitab. (Véase el ejercicio 11, donde se presenta la corrección a utilizar cuando hay muchos valores muestrales empatados).

Minitab

```
H = 2.78  DF = 6  P = 0.836
H = 3.85  DF = 6  P = 0.697 (adjusted for ties)
```

No rechazamos la hipótesis nula de distribuciones idénticas, puesto que el valor *P* de Minitab es mayor que un nivel razonable de significancia de 0.05. El STATDISK produjo un estadístico de prueba $H = 2.7806$ y un valor crítico de 12.592; además, incluye la conclusión de "no rechazo de la hipótesis nula". No hay suficiente evidencia para sustentar la aseveración de que las cantidades de lluvia durante los siete días de la semana tienen distribuciones distintas. Las cantidades de lluvia parecen ser las mismas durante los diferentes días de la semana.

INTERPRETACIÓN Con base en las cantidades de lluvia de Boston, no parece haber evidencia que sustente la aseveración de que llueve más durante los fines de semana. Por lo tanto, ¿cómo es que los reportes del periódico, las revistas y la televisión nos hagan creer que llueve más los fines de semana? Los científicos Randall S. Cerveny y Robert C. Balling, de la Arizona State University, realizaron el estudio original. ¿Es posible culparlos de brindar información falsa? No. El autor contactó a Randall Cerveny, quien afirmó que el documento original incluyó la lluvia lejos de la costa del Atlántico, no la lluvia que se asocia con alguna ciudad en particular. Cerveny y Balling utilizaron estimados de lluvia satelitales y encontraron que ciertas áreas en el océano y cerca de la costa tuvieron mayor precipitación pluvial los fines de semana, es decir, explican este fenómeno por su relación con la contaminación que viene de las regiones costeras. Sus descubrimientos son interesantes y significativos. Los medios de comunicación malinterpretaron las conclusiones de Cerveny y Balling, pues dieron reportes que indicaban que llueve más los fines de semana para quienes vivimos en tierra, junto a la costa del Atlántico. Se trata de un caso interesante de una mala interpretación hecha por los medios de comunicación de los resultados de los estudios.

Fundamentos: El estadístico de prueba *H*, como se presentó, es la versión con rangos del estadístico de prueba *F* que se utiliza en el análisis de varianza que se estudió en el capítulo 11. Cuando tratamos con rangos *R*, en lugar de valores *x* originales, muchos componentes están predeterminados. Por ejemplo, la suma de todos los rangos se expresa como $N(N + 1)/2$, donde *N* es el número total de valores en todas las muestras combinadas. La expresión

$$H = \frac{12}{N(N+1)} \Sigma n_i(\overline{R}_i - \overline{\overline{R}})^2$$

donde $$\overline{R}_i = \frac{R_i}{n_i} \qquad \overline{\overline{R}} = \frac{\Sigma R_i}{\Sigma n_i}$$

combina varianzas ponderadas de rangos para producir el estadístico de prueba *H* que se dio aquí. Tal expresión de *H* es equivalente algebraicamente a la expresión de *H* que se dio antes como estadístico de prueba. La forma anterior de *H* (no la que se dio aquí) es más sencilla de aplicar. Al comparar los procedimientos de la prueba paramétrica *F* con la prueba no paramétrica de Kruskal-Wallis, veremos que, en ausencia de programas de cómputo, la prueba de Kruskal-Wallis es mucho más simple de aplicar. No necesitamos calcular las varianzas muestrales ni las medias muestrales. Tampoco requerimos distribuciones poblacionales normales. La vida se vuelve mucho más fácil. Sin embargo, la prueba de Kruskal-Wallis no es tan eficiente como la prueba F, de manera que requeriría de diferencias más marcadas para el rechazo de la hipótesis nula.

Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **Kruskal-Wallis Test** y proceda a ingresar una copia de los datos muestrales en el cuadro de diálogo. STATDISK mostrará en la pantalla la suma de los rangos para cada muestra, el estadístico de prueba H, el valor crítico y la conclusión.

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 facilidad de lectura de la tabla 12-5, ingrese las 36 calificaciones en la columna C1 del Minitab y los 12 valores de Clancy, seguidos por los 12 valores de Rowling, y después los 12 valores de Tolstoi. En la columna C2, ingrese 12 números 1 seguidos por 12 números 2 y después por 12 números 3. Ahora seleccione **Stat**, **Nonparametrics** y **Kruskal-Wallis**. En el cuadro de diálogo, ingrese C1 como respuesta, C2 como 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.

TI-83 Plus La calculadora TI-83 Plus no está programada para la prueba de Kruskal-Wallis.

12-5 Destrezas y conceptos básicos

Interpretación de resultados de la prueba de Kruskal-Wallis. *En los ejercicios 1 y 2, interprete los resultados de la prueba de Kruskal-Wallis y resuelva la pregunta.*

1. Tiempos de maratón Los tiempos de carrera de los hombres en la carrera de maratón de la ciudad de Nueva York se listan en el conjunto de datos 8 del Apéndice B. Cuando estos tiempos de carrera se dividen en categorías, con edades de 21 a 29, 30 a 39 y 40 o mayores, los resultados de la prueba de Kruskal-Wallis de Minitab son los que se presentan más adelante. ¿Los tiempos de carrera para los diferentes grupos de edad parecen provenir de poblaciones idénticas?

Minitab

```
H = 0.58   DF = 2   P = 0.747
```

2. ¿El Old Faithful cambia con el tiempo? Se registraron 12 diferentes intervalos de tiempo (en minutos) entre las erupciones del géiser Old Faithful en 1951, 1985 y 1996. (Los datos provienen del geólogo Rick Hutchinson y del National Park Service). Cuando se utiliza Minitab con la prueba de Kruskal-Wallis, los resultados son los que se presentan más adelante. ¿Tienen los diferentes años intervalos de tiempo con poblaciones idénticas? ¿Parece que el comportamiento de las erupciones del Old Faithful cambia con el tiempo?

Minitab

```
H = 11.93   DF = 2   P = 0.003
```

Uso de la prueba de Kruskal-Wallis. *En los ejercicios 3 a 8, utilice la prueba de Kruskal-Wallis.*

3. ¿Afecta el peso de un automóvil las heridas en la cabeza que se producen en un choque? Se obtuvieron datos de experimentos de choques realizados por la National Transportation Safety Administration. Se compraron automóviles nuevos y se hicieron estrellar contra una barrera fija a 35 millas/hora. Las mediciones se registraron en un maniquí que se colocó en el asiento del conductor. Utilice los datos muestrales que se listan más adelante para probar las diferencias en las mediciones de heridas en la cabeza en cuatro categorías de peso. ¿Hay 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

4. ¿La energía solar es la misma todos los días? Una alumna del autor vive en una casa con sistema eléctrico solar. A la misma hora de cada día, ella reúne lecturas de voltaje con un medidor conectado al sistema; los resultados se listan en la tabla adjunta. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las lecturas de voltaje son las mismas para los tres diferentes tipos de día. ¿Hay suficiente evidencia para sustentar una aseveración de distribuciones poblacionales diferentes? Esperaríamos que un sistema solar proporcione más energía eléctrica en días soleados que en días nublados o lluviosos. ¿Concluiríamos que los días soleados dan como resultado mayores cantidades de energía eléctrica?

5. Prueba de diferencias de amplitud craneana en distintas épocas Los valores adjuntos son amplitudes máximas medidas de cráneos de hombres egipcios de diferentes épocas (datos tomados 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?

6. 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, consistente en la quema de piezas de tela en condiciones bajo control. Después de detener la combustión, se midió y registró la longitud de la porción que se quemó. Al margen se presentan los resultados para la misma tela que se probó en laboratorios diferentes. Puesto que se utilizó la misma tela, los diferentes laboratorios deberían obtener los mismos resultados. ¿Fue así?

T **7.** ¿El peso de todos los colores de los dulces M&M es el mismo? Remítase al conjunto de datos 19 del Apéndice B. Con un nivel de significancia de 0.05, pruebe la aseveración de que los pesos de los dulces M&M son los mismos 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?

T **8.** Distancias de jonrones Remítase al conjunto de datos 30 del Apéndice B. Considere que las distancias de jonrones son muestras seleccionadas aleatoriamente a partir de poblaciones. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que las poblaciones de distancias de jonrones que conectaron Barry Bonds, Mark McGwire y Sammy Sosa son idénticas.

Datos para el ejercicio 4

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 para el ejercicio 5

4000 AC	1850 AC	150 DC
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

Datos para el ejercicio 6

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			

12-5 Más allá de lo básico

9. Prueba del efecto de transformar los datos muestrales
   a. En general, ¿cómo se afecta el valor del estadístico de prueba H si se suma (o se resta) una constante a cada valor muestral?
   b. En general, ¿cómo se afecta el valor del estadístico de prueba H si cada valor muestral se multiplica (o divide) por una constante positiva?
   c. En general, ¿cómo se afecta el valor del estadístico de prueba H si un solo valor muestral se cambia para convertirse en un dato distante?

10. Cálculo de valores del estadístico de prueba Para tres muestras, cada una de tamaño 5, calcule los valores máximo y mínimo posibles del estadístico de prueba H.

11. Corrección del estadístico de prueba H por empates En el uso de la prueba de Kruskal-Wallis hay un factor de corrección, que debe aplicarse siempre que existan muchos empates: divida H entre

$$1 - \frac{\Sigma T}{N^3 - N}$$

Para cada grupo de observaciones empatadas en el conjunto de datos muestrales que se combinó, calcule $T = t^3 - t$, donde t es el número de observaciones que están empatadas en el grupo individual. Calcule t para cada grupo de valores empatados, luego el valor de T para cada grupo. Entonces, sume los valores T para obtener ΣT. 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 4. ¿Difiere el valor corregido de H sustancialmente del valor que se calculó en el ejercicio 4?

12. 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 se logra 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.

12-6 Correlación de rangos

En esta sección describimos cómo el método no paramétrico de correlación de rangos se utiliza con datos apareados para probar una asociación entre dos variables. En el capítulo 9 manejamos datos muestrales apareados para calcular valores del coeficiente de correlación lineal r, pero en esta sección emplearemos *rangos* como base para medir la fuerza de la correlación entre dos variables.

Definición

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 usa 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: $\rho_s = 0$ (*No* existe correlación entre las dos variables).
H_1: $\rho_s \neq 0$ (Existe una correlación entre las dos variables).

Ventajas: La correlación de rangos tiene varias ventajas sobre los métodos paramétricos que se analizaron en el capítulo 9:

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, analizamos datos apareados que sean rangos o puedan convertirse a rangos. Por ejemplo, si dos jueces califican el rango de 30 gimnastas diferentes, utilizaríamos la correlación de rangos, pero no la correlación lineal. A diferencia de los métodos paramétricos del capítulo 9, el método de correlación de rangos *no* requiere una distribución normal de la población.
2. La correlación de rangos puede utilizarse para detectar algunas de las relaciones (no todas) que no son lineales. (Se dará un ejemplo más adelante en esta sección).

Desventaja: Una desventaja de la correlación de rangos es su tasa de eficacia de 0.91, como se describe en la sección 12-1. Esta tasa de eficacia muestra que, con todas las demás circunstancias iguales, el método no paramétrico de correlación de rangos requiere de 100 pares de datos muestrales para obtener los mismos resultados que sólo 91 pares de observaciones muestrales que se analizan a través del método paramétrico, suponiendo que los requisitos más estrictos del método paramétrico se satisfacen.

Los supuestos, la notación, el estadístico de prueba y los valores críticos se resumen en el siguiente recuadro. Utilizamos la notación r_s para el coeficiente de correlación de rangos, con la finalidad de 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 desarrolló 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 12-4.

Correlación de rangos

Supuestos

1. Los datos muestrales apareados se seleccionaron aleatoriamente.
2. A diferencia de los métodos paramétricos de la sección 9-2, *no* existe el requisito de que los datos muestrales apareados tengan una distribución normal bivariada (como se describe en la sección 9-2). *No* existe el requisito de una distribución normal para la 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

continúa

Vínculo directo entre el cigarro y el cáncer

Cuando hallamos una correlación estadística entre dos variables, debemos ser extremadamente cuidadosos para evitar el error de concluir que hay una conexión de causa y efecto. La industria tabacalera ha puesto énfasis, una y otra vez, en que la correlación no implica causalidad. Sin embargo, el doctor David Sidransky, de la John Hopkins University, dice que "tenemos pruebas moleculares tan fuertes que podemos tomar un cáncer individual y potencialmente, con base en los patrones de cambio genético, determinar si fumar cigarrillos fue la causa de ese cáncer". Según sus hallazgos, él agregó que "el fumador tuvo una incidencia mucho más alta de mutación, pero el segundo hecho que lo confirmó fue el patrón tan claro de mutaciones... así que ya teníamos la pistola humeante". Aunque los métodos estadísticos no pueden probar que fumar *causa* cáncer, con evidencia física del tipo descrito por el doctor Sidransky, es posible establecer demostraciones como ésta.

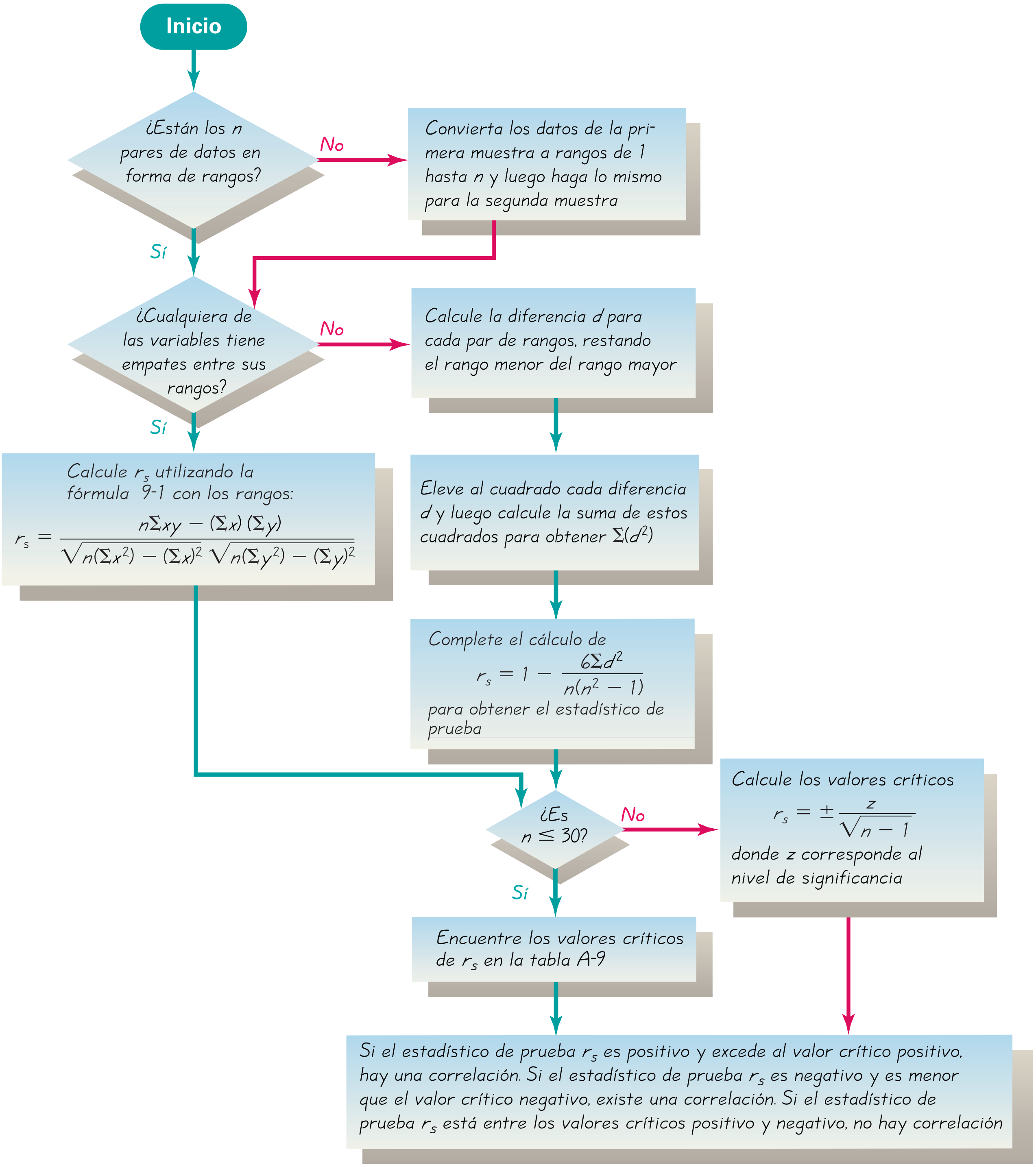


FIGURA 12-4 **Procedimiento de correlación de rangos para probar H_0: $\rho_s = 0$**

Estadístico de prueba

Sin empates: Después de convertir los datos de cada muestra a rangos, si no hay empates entre los rangos para la primera variable ni entre los rangos para la segunda variable, el valor exacto del estadístico de prueba se calcula utilizando esta fórmula:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

Empates: Después 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 se calcula utilizando la fórmula 9-1 con los rangos:

$$r_s = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\,\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

Valores críticos

1. Si $n \le 30$, 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 12-1.

Fórmula 12-1 $$r_s = \frac{\pm z}{\sqrt{n - 1}}$$ (valores críticos cuando $n > 30$)

donde el valor de z corresponde al nivel de significancia.

EJEMPLO Percepción de belleza La revista *marie claire* pidió a hombres y mujeres que calificaran el grado de belleza de 10 mujeres diferentes, todas ellas bastante atractivas. ("¿Suele sorprenderse por lo que los hombres y las mujeres encuentran atractivo? Pedimos a 100 hombres y a 100 mujeres que calificaran el grado de belleza de estos rostros sumamente atractivos y explicaran exactamente lo que ellos encuentran atractivo"). La tabla 12-6 lista los rangos resultantes. ¿Hay una correlación entre las calificaciones de los hombres y las mujeres? La revista preguntó: "¿Los hombres y las mujeres están de acuerdo?". ¿Lo están? Utilice un nivel de significancia de $\alpha = 0.05$.

Tabla 12-6 Grado de belleza de 10 mujeres

Hombres	**4**	**2**	**5**	**1**	**3**	**6**	**7**	**8**	**9**	**10**	
Mujeres	**2**	**6**	**7**	**3**	**1**	**10**	**4**	**8**	**5**	**9**	
d	2	4	2	2	2	4	3	0	4	1	
d^2	4	16	4	4	4	16	9	0	16	1	→ Total = 74

SOLUCIÓN El coeficiente de correlación lineal r (sección 9-2) no debe utilizarse puesto que requiere de distribuciones normales y los datos consisten en rangos que no se distribuyen normalmente. En su lugar, utilizamos el coeficiente de correlación de rangos para probar una relación entre los rangos de hombres y de mujeres.

Las hipótesis nula y alternativa son las siguientes:

$$H_0: \quad \rho_s = 0$$
$$H_1: \quad \rho_s \neq 0$$

continúa

Siguiendo el procedimiento de la figura 12-4, los datos están en forma de rangos; ninguna de las dos variables (hombres y mujeres) tuvo empates entre los rangos; por lo tanto, el valor exacto del estadístico de prueba se calcula como se indica abajo. Utilizamos $n = 10$ (para 10 pares de datos) y $\Sigma d^2 = 74$ (como se indica en la tabla 12-4) para obtener

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} = 1 - \frac{6(74)}{10(10^2 - 1)}$$
$$= 1 - \frac{444}{990} = 0.552$$

Ahora nos remitimos a la tabla A-9 para determinar que los valores críticos son ± 0.648 (con base en $\alpha = 0.05$ y $n = 10$). Puesto que el estadístico de prueba $r_s = 0.552$ no excede al valor crítico de 0.648, no rechazamos la hipótesis nula. No hay suficiente evidencia para sustentar una aseveración de correlación entre las calificaciones de los hombres y de las mujeres. Parece que en asuntos de belleza los hombres y las mujeres no están de acuerdo. (Si estuvieran de acuerdo, habría una correlación significativa, pero no es así).

EJEMPLO Caso de muestra grande Suponga que el ejemplo anterior se expande, incluyendo un total de 40 mujeres, y que se encuentra que el estadístico de prueba r_s es 0.291. Si el nivel de significancia es $\alpha = 0.05$, ¿qué concluye acerca de la correlación?

SOLUCIÓN Puesto que hay 40 pares de datos, tenemos $n = 40$. Puesto que n excede de 30, calculamos los valores críticos con la fórmula 12-1, en lugar 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 $r_s = 0.291$ no excede al valor crítico de 0.314; por lo tanto, no rechazamos la hipótesis nula. No hay suficiente evidencia para sustentar la aseveración de una correlación entre hombres y mujeres.

El siguiente ejemplo tiene la intención de ilustrar el principio de que la correlación de rangos algunas veces se utiliza 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 12-7 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 ahí haya una asociación entre el número de juegos que se realizaron y la puntuación del pinball. ¿Existe suficiente evidencia para sustentar la aseveración de que hay una asociación como ésta?

SOLUCIÓN Probaremos la hipótesis nula de no correlación de rangos ($\rho_s = 0$).

$$H_0: \quad \rho_s = 0 \text{ (sin correlación)}$$
$$H_1: \quad \rho_s \neq 0 \text{ (correlación)}$$

Tabla 12-7 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

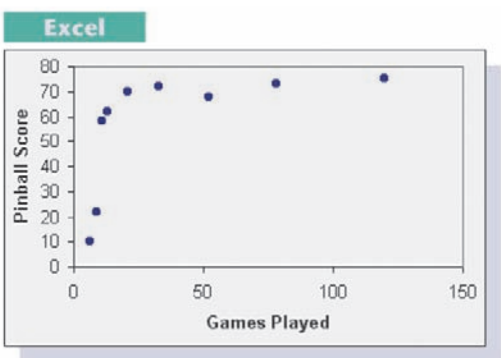
Remítase a la figura 12-4, la cual seguimos para esta solución. Las puntuaciones originales no son rangos, por lo cual las convertimos a rangos e introdujimos los resultados entre paréntesis en la tabla 12-7. (La sección 12-1 describe el procedimiento para convertir puntuaciones en rangos).

Después de expresar todos los datos como rangos, calculamos las diferencias d y luego las elevamos al cuadrado. La suma de los valores de d^2 es 6. Ahora calculamos

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} = 1 - \frac{6(6)}{9(9^2 - 1)}$$
$$= 1 - \frac{36}{720} = 0.950$$

Continuando con la figura 12-4, 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 ± 0.683. Finalmente, el estadístico muestral de 0.950 excede a 0.683; por lo tanto, concluimos que hay una correlación significativa. Los números más altos de juegos parecen asociarse 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 (mediante la fórmula 9-1) para los datos originales, obtendremos $r = 0.586$, lo que nos lleva a la conclusión de que no hay 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, veremos 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.



Utilizando la tecnología

STATDISK Seleccione **Analysis** de la barra del menú principal, luego **Rank Correlation**. Ingrese los datos muestrales apareados en el cuadro de diálogo, después haga clic en **Evaluate**. Los resultados del 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** del Minitab para convertir los datos a rangos, luego seleccione **Stat**, seguido por **Basic Statistics** y **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 9-2, en realidad se trata del coeficiente de correlación de Spearman, descrito en esta sección (puesto que se basa 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 calcula 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* que se localiza en la barra del menú principal. Seleccione la categoría de función **Statistical** y el nombre de función **CORREL**, luego haga clic en **OK**. En el cuadro de diálogo, ingrese en la celda el rango de valores para x, como es A1:A10. También ingrese en la celda el rango de valores para y, como es B1:B10. Excel mostrará en la pantalla el valor exacto del coeficiente de correlación de rango r_s.

TI-83 Plus Si utilizamos una calculadora TI-83 Plus o cualquiera otra calculadora con estadística de dos variables, será posible calcular el valor exacto de r_s como sigue: 1. reemplace cada valor muestral por su rango correspondiente, 2. calcule el valor del coeficiente de correlación lineal r con los mismos procedimientos que se utilizaron en la sección 9-2. Ingrese los rangos apareados en las listas L1 y L2, después oprima **STAT** y elija **TESTS**. El uso de la opción **LinRegTTest** dará como resultado diversos valores que se muestran en la pantalla, incluyendo el valor exacto del coeficiente de correlación de rangos r_s.

12-6 Destrezas y conceptos básicos

1. Cálculo del estadístico de prueba y el valor crítico Para cada una de las siguientes muestras de rangos apareados, dibuje un diagrama de dispersión, estime el valor de r_s, calcule el valor de r_s y establezca si parece haber una correlación entre x y y.

a.

x	1	3	5	4	2
y	1	3	5	4	2

b.

x	1	2	3	4	5
y	5	4	3	2	1

c.

x	1	2	3	4	5
y	2	5	3	1	4

2. Cálculo de valores críticos Calcule el(los) valor(es) crítico(s) para r_s mediante la tabla A-9 o la fórmula 12-1, según resulte adecuado. Suponga casos de dos colas, donde α representa el nivel de significancia y n representa el número de pares de datos.

a. $n = 20, \alpha = 0.05$ **b.** $n = 50, \alpha = 0.05$
c. $n = 40, \alpha = 0.02$ **d.** $n = 15, \alpha = 0.01$
e. $n = 82, \alpha = 0.04$

Prueba para correlación de rangos. *En los ejercicios 3 a 12, utilice el coeficiente de correlación de rangos para probar una correlación entre las dos variables. Utilice un nivel de significancia de $\alpha = 0.05$.*

3. Correlación entre salario y estrés La tabla adjunta lista rangos de salario y rangos de estrés de trabajos que se seleccionaron al azar (datos tomados de *The Jobs Rated Almanac*). ¿Parece que el salario se incrementa a medida que se incrementa el estrés?

Trabajo	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 ocupacional	9	9
Economista doméstico	8	10
Psicólogo	4	3
Piloto de línea aérea	1	1

4. Correlación entre salario y demanda física El ejercicio 3 incluye rangos apareados de salario y nivel de estrés para 10 empleos que se seleccionaron al azar. Las demandas físicas de los empleos también se ordenaron en rangos; los rangos de salario y demanda física se presentan abajo (según datos de *The Jobs Rated Almanac*). ¿Parece haber 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

5. Rangos de orden de escuelas de negocios La revista *Business Week* ordenó en rangos escuelas de negocios de dos formas diferentes. Los rangos de orden institucional se basaron en encuestas a reclutadores de la institución, y los rangos de orden de graduados se basaron en encuestas a graduados de la maestría en negocios. La tabla de abajo se basa en los resultados para 10 escuelas. ¿Hay una correlación entre los rangos institucionales y los rangos de los graduados? Utilice un nivel de significancia de $\alpha = 0.05$.

Escuela	PA	NW	Chi	Sfd	Hvd	MI	IN	Clb	UCLA	MIT
Rango institucional	1	2	4	5	3	6	8	7	10	9
Rango de graduados	3	5	4	1	10	7	6	8	2	9

6. Correlación entre cuentas de restaurante y propinas Los alumnos del autor reunieron datos muestrales consistentes en cantidades en cuentas de restaurante y la cantidad correspondiente de propina. Los datos se listan más adelante. Utilice la correlación de rangos para determinar si hay una correlación entre la cantidad de la cuenta y la cantidad de la propina.

Cuenta (dólares)	33.46	50.68	87.92	98.84	63.60	107.34
Propina (dólares)	5.50	5.00	8.08	17.00	12.00	16.00

7. Correlación entre estaturas y pesos de supermodelos Más adelante se listan estaturas (en pulgadas) y pesos (en libras) de las supermodelos Niki Taylor, Diana Auermann, Claudia Schiffer, Elle MacPherson, Christy Turlington, Bridget Hall, Kate Moss, Valerie Mazza y Kristy Hume.

Estatura	71	70.5	71	72	70	70	66.5	70	71
Peso	125	119	128	128	119	127	105	123	115

8. Precio de una audiencia televisiva El *New York Post* publicó los salarios anuales (en millones) y el número de televidentes (en millones), con los resultados que se incluyen abajo, para Oprah Winfrey, David Letterman, Jay Leno, Kelsey Grammer, Barbara Walters, Dan Rather, James Gandolfini y Susan Lucci, respectivamente. ¿Hay una correlación entre el salario y el número de televidentes?

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

T **9.** Cereales asesinos Remítase al conjunto de datos 16 del Apéndice B. Utilice las cantidades de grasa y los conteos calóricos medidos. ¿Hay una correlación?

T **10.** Colesterol e índice de masa corporal Remítase al conjunto de datos 1 del Apéndice B. Utilice los niveles de colesterol y los valores de índice de masa corporal de las 40 mujeres. ¿Hay una correlación entre el nivel de colesterol y el índice de masa corporal?

T **11.** Lo malo de los cigarros Remítase al conjunto de datos 5 del Apéndice B.

a. Utilice los datos apareados referentes a alquitrán y nicotina. Con base en el resultado, ¿parece haber ahí una correlación significativa entre el alquitrán y la nicotina de los cigarros? Si es así, ¿pueden los investigadores reducir sus gastos de laboratorio midiendo sólo una de estas dos variables?

b. Utilice los datos apareados consistentes en monóxido de carbono y nicotina. Con base en el resultado, ¿parece haber una correlación significativa entre el monóxido de carbono y la nicotina de los cigarros? 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 desean medir sólo otro elemento. Al elegir entre alquitrán y monóxido de carbono, ¿cuál es la mejor opción? ¿Por qué?

T **12.** Pronósticos del clima Remítase al conjunto de datos 10 del Apéndice B.

a. Utilice las temperaturas máximas que se pronosticaron para cinco días y las temperaturas máximas reales. ¿Hay una correlación? ¿Una correlación significativa implica que las temperaturas del pronóstico de cinco días son precisas?

b. Utilice las temperaturas máximas que se pronostican para un día y las temperaturas máximas reales. ¿Hay una correlación? ¿Una correlación significativa implica que las temperaturas de pronóstico para un día son precisas?

c. ¿Cómo esperaría 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* concuerdan con lo que esperaría? Si hay una correlación muy alta entre las temperaturas de pronóstico y las temperaturas reales, ¿se deduce que las temperaturas de pronóstico son precisas?

12-6 Más allá de lo básico

13. Cálculo de valores críticos Una alternativa al uso de la tabla A-9 para encontrar valores críticos es calcularlos utilizando esta aproximación:

$$r_s = \pm\sqrt{\frac{t^2}{t^2 + n - 2}}$$

Aquí t es la puntuación t de la tabla A-3, correspondiente al nivel de significancia y a $n - 2$ grados de libertad. Aplique esta aproximación para calcular los valores críticos de r_s en los casos siguientes.

a. $n = 8, \alpha = 0.05$
b. $n = 15, \alpha = 0.05$
c. $n = 30, \alpha = 0.05$
d. $n = 30, \alpha = 0.01$
e. $n = 8, \alpha = 0.01$

14. Efecto de empates en r_s Remítase al conjunto de datos 7 del Apéndice B para los tiempos (en segundos) de consumo de tabaco y consumo de alcohol que se presentan en películas de dibujos animados para niños. Calcule el valor del estadístico de prueba r_s utilizando cada una de las dos fórmulas presentadas en esta sección. ¿Hay una diferencia sustancial entre los dos resultados? ¿Cuál resultado es mejor? ¿La conclusión se ve afectada por la fórmula utilizada?

12-7 Prueba de rachas para detectar aleatoriedad

El objetivo principal de esta sección es introducir la prueba de rachas para detectar aleatoriedad, que permite determinar si los datos muestrales en una secuencia están en un orden aleatorio. En la importancia de la aleatoriedad se ha puesto énfasis a lo largo de este libro; ahora nos enfocamos en un método para determinar si esta característica está presente.

Definiciones

Racha: Es una secuencia de datos que tiene 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; por lo tanto, el número de rachas es muy *bajo*.
- Ejemplo: La secuencia de género MHMHMHMHMH no es aleatoria, puesto que hay 10 rachas, lo cual es muy *alto*.

El criterio exacto para determinar si un número de rachas es muy alto o muy bajo se encuentra en el recuadro adjunto, que resume los elementos clave de la prueba de rachas para detectar aleatoriedad. Además, el procedimiento de la prueba de rachas para detectar aleatoriedad se resume en la figura 12-5.

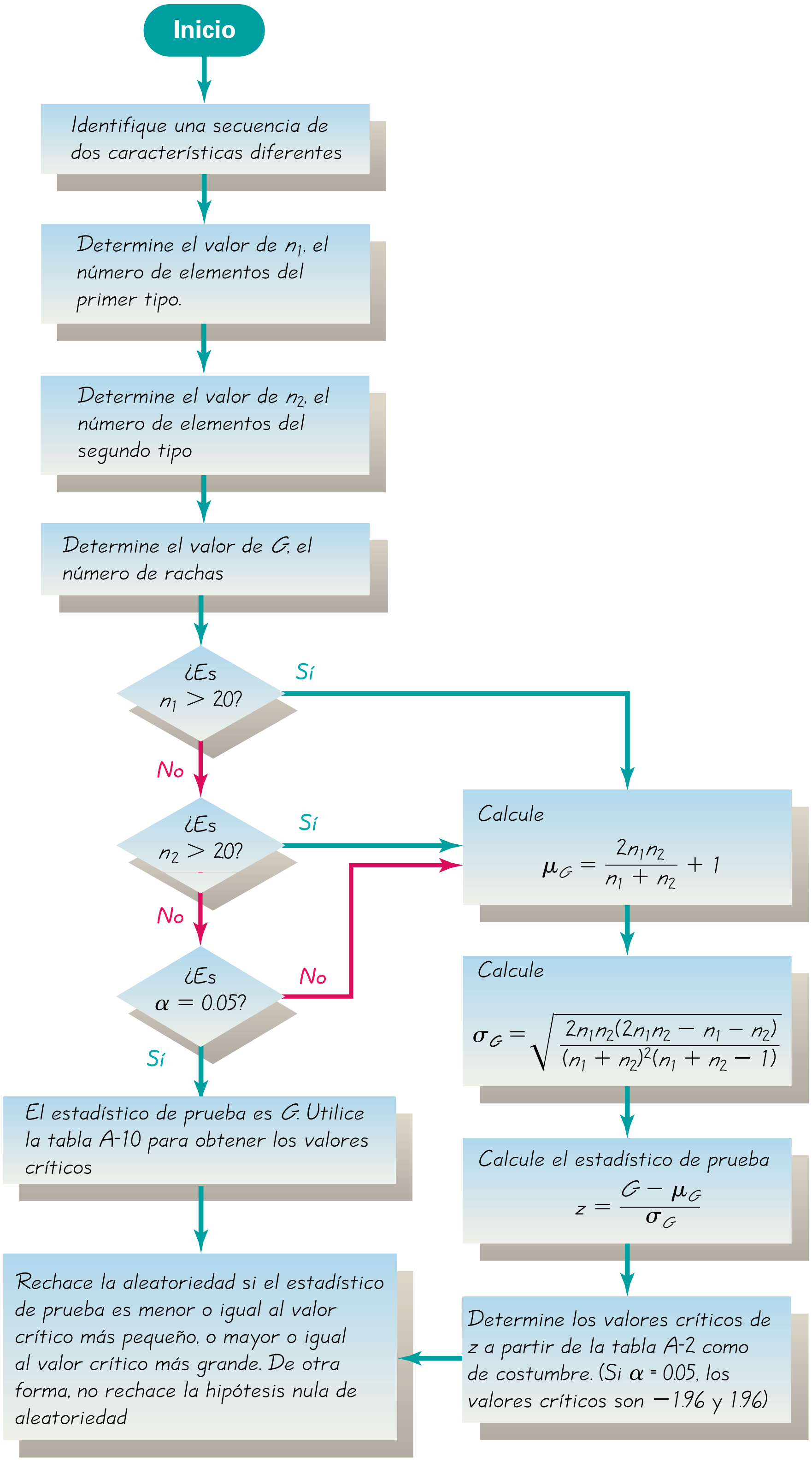


FIGURA 12-5 **Prueba de rachas para detectar aleatoriedad**

Prueba de rachas para detectar aleatoriedad

Supuestos

1. Los datos muestrales se acomodan de acuerdo con algún esquema de orden, como 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.
3. La prueba de rachas paramétrica para detectar aleatoriedad se basa en el *orden* en el que los datos ocurren, *no* se basa en la *frecuencia* de los datos. (Por ejemplo, una secuencia de tres hombres y 20 mujeres parecería aleatoria, aunque el punto de si 3 hombres y 20 mujeres constituyen una muestra *sesgada* no se conoce por la prueba de rachas).

Notación

n_1 = número de elementos en la secuencia con 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. A continuación el criterio de decisión:

Rechace la aleatoriedad si el número de rachas G es

- menor o igual al valor crítico más chico que se encuentra en la tabla A-10.
- mayor o igual al valor crítico más grande que se encuentra en la tabla A-10.

Para muestras grandes o $\alpha \neq 0.05$: Si $n_1 > 20$ o $n_2 > 20$ o $\alpha \neq 0.05$, utilice el estadístico de prueba y los valores críticos siguientes.

Estadístico de prueba: $$z = \frac{G - \mu_G}{\sigma_G}$$

donde $$\mu_G = \frac{2n_1n_2}{n_1 + n_2} + 1$$

y $$\sigma_G = \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

Valores críticos de z: Utilice la tabla A-2.

Rachas en los deportes

Es una creencia común que los deportistas suelen tener "buenas rachas", esto es, breves periodos de extraordinario éxito. El psicólogo Amos Tversky, de Stanford University, y otros investigadores utilizaron la estadística para analizar los miles de tiros intentados por los jugadores del equipo de basquetbol profesional los 76 de Filadelfia durante toda una temporada y la mitad de otra. Encontraron que el número de "rachas" no era diferente del que usted esperaría de intentos aleatorios en los que el resultado de cada intento es independiente de cualquiera de los resultados anteriores. Es decir, la probabilidad de un acierto no depende del acierto o falla anterior.

EJEMPLO Muestras pequeñas: Tiros con falta en basquetbol
En el transcurso de un juego, Cynthia Cooper realizó 12 tiros libres. Si denotamos los tiros certeros por A (para "anotación") y los tiros fallados por F, sus resultados son los siguientes: A, A, A, F, A, A, A, A, F, F, F y A. Utilice un nivel de significancia de 0.05 y pruebe la aleatoriedad en la secuencia de anotaciones y fallas.

continúa

SOLUCIÓN Remítase al procedimiento que se resume en la figura 12-5. Ya se identificó la secuencia de dos características (anotar y fallar). 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.

$$\underbrace{\text{HHH}}_{1^{\text{a}}\text{ racha}}\quad \underbrace{\text{M}}_{2^{\text{a}}\text{ racha}}\quad \underbrace{\text{HHHH}}_{3^{\text{a}}\text{ racha}}\quad \underbrace{\text{MMM}}_{4^{\text{a}}\text{ racha}}\quad \underbrace{\text{H}}_{5^{\text{a}}\text{ racha}}$$

Puesto que hay ocho anotaciones, cuatro fallas y cinco rachas, tenemos

$$n_1 = \text{número de tiros anotados (A)} = 8$$
$$n_2 = \text{número de tiros fallados (F)} = 4$$
$$G = \text{número de rachas} = 5$$

Puesto que $n_1 \leq 20$, $n_2 \leq 20$ y $\alpha = 0.05$, el estadístico de prueba es $G = 5$; nos remitimos a la tabla A-10 para encontrar los valores críticos de 3 y 10. Puesto que $G = 5$ no es menor o igual a 3, ni tampoco es mayor o igual a 10, no rechazamos la aleatoriedad. No hay evidencia suficiente para fundamentar el rechazo de la aseveración de que las anotaciones y las fallas ocurren aleatoriamente. Parece que la secuencia de anotaciones y fallas es aleatoria.

EJEMPLO Muestras grandes: lluvia en Boston los lunes Remítase a las cantidades de lluvia en Boston que se listan en el conjunto de datos 11 del Apéndice B. ¿Hay 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 Permita que S (para seco) represente los lunes sin lluvia (indicados por valores de 0.00), y permita que L represente los lunes con alguna lluvia (cualquier valor mayor que 0.00). Los 52 lunes consecutivos se representan con esta secuencia:

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 L S S S L

Las hipótesis nula y alternativa son las siguientes:

H_0: La secuencia es aleatoria.
H_1: La secuencia no es aleatoria.

El estadístico de prueba se obtiene buscando primero el número de S, el número de L y el número de rachas. Es fácil examinar la secuencia para encontrar que

$$n_1 = \text{número de S} = 33$$
$$n_2 = \text{número de L} = 19$$
$$G = \text{número de rachas} = 30$$

Ya que seguimos el procedimiento de la figura 12-5, contestamos sí a la pregunta "¿Es $n_1 > 20$?". Por lo tanto, necesitamos evaluar el estadístico de prueba z

que aparece 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

$$\mu_G = \frac{2n_1n_2}{n_1 + n_2} + 1 = \frac{2(33)(19)}{33 + 19} + 1 = 25.115$$

$$\sigma_G = \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

$$= \sqrt{\frac{(2)(33)(19)[2(33)(19) - 33 - 19]}{(33 + 19)^2(33 + 19 - 1)}} = 3.306$$

Ahora calculemos 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 parece ser aleatoria.

Datos numéricos: aleatoriedad por encima 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, aunque también probamos la aleatoriedad con que los datos numéricos fluctúan por encima o por debajo de una media o una mediana. Para probar la aleatoriedad por encima 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 E si está por *encima* de la mediana, ahora reemplácelo con D si está por *debajo* de la mediana. Excluya cualquier valor que sea igual a la mediana. Es útil escribir las E 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 de encontrar la secuencia de las letras E y D, procedamos a aplicar la prueba de rachas tal como se describió. Los economistas utilizan la prueba de rachas para detectar aleatoriedad por encima y por debajo de la media en un intento de identificar tendencias o ciclos. Una tendencia económica a la alza contendría una predominancia de letras D al principio y de E al final, por lo cual el número de rachas sería pequeño. Una tendencia a la baja tendría predominio de las letras E al principio y las D al final, con un número bajo de rachas. Un patrón cíclico produciría una secuencia que cambia sistemáticamente; por lo tanto, el número de rachas tendería a ser grande. (Véase el ejercicio 11).

Utilizando la tecnología

STATDISK El STATDISK está programado para la prueba de rachas; pero, por la naturaleza de los datos, usted debe determinar primero los valores de n_1 y n_2, así como el número de rachas G. Seleccione **Analysis** de la barra del menú principal, luego **Runs Test** y proceda a ingresar los datos que se requieren en el cuadro de diálogo. La pantalla del STATDISK incluirá el estadístico de prueba (G o z según lo propio), los valores críticos y la conclusión.

Minitab Minitab efectuará una prueba de rachas únicamente con una secuencia de datos numéricos, pero véase *Minitab Student Laboratory Manual and Workbook* para evitar tal 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 encima 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 Plus La calculadora TI-83 Plus no está programada para la prueba de rachas para detectar aleatoriedad.

12-7 Destrezas y conceptos básicos

Identificación de rachas y cálculo de valores críticos. *En los ejercicios 1 a 4, 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; utilice esos resultados para determinar si la secuencia parece ser aleatoria.*

1. H H H H M M M M M M M H H H H H H H

2. M M M M F F F M M M F F M M F M

3. A A B B A A B B A A B B A A B B A A B B

4. T T T T T F F F F F T T T T T F F F F F

Uso de la prueba de rachas para detectar aleatoriedad. *En los ejercicios 5 a 12, utilice la prueba de rachas de esta sección para determinar si la secuencia que se indica es aleatoria. Emplee un nivel de significancia de $\alpha = 0.05$. (Todos los datos se listan en orden por renglón).*

5. Aleatoriedad de los resultados de la ruleta Al realizar la investigación para este libro, el autor registró los resultados de una ruleta en el Stardust Casino. (Sí, fue un trabajo duro, pero alguien tenía que hacerlo). Pruebe la aleatoriedad de números impares (I) y pares (P) en los resultados dados en la siguiente secuencia. ¿Qué significaría para el autor una carencia de aleatoriedad? ¿Para el casino?

I I P P P P I I P I P I I I I I I P I P

6. Prueba de aleatoriedad de sujetos de encuesta En la selección de sujetos a encuestarse acerca del juego *Roller Coaster Tycoon* de Infogrames, los sujetos se seleccionaron en una secuencia con los géneros que se listan más adelante. ¿Parece que los sujetos se seleccionaron aleatoriamente de acuerdo con el género?

H H M M M H M H H H H M M H H M M M M H M

7. Prueba de aleatoriedad en prospectos para citas Fred tiene dificultades para obtener citas con las mujeres; por lo tanto, él está abandonando su estrategia de una selección cuidadosa y la está reemplazando por una estrategia desesperada de selección aleatoria. Al buscar citas con mujeres que selecciona al azar, Fred encuentra que algunas de ellas no están disponibles porque son casadas. Fred, quien cuenta con mucho tiempo para dichas actividades, registra y analiza sus observaciones. A partir de los resultados que se listan más adelante (donde C denota casada y S denota soltera), ¿qué debería concluir Fred acerca de la aleatoriedad de las mujeres que él selecciona?

C C C C S S S S S S C C C C C S
S S C C C C C C C C C C S S S S

8. Prueba de la aleatoriedad de las victorias en las series mundiales de beisbol 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, que se representan por A y N, respectivamente. ¿Qué sugieren los resultados acerca de las habilidades de las dos ligas?

A N A 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

9. 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

10. Prueba de aleatoriedad de fechas obtenidas para selección militar En una ocasión a los hombres se les reclutó en el ejército de Estados Unidos utilizando un proceso que se suponía seleccionaba fechas de cumpleaños al azar. Suponga que las primeras selecciones son las que se listan más adelante. Pruebe la aleatoriedad de la secuencia para antes y después de la mitad del año.

27 Nov.	7 Julio	3 Ago.	19 Oct.	19 Dic.	21 Sept.	3 Mayo
5 Mar.	10 Junio	15 Mayo	27 Junio	5 Ene.		

11. Mercado bursátil: prueba de aleatoriedad por encima y por debajo de la mediana Las tendencias de las aplicaciones en negocios y economía suelen analizarse con la prueba de rachas. El conjunto de datos 25 del Apéndice B lista 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 E si está por encima de la mediana y por D si está por debajo de la mediana. Después aplique la prueba de rachas a la secuencia resultante de letras E y D. ¿Qué sugiere el resultado acerca del mercado bursátil como una consideración para invertir? (Los actos de terrorismo y las condiciones económicas adversas causaron una caída importante en el promedio industrial DJ en 2001).

12. Prueba de aleatoriedad de muertes en vehículos con motor Remítase al conjunto de datos 25 del Apéndice B para los números de muertes en vehículos con motor en Estados Unidos durante las dos décadas anteriores. Pruebe la aleatoriedad por encima y por debajo de la media. ¿Los números de muertes en vehículos con motor parecen ser aleatorios? Si no es así, ¿hay una tendencia? ¿Puede explicarse la tendencia?

13. Muestra grande: prueba de aleatoriedad de dígitos impares y pares en pi Un artículo del *The 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

14. Muestra grande: prueba de aleatoriedad de victorias en las series mundiales de beisbol 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, que se representan 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 A N
N A A A A A N A N A N A N A A A A A A A N N A N A N N A A N
N N A N A N A 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

15. Muestra grande: prueba de aleatoriedad de corredores de maratón Remítase al conjunto de datos 8 del Apéndice B para la muestra aleatoria de corredores que terminaron la carrera de maratón de la ciudad de Nueva York. Los corredores se listan en el orden en el que terminaron. Pruebe la aleatoriedad de la secuencia del género. ¿Hay evidencia suficiente para sustentar la aseveración de un reportero que escribe que los corredores hombres tienden a terminar primero que las mujeres corredoras?

16. Muestra grande: prueba de aleatoriedad de corredores de maratón Remítase al conjunto de datos 8 del Apéndice B para la muestra aleatoria de corredores que terminaron la carrera de maratón de la ciudad de Nueva York. Pruebe la aleatoriedad de las edades por encima y por debajo de la edad media. ¿Hay evidencia suficiente para sustentar la aseveración de un reportero que escribe que los corredores más jóvenes tienden a terminar primero que los más viejos?

12-7 Más allá de lo básico

17. 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 con la finalidad de encontrar los valores críticos G para $n_1 = n_2 = 2$. ¿Qué concluye acerca de este caso?

18. Cálculo de valores críticos

a. Utilizando todos los elementos A, A, A, B, B, B, B, B, B, haga una lista de las 84 diferentes secuencias posibles.

b. Calcule el número de rachas para cada una de las 84 secuencias.

c. Emplee los resultados de los incisos a) y b) para determinar sus propios valores críticos para G.

d. Compare los resultados con los que se incluyeron en la tabla A.10.

Repaso

En este capítulo examinamos seis pruebas no paramétricas diferentes para analizar datos muestrales. Las pruebas no paramétricas también se conocen como pruebas de distribución libre, puesto que no requieren que las poblaciones tengan una distribución particular, como la distribución normal. Sin embargo, las pruebas no paramétricas no son tan eficaces como las pruebas paramétricas, de forma que generalmente necesitamos una evidencia más fuerte antes de rechazar la hipótesis nula.

La tabla 12-8 lista las pruebas no paramétricas que se presentan en este capítulo, junto con sus funciones. La tabla lista además las pruebas paramétricas correspondientes.

Tabla 12-8 Resumen de pruebas no paramétricas

Prueba no paramétrica	Función	Prueba paramétrica
Prueba del signo (sección 12-2)	Prueba del valor aseverado del promedio con una muestra	Prueba z o prueba t (secciones 7-4, 7-5)
	Prueba de las diferencias entre datos apareados	Prueba t (sección 8-4)
	Prueba del valor aseverado de una proporción	Prueba z (sección 7-3)
Prueba de rangos con signo de Wilcoxon (sección 12-3)	Prueba de las diferencias entre datos apareados	Prueba t (sección 8-4)
Prueba de la suma de rangos de Wilcoxon (sección 12-4)	Prueba de la diferencia entre dos muestras independientes	Prueba t o prueba z (sección 8-3)
Prueba de Kruskal-Wallis (sección 12-5)	Prueba si más de dos muestras independientes provienen de poblaciones idénticas	Análisis de varianza (sección 11-2)
Correlación de rangos (sección 12-6)	Prueba de la relación entre dos variables	Correlación lineal (sección 9-2)
Prueba de rachas (sección 12-7)	Prueba de la aleatoriedad de datos muestrales	(No hay prueba paramétrica)

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 que se indica. Si no se especifica una prueba en particular, utilice la prueba no paramétrica adecuada de este capítulo.*

1. Prueba de eficacia de cursos preparatorios para el SAT ¿Conviene tomar cursos preparatorios para pruebas estandarizadas como el SAT? Utilizando un nivel de significancia de 0.05, pruebe la aseveración de que el curso de preparación Allan no surte efecto en las calificaciones del SAT. Emplee la prueba del signo con los datos muestrales de la tabla adjunta (datos tomados del College Board y "An Analysis of the Impact of Commercial Test Preparation Courses on SAT Scores", de Sesnowitz, Bernhardt y Knain, *American Educational Research Journal*, vol. 19, núm. 3).

Sujeto	A	B	C	D	E	F	G	H	I	J
Puntaje SAT antes del curso	700	840	830	860	840	690	830	1180	930	1070
Puntaje SAT después del curso	720	840	820	900	870	700	800	1200	950	1080

2. Prueba de eficacia de cursos preparatorios para el SAT Realice el ejercicio 1 utilizando la prueba de rangos con signo de Wilcoxon.

3. Prueba de discriminación por género La Tektronics Internet Company afirma que la contratación se realiza sin ningún sesgo de género. De los últimos 66 empleados nuevos que se contrató, 1/3 son mujeres. Casi la mitad de los aspirantes al empleo son hombres y la otra mitad mujeres, que califican para el empleo. ¿Hay suficiente evidencia para acusar de un sesgo a favor de los hombres? Utilice un nivel de significancia de 0.01, ya que no queremos hacer un cargo tan grave, a menos de que exista evidencia muy fuerte.

4. ¿Tienen niveles de CAS diferentes los bebedores de cerveza y los bebedores de licor? Los datos muestrales de la siguiente lista indican los niveles de CAS (concentración de alcohol en sangre) en el momento del arresto de personas que se seleccionaron al azar, a quienes se encarceló por conducir en estado de ebriedad (CEE) o bajo la influencia del alcohol (CIA). Los datos se categorizaron por el tipo de bebida que se consumió (de acuerdo con 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 tales resultados, ¿parecen ambos grupos ser igualmente peligrosos o hay un grupo más peligroso que el otro?

Cerveza

0.129	0.146	0.148	0.152
0.154	0.155	0.187	0.212
0.203	0.190	0.164	0.165

Licor

0.220	0.225	0.185	0.182
0.253	0.241	0.227	0.205
0.247	0.224	0.226	0.234
0.190	0.257		

5. Correlación entre el peso del automóvil y el consumo de combustible La tabla adjunta lista los pesos (en cientos de libras) y las cantidades de consumo de combustible en carretera (en millas/galón), de una muestra de automóviles estadounidenses nuevos (datos de la Agencia para la Protección Ambiental). Con base en el resultado, ¿esperaría gastar más en gasolina si compra un automóvil más pesado? ¿Cómo cambian los resultados si los pesos se introducen como 2900, 3500, . . . , 2400?

x Peso	29	35	28	44	25	34	30	33	28	24
y Combustible	31	27	29	25	31	29	28	28	28	33

6. ¿Es aleatoria la lotería? A continuación se presentan los primeros dígitos que se seleccionaron de 40 tomas consecutivas de la urna del juego de lotería Win 4 del estado de Nueva York. (Véase el conjunto de datos 26 del Apéndice B). ¿Los dígitos impares y pares parecen tomarse 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

7. ¿El peso de un automóvil afecta las heridas en la pierna en un choque? Se obtuvieron datos de experimentos de choques de automóviles que realizó la National Transportation Safety Administration. Se adquirieron automóviles nuevos y se chocaron contra una barrera fija a 35 millas/hora. Las mediciones se registraron con un maniquí en el asiento del conductor. Utilice los datos muestrales que se listan más adelante para probar las diferencias en las mediciones de carga (en libras) del fémur izquierdo entre las cuatro categorías de peso. ¿Hay 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

8. Prueba de correlación entre rendimiento y precio *Consumer Reports* probó las cintas VHS que se utilizan en las reproductoras de vídeo. A continuación se presentan las puntuaciones de rendimiento y precios (en dólares) de cintas que se seleccionaron al azar. ¿Hay una correlación entre el rendimiento y el precio? ¿Qué sugiere la conclusión acerca de la compra de cintas VHS?

Rendimiento	91	92	82	85	87	80	94	97
Precio	4.56	6.48	5.99	7.92	5.36	3.32	7.32	5.27

Ejercicios de repaso acumulativo

1. Análisis de resultados de encuesta Un investigador de mercado para American Airlines recibió instrucciones de seleccionar al azar a pasajeros en espera de abordar (el autor fue uno de los sujetos que se seleccionaron). A los pasajeros se les hicieron varias preguntas acerca del servicio de la aerolínea. Se registraron las respuestas junto con sus géneros, los cuales se listan en el orden en el que se seleccionó a las personas.
   a. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que la secuencia es aleatoria.
   b. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que la proporción de mujeres es diferente de 0.5. Use la prueba paramétrica descrita en la sección 7-3.
   c. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que la proporción de mujeres es diferente de 0.5. Use la prueba del signo descrita en la sección 12-2.
   d. Utilice los datos muestrales para construir un intervalo de confianza del 95% para la población de mujeres.
   e. ¿Qué sugieren los resultados anteriores? ¿La muestra se sesgó en contra de algún género? ¿Se obtuvo la muestra en una secuencia aleatoria? Si usted fuera el director, ¿tendría algún problema con estos resultados?

   H H H H H M H H H M H H M H M H H M H H M H H M H H H

2. Estaturas de ganadores y perdedores presidenciales La tabla adjunta indica las estaturas de los presidentes, apareadas con las estaturas de los candidatos a quienes ellos derrotaron. Todas las estaturas están en pulgadas y sólo se incluyen los candidatos que quedaron en segundo lugar. Use un nivel de significancia de 0.05 para lo siguiente.
   a. Utilice el coeficiente de correlación lineal r para probar si hay una correlación lineal significativa entre las estaturas de los ganadores y las estaturas de los candidatos que derrotaron. (Véase la sección 9-2). ¿Parece existir una correlación?
   b. Utilice el coeficiente de correlación de rangos r_s para probar si hay una correlación lineal significativa entre las estaturas de los ganadores y las estaturas de los candidatos que derrotaron. (Véase la sección 2-6). ¿Parece existir una correlación?
   c. Utilice la prueba del signo para probar la aseveración de que hay una diferencia entre las estaturas de los candidatos ganadores y las estaturas de los candidatos perdedores.
   d. Utilice la prueba de rangos con signo de Wilcoxon para probar la aseveración de que hay una diferencia entre las estaturas de los candidatos ganadores y las estaturas de los candidatos perdedores correspondientes.
   e. Utilice la prueba paramétrica t (véase la sección 8-4) para probar la aseveración de que hay una diferencia entre las estaturas de los candidatos ganadores y las estaturas de los candidatos perdedores correspondientes.
   f. ¿Qué sugieren los resultados anteriores acerca de las estaturas de los candidatos presidenciales ganadores y las estaturas de los candidatos presidenciales perdedores correspondientes?

Ganador	76	66	70	70	74	71.5	73	74
Segundo lugar	64	71	72	72	68	71	69.5	74

Actividades de cooperación en equipo

1. 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 realizará el análisis en subgrupos de tres o cuatro estudiantes.

2. Actividad en clase Forme grupos de ocho a 12 personas. Para cada miembro del grupo, *mida* su estatura y el largo de sus brazos. Para medir el largo de los brazos, el sujeto debe pararse con los brazos extendidos, como las alas de un avión. Es fácil marcar la altura y la apertura de brazos en un pizarrón. Mida las distancias ahí. Divida las tareas siguientes en subgrupos de tres o cuatro personas.
   a. Utilice la correlación de rangos con los datos muestrales apareados para determinar si hay una correlación entre la estatura y la apertura de brazos.
   b. Utilicen la prueba del signo para probar la diferencia entre las dos variables.
   c. Utilicen la prueba de rangos con signo de Wilcoxon para probar la diferencia entre las dos variables.

3. Actividad en clase Realice la actividad 2 utilizando el pulso en lugar del largo de los brazos. Mida los pulsos contando el número de latidos cardiacos en un minuto.

4. Actividad fuera de clase Forme grupos de tres o cuatro estudiantes. Investigue la relación entre dos variables reuniendo sus propios datos muestrales apareados y utilizando los métodos de la sección 12-6 para determinar si hay una correlación significativa. Temas que se sugieren:
   - ¿Hay 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 beisbol (o basquetbol o futbol) y sus logros en la temporada?
   - Tasas contra pesos: ¿Hay una relación entre las tasas de consumo de combustible de los automóviles y los pesos de los automóviles?
   - ¿Existe una relación entre las longitudes de los pies de los hombres (o de las mujeres) y sus estaturas?
   - ¿Hay una relación entre el promedio de las calificaciones de los estudiantes y la cantidad de tiempo que destinan a ver televisión?
   - ¿Existe una relación entre las estaturas de los padres (o de las madres) y las estaturas de sus hijos (o hijas) primogénitos?

5. Actividad fuera de clase Consulte el proyecto “De los datos a la decisión”, relativo al análisis del sorteo de 1970 utilizado para reclutar hombres en el ejército estadounidense. Puesto que los resultados de 1970 elevaron el interés acerca de 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 que se obtuvieron en 1970? ¿Su procedimiento de selección aleatoria parece ser mejor que el que se usó en 1970? Elabore un reporte que describa con claridad el proceso que diseñó. También incluya su análisis y conclusiones.

6. Actividad fuera de clase Forme grupos de tres o cuatro. Encueste a estudiantes, pidiéndoles que identifiquen su área de estudios y su género. Para cada sujeto que se entreviste, determine la precisión de la hora en su reloj. Primero ajuste su propio reloj a la hora correcta utilizando una fuente precisa y confiable (“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 dichas preguntas:
   - ¿Parecen los errores ser los mismos para ambos géneros?
   - ¿Parecen los errores ser los mismos para las diferentes áreas de estudio?

7. Actividad en clase Forme grupos de ocho a 12 personas. Para cada miembro del grupo, mida la estatura de la persona y mida también 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 hay una correlación entre la estatura y la altura del ombligo.

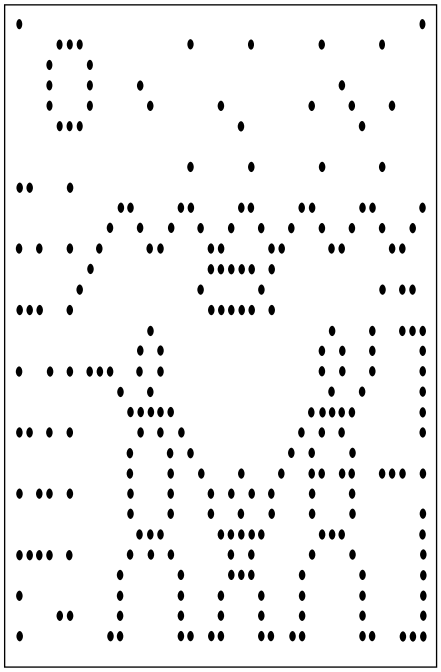
8. Actividad en clase Forme grupos de tres o cuatro personas. El Apéndice B incluye muchos conjuntos de datos que todavía no están resueltos con los métodos de este capítulo. Por ejemplo, si utilizamos el conjunto de datos 25, investigaríamos la correlación entre los valores máximos del promedio industrial Dow-Jones y los números de ventas de automóviles en Estados Unidos. Revise el Apéndice B y busque las variables de interés, luego investigue con el uso de los métodos apropiados de estadística no paramétrica. Enuncie sus conclusiones y trate de identificar aplicaciones prácticas.

Proyecto tecnológico

Intentos pasados de identificar vida inteligente extraterrestre incluyeron esfuerzos para enviar mensajes de radio llevando información acerca de nosotros los terrícolas. El doctor Frank Drake, de Cornell University, desarrolló un mensaje de radio de este tipo, que se transmitiría en series de 1271 pulsos y silencios. Se piensa en los pulsos y silencios como ceros y unos. Si factorizamos 1271 en los números primos 41 y 31, y luego hacemos una cuadrícula de 41×31 y ponemos un punto en aquellas posiciones correspondientes a un pulso o 1, obtendremos el patrón que se muestra en la figura adjunta. Dicho patrón contiene información que incluye la posición de la Tierra en el sistema solar; los símbolos del hidrógeno, el carbono y el oxígeno, así como dibujos de un hombre, una mujer, un niño, un pez y agua. Trate de identificar en el patrón al hombre, la mujer, el pez y el agua.

Suponga que la secuencia de 1271 unos y ceros se envía como un mensaje de radio que se intercepta por vida extraterrestre con suficiente inteligencia como para haber estudiado este libro. Si el mensaje de radio se prueba utilizando los métodos de este capítulo, ¿la secuencia parecerá ser “ruido aleatorio” o se identificará como un patrón que no es aleatorio? Utilice el STATDISK o el Minitab para el análisis.

```
10000000000000000000000000000000000000001
00001110000000000100000100000010000010000
00010001000000000000000000000000000000000
00010001000010000000000000000000100000000
00010001000001000000100000000100010001000
00001110000000000000001000000000001000000
00000000000000000000000000000000000000000
00000000000000000100000100000010000010000
11000100000000000000000000000000000000000
00000000001100001100001100001100001100001
00000000010010010010010010010010010010010
10100100100001100001100001100001100001100
00000001000000000001111101000000000000000
00000010000000000010000010000000000010110
11100100000000000000111110100000000000000
00000000000001000000000000000001000100111
00000000000010100000000000000010100100001
10010101110010100000000000000010100100001
00000000001001000000000000000001001000001
00000000000111110000000000000111110000001
11010100000010101000000000001010100000001
00000000000100010100000000010100010000000
00000000000100010010001000100110110011101
10110100000100010001010101000100010000000
00000000000100010001001001000100010000001
00000000000011100000111110000011100000001
11110100000101010000010100000100010000001
00000000001000001000011100001000001000001
10000000001000001000100010001000001000001
00001100001000001000100010001000001000001
10000000011000001101100011011000001100111
```



de los DATOS *a la* DECISIÓN

Pensamiento crítico: ¿Fue aleatorio el sorteo?

En 1970 se utilizó un sorteo para determinar quién se reclutaría 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. Entonces 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 que se seleccionó fue el 14 de septiembre, así que los hombres que nacieron en esa fecha se reclutaron primero. La lista adjunta muestra las 366 fechas en el orden de su selección.

Análisis de los resultados

a. Utilice la prueba de rachas para probar la aleatoriedad de la secuencia por encima y por debajo de la mediana de 183.5.

b. Utilice la prueba de Kruskal-Wallis para probar la aseveración de que los 12 meses tienen números prioritarios que se obtuvieron de la misma población.

c. Calcule las 12 medias mensuales. Luego registre estas 12 medias en una gráfica (donde la escala horizontal liste los 12 meses y la escala vertical vaya desde 100 hasta 260). Observe cualquier patrón que sugiera que los números

continúa

prioritarios originales no se seleccionaron aleatoriamente.

d. 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 explique por qué cree que no fue justo. Si decide que este sorteo no fue justo, describa un procedimiento para seleccionar los números 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

Este capítulo introdujo métodos de prueba de hipótesis de la variedad no paramétrica o de distribución libre. Los métodos no paramétricos permiten probar hipótesis sin hacer supuestos al respecto de la distribución poblacional subyacente que se está muestreando. Para continuar su trabajo con este importante tipo de métodos estadísticos de prueba, vaya al sitio Web de *Estadística*:

http://www.pearsoneducacion.net/triola

El proyecto de Internet para este capítulo le pide que consulte de nuevo algunos de los conjuntos de datos de proyectos anteriores; en específico, los conjuntos de datos utilizados en pruebas paramétricas. Esta vez, sin embargo, utilizará la prueba no paramétrica adecuada y comparará los resultados. Además, realizará una investigación de aleatoriedad aplicando la prueba de rachas.

La estadística @ en el trabajo

"Si no hubiera tenido una formación en estadística, no sería capaz de entender por completo los datos que produce mi compañía. . . ayuda para proteger a nuestros trabajadores y clientes".

Jeffrey Foy

Jeffrey Foy es un toxicólogo que trabaja para la Cabot Corporation, una empresa de químicos.

Jeffrey Foy también es el responsable de la evaluación de la peligrosidad de los químicos que produce la Cabot Corporation. Su trabajo consiste en entender la forma en que los productos de la compañía pueden afectar a los seres humanos o al ambiente, así como ayudar a decidir sobre las mejores maneras para proteger a ambos.

¿Qué hace usted en su trabajo?

Mis responsabilidades incluyen organizar y evaluar los estudios toxicológicos, escribiendo las hojas de especificaciones de seguridad de los materiales, y ayudando a que nuestros grupos de investigación y desarrollo produzcan materiales que sean seguros, tanto para las personas como para el ambiente. También investigo qué peligros potenciales pueden tener los materiales.

¿Qué conceptos de estadística utiliza usted?

El concepto primario que utilizo es la prueba de hipótesis (prueba de probabilidad).

¿Cómo utiliza la estadística en el trabajo?

Utilizo la estadística a diario. En mi trabajo los métodos estadísticos se emplean en dos formas. Primero, la estadística se utiliza para determinar la forma en que diseño mis experimentos. Segundo, la estadística sirve para determinar si los datos que se generan son significativos o, algunas veces, incluso si son suficientemente buenos como para trabajar con ellos.

Los estudios en los que intervengo pueden costar tan poco como $1000 hasta tanto como $500,000 o más; si usted no determina adecuadamente cómo va a evaluar los datos, podría costarle a su empresa un gran desperdicio de tiempo y dinero. Si el experimento se realiza adecuadamente, entonces nos movemos sobre el análisis de los datos. Los datos de los estudios que realizamos se utilizan para valorar cualquier efecto potencial que nuestros productos tuvieran en la salud de nuestros trabajadores, clientes o en el ambiente. Los resultados se utilizan para determinar cómo pueden venderse o manejarse los químicos. Cuando se realizan experimentos en un laboratorio de pruebas o una compañía de fármacos, quiere determinar si sus materiales tienen un efecto, si es el que se desea (un fármaco que cura una enfermedad) o el que no se desea (que ese mismo fármaco sea tóxico). La estadística juega un papel crucial en nuestra evaluación de la significancia de los efectos.

Por favor, describa un ejemplo específico que ilustre la forma en que el uso de la estadística tuvo éxito en la mejora de un producto o servicio.

Recientemente se realizó un estudio toxicológico que costó cerca de $300,000. Los datos del estudio iban a utilizarse para ayudar a determinar si un químico en particular causaba algún efecto en los sujetos que se estudiaron. Después de que se realizó el estudio, se encontraron defectos en los datos y en la estadística que se empleó. Tomó dos años más revisar adecuadamente los datos y terminar la evaluación de salud. Si se hubieran elegido los métodos y los fines adecuados, el tiempo y el dinero adicionales no hubieran sido necesarios. El conocimiento de los datos y la evaluación estadística correcta fue lo que ayudo a prevenir el fracaso y la potencial repetición del estudio.

13

Control estadístico de procesos

PROBLEMA DEL CAPÍTULO

¿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 precisión 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 precisión de los altímetros es especialmente importante cuando los pilotos van a aterrizar sin ver el suelo. En el pasado, pilotos y pasajeros han muerto en accidentes ocasionados 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 a una distancia baja.

Puesto que los altímetros de aviones son sumamente importantes para la seguridad de la aviación, su precisión se controla con mucho cuidado a través de 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 cuatro altímetros al azar de la producción diaria durante 20 días hábiles consecutivos; en la tabla 13-1 se muestran 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 son 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 el proceso de fabricación de altímetros, analizando el comportamiento de los errores al paso del tiempo. Estudiaremos, también, la forma en que se utilizan los métodos estadísticos para verificar un proceso de fabricación, con la meta de identificar y corregir cualquier problema grave. Además de ayudar a que las empresas permanezcan abiertas, los métodos de estadística pueden afectar de manera positiva nuestra seguridad de forma muy significativa.

Tabla 13-1 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

13-1 Panorama general

En el capítulo 2 señalamos que al describir, explorar o comparar conjuntos de datos, las siguientes características suelen ser extremadamente importantes. (Sugerimos que la frase "Cuidado con los Virus que Destruyen Datos y Trabajo" puede utilizarse como un recurso mnemotécnico para recordar CVDDT, que resume dichas características).

1. *Centro:* Medida de tendencia central representativa del valor promedio que nos indica en dónde se localiza la parte media del conjunto de datos.
2. *Variación:* Medida de la cantidad en que los valores varían entre ellos.
3. *Distribución:* Naturaleza o forma de la distribución de los datos, tal como en forma de campana, uniforme o sesgada.
4. *Datos distantes:* Valores muestrales que se encuentran muy alejados de la gran mayoría de los otros datos muestrales.
5. *Tiempo:* Características cambiantes de los datos a lo largo del tiempo.

El principal objetivo de este capítulo es poner énfasis en el quinto 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 con el paso del tiempo.

Actualmente hay una fuerte tendencia a tratar de mejorar la calidad de los bienes y servicios estadounidenses, a la vez que un número creciente de empresas están utilizando los métodos que se presentan en este capítulo. La evidencia de la creciente importancia de la calidad se encuentra en la publicidad, así como en el gran número de libros y artículos que cada vez destacan más el tema. En muchos casos, quienes solicitan empleo (¿usted?) poseen una ventaja definitiva cuando son capaces de 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 estas representaciones gráficas. Las gráficas de control, los histogramas, las gráficas de cuadro y los diagramas de dispersión son algunos 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 comprender. El mundo necesita más personas capaces de construir e interpretar gráficas importantes, tales como las gráficas de control descritas en este capítulo.

13-2 Gráficas de control para la variación y la media

El principal objetivo de esta sección es controlar características importantes de datos a lo largo del tiempo. Este tipo de datos suelen denominarse *datos de proceso*.

Definición

Datos de proceso: 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 13-1 incluye datos de proceso consistentes 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 13-1 se ordenan 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 llegan cambiar a lo largo del tiempo.

Al producir altímetros, el fabricante puede emplear personal competente y bien entrenado, además de buenas máquinas correctamente calibradas; no obstante, si el personal es reemplazado o las máquinas se estropean con el uso, los altímetros empezarían a resultar defectuosos. Hay compañías que fueron a la bancarrota por permitir, involuntariamente, que el proceso de fabricación se deteriorara al no tener un control constante.

Gráficas de rachas

Hay varios métodos que permiten controlar un proceso y así asegurar que las características importantes que se desean no cambien; el análisis de una *gráfica de rachas* es un método de este tipo.

Definición

Gráfica de rachas: Una gráfica secuencial de valores de datos *individuales* a lo largo del tiempo. Un eje (generalmente el vertical) se utiliza para los valores de los datos, en tanto que el otro eje (generalmente el horizontal) se emplea para la secuencia de tiempo.

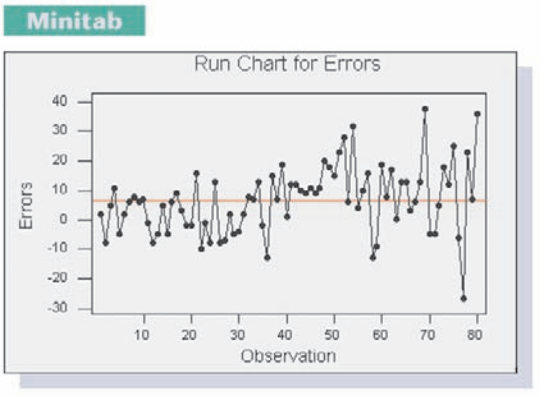
PROBLEMA 13 DEL CAPÍTULO

EJEMPLO Fabricación de altímetros para aviones Trate los 80 errores de los altímetros de la tabla 13-1 como una secuencia de mediciones consecutivas, construya una gráfica de rachas, utilice el eje vertical para los errores y el eje horizontal para identificar el orden de los datos muestrales.

SOLUCIÓN La figura 13-1 es la gráfica de rachas generada por Minitab, para los datos de la tabla 13-1. La escala vertical se diseñó para ajustarse a los errores de los altímetros, que van desde –27 hasta 38 pies, que son los valores mínimo y máximo de la tabla 13-1. La escala horizontal se diseñó 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.

continúa

FIGURA 13-1 **Gráfica de rachas de los errores individuales de altímetros de la tabla 13-1**



En la figura 13-1, la escala horizontal identifica el número de muestra, de forma que el número 20 indica el artículo 20o. La escala vertical representa el error del altímetro (en pies). Ahora examine la figura 13-1 y trate de identificar cualquier *patrón* que resalte a la vista. La figura 13-1 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 están correctos, 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 que se deterioró con el paso del tiempo. Si se deja como está, dicho proceso de fabricación provocará que la empresa cierre.

Interpretación de las gráficas de rachas Únicamente cuando un proceso es *estadísticamente estable*, sus datos se tratan como si provinieran de una población con una media, una desviación estándar, una distribución y otras características constantes.

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 13-2 ilustra los patrones típicos que indican formas en las cuales el proceso de llenado de latas de sopa de 16 onzas puede no ser estadísticamente estable.

- **Figura 13-2(a):** Hay una evidente *tendencia creciente*, que corresponde a valores que se incrementan al paso del tiempo. Si el proceso de llenado continúa con este tipo de patrón, las latas se llenarían con más y más sopa hasta empezar a derramarse y eventualmente los empleados terminarían nadando en sopa.
- **Figura 13-2(b):** Hay una evidente *tendencia descendente* que corresponde a valores que disminuyen de manera estable. Las latas se llenarían con me-

nos y menos sopa hasta estar 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.

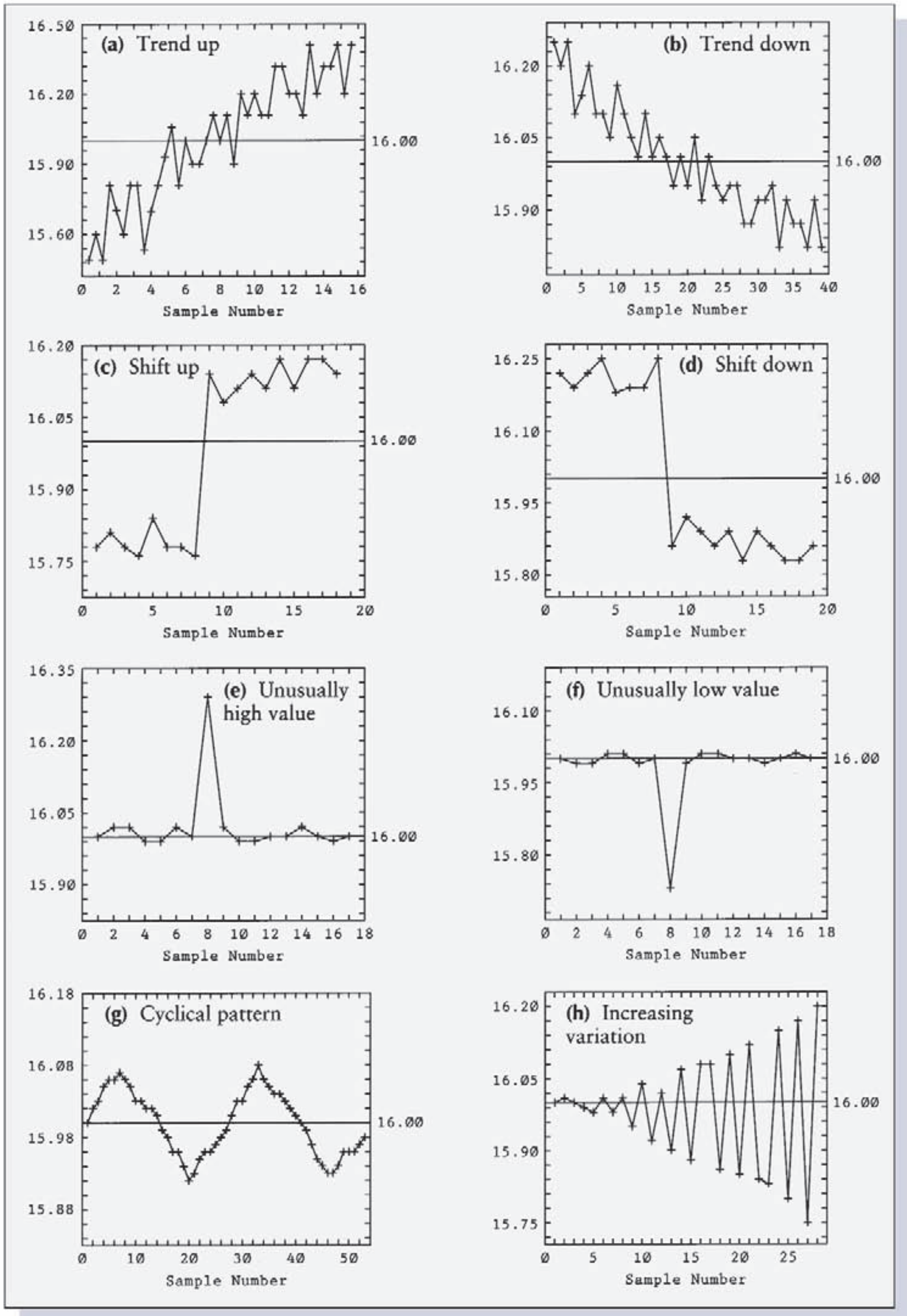


FIGURA 13-2 **Procesos que no son estadísticamente estables**

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 exhiben una tendencia a incrementarse, ya que las puntuaciones de CI están aumentando de forma estable desde que empezaron a utilizarse hace casi 70 años. Dicha 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 la tendencia al incremento se le llama *efecto Flynn*, porque el científico político James R. Flynn la descubrió 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.

- **Figura 13-2(c):** Existe un *cambio hacia arriba*. Una gráfica de rachas como ésta resultaría de un ajuste en el proceso de llenado, provocando que los valores subsecuentes sean más altos.
- **Figura 13-2(d):** Hay un *cambio hacia abajo*. Los primeros valores son relativamente estables, pero después algo sucede, ya que los últimos valores son relativamente estables, aunque a un nivel mucho más bajo.
- **Figura 13-2(e):** El proceso es estable, excepto 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.
- **Figura 13-2(f):** Existe un *valor excepcionalmente bajo*.
- **Figura 13-2(g):** Hay un patrón cíclico (o ciclo repetitivo). Evidentemente, este patrón no es aleatorio; por lo tanto, revela un proceso estadísticamente inestable. Quizá se hagan reajustes periódicos a la maquinaria, con el efecto de que se busca de continuo algún valor deseado, pero nunca se logra bien.
- **Figura 13-2(h):** La variación está aumentando al 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 sopa se derramarán, desperdiciando sopa, 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 se construyeron dentro de los límites permitidos, las transmisiones Mazda eran más confiables por su menor variación. La variación en un proceso a veces resulta por dos causas.

Definiciones

Variación aleatoria: El tipo de variación inherente a cualquier proceso que no es capaz de producir cada bien o servicio exactamente de la misma forma cada vez.

La **variación asignable** resulta de causas que pueden identificarse (factores tales como maquinaria defectuosa, empleados sin entrenamiento, etcétera).

Más adelante, en este capítulo, consideraremos formas de distinguir entre la variación asignable y la variación aleatoria.

La gráfica de rachas es una herramienta para controlar la estabilidad de un proceso. Ahora estudiaremos las *gráficas de control*, que también son sumamente útiles para los mismos propósitos.

Gráfica de control para verificar la variación: la gráfica *R*

En el artículo "The State of Statistical Process Control as We Proceed into the 21st Century" (de Stoumbos, Reynolds, Ryan y Woodall, *Journal of the American Statistical Association,* vol. 95, núm. 451), los autores afirman que "las gráficas de control son de las herramientas más importantes y que más se utilizan en la estadística. Sus aplicaciones pasaron 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

Gráfica de control de una característica de proceso (como la media o la variación): consiste en valores que se grafican secuencialmente 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 cualquier punto que se considera *fuera de lo común*.

Asumiremos 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 verificar la variación, y 2. las gráficas $\overline{x}$, que se emplean para verificar medias. Al manejar gráficas de control para verificar procesos, es común que se tomen en cuenta las gráficas R y las gráficas $\overline{x}$ al mismo tiempo, ya que un proceso estadísticamente inestable puede ser el resultado de un aumento en la variación, de cambios en las medias o de ambos.

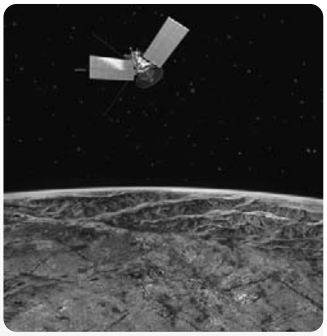
Una gráfica R (o gráfica de rangos) es una gráfica de los rangos muestrales, en lugar de valores muestrales individuales; se aplica para verificar la *variación* en un proceso. (Parecería más sensato utilizar desviaciones estándar, pero las gráficas de rangos se emplean con mayor frecuencia en la práctica. Esto es una consecuencia de los tiempos en que no se disponía de calculadoras ni de computadoras. Véase el ejercicio 13, donde se incluye una gráfica de control que se basa en desviaciones estándar). Además de graficar los valores de los rangos, incluimos una línea central que se localiza en $\overline{R}$, 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

Considere: Los datos de proceso consisten en una secuencia de muestras, todas del mismo tamaño, y la distribución de los datos de proceso es esencialmente normal.

n = tamaño de cada muestra o *subgrupo*

$\overline{R}$ = media de los rangos muestrales (es decir, la suma de los rangos muestrales, dividida entre el número de muestras)



Variación asignable costosa

La NASA envió a Marte el *Mars Climate Orbiter*, aunque se destruyó cuando voló muy cerca del planeta destino. La pérdida se calculó en 125 millones de dólares. 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 se dieron en las unidades inglesas de libras de fuerza, pero el Jet Propulsion Laboratory asumió que las unidades eran "newtons" métricos en lugar de libras. Quienes dirigían la nave espacial proporcionaron subsecuentemente cantidades erróneas de la fuerza para ajustar la posición de la nave. Los errores que causó la discrepancia fueron muy pequeños al principio, pero el error que se acumuló a lo largo de los meses de travesía de la nave espacial probó ser 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 por la falta de un signo menos en un programa de cómputo.

Verificación de un proceso de variación: gráfica de control para R

Puntos graficados: rangos muestrales

Línea central: $\overline{R}$

Límite de control superior (LCS): $D_4\overline{R}$ (donde D_4 se encuentra en la tabla 13-2)

Límite de control inferior (LCI): $D_3\overline{R}$ (donde D_3 se encuentra en la tabla 13-2)

Tabla 13-2 Constantes de una gráfica de control

	$\overline{x}$		s		R	
n: Número de observaciones en subgrupo	A_2	A_3	B_3	B_4	D_3	D_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. Se reproduce con autorización de American Society for Testing and Materials.

Los valores D_4 y D_3, que fueron calculados por expertos en control de calidad, sirven para simplificar los cálculos. Los límites de control superior e inferior de $D_4\overline{R}$ y $D_3\overline{R}$ 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 tales límites. Si un valor cae fuera de esos límites, es muy probable que el proceso no sea estadísticamente estable.

EJEMPLO Fabricación de altímetros para aviones Remítase a los errores de los altímetros en la tabla 13-1. Con el uso de muestras de tamaño $n = 4$, que se reúnen cada día de fabricación, construya una gráfica de control para R.

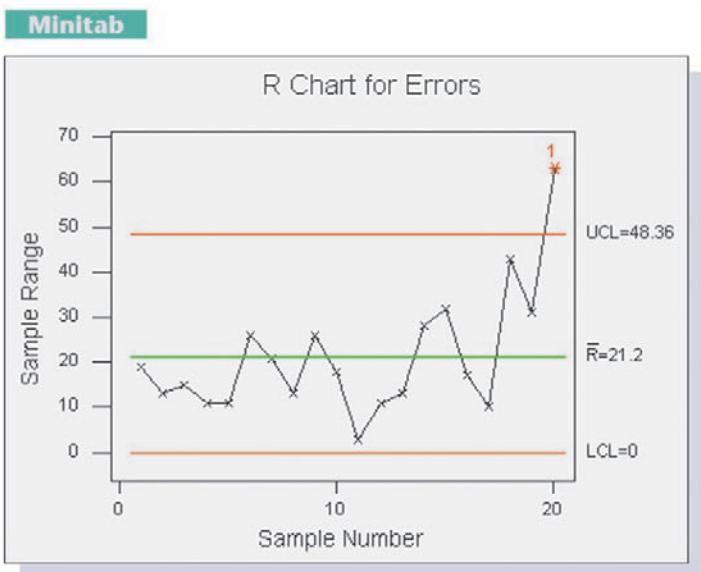
SOLUCIÓN Iniciamos con el cálculo del valor de $\overline{R}$, la media de los rangos muestrales.

$$\overline{R} = \frac{19 + 13 + \cdots + 63}{20} = 21.2$$

Por lo tanto, la línea central de nuestra gráfica está en $\overline{R} = 21.2$. 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 13-2, para $n = 4$, obtenemos $D_3 = 0.000$ y $D_4 = 2.282$, de manera que los límites de control son los siguientes:

$$\text{Límite de control superior: } D_4\overline{R} = (2.282)(21.2) = 48.4$$
$$\text{Límite de control inferior: } D_3\overline{R} = (0.000)(21.2) = 0.0$$

Con un valor de línea central de $\overline{R} = 21.2$, así como con límites de control de 48.4 y 0.0, procedemos a graficar los rangos muestrales. Los resultados se presentan en la pantalla de Minitab.



¡No manosear!

La empresa Nashua Corp., que tuvo problemas con su máquina para recubrimiento de papel, 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, que se denominan *manoseo*, causaron desviaciones de la distribución que hasta entonces había sido buena. El efecto fue un incremento en los defectos. Cuando el estadístico y experto en 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 el manoseo realizado.

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 se basan 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 decretadas por el fabricante.**

Al investigar la calidad de algún proceso, hay comúnmente dos preguntas importantes que necesitan plantearse:

1. Con base en el comportamiento actual del proceso, ¿concluiremos 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 desarrollaron para responder la primera pregunta, aunque no la segunda. Es decir, nos enfocamos en el comportamiento del proceso, con el objetivo de determinar si está 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 de este capítulo. Por ejemplo, la gráfica *R* de Minitab que se muestra aquí 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 13-1. Las normas gubernamentales requieren que los altímetros tengan errores entre −20 pies y 20 pies, sin embargo, las especificaciones que se desean (o requieren) 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 considerado que un proceso no es estadísticamente estable si su patrón se asemeja a cualquiera de los que se presentan en la figura 13-2. Este criterio se incluye con algunos otros de la siguiente lista.

Criterios para determinar cuando un proceso no es estadísticamente estable (fuera de control estadístico)

1. Hay un patrón, una tendencia o un ciclo que evidentemente no son aleatorios (tales como los que se incluyen en la figura 13-2).
2. Existe un punto fuera de la región entre los límites superior e inferior. (Esto es, hay 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, hay 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 antes mencionados 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 lejos de los límites de control y a dos desviaciones estándar de la línea central.
- Cuatro de cada cinco puntos consecutivos están lejos de los límites de control y a una desviación estándar de la línea central.

EJEMPLO Control estadístico de procesos Examine la gráfica *R* del ejemplo anterior, que se muestra en la pantalla de Minitab, y determine si la variación del proceso está bajo control estadístico.

SOLUCIÓN Interpretamos 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, concluiremos que la variación del proceso está fuera de control estadístico. No hay ocho puntos consecutivos por encima o por debajo de la línea central, de forma que no se viola la tercera condición, pero las primeras dos condiciones no se cumplen.

1. Existe un patrón, una tendencia o un ciclo que evidentemente no son aleatorios: de izquierda a derecha hay un patrón de tendencia creciente, como en la figura 13-2a.
2. Existe 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 la variación parece incrementarse con el tiempo, tiene que hacerse una corrección inmediata para fijar la *variación* entre los errores de los altímetros.

Gráfica de control para verificación de medias: la gráfica $\bar{x}$

Una **gráfica** $\bar{x}$ es una gráfica de las medias muestrales que se utiliza para verificar el *centro* en un proceso. Además de graficar las medias muestrales, incluimos una línea central que se localiza en $\bar{\bar{x}}$, lo cual denota la media de todas las medias muestrales (igual a la media de todos los valores muestrales que se combinan), 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 la industria, la línea central y los límites de control se basan en rangos, en lugar de desviaciones estándar. Véase el ejercicio 14, que incluye una gráfica $\bar{x}$ que se basa en desviaciones estándar.

Verificación 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 13-2)

Límite de control inferior (LCI): $\bar{\bar{x}} - A_2\bar{R}$ (donde A^2 se encuentra en la tabla 13-2)

Detección de soborno con gráficas de control

Las gráficas de control se utilizaron para ayudar a sentenciar 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 Statician,* 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 dichas apuestas. En la Corte se utilizaron gráficas *R* y $\bar{x}$ como evidencia de patrones sumamente raros de apuestas. El examen de las gráficas de control muestra claramente puntos que se encuentran muy lejos del límite de control superior, lo que indica que el proceso de apuestas estaba fuera de control estadístico. El estadístico fue capaz de identificar un dato en el cual la variación asignable parecía detenerse, aunque los fiscales saben que se trata de la fecha de arresto del acusado.

EJEMPLO Fabricación de altímetros para aviones Remítase a los errores en los altímetros en la tabla 13-1. Con el uso de las muestras de tamaño $n = 4$, que se reunieron 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 hay que calcular el valor de la línea central y los valores de los límites de control. Obtenemos

$$\bar{\bar{x}} = \frac{2.50 + 2.75 + \cdots + 9.75}{20} = 6.45$$

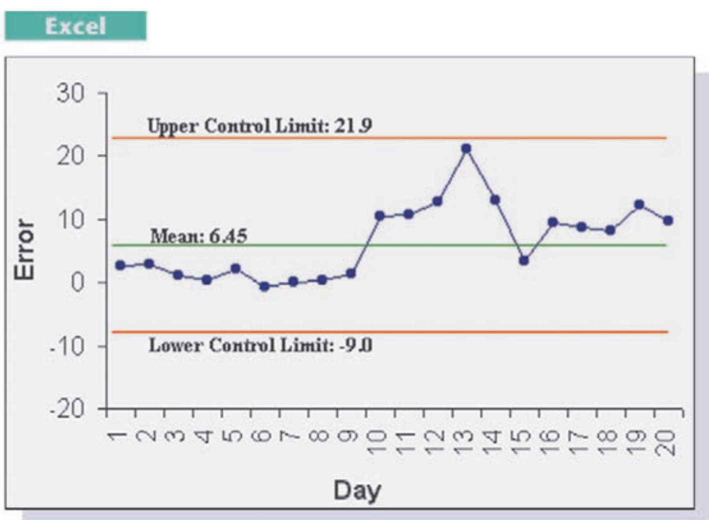
$$\bar{R} = \frac{19 + 13 + \cdots + 63}{20} = 21.2$$

Si nos remitimos a la tabla 13-2, encontramos que para $n = 4$, $A_2 = 0.729$. Conociendo los valores de $\bar{\bar{x}}$, A_2 y $\bar{R}$, evaluaremos los límites de control.

Límite de control superior: $\bar{\bar{x}} + A_2\bar{R} = 6.45 + (0.729)(21.2) = 21.9$

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 se muestra 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, el tercer criterio no está satisfecho porque hay ocho (o más) puntos consecutivos por debajo de la línea central. Además, parece existir un patrón de tendencia creciente. Nuevamente, se requieren acciones correctivas para fijar el proceso de producción.



Utilizando 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 que se aprecia en la figura 13-1, inicie introduciendo todos los datos muestrales en la columna C1. Seleccione la opción **Stat**, luego **Quality Tools** y después **Run Chart.** En los recuadros que se indican, introduzca C1 para la columna única de variable y 1 para el tamaño del subgrupo, después haga clic en **OK**.

Gráfica *R*: Primero introduzca los valores muestrales individuales de manera secuencial en la columna C1. Después, seleccione las opciones **Stat, Control Charts** y **R**. Ingrese C1 en el recuadro de "single column", el tamaño de la muestra en el recuadro del tamaño del subgrupo. Haga clic en **estimate**. Seleccione **Rbar.** (La selección del estimado *R*bar hace que la variación de la distribución poblacional se estime con los rangos muestrales, en lugar 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 y Xbar. Introduzca C1 en el recuadro de *single column*, el tamaño de cada muestra en el recuadro de *subgroup size box* y haga clic en estimate; después, seleccione Rbar. Haga clic en OK dos veces.

Excel Para utilizar el complemento Data Desk XL, haga clic en **DDXL** y seleccione **Process Control**. Seleccione el tipo de gráfica que desea. (Primero debe introducir los datos en la columna A con los códigos identificadores de muestra que se indican en la columna B. Por ejemplo, para los datos de la tabla 13-1, 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 lugar de Data Desk XL, haga lo siguiente:

Gráfica de rachas: Anote todos los datos muestrales en la columna A. En la barra del menú principal haga clic en el icono **Chart Wizard**, que aparece como una gráfica de barras. Para el tipo de gráfica, seleccione **Line**. Para el subtipo de gráfica, la primera gráfica del segundo renglón, luego haga clic en **Next**. Continúe haciendo clic en **Next** y luego en **Finish**. La gráfica se editará para incluir etiquetas, borrar líneas, etcétera.

Gráfica R *Paso 1:* Ingrese los datos muestrales en los renglones y las columnas correspondientes al conjunto del datos. Por ejemplo, ingrese los datos de la tabla 13-1 en cuatro columnas (A, B, C, D) y 20 renglones, como aparecen en la tabla.

Paso 2: Después, cree una columna para el rango de valores, por medio del siguiente procedimiento. Posicione el cursor en la primera celda vacía a la derecha del bloque de datos muestrales, después introduzca dicha expresión en el recuadro de la fórmula: = MAX(A1:D1)–MIN(A1:D1), donde el rango A1:D1 tiene que 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 forma 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, así como los límites superior e inferior que se requieren editando la gráfica. Haga clic sobre la línea al final de la pantalla, después haga de nuevo 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 13-1 en cuatro 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 tiene que modificarse para describir el primer renglón de su conjunto de datos. Luego de presionar la tecla **Enter,** debe aparecer la media del primer renglón. Utilice el ratón para deslizar la esquina derecha inferior de esta celda, de forma 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, así como 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 de nuevo clic y deslice para colocar la línea correctamente. Esto no es sencillo.

13-2 Destrezas y conceptos básicos

1. a. ¿Qué son *datos de proceso*?
b. ¿Qué significa que un proceso esté fuera de control estadístico?
c. ¿Cuáles son los tres criterios para determinar si un proceso está fuera de control estadístico?
d. ¿Cuál es la diferencia entre variación aleatoria y variación asignable?
e. ¿Cuál es la diferencia entre una gráfica R y una gráfica $\bar{x}$?

Verificación de consumo de energía doméstica. *En los ejercicios 2 a 4, utilice la siguiente información: el autor registró el consumo de energía eléctrica (en kilowatt-hora) en su casa al norte de Nueva York, durante intervalos de dos meses por cuatro años. Los resultados se listan en la tabla.*

	Ene.-Feb.	Mar.-Abr.	May.-Jun.	Jul.-Ago.	Sep.-Oct.	Nov.-Dic.
Año 1	4762	3875	2657	4358	2201	3187
Año 2	4504	3237	2198	2511	3020	2857
Año 3	3952	2785	2118	2658	2139	3071
Año 4	3863	3013	2023	2953	3456	2647

2. Consumo de energía: construcción de una gráfica de rachas Construya una gráfica de rachas para los 24 valores. ¿Parece haber un patrón que sugiera que el proceso no está bajo control estadístico? ¿Existe algún patrón de variación que pueda explicarse?

3. Consumo de energía: construcción de una gráfica R Utilice muestras de tamaño 3, combinando los primeros tres valores de cada año y los últimos tres valores de cada año. Con las ocho muestras de tamaño 3, 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.

4. Consumo de energía: construcción de una gráfica $\bar{x}$ Utilice muestras de tamaño 3, combinando los primeros tres valores de cada año y los últimos tres valores de cada año. Con las ocho muestras de tamaño 3, 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 conducen al rechazo de una media estadísticamente estable. ¿Cuál es un efecto práctico de no tener bajo control estadístico tal proceso? Dé un ejemplo de una causa que pondría a un proceso fuera de control estadístico.

Construcción de gráficas de control para latas de aluminio. *Los ejercicios 5 y 6 se basan en las cargas axiales (en libras) de latas de aluminio con un grosor de 0.0109 pulgadas, tal como se listan en el conjunto de datos 20 del Apéndice B. La carga axial de una lata es el peso máximo que soporta por su costado, por lo cual es importante tener una carga axial suficientemente alta para que la lata no se destruya cuando la tapa superior se coloque en su lugar. Los datos provienen de un proceso de fabricación real y fueron proporcionados por un estudiante que utilizó una edición anterior de este libro.*

5. Durante cada día de producción, se seleccionaron siete latas de aluminio con un grosor de 0.0109 pulgadas, luego se midieron sus cargas axiales. A continuación se presentan los rangos de los diferentes días, aunque también se encuentran los valores en el conjunto de datos 20 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	

6. Durante cada día de producción, se seleccionaron 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, aunque también se encuentran los valores en el conjunto de datos 20 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 de dólar**. En los ejercicios 7 a 9, 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, sin embargo, 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 que se selecciona aleatoriamente cada 12 minutos durante 20 horas consecutivas. Los resultados se listan en la tabla adjunta.*

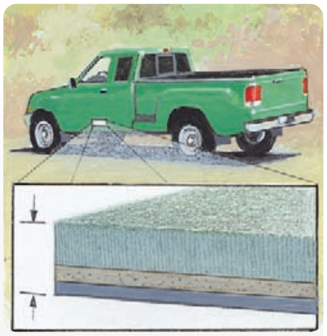
7. Acuñación de monedas: construcción de una gráfica de rachas Construya una gráfica de rachas para los 100 valores. ¿Parece haber un patrón que sugiera que el proceso no está bajo control estadístico? ¿Cuáles son las implicaciones prácticas de la gráfica de rachas?

8. Acuñación de monedas: construcción de una 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.

9. Acuñación de monedas: construcción de una 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



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 Ford Rangers y Jeep Grand Cherokees. El costo de la computadora de $20,000 se pagó con los ahorros de $40,000 del primer año de operaciones, que se emplearon para generar gráficas de control manuales que aseguraban que el grosor del aislamiento cumpliera con las especificaciones de medir entre 2.912 mm y 2.988 mm. Por medio del uso de gráficas de control y de otros métodos de control de calidad, Perstorp redujo sus mermas en más de dos tercios.

Construcción de gráficas de control para la lluvia en Boston. *En los ejercicios 10 a 12, remítase a las cantidades diarias de lluvia en Boston en un año, del conjunto de datos a 11 del Apéndice B. Omita el último dato de los miércoles, de manera que cada día de la semana tenga exactamente 52 valores.*

T **10.** 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?

T **11.** Lluvia en Boston: construcción de una gráfica R Utilice las 52 muestras, con siete valores cada una, para construir una gráfica R, luego 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.

T **12.** Lluvia en Boston: construcción de una gráfica $\bar{x}$ Con las 52 muestras, con siete valores cada una, construya una gráfica $\bar{x}$, luego 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.

13-2 Más allá de lo básico

13. Construcción de una gráfica s En esta sección describimos las gráficas de control de R y $\bar{x}$**,** que se basan en rangos. Las gráficas de control para verificar la variación y el centro (media) también pueden basarse en desviaciones estándar. Una *gráfica s* para verificar 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 en la tabla 13-2. Construya una gráfica s para los datos de la tabla 13-1. Compare el resultado con la gráfica R dada en esta sección.

14. Construcción de una gráfica $\bar{x}$ basada en desviaciones estándar Una gráfica $\bar{x}$ que se basa en desviaciones estándar (en lugar de 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 en la tabla 13-2 y es la media de las desviaciones estándar muestrales. Utilice los datos de la tabla 13-1 para construir una gráfica $\bar{x}$ basada en desviaciones estándar. Compare el resultado con la gráfica $\bar{x}$ que se basa en rangos muestrales (presentada en esta sección).

13-3 Gráficas de control de atributos

El principal objetivo de esta sección es desarrollar la habilidad para verificar un *atributo* construyendo e interpretando una gráfica de control propia. En la sección 13-2 verificamos datos *cuantitativos*, pero ahora consideraremos datos *cualitativos*, al investigar situaciones tales como si un artículo es defectuoso, si un artículo pesa menos que la cantidad prescrita o si un artículo no cumple con las normas. (Un bien o servicio no cumple con las normas si no satisface las especificaciones o requisitos; en ocasiones, los artículos que no cumplen con las normas se descartan, reparan o se denominan “de segunda”, por lo que se venden a precios bajos). Igual que en la sección 13-2, seleccionamos muestras de tamaño n en intervalos regulares de tiempo y dibujamos los puntos en una gráfica secuencial, con una línea central y límites de control. (Existen formas de manejar muestras con tamaños diferentes, pero por ahora no las consideraremos aquí).

La **gráfica de control de *p*** (o **gráfica *p***) es una gráfica de control que se utiliza para verificar la proporción p de algún atributo. La notación y los valores de una gráfica de control son los siguientes (donde el atributo de "defectuoso" puede reemplazarse por cualquiera otro relevante).

Notación

$\bar{p}$ = grado estimado de la proporción de artículos defectuosos en el proceso

$$= \frac{\text{número total de defectos que se encontraron en todos los artículos muestrados}}{\text{número total de artículos muestrados}}$$

$\bar{q}$ = estimado agrupado de la proporción de artículos del proceso que *no* son defectuosos

$= 1 - \bar{p}$

n = tamaño de cada muestra (no el número de muestras)

Gráfica de control de *p*

Línea central: $\bar{p}$

Límite de control superior: $\bar{p} + 3\sqrt{\dfrac{\bar{p}\,\bar{q}}{n}}$

Límite de control inferior: $\bar{p} - 3\sqrt{\dfrac{\bar{p}\,\bar{q}}{n}}$

(Si el cálculo del límite de control inferior da como resultado un valor negativo, utilice el 0 en su lugar. Si el cálculo del límite de control superior excede a 1, utilice el 1 en su lugar.)

Sea $\bar{p}$ la línea central, ya que 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 6-2.

EJEMPLO Muertes por enfermedad infecciosa Los médicos reportan que las enfermedades infecciosas deben verificarse cuidadosamente a lo largo del tiempo, ya que tienen mucho más posibilidades de sufrir cambios súbitos en las tendencias que enfermedades tales como el cáncer. En cada uno de 13 años consecutivos recientes, se seleccionaron al azar 100,000 sujetos y se registró el número de los que murieron de infecciones del tracto respiratorio; 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 p 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.

Número de muertes: 25 24 22 25 27 30 31 30 33 32 33 32 31

coninúa

Seis Sigma en la industria

Seis Sigma es el término utilizado en la industria para describir un proceso que da una proporción 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 de tres o cuatro defectos por millón.

Los programas Seis Sigma, que iniciaron en 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 de dólares en tres años. Allied Signal reportó ahorros de $1500 millones. GE, Polaroid, Ford, Honeywell, Sony y Texas Instruments son otras compañías grandes que adoptaron la meta Seis Sigma.

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 *The New York Times*, de Melody Patersen, "algunos de los problemas se relacionan con la falta de controles que identifican medicamentos defectuosos, mientras otros provienen de equipos muy viejos. Tales problemas se detectaron en alrededor de 200 medicamentos, incluido Claritin, el fármaco contra alergias que es el producto de mayor venta de Schering".

SOLUCIÓN La línea central de nuestra gráfica de control se localiza en el valor de $\bar{p}$:

$$\bar{p} = \frac{\text{número total de muertes de todas las muestras combinadas}}{\text{número total de sujetos muestrados}}$$

$$= \frac{25 + 24 + 22 + \cdots + 31}{13 \cdot 100{,}000} = \frac{375}{1{,}300{,}000} = 0.000288$$

Puesto que $\bar{p} = 0.000288$, se infiere que $\bar{q} = 1 - = 0.999712$. Al utilizar $\bar{p} = 0.000288$, $\bar{q} = 0.999712$ y $n = 100{,}000$, calculamos los límites de control de la siguiente manera:

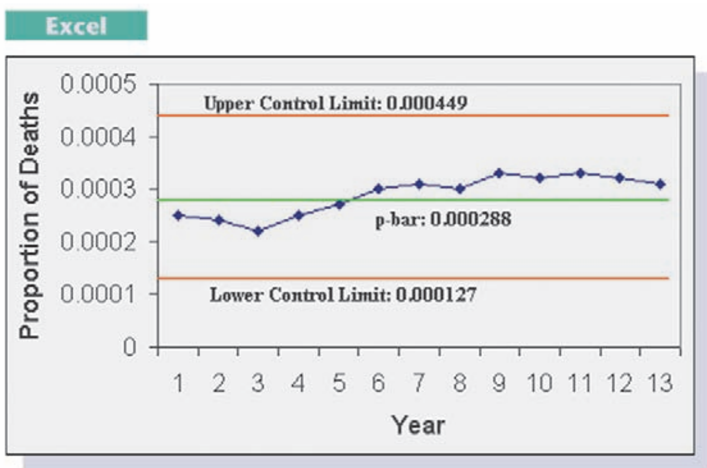
Límite de control superior:

$$\bar{p} + 3\sqrt{\frac{\bar{p}\,\bar{q}}{n}} = 0.000288 + 3\sqrt{\frac{(0.000288)(0.999712)}{100{,}000}} = 0.000449$$

Límite de control inferior:

$$\bar{p} - 3\sqrt{\frac{\bar{p}\,\bar{q}}{n}} = 0.000288 - 3\sqrt{\frac{(0.000288)(0.999712)}{100{,}000}} = 0.000127$$

Una vez que encontramos los valores de la línea central y los límites de control, procedemos a graficar la proporción anual de muertes por infecciones del tracto respiratorio. La gráfica de control de p de Excel se presenta en la pantalla adjunta.



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 13-2. Con esos criterios concluiríamos que dicho proceso está fuera de control estadístico por las siguientes razones: parece haber una tendencia creciente y existen ocho puntos consecutivos que se ubican por arriba de la línea central (regla de la racha de 8). Con base en tales datos, las políticas de salud pública que afectan las infecciones del tracto respiratorio deben modificarse para disminuir la tasa de muertes.

Utilizando 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** y después **P**. Introduzca C1 en el recuadro que se identifica como variable y el tamaño de las muestras en el recuadro que se identifica como tamaño del subgrupo, después haga clic en **OK**.

Excel **Uso de DDXL:** Para utilizar el complemento DDXL, inicie anotando los números de defectos o éxitos en la columna A y registre los tamaños de muestra 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 muestra abajo.

	A	B
1	25	100000
2	24	100000
3	22	100000

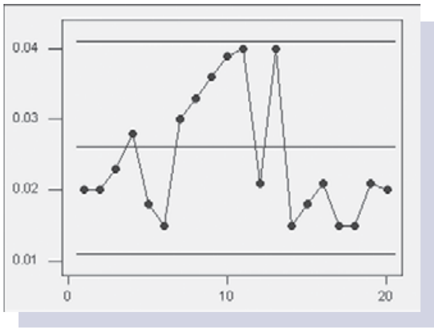
Haga clic en **DDXL**, seleccione **Process Control**, después **Summ Prop Control Chart** (para gráfica de control de resumen de proporciones). Debe aparecer un cuadro de diálogo. Haga clic en el icono del lápiz de "Success Variable" e introduzca el rango de valores para la columna A, tal como A1:A13. De nuevo, haga clic en el icono del lápiz de "Totals Variable" e introduzca el rango de valores para la columna B, tal como B1:B13. Ahora haga clic en **OK** y después en la barra **Open Control Chart**. Aparecerá la gráfica de control.

Uso del Chart Wizard de Excel: Introduzca las proporciones muestrales en la columna A. (Puede ingresar los números reales de defectos en la columna A, después utilizar Excel para crear una columna B consistente en las proporciones. En el recuadro de la fórmula, introduzca $=\text{A}1/n$, donde n se reemplaza por el tamaño de cada muestra. Después de presionar **Enter**, la celda B1 debe contener la primera proporción muestral. Haga clic y deslice la esquina inferior derecha de la celda B1, de manera que toda la columna B contenga las proporciones muestrales correspondientes al número real de defectos en la columna A). Una vez que se ingresaron los datos, proceda a generar la gráfica haciendo clic primero en el icono de **Chart Wizard**, que tiene la apariencia de una gráfica de barras. Para el tipo de gráfica, seleccione **Line**. Para el subtipo de gráfica, la primera gráfica del segundo renglón, y luego haga clic en **Next**. Continúe haciendo clic en **Next** y luego en **Finish**. Es posible editar la gráfica para agregar etiquetas, borrar líneas, etcétera. Inserte la línea central y los límites de control inferior y superior que se requieren, 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.

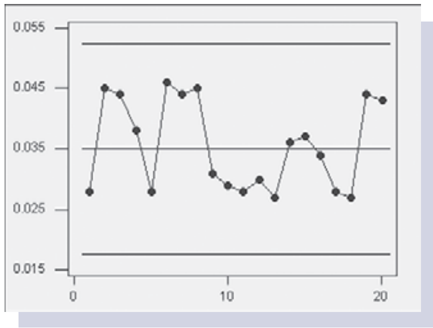
13-3 Destrezas y conceptos básicos

Determine si un proceso está bajo control. *En los ejercicios 1 a 4, examine la gráfica de control de p 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.*

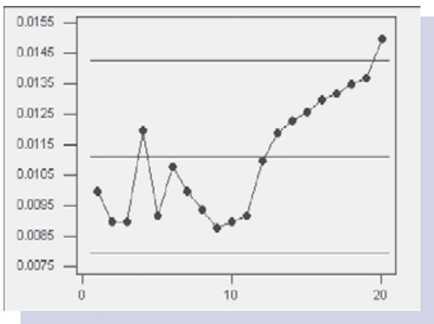
1.



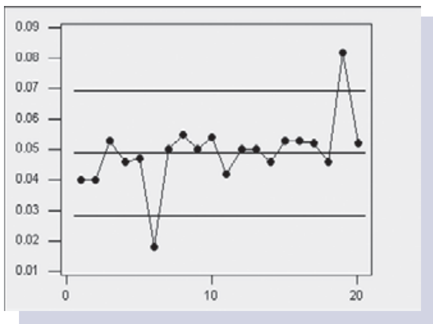
2.



3.



4.



Construcción de gráficas de control de p. *En los ejercicios 5 a 8, utilice los datos de proceso dados para construir una gráfica de control de p. En cada caso, considere los tres criterios para establecer una falta de control que se listan en la sección 13-2, 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.*

5. Gráfica *p* para muertes por enfermedades infecciosas En cada uno de 13 años consecutivos recientes, se seleccionaron 100,000 niños en el rango de 0 a 4 años de edad; en consecuencia, se registró el número de muertes por enfermedades infecciosas; los resultados se presentan abajo (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). ¿Sugieren los resultados un problema que hay que corregir?

Número de muertes: 30 29 29 27 23 25 25 23 24 25 25 24 23

6. Gráfica *p* para víctimas del crimen En cada uno de 20 años consecutivos recientes, se seleccionaron aleatoriamente 1000 adultos para aplicarles una encuesta. Cada uno de los siguientes valores es el número de personas que se encuestaron que fueron víctimas de un crimen violento (datos del Departamento de Justicia, Bureau of Justice Statistics). ¿Sugieren los datos un problema que debe corregirse?

29	33	24	29	27	33	36	22	25	24
31	31	27	23	30	35	26	31	32	24

7. Gráfica *p* para la lluvia en Boston Remítase a las cantidades de lluvia en Boston en el conjunto de datos 11 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 53o. de los miércoles). Durante la primera semana, por ejemplo, la proporción muestral es $3/7 = 0.429$. ¿Representan los datos un proceso estadísticamente estable?

8. Gráfica *p* para las tasas de matrimonios Utilice gráficas *p* para comparar la estabilidad estadística de la tasa de matrimonios de Japón y Estados Unidos. De cada año, se seleccionaron aleatoriamente 10,000 personas de cada país; el número de matrimonios que se obtuvo corresponde a ocho años consecutivos recientes (datos de las Naciones Unidas).

Japón:	58	60	61	64	63	63	64	63
Estados Unidos:	98	94	92	90	91	89	88	87

13-3 Más allá de lo básico

9. 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 lugar de las *proporciones* de defectos. La gráfica $n\bar{p}$ tiene un valor de línea central *n* y los límites de control, 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 que se emplea en el eje vertical. Construya una gráfica *np* para el ejemplo que se presenta en esta sección. Compare el resultado con la gráfica de control de *p* que se obtuvo en la sección.

10. Identificación del efecto del tamaño de muestra en las gráficas *p*

a. Identifique la ubicación de la línea central y de los límites de control de una gráfica *p*, que represente un proceso con una tasa del 5% de artículos que no cumplen la norma, con base en muestras de tamaño 100.

b. Repita el inciso *a* después de cambiar el tamaño de la muestra a 300.

c. Compare los dos conjuntos de resultados. Nombre una ventaja y una desventaja del uso de tamaños de muestra grandes. ¿Qué gráfica sería mejor para detectar un cambio del 5% al 10%?

Repaso

Mientras que los capítulos anteriores de este libro se enfocan en las características importantes de los datos del centro, la variación, la distribución y los datos distantes, este capítulo se centró en un patrón a lo largo del tiempo. Los datos de proceso se definieron como datos que se ordenan de acuerdo con alguna secuencia temporal; datos como los que se mencionan 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 variación natural sin patrones, ciclos o puntos poco comunes. Las decisiones sobre la estabilidad estadística se basan en la forma en que el proceso se comporta en realidad y no en la forma en que nos gustaría que se comportara, por factores tales 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 verificar la *variación* en un proceso

- *Gráfica $\bar{x}$:* gráfica de control que se utiliza para determinar si la media del proceso está bajo control estadístico
- *Gráfica p:* gráfica de control utilizada para verificar la proporción de algún *atributo* del proceso, como por ejemplo si los artículos son defectuosos

Ejercicios de repaso

Construcción de gráficas de control de la lluvia ácida. *En los ejercicios 1 a 3, utilice la siguiente información. Como parte de un estudio para verificar la lluvia ácida, se registran mediciones de depósitos de sulfato (kilogramos/hectárea) en diversos lugares de la Costa Este (según datos del Departamento de Agricultura de Estados Unidos). Los resultados se incluyen en la siguiente tabla de 11 años recientes consecutivos.*

Lluvia ácida: depósitos de sulfato (kilogramos/hectárea)

Año	Lugar 1	Lugar 2	Lugar 3	Lugar 4	Lugar 5
1	11.94	13.09	7.96	17.29	12.12
2	11.28	10.88	12.84	13.87	11.21
3	10.38	12.19	7.38	13.64	9.95
4	8.00	10.75	7.26	12.37	8.77
5	12.12	17.21	10.12	15.73	11.68
6	10.27	10.26	8.89	13.21	9.71
7	14.80	15.49	11.60	17.94	15.59
8	13.52	11.61	9.02	11.22	13.05
9	10.55	10.53	7.78	10.57	11.77
10	9.81	12.50	8.70	13.29	9.37
11	11.27	9.94	10.50	11.28	10.54

1. Depósitos de sulfato: construcción de una gráfica de rachas Construya una gráfica de rachas con los 55 valores. ¿Parece haber un patrón que sugiera que el proceso está bajo control estadístico?

2. Depósitos de sulfato: construcción de una gráfica *R* Construya una gráfica *R* y determine si el proceso de variación está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para establecer una falta de control condujo al rechazo de una variación estadísticamente estable.

3. Depósitos de sulfato: construcción de una gráfica Construya una gráfica $\bar{x}$ y determine si la media del proceso está bajo control estadístico. ¿Parece que el proceso es estadísticamente estable? ¿Cómo debería comportarse este proceso si implementáramos programas eficaces para reducir la cantidad de lluvia ácida?

4. Construcción de una gráfica de control para enfermedades infecciosas En cada uno de 13 años consecutivos recientes se seleccionaron 100,000 adultos de 65 años de edad o mayores, luego se registró el número de muertes por enfermedades infecciosas; los resultados se presentan abajo (datos que se tomaron 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 propia 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

5. Construcción de una gráfica de control de cantidad de votantes En un estudio continuo de cantidad de votantes, cada año en que hubo una elección nacional se seleccionaron aleatoriamente 100 personas en edad de votar; el número de personas que en realidad votaron se presentan abajo (según datos del *Time Almanac*). Construya una gráfica de control adecuada y determine si el proceso está bajo control estadístico. Si no es así, identifique los criterios que condujeron al rechazo de la estabilidad estadística.

Número de votantes:	608	466	552	382	536	372	526	398
	531	364	501	365	551	388	491	

Ejercicios de repaso acumulativo

1. Análisis del proceso de producción de fusibles La Telektronic Company produce fusibles de 20 amperes para proteger los radios de un exceso de energía eléctrica. Diariamente se seleccionan 400 fusibles aleatoriamente, para luego probarse; los resultados (número de defectos por 400 fusibles probados) de 20 días consecutivos son los siguientes:

10 8 7 6 6 9 12 5 4 7 9 6 11 4 6 5 10 5 9 11

a. Utilice una gráfica de control de p para verificar que el proceso está bajo control estadístico, de manera que los datos puedan tratarse como provenientes de una población con variación y media fijas.
b. Utilice todos los datos combinados y construya un intervalo de confianza del 95% para la proporción de defectos.
c. Utilice un nivel de significancia de 0.05 y pruebe la aseveración de que la proporción de defectos es mayor que el 1%.

2. Uso de la probabilidad en las gráficas de control Al interpretar gráficas de control, uno de los tres criterios para determinar que no hay control es que hay ocho puntos consecutivos por encima o por debajo de la línea central. Para un proceso estadísticamente estable existe una probabilidad de 0.5 de que un punto esté por encima de la línea central y una probabilidad de 0.5 de que un punto esté por debajo de la línea central. En los siguientes planteamientos, suponga que los valores muestrales son independientes y que el proceso es estadísticamente estable.
a. Calcule la probabilidad de que, cuando se seleccionan aleatoriamente ocho puntos consecutivos, todos estén por arriba de la línea central.
b. Calcule la probabilidad de que, cuando se seleccionan aleatoriamente ocho puntos consecutivos, todos estén por debajo de la línea central.
c. Calcule la probabilidad de que, cuando se seleccionan aleatoriamente ocho puntos consecutivos, todos estén por encima o por debajo de la línea central.

3. Uso de las gráficas de control para temperaturas En los ejercicios 2 a 4 de la sección 13-2, se listaron las cantidades del consumo de energía eléctrica de la casa del autor, durante un periodo de cuatro años recientes. La tabla adjunta incluye la temperatura promedio (en grados Fahrenheit) del mismo periodo. Utilice gráficas de control o de rachas adecuadas para determinar si los datos parecen formar parte de un proceso estadísticamente estable.

	Ene.-Feb.	Mar.-Abr.	Mayo-Jun.	Jul.-Ago.	Sep.-Oct.	Nov.-Dic.
Año 1	32	35	59	76	66	42
Año 2	22	33	56	70	63	42
Año 3	30	38	55	71	61	38
Año 4	32	40	57	72	65	45

4. Relación entre consumo de energía y temperatura Remítase a los datos del ejercicio 3 y a los datos que se utilizaron para los ejercicios 2 a 4 de la sección 13-2. Realice un apareamiento de los datos de acuerdo con los periodos correspondientes.
a. ¿Hay una correlación lineal significativa entre las cantidades de consumo de energía eléctrica y las temperaturas? Explique.

b. Identifique la ecuación de correlación lineal que relaciona el consumo de energía eléctrica (y) y la temperatura (x).

c. ¿Cuál es la mejor cantidad predicha del consumo de energía eléctrica para un periodo de dos meses con una temperatura promedio de 60°F?

Actividades de cooperación en equipo

1. 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 sea difícil de lograr. Si es así, considere una simulación o remítase a datos que ya se publicaron, tales como los que se encuentran en un almanaque. He aquí algunas sugerencias:

- Lance cinco tiros libres de basquetbol (o lance cinco papeles arrugados en un basurero) y registre el número de canastas que anotó; repita este procedimiento 20 veces y utilice una gráfica *p* para probar la estabilidad estadística de la proporción de tiros realizados.
- Puede medir su pulso contando el número de latidos de su corazón en un minuto. Mida su pulso cuatro veces cada ahora durante varias horas, después construya una gráfica de control adecuada. ¿Qué factores contribuyen a la variación aleatoria? ¿Y a la variación asignable?
- Busque periódicos de las dos últimas semanas y registre el cierre del promedio industrial Dow-Jones. Utilice gráficas de rachas y de control para explorar la estabilidad estadística del promedio industrial Dow-Jones. Identifique al menos una consecuencia práctica de que este proceso sea estadísticamente estable; identifique al menos una consecuencia práctica de que dicho proceso esté fuera de control estadístico.
- Calcule la tasa de divorcios en términos de divorcios por 1000 habitantes durante varios años. (Véase el *Information Please Almanac* o el *Statistical Abstract of the United States*). Suponga que se seleccionaron 1000 personas cada año y que se entrevistaron para determinar si estaban divorciadas. Utilice una gráfica *p* para probar la estabilidad estadística de la tasa de divorcios. (Otras tasas posibles: matrimonio, nacimiento, muerte, muerte en accidentes).

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 puede distribuir el número de ausencias en cada clase, grupos de tres o cuatro estudiantes las analizarán para verificar su estabilidad estadística y hacer recomendaciones con base en las conclusiones.

3. Actividad fuera de clase Realice una investigación para identificar el *Deming's funnel experiment*, después utilice un embudo y canicas para reunir datos respecto de las diferentes reglas, con la finalidad de ajustar la ubicación del embudo. Construya gráficas de control adecuadas 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 marcapasos cardiacos con una tasa del 1% de unidades defectuosas; la proporción de defectos se registra durante cada uno de los 20 días. Los marcapasos de un día se simulan generando aleatoriamente 200 números, donde cada número está entre 1 y 100. Considere que un resultado de 1 es un defecto, mientras que del 2 al 100 son aceptables. Esto corresponde a una tasa del 1% de defectos. (Partes *b*, *c* y *d* siguen las instrucciones de las herramientas tecnológicas).

STATDISK Seleccione **Data** de la barra del menú principal, luego **Uniform Generator**. Ahora proceda a generar 200 valores con un mínimo de 1 y un máximo de 100. Para que aparezcan los datos que se generen en la pantalla, utilice el menú **Format/Sort** en **Sample Editor**. Repita el procedimiento hasta obtener los 20 días simulados.

Minitab En la barra del menú principal, seleccione **Calc**, luego **Random Data** e **Integer**. Introduzca 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 icono ***fx*** en la barra del menú principal, luego seleccione la categoría de la función **Math & Trig**, seguida por **RANDBETWEEN**. En el cuadro de diálogo introduzca 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 mouse 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 mouse a la derecha, de forma que obtenga 20 columnas de 200 números cada una. Las diferentes columnas representan los distintos días de fabricación.

TI-83 Plus Presione la tecla **MATH**, seleccione **PRB**, luego el 5o. elemento del menú, **randInt(,** y proceda a teclear 1, 100, 200; luego oprima la tecla **ENTER**. Luego, haga clic en **STO** y **L1** para almacenar los datos en la lista L1. Después, registre el número de defectos; repita el procedimiento hasta obtener resultados para 20 días simulados.

a. Construya una gráfica *p* para la proporción de marcapasos defectuosos y determine si el proceso está bajo control estadístico. Como sabemos que el proceso en realidad es estable, con $p = 0.01$, 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 tiene que dejar como está.

b. El resultado del inciso *a* es una simulación de 20 días. Ahora simule otros 10 días de fabricación de marcapasos, pero modifique estos últimos días de manera que la tasa de incumplimiento de la norma sea del 3% en lugar del 1%.

c. Combine los datos que se generaron 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 estaba fuera de control, cometeríamos 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 del 3% de incumplimiento de la norma.

de los DATOS *a la* DECISIÓN

Pensamiento crítico: ¿Están las cargas axiales bajo control estadístico?

¿Funciona como debe el proceso de fabricación de latas?

Los ejercicios 5 y 6 de la sección 13-2, 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 20 del Apéndice B y realice un análisis de los datos de proceso para las latas con 0.0111 pulgadas de grosor. Los valores en el conjunto de datos son las cargas axiales medidas de las latas, en tanto que las tapas superiores se colocan en su lugar con presiones que varían entre 158 libras y 165 libras.

Análisis de los resultados

¿Debe tomar acciones correctivas? Escriba un reporte que resuma sus conclusiones. Ponga énfasis no sólo en el tema de la estabilidad estadística, sino también en la capacidad de las latas para soportar la presión que se aplica cuando se colocan las tapas 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 tiene que utilizarse.

PROYECTO DE INTERNET **Gráficas de control**

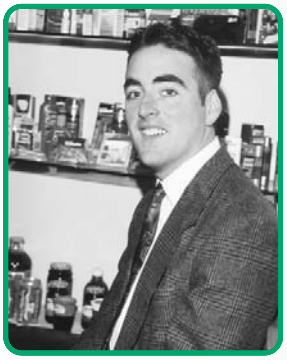
Este capítulo introduce diferentes técnicas de graficación utilizadas para resumir y estudiar datos que se asocian con un proceso, junto con métodos para analizar la estabilidad de ese proceso. Con excepción de la gráfica de rachas, no se requieren datos individuales para construir una gráfica. Por ejemplo, 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. Vaya al sitio Web de *Estadística*:

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 datos que utilizará en la construcción de gráficas de control. De las gráficas resultantes, se le pedirá que interprete y discuta las tendencias en los procesos subyacentes.

La estadística @ en el trabajo

"Se le brinda cierto respeto a quien sabe de estadística y puede explicarla a alguien que no sabe".



Dan O´Toole

Ejecutivo de cuenta: A. C. Nielsen

En su trabajo en el Advanced Analytics Group en A. C. Nielsen, Dan crea soluciones estadísticas para ayudar a que clientes como Polaroid, Ocean Spray y Gillette comprendan cuáles de sus vehículos de venta 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 como la regresión múltiple, análisis factorial, análisis de correspondencia y análisis cluster.

¿Cómo utiliza la estadística en el trabajo?

Mi trabajo consiste en descubrir o cubrir problemas de los clientes, y después encontrar si podemos aplicar una 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: un cliente diría: "yo vendo el producto X", ya sea jugo, pan o una cámara. Ellos tal vez controlen el 20% del mercado. Pueden venir a nosotros para ver si logran incrementar sus ventas en el mercado bajando su precio. Mi trabajo consistiría en diseñar un estudio para analizar dicha cuestión. Para hacerlo, debo diseñar un estudio que tome en cuenta todo lo que afecte las ventas de un producto. Con el uso de 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, debo 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 (para mí lo hace).

¿Podría describir un ejemplo específico que ilustre cómo el uso de la estadística tuvo éxito al mejorar un producto o servicio?

Estamos haciendo un modelo variado para un producto de jugo. El cliente consideró que las marcas propias o las marcas privadas estaban afectando sus ventas, por lo cual necesitaba bajar el precio para mantener su mercado. Cuando terminamos de hacer el modelo, parecía que los dos productos no competían entre sí por el precio. Así, si las marcas propias disminuían sus precios, eso no afectaría sus ventas, lo cual parecía carecer de sentido. ¿Cómo sería verdad esto? Lo que descubrimos fue que, cuando la marca privada entraba a un mercado, robaba a todos los clientes que compran los productos de menor precio a toda costa. Sin embargo, el resto permanecería cautivo. Por lo tanto, aunque el cliente perdió a algunos de sus consumidores, el hecho de bajar sus precios no lograría que ganara más ventas o compradores.

¿Cree que las personas que buscan un empleo se consideran mejores si tienen estudios de estadística?

Claro que sí. Se le brinda cierto respeto a quien sabe de estadística y puede explicarla a alguien que no sabe, porque esto significa que uno realmente sabe y no sólo recita las páginas de un libro de texto. Casi cualquier empleo utiliza estadística, particularmente correlaciones y regresiones. La gente dice cosas 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 (consultoría) constantemente uno se enfrenta al reto de aprender una nueva técnica o de recurrir a una vieja técnica para mejorarla. Además, porque constantemente sacamos 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 podamos explicarla a nuestros clientes en términos “populares”.

En términos de la estadística, ¿qué recomendaría a futuros empleados?

Mucha gente toma cursos de estadística, pero el nivel de retención de la mayoría de los conceptos es muy bajo. Si usted se tiene que enfocar en ciertos conceptos centrales, yo diría que necesita comprender la correlación y la regresión; la comprensión de tales conceptos le ayudará a interpretar otros conceptos con los que se tope, como la regresión múltiple. También aconsejaría el análisis factorial, tan sólo una comprensión general lo pondría por arriba de la curva en muchos aspectos.

14 Proyectos, procedimientos y perspectivas

14-1 Proyectos

El objetivo principal de esta sección es proporcionar algunas sugerencias para un estudio que sirva como proyecto final en el curso de introducción a la estadística. Una ventaja fantástica de este curso es que trata con destrezas y conceptos aplicables al mundo real de forma inmediata. Después de cursar sólo un divertido semestre, los estudiantes son capaces de realizar sus propios estudios. Algunos de los temas que se sugieren implica realizar experimentos, mientras que otros son estudios observacionales que requieren investigar resultados ya disponibles. Por ejemplo, no se recomienda en absoluto probar la eficacia de las bolsas de aire chocando automóviles en la realidad, pero las pruebas del sabor destructivo de galletas con chispas de chocolate son un experimento sencillo y hasta agradable. A continuación se presenta una sugerencia de formato, a la que le sigue una lista de temas que se sugieren.

Trabajo en equipo o individual Los temas se pueden asignar en forma individual, pero los proyectos en equipo son particularmente efectivos, puesto que ayudan a desarrollar las destrezas interpersonales que son tan necesarias en el ambiente de trabajo real. Un estudio mostró que la "incompetencia para realizar tareas junto con otros" es el motivo principal de que se despidan empleados, por lo que un proyecto grupal resultará muy útil para preparar a los estudiantes para sus ambientes de trabajo futuros.

Reporte oral Una presentación en clase de 10 a 15 minutos de duración debe incluir a todos los miembros del grupo en un esfuerzo combinado para describir claramente los componentes importantes del estudio. Los estudiantes, por lo regular, tienen cierta reticencia para hablar en público, así que un breve reporte oral será muy útil para desarrollar la confianza que ellos bien se merecen. De nuevo, el reporte oral es una actividad que ayuda a los estudiantes a prepararse mejor para actividades profesionales futuras.

Reporte escrito El objetivo principal del proyecto no es producir un trabajo escrito, equivalente a un trabajo final, pero es necesario presentar un reporte escrito que incluya los siguientes componentes:

1. Lista de los datos reunidos
2. Descripción del método de análisis
3. Gráficas y estadísticos relevantes, incluyendo pantallas de resultados de STATDISK, Minitab, Excel o la calculadora TI-83 Plus
4. Conclusiones
5. Las razones por las que los resultados no serían correctos, junto con una descripción de las formas en las cuales el estudio se mejoraría, con el tiempo y el dinero suficientes.

Temas sugeridos Además de los temas que se sugieren en la lista siguiente, véanse también las *Actividades de cooperación en equipo* que se encuentran cerca del final de cada capítulo.

1. Rehacer una gráfica de un periódico o una revista para una mejor descripción de los datos.
2. Reescribir un artículo de periódico acerca de una encuesta para informar mejor al lector.
3. Utilizar lanzamientos de monedas para obtener mejores resultados de una encuesta con preguntas sensibles.
4. Antigüedad de los automóviles de los estudiantes en comparación con la de los automóviles de los profesores y el personal administrativo.
5. Proporción de automóviles extranjeros conducidos por estudiantes en comparación con la proporción de automóviles extranjeros conducidos por profesores.
6. Antigüedad de los automóviles en el estacionamiento de una tienda de descuento en comparación con la antigüedad de los automóviles en el estacionamiento de un almacén departamental de gran escala.
7. ¿Los maridos son mayores en 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 la universidad con la de los libros en 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. ¿Hay una relación entre las horas de estudio y las calificaciones obtenidas?
14. ¿Existe una relación entre las horas de trabajo escolar y las calificaciones obtenidas?
15. Un estudio de estaturas *reportadas* comparadas con estaturas *medidas*.
16. Un estudio de la precisión de los relojes.
17. ¿Hay una relación entre el sabor y el costo de marcas diferentes de galletas con chispas de chocolate?
18. ¿Existe una relación entre el sabor y el costo de marcas diferentes de mantequilla de cacahuate?

19. ¿Hay una relación entre el sabor y el costo de marcas diferentes de bebidas de cola?

20. ¿Existe una relación entre los salarios de los jugadores profesionales de beisbol (o basquetbol o futbol) 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. ¿Hay una relación entre la longitud de los pies de los hombres (o de las mujeres) y sus estaturas?

23. ¿Existen diferencias entre el sabor del agua común de la llave y el de las diferentes marcas de agua embotellada?

24. ¿Se afectaron las tasas de fatalidad por las leyes que requieren el uso de cinturones de seguridad?

25. ¿Se afectaron las tasas de fatalidad cuando se eliminó el límite de velocidad nacional de 55 millas/hora?

26. ¿Se afectaron las tasas de fatalidad con la presencia de las bolsas de aire?

27. ¿Existe una diferencia de sabor entre la Coca Cola y la Pepsi Cola?

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 pies cuadrados), el tamaño del terreno (en acres), 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. ¿Existe sustento para la teoría de que la proporción que guarda la estatura de una persona con la altura de su ombligo es la razón áurea de alrededor de 1.6:1?

32. Una comparación del número de llaves que llevan los hombres y las mujeres.

33. Una comparación del número de tarjetas de crédito que llevan los hombres y las mujeres.

34. ¿Son ahora los homicidas más jóvenes que antes?

35. ¿Tienden las personas que practican ejercicio vigoroso a tener el pulso más bajo que las que no lo hacen?

36. ¿Tienden las personas que practican ejercicio vigoroso a tener tiempos de reacción diferentes a las personas que no lo hacen?

37. ¿Tienden las personas que fuman a tener tasas más altas de pulso que aquellas que no lo hacen?

38. En las personas que no practican ejercicio, ¿cómo es afectado su pulso al subir un tramo de escaleras?

39. ¿Tienden los estudiantes de estadística a tener pulsos diferentes de los de 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. ¿Tienden las personas zurdas a ser protagonistas de más choques de automóviles?

42. ¿Se ven más implicados los hombres en choques de automóvil que las mujeres?

43. ¿Se relacionan más los conductores jóvenes en choques de automóvil que los conductores de más edad?

44. ¿Tienen los conductores que reciben multas mayores posibilidades de verse implicados en choques?

45. ¿Tienden los fumadores a relacionarse más en choques automovilísticos?

46. ¿Tienden las personas con un pulso más alto a relacionarse en más o en menos choques de automóviles?

47. Una comparación de tiempos de reacción medidos de las manos derecha e izquierda.

48. ¿Son iguales las proporciones de fumadores hombres y mujeres?

49. ¿Tienden los estudiantes de estadística a fumar más (o menos) que la población general?

50. ¿Es más probable que las personas fumen si sus padres fumaron?

51. Evidencia para sustentar o refutar la creencia de que fumar tiende a frenar el crecimiento.

52. ¿Tiene ventaja un equipo deportivo por jugar en casa en lugar de ser visitante?

53. Análisis de los tiempos de servicio (en segundos) de una ventanilla bancaria con servicio al automóvil.

54. Una comparación de los tiempos de servicio de las ventanillas de servicio al automóvil de dos bancos diferentes.

55. Análisis de los tiempos en que los clientes de McDonald's se sientan a una mesa.

56. Análisis de los tiempos en que los clientes de McDonald's esperan en la fila.

57. Análisis de los tiempos que los automóviles requieren para relleno de combustible.

58. ¿Es una inversión sensata la lotería estatal?

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 si usted apuesta \$1 por juego que si apuesta 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 las 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 Díaz, 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 o tenis, nadar, ver un partido de beisbol y un partido de futbol.

66. ¿Hay sustento para la teoría de que los cereales con alto contenido de azúcar se ponen en estantes al nivel visual de los niños?

67. ¿Existe sustento para la aseveración de que la temperatura corporal media es menor que 98.6°F?

68. ¿Hay una relación entre fumar y beber café?

69. ¿Existe una relación entre las calificaciones del curso y el tiempo que se emplea en entretenerse con juegos de vídeo?

70. ¿Hay sustento para la teoría de que el viernes es de mala suerte si es el día 13 del mes?

14-2 Procedimiento

Recolección de datos Usted puede reunir sus propios datos a través de experimentos o estudios observacionales. Es absolutamente esencial criticar el método que se utiliza para reunir los datos, puesto que los datos reunidos con descuido son tan inútiles que ninguna cantidad de tortura estadística es capaz de salvarlos. Revise con cuidado, tanto para identificar prejuicios, en la forma en que se reúnen los datos, como para identificar favoritismos en la persona o el grupo que reúne los datos. Muchos de los procedimientos de este libro se basan en la suposición de que estamos trabajando con una muestra aleatoria simple, lo que significa que cada muestra posible del mismo tamaño tiene la misma oportunidad de seleccionarse. Una muestra autoseleccionada (de respuesta voluntaria) es inútil para hacer inferencias acerca de una población.

Exploración, comparación y descripción Después de reunir los datos, primero considere la exploración, la descripción y la comparación de los conjuntos de datos utilizando las herramientas básicas que se incluyen en el capítulo 2. Asegúrese de aplicar 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 dan una indicación de 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, y determine si la distribución tiene forma de campana, es uniforme o sesgada.
4. *Datos distantes:* Identifique cualquier valor muestral que se encuentre muy lejano a la gran mayoría de los otros valores muestrales.
5. *Tiempo:* Determine si la proporción es estable o si sus características cambian con el tiempo.

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 escoger el procedimiento en particular que hay que aplicar. Este texto incluye una amplia variedad de procedimientos aplicables a muchas circunstancias diferentes. A continuación se plantean algunas preguntas clave que deben responderse:

- ¿Cuál es el nivel de medición (nominal, ordinal, de intervalo o de razón) de los datos?
- ¿El estudio considera 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”).
- ¿Hay una razón para creer que la población se distribuye normalmente?
- ¿Cuál es la pregunta básica o el tema al que usted se quiere dirigir?

En la figura 14-1 listamos los principales métodos que se incluyen en este libro, junto con un esquema para determinar cuáles de estos métodos conviene utilizar según el caso. Para hacer uso de la figura 14-1, inicie en el extremo izquierdo de la figura e identifique el nivel de medición de los datos. Proceda siguiendo el flujo que se sugiere 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 cambiaría con el tiempo, construya una gráfica de control (véase el capítulo 13) para determinar si el proceso es estadísticamente estable. Dicha figura se aplica a datos de proceso sólo si el proceso es estadísticamente estable.

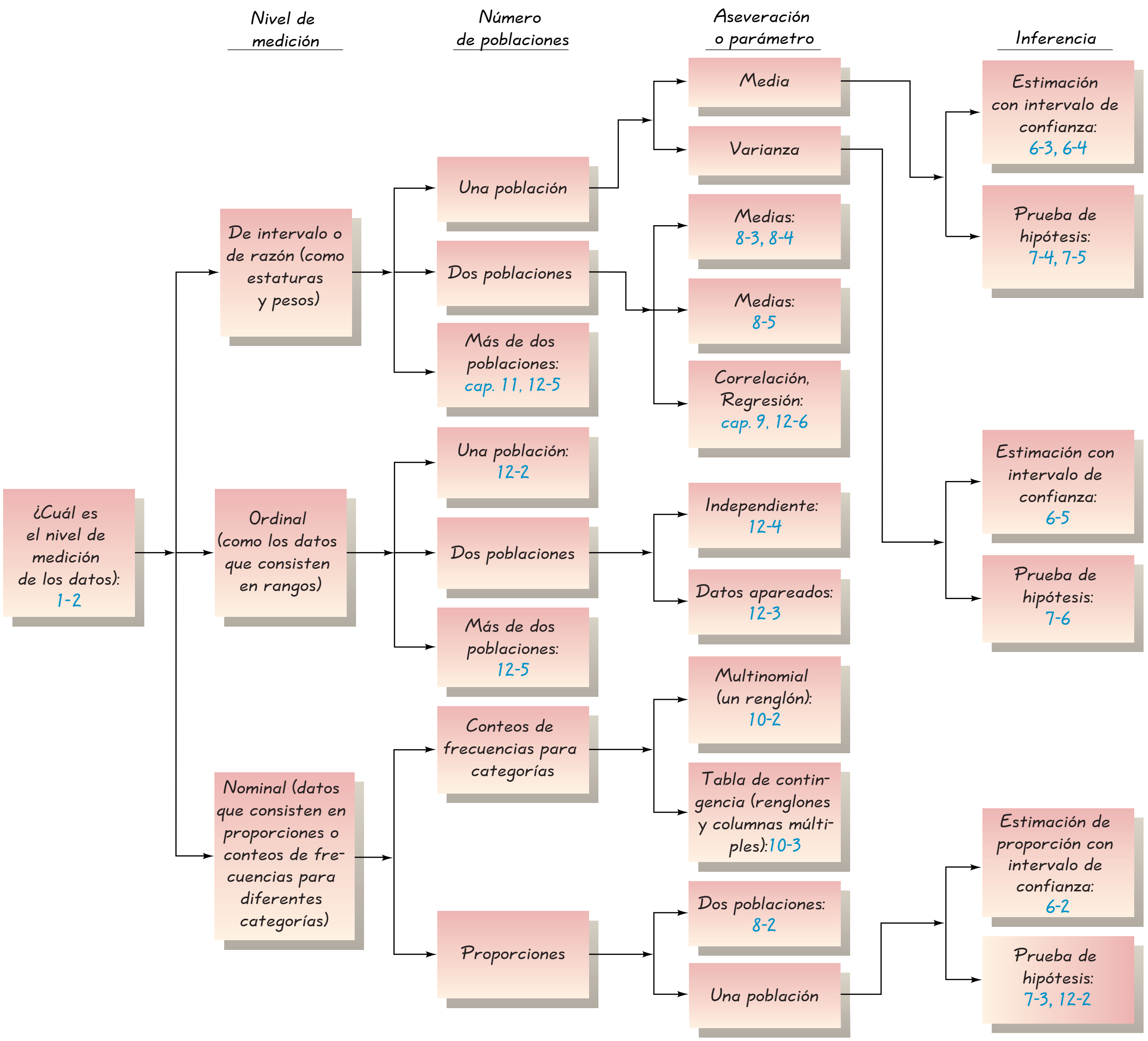


FIGURA 14-1 **Selección del procedimiento adecuado**

La figura 14-1 puede utilizarse con los métodos estadísticos que se presentan en este libro, pero quizás existan otros métodos que resulten más adecuados para un análisis estadístico en particular. Para recibir ayuda con otros métodos, consulte al estadístico profesional de su preferencia.

14-3 Perspectiva

Nadie espera que un simple curso de introducción a la estadística transforme a alguien en un estadístico experto. Después de estudiar varios capítulos de este libro, es natural que los estudiantes sientan que no dominan el material al nivel necesario para utilizar la estadística con confianza en aplicaciones reales. Muchos temas importantes (como el análisis factorial y el análisis discriminante) no se incluyen en este texto, ya que son muy avanzados para este nivel introductorio. Algunos temas más sencillos (como las series de tiempo) se excluyeron por otras razones. Es importante saber que la ayuda profesional de estadísticos expertos está disponible y que este curso de introducción a la estadística les ayudará en los análisis con alguno de esos expertos.

Aunque este curso no se diseñó para convertirle en un estadístico experto, sí se pensó para convertirlo en una persona con mejor educación, con una mayor posibilidad de empleo. Debe conocer y entender los conceptos básicos de probabilidad y posibilidad. También, saber que al intentar comprender con mayor profundidad 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 una gráfica), la presencia de datos distantes, y si la población es estable o si cambia con el tiempo. Usted deberá conocer y entender la importancia de la estimación de parámetros poblacionales (como una media, una desviación estándar y una proporción), así como probar aseveraciones hechas acerca de parámetros poblacionales. Comprenderá que la naturaleza y configuración de los datos produce un efecto importante en los procedimientos estadísticos particulares que se utilicen.

A lo largo de este texto hemos puesto énfasis en la importancia de un buen muestreo. Hay que reconocer que una muestra mala puede ser muy difícil de reparar, aun por los estadísticos más expertos y utilizando las técnicas más complejas. Existen muchas encuestas por correo, de revistas y de respuesta telefónica, que permiten que los que responden sean “autoseleccionados”. Los resultados de encuestas de este tipo generalmente son inútiles cuando se juzgan de acuerdo con los criterios de la metodología estadística sana. Tenga esto en mente cuando encuentre encuestas de respuesta voluntaria (autoseleccionadas); no permita que afecten sus creencias y decisiones. Sin embargo, también tiene que reconocer que muchas encuestas y entrevistas arrojan muy buenos resultados, aun cuando los tamaños de las muestras parezcan relativamente pequeños. Aunque muchas personas se rehúsan a creerlo, una encuesta a nivel nacional de sólo 1700 votantes llega a proporcionar buenos resultados si el muestreo se planea y ejecuta cuidadosamente.

A lo largo de este texto pusimos énfasis en 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 proporcionar resultados, pero dichos resultados por lo regular requieren de la interpretación cuidadosa que les da vida, de otra forma, carecerían de significado.

Hay que reconocer que un resultado no es automáticamente válido y bueno simplemente porque fue generado por computadora. Las computadoras no piensan, aunque son capaces de proporcionar resultados que son 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!*

En otros tiempos, se consideraba que una persona era educada por el solo hecho de saber leer. Pero ahora estamos en una era que demanda mucho más. Hoy, una persona educada debe ser capaz de leer, escribir, utilizar programas de cómputo, hablar una lengua extranjera y conocer álgebra básica. Una persona verdaderamente educada es capaz de combinar las disciplinas con metas comunes, incluyendo la búsqueda de la verdad. El estudio de la estadística nos ayuda a ver la verdad que en ocasiones otros distorsionan o es encubierta por datos en desorden o que tal vez ni siquiera se reunieron. Ahora el entendimiento de los principios de la estadística es esencial para cada persona educada. H. G. Wells dijo que “el pensamiento estadístico algún día será tan necesario para una ciudadanía eficiente como la habilidad de leer y escribir”. Ese día ha llegado.

Apéndice A: Tablas

TABLA A-1 Probabilidades binomiales

n	x	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	x
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+	.001	.015	.060	.138	.234	.311	.324	.246	.098	.031	.001	4
	5	0+	0+	0+	.002	.010	.037	.094	.187	.303	.393	.354	.232	.057	5
	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+	.003	.029	.097	.194	.273	.290	.227	.115	.023	.004	0+	4
	5	0+	0+	0+	.004	.025	.077	.164	.261	.318	.275	.124	.041	.002	5
	6	0+	0+	0+	0+	.004	.017	.055	.131	.247	.367	.372	.257	.066	6
	7	0+	0+	0+	0+	0+	.002	.008	.028	.082	.210	.478	.698	.932	7
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+	.005	.046	.136	.232	.273	.232	.136	.046	.005	0+	0+	4
	5	0+	0+	0+	.009	.047	.124	.219	.279	.254	.147	.033	.005	0+	5
	6	0+	0+	0+	.001	.010	.041	.109	.209	.296	.294	.149	.051	.003	6
	7	0+	0+	0+	0+	.001	.008	.031	.090	.198	.336	.383	.279	.075	7
	8	0+	0+	0+	0+	0+	.001	.004	.017	.058	.168	.430	.663	.923	8

NOTA: 0+ representa una probabilidad positiva menor que 0.0005.

(continúa)

TABLA A-1 Probabilidades binomiales (*continuación*)

n	*x*	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	*x*
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+	.001	.009	.042	.117	.215	.267	.201	.057	.010	0+	7
	8	0+	0+	0+	0+	.001	.011	.044	.121	.233	.302	.194	.075	.004	8
	9	0+	0+	0+	0+	0+	.002	.010	.040	.121	.268	.387	.315	.091	9
	10	0+	0+	0+	0+	0+	0+	.001	.006	.028	.107	.349	.599	.904	10
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+	.002	.017	.070	.161	.236	.220	.111	.016	.001	0+	7
	8	0+	0+	0+	0+	.004	.023	.081	.177	.257	.221	.071	.014	0+	8
	9	0+	0+	0+	0+	.001	.005	.027	.089	.200	.295	.213	.087	.005	9
	10	0+	0+	0+	0+	0+	.001	.005	.027	.093	.236	.384	.329	.099	10
	11	0+	0+	0+	0+	0+	0+	0+	.004	.020	.086	.314	.569	.895	11
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+	.003	.029	.101	.193	.227	.158	.053	.004	0+	0+	7
	8	0+	0+	0+	.001	.008	.042	.121	.213	.231	.133	.021	.002	0+	8
	9	0+	0+	0+	0+	.001	.012	.054	.142	.240	.236	.085	.017	0+	9
	10	0+	0+	0+	0+	0+	.002	.016	.064	.168	.283	.230	.099	.006	10
	11	0+	0+	0+	0+	0+	0+	.003	.017	.071	.206	.377	.341	.107	11
	12	0+	0+	0+	0+	0+	0+	0+	.002	.014	.069	.282	.540	.886	12

NOTA: 0+ representa una probabilidad positiva menor que 0.0005.

(*continúa*)

TABLA A-1 Probabilidades binomiales (*continuación*)

		p													
n	*x*	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	*x*
13	0	.878	.513	.254	.055	.010	.001	0+	0+	0+	0+	0+	0+	0+	0
	1	.115	.351	.367	.179	.054	.011	.002	0+	0+	0+	0+	0+	0+	1
	2	.007	.111	.245	.268	.139	.045	.010	.001	0+	0+	0+	0+	0+	2
	3	0+	.021	.100	.246	.218	.111	.035	.006	.001	0+	0+	0+	0+	3
	4	0+	.003	.028	.154	.234	.184	.087	.024	.003	0+	0+	0+	0+	4
	5	0+	0+	.006	.069	.180	.221	.157	.066	.014	.001	0+	0+	0+	5
	6	0+	0+	.001	.023	.103	.197	.209	.131	.044	.006	0+	0+	0+	6
	7	0+	0+	0+	.006	.044	.131	.209	.197	.103	.023	.001	0+	0+	7
	8	0+	0+	0+	.001	.014	.066	.157	.221	.180	.069	.006	0+	0+	8
	9	0+	0+	0+	0+	.003	.024	.087	.184	.234	.154	.028	.003	0+	9
	10	0+	0+	0+	0+	.001	.006	.035	.111	.218	.246	.100	.021	0+	10
	11	0+	0+	0+	0+	0+	.001	.010	.045	.139	.268	.245	.111	.007	11
	12	0+	0+	0+	0+	0+	0+	.002	.011	.054	.179	.367	.351	.115	12
	13	0+	0+	0+	0+	0+	0+	0+	.001	.010	.055	.254	.513	.878	13
14	0	.869	.488	.229	.044	.007	.001	0+	0+	0+	0+	0+	0+	0+	0
	1	.123	.359	.356	.154	.041	.007	.001	0+	0+	0+	0+	0+	0+	1
	2	.008	.123	.257	.250	.113	.032	.006	.001	0+	0+	0+	0+	0+	2
	3	0+	.026	.114	.250	.194	.085	.022	.003	0+	0+	0+	0+	0+	3
	4	0+	.004	.035	.172	.229	.155	.061	.014	.001	0+	0+	0+	0+	4
	5	0+	0+	.008	.086	.196	.207	.122	.041	.007	0+	0+	0+	0+	5
	6	0+	0+	.001	.032	.126	.207	.183	.092	.023	.002	0+	0+	0+	6
	7	0+	0+	0+	.009	.062	.157	.209	.157	.062	.009	0+	0+	0+	7
	8	0+	0+	0+	.002	.023	.092	.183	.207	.126	.032	.001	0+	0+	8
	9	0+	0+	0+	0+	.007	.041	.122	.207	.196	.086	.008	0+	0+	9
	10	0+	0+	0+	0+	.001	.014	.061	.155	.229	.172	.035	.004	0+	10
	11	0+	0+	0+	0+	0+	.003	.022	.085	.194	.250	.114	.026	0+	11
	12	0+	0+	0+	0+	0+	.001	.006	.032	.113	.250	.257	.123	.008	12
	13	0+	0+	0+	0+	0+	0+	.001	.007	.041	.154	.356	.359	.123	13
	14	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
	1	.130	.366	.343	.132	.031	.005	0+	0+	0+	0+	0+	0+	0+	1
	2	.009	.135	.267	.231	.092	.022	.003	0+	0+	0+	0+	0+	0+	2
	3	0+	.031	.129	.250	.170	.063	.014	.002	0+	0+	0+	0+	0+	3
	4	0+	.005	.043	.188	.219	.127	.042	.007	.001	0+	0+	0+	0+	4
	5	0+	.001	.010	.103	.206	.186	.092	.024	.003	0+	0+	0+	0+	5
	6	0+	0+	.002	.043	.147	.207	.153	.061	.012	.001	0+	0+	0+	6
	7	0+	0+	0+	.014	.081	.177	.196	.118	.035	.003	0+	0+	0+	7
	8	0+	0+	0+	.003	.035	.118	.196	.177	.081	.014	0+	0+	0+	8
	9	0+	0+	0+	.001	.012	.061	.153	.207	.147	.043	.002	0+	0+	9
	10	0+	0+	0+	0+	.003	.024	.092	.186	.206	.103	.010	.001	0+	10
	11	0+	0+	0+	0+	.001	.007	.042	.127	.219	.188	.043	.005	0+	11
	12	0+	0+	0+	0+	0+	.002	.014	.063	.170	.250	.129	.031	0+	12
	13	0+	0+	0+	0+	0+	0+	.003	.022	.092	.231	.267	.135	.009	13
	14	0+	0+	0+	0+	0+	0+	0+	.005	.031	.132	.343	.366	.130	14
	15	0+	0+	0+	0+	0+	0+	0+	0+	.005	.035	.206	.463	.860	15

NOTA: 0+ representa una probabilidad positiva menor que 0.0005.

Puntuaciones *z* NEGATIVAS

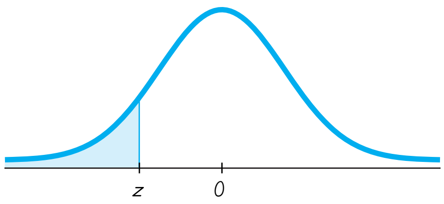
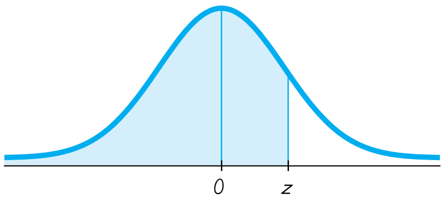


TABLA A-2 Distribución normal estándar (*z*): Área acumulativa de la IZQUIERDA

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 de 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									

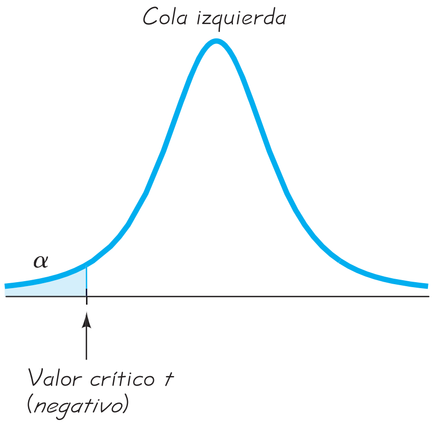
NOTA: Para valores de *z* por encima de 3.49, utilice 0.9999 para el área.
*Utilice estos valores comunes, que resultan por interpolación:

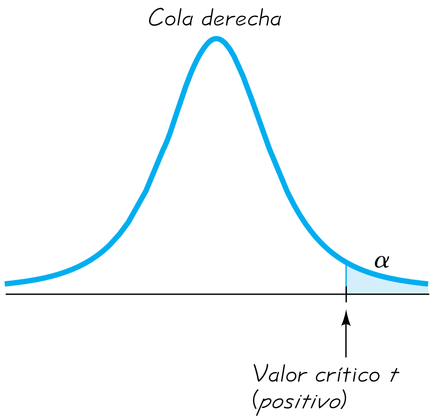
Puntuación *z*	Área
1.645	0.9500
2.575	0.9950

Valores comunes críticos

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
			Área en dos colas		
Grados de libertad	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





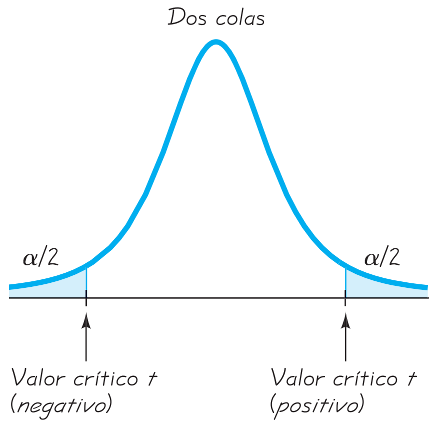


TABLA A-4 Distribución chi cuadrada (χ^2)

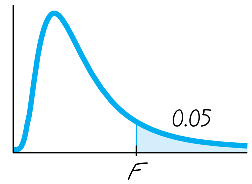
Grados de libertad	Área a la derecha del valor crítico									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	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. Reimpreso bajo 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 (gl_1)

Grados de libertad del denominador (gl_2)	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

TABLA A-5 Distribución F ($\alpha = 0.025$ en la cola derecha) (*continuación*)

Grados de libertad del numerador (gl_1)

Grados de libertad del denominador (gl_2)	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

De Maxine Merrington y Catherine M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution", *Biometrika* 33 (1943): 80-84.

(*continúa*)

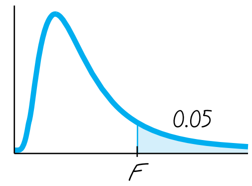


Tabla A-5 Distribución F ($\alpha = 0.025$ en la cola derecha)

Grados de libertad del numerador (gl_1)

Grados de libertad del denominador (gl_2)	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

(continúa)

TABLA A-5 Distribución F ($\alpha = 0.025$ en la cola derecha) *(continuación)*

Grados de libertad del numerador (gl_1)

Grados de libertad del denominador (gl_2)	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

De Maxine Merrington y Catherine M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution", *Biometrika* 33 (1943): 80-84. Reimpreso bajo 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

	α			
n	.005 (una cola) .01 (dos colas)	.01 (una cola) .02 (dos colas)	.025 (una cola) .05 (dos colas)	.05 (una cola) .10 (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

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

TABLA A-8 Valores críticos de T para la prueba de rangos con signo de Wilcoxon

	α			
n	.005 (una cola) .01 (dos colas)	.01 (una cola) .02 (dos colas)	.025 (una cola) .05 (dos colas)	.05 (una cola) .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 que se encontró en esta tabla. No rechace la hipótesis nula si el estadístico de prueba T es mayor que el valor crítico que se encontró en la tabla.

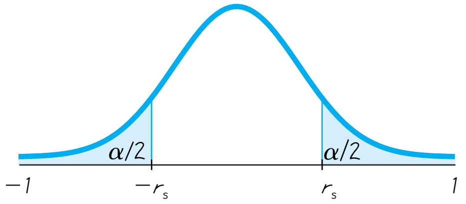


TABLA A-9 Valores críticos del coeficiente de correlación de rangos de Spearman r_s

n	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	—
8	.643	.738	.833	.881
9	.600	.683	.783	.833
10	.564	.648	.745	.794
11	.523	.623	.736	.818
12	.497	.591	.703	.780
13	.475	.566	.673	.745
14	.457	.545	.646	.716
15	.441	.525	.623	.689
16	.425	.507	.601	.666
17	.412	.490	.582	.645
18	.399	.476	.564	.625
19	.388	.462	.549	.608
20	.377	.450	.534	.591
21	.368	.438	.521	.576
22	.359	.428	.508	.562
23	.351	.418	.496	.549
24	.343	.409	.485	.537
25	.336	.400	.475	.526
26	.329	.392	.465	.515
27	.323	.385	.456	.505
28	.317	.377	.448	.496
29	.311	.370	.440	.487
30	.305	.364	.432	.478

NOTA: Para $n > 30$, utilice $r_s = \pm z/\sqrt{n-1}$, donde z corresponde al nivel de significancia. Por ejemplo, si $\alpha = 0.05$, entonces $z = 1.96$.

Para probar H_0: $\rho_s = 0$

contra H_1: $\rho_s \neq 0$

De "Distribution of sums of squares of rank differences to small numbers of individuals", *The Annals of Mathematical Statistics*, vol. 9, núm. 2. Reimpreso bajo permiso del Institute of Mathematical Statistics.

TABLA A-10 Valores críticos para el número de rachas G

Valor de n_1 \ Valor de n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	1	1	1	1	1	1	1	1	1	1	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
3	1	1	1	1	2	2	2	2	2	2	2	2	2	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
4	1	1	1	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
	6	8	9	9	9	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	5	5	5
	6	8	9	10	10	11	11	12	12	12	12	12	12	12	12	12	12	12	12
6	1	2	2	3	3	3	3	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
7	1	2	2	3	3	3	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	16	16	16	16	16
8	1	2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
	6	8	10	11	12	13	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
	6	8	10	12	13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10	1	2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
	6	8	10	12	13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11	1	2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
	6	8	10	12	13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
	6	8	10	12	13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13	2	2	3	4	5	5	6	6	7	7	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
14	2	2	3	4	5	5	6	7	7	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
15	2	3	3	4	5	6	6	7	7	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
16	2	3	4	4	5	6	6	7	8	8	9	9	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
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
	6	8	10	12	14	16	17	18	19	20	21	22	23	23	24	25	25	26	26
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
	6	8	10	12	14	16	17	18	19	20	21	22	23	24	25	25	26	26	27
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
	6	8	10	12	14	16	17	18	20	21	22	23	23	24	25	26	26	27	27
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14
	6	8	10	12	14	16	17	18	20	21	22	23	24	25	25	26	27	27	28

NOTA:

1. Los valores en esta tabla son los valores críticos G, suponiendo una prueba de dos colas con un nivel de significancia de $\alpha = 0.05$.
2. 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. Reimpreso bajo permiso del Institute of Mathematical Statistics.

Apéndice B: Conjuntos de datos

Conjunto de datos 1:	Resultados de examen de salud
Conjunto de datos 2:	Estaturas de padres e hijos
Conjunto de datos 3:	Circunferencias de la cabeza
Conjunto de datos 4:	Temperaturas corporales de adultos saludables
Conjunto de datos 5:	Alquitrán, nicotina y monóxido de carbono de cigarros
Conjunto de datos 6:	Fumadores activos y pasivos
Conjunto de datos 7:	Consumo de alcohol y tabaco en películas de dibujos animados para niños
Conjunto de datos 8:	Finalistas de la carrera de maratón de la ciudad de Nueva York
Conjunto de datos 9:	Osos (osos salvajes anestesiados)
Conjunto de datos 10:	Temperaturas reales y pronosticadas
Conjunto de datos 11:	Precipitación pluvial en Boston durante un año
Conjunto de datos 12:	Temperatura, lluvia y conductividad en los Everglades
Conjunto de datos 13:	Géiser Old Faithful
Conjunto de datos 14:	Libros de Clancy, Rowling y Tolstoi
Conjunto de datos 15:	Edades de los polizones del *Queen Mary*
Conjunto de datos 16:	Cereal
Conjunto de datos 17:	Pesos y volúmenes de bebidas de cola
Conjunto de datos 18:	Diamantes
Conjunto de datos 19:	Pesos de una muestra de dulces M&M clásicos
Conjunto de datos 20:	Cargas axiales de latas de aluminio
Conjunto de datos 21:	Películas
Conjunto de datos 22:	Automóviles
Conjunto de datos 23:	Pesos de basura desechada en una semana
Conjunto de datos 24:	Casas vendidas en el condado Dutchess
Conjunto de datos 25:	Misceláneos: DJ, ventas de automóviles, muertes en vehículos motorizados, homicidios, manchas de sol y Súper Bowl
Conjunto de datos 26:	Lotería estatal de Nueva York
Conjunto de datos 27:	Resultados del solitario
Conjunto de datos 28:	Pesos de sobres de azúcar Dominó
Conjunto de datos 29:	Pesos de monedas de 25 centavos de dólar
Conjunto de datos 30:	Distancias de jonrones

Conjunto de datos 1: Resultados de examen de salud

EDAD en años, EST es estatura (pulgadas), PE es peso (libras), CINT es circunferencia de la cintura (cm), PULSO es frecuencia del pulso (latidos por minuto), SIS es presión sanguínea sistólica (mmHg), DIA 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 Department of Health and Human Services de EUA, National Center for Health Statistics, Third National Health and Nutrition Examination Survey.

STATDISK Los nombres de los archivos de STATDISK y de los archivos de texto para hombres son MAGE, MHT, MWT, MWAST, MPULS, MSYS, MDIAS, MCHOL, MBMI, MLEG, MELBW, MWRST, MARM.
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 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 y de STATDISK.

Hombre	Edad	Est	Pe	Cint	Pulso	SIS	DIA	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 Los nombres de los archivos de STATDISK y de los archivos de texto para mujeres son FAGE, FHT, FWT, FWAST, FPULS, FSYS, FDIAS, FCHOL, FBMI, FLEG, FELBW, FWRST, FARM.
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 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 y de STATDISK.

Mujer	Edad	Est	Pe	Cint	Pulso	SIS	DIA	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: Estaturas de padres e hijos (pulgadas)

Los datos son del Department of Health and Human Services de EUA, National Center for Health Statistics, Third National Health and Nutrition Examination Survey.

STATDISK Nombres de los archivos de STATDISK y de los archivos de texto: CHDHT, MOMHT, DADHT.
Minitab: El nombre de la hoja de cálculo es PARENTHT.MTW.
Excel: El nombre del libro de trabajo es PARENTHT.XLS.
TI-83 Plus: El nombre de la App es PARENTHT y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Género	Estatura	Estatura de la madre	Estatura del padre
H	62.5	66	70
H	64.6	58	69
H	69.1	66	64
H	73.9	68	71
H	67.1	64	68
H	64.4	62	66
H	71.1	66	74
H	71.0	63	73
H	67.4	64	62
H	69.3	65	69
H	64.9	64	67
H	68.1	64	68
H	66.5	62	72
H	67.5	69	66
H	66.5	62	72
H	70.3	67	68
H	67.5	63	71
H	68.5	66	67
H	71.9	65	71
H	67.8	71	75
M	58.6	63	64
M	64.7	67	65
M	65.3	64	67
M	61.0	60	72
M	65.4	65	72
M	67.4	67	72
M	60.9	59	67
M	63.1	60	71
M	60.0	58	66
M	71.1	72	75
M	62.2	63	69
M	67.2	67	70
M	63.4	62	69
M	68.4	69	62
M	62.2	63	66
M	64.7	64	76
M	59.6	63	69
M	61.0	64	68
M	64.0	60	66
M	65.4	65	68

Conjunto de datos 3: Circunferencias de la cabeza (cm) de bebés de dos meses de edad

Los datos son del Department of Health and Human Services de EUA, National Center for Health Statistics, Third National Health and Nutrition Examination Survey.

STATDISK	**Nombres de los archivos de STATDISK y de los archivos de texto: MHED, FHED.**
Minitab:	**El nombre de la hoja de cálculo es HEADCIRC.MTW.**
Excel:	**El nombre del libro de trabajo es HEADCIRC.XLS.**
TI-83 Plus:	**El nombre de la App es HEADCIRC y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.**

Hombre

40.1	39.8	42.3	41.0	42.5	40.9	35.5	35.7	41.1	41.4
42.2	42.3	43.2	42.2	42.4	43.2	39.9	40.9	40.7	41.7
41.7	41.0	40.4	42.0	41.2	39.7	41.9	41.3	40.2	41.0
41.1	40.4	39.2	42.8	41.9	42.8	41.0	40.9	42.0	42.6
41.0	39.6	40.2	40.9	40.2	41.8	41.7	41.7	40.9	42.8

Mujer

39.3	40.2	41.3	38.1	39.6	40.6	38.6	40.5	40.5	40.3
39.5	40.7	40.2	38.2	40.3	42.6	39.9	40.0	40.7	38.6
41.0	43.7	40.0	40.1	41.0	40.8	41.0	40.3	40.2	39.2
34.4	41.0	39.6	40.9	36.9	43.6	40.2	40.8	37.8	41.2
42.0	38.3	39.6	38.9	36.3	39.9	40.3	40.1	42.0	41.6

Conjunto de datos 4: Temperaturas corporales (en grados Fahrenheit) de adultos saludables

Datos proporcionados por los doctores Steven Wasserman, Philip Mackowiak y Myron Levine de la Universidad de Maryland.

				Temperatura día 1		Temperatura día 2	
Sujeto	Edad	Género	Fuma	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	M	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
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

(continúa)

Conjunto de datos 4: Temperaturas corporales (*continuación*)

Sujeto	Edad	Género	Fuma	Temperatura día 1 8 AM	Temperatura día 1 12 AM	Temperatura día 2 8 AM	Temperatura día 2 12 AM
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
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

(continúa)

Conjunto de datos 4: Temperaturas corporales (*continuación*)

Sujeto	Edad	Género	Fuma	Temperatura día 1 8 AM	Temperatura día 1 12 AM	Temperatura día 2 8 AM	Temperatura día 2 12 AM
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	N	—	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 5: Alquitrán, nicotina y monóxido de carbono de cigarros

Todas las mediciones son en miligramos por cigarro, y todos los cigarros son de 100 mm de largo, con filtro, y no son del tipo mentolado ni *light*. Los datos son de la Federal Trade Commission.

STATDISK Nombres de los archivos de STATDISK y de los archivos de texto: TAR, NICOT, CO.
Minitab: El nombre de la hoja de cálculo es CIGARET.MTW.
Excel: El nombre del libro de trabajo es CIGARET.XLS.
TI-83 Plus: El nombre de la App es CIGARET y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

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 6: 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 Department of Health and Human Services de EUA, National Center for Health Statistics, Third National Health and Nutrition Examination Survey.

STATDISK	**Nombres de los archivos de STATDISK y de los archivos de texto: NOETS, ETS, SMKR.**
Minitab:	**El nombre de la hoja de cálculo es COTININE.MTW.**
Excel:	**El nombre del libro de trabajo es COTININE.XLS.**
TI-83 Plus:	**TI-83 Plus: El nombre de la App es COTININE y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.**

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)

0	0	0	0	0	0	0	0	0	0
0	9	0	0	0	0	0	0	244	0
1	0	0	0	90	1	0	309	0	0
0	0	0	0	0	0	0	0	0	0

Conjunto de datos 7: 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 están en segundos, y los tiempos de consumo de alcohol están 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 Nombres de los archivos de STATDISK y de los archivos de texto: CHLEN, CHTOB, CHALC.
Minitab: El nombre de la hoja de cálculo es CHMOVIE.MTW.
Excel: El nombre del libro de trabajo es CHMOVIE.XLS.
TI-83 Plus: El nombre de la App es CHMOVIE y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Película	Compañía	Duración (min)	Consumo de tabaco (seg)	Consumo de alcohol (seg)
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
Fun and Fancy Free	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
Rescuers	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
Policías y ratones	Disney	73	165	414
Oliver y su pandilla	Disney	72	74	0
La sirenita	Disney	82	9	0
Rescuers Down Under	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 and the Giant Peach	Disney	79	206	38
Hércules	Disney	92	9	13
Secret of NIMH	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
Babes in Toyland	MGM	74	0	0
Pulgarcita	Warner Bros	86	6	5
Troll en el Parque Central	Warner Bros	76	1	0
Space Jam	Warner Bros	81	117	0
Pippi Longstocking	Warner Bros	75	5	0
Los gatos no bailan	Warner Bros	75	91	0
An American Tail	Universal	77	155	74
Land Before Time	Universal	70	0	0
Fievel Goes West	Universal	75	24	28
We’re Back: Dinosaur Story	Universal	64	55	0
Land Before Time 2	Universal	73	0	0
Balto	Universal	74	0	0
Once Upon a Forest	20th Century Fox	71	0	0
Anastasia	20th Century Fox	94	17	39

Conjunto de datos 8: Finalistas de la carrera de maratón de la ciudad de Nueva York

La muestra es de 150 corredores seleccionados al azar de la población de 29,373 corredores finalistas de la maratón de la ciudad de Nueva York en un año reciente.

STATDISK Nombres de los archivos de STATDISK y de los archivos de texto: MRORD, MREDAD, MRTIM.
Minitab: Minitab: El nombre de la hoja de cálculo es MARATHON. MTW.
Excel: Excel: El nombre del libro de trabajo es MARATHON.XLS.
TI-83 Plus: El nombre de la App es MARATHON y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Orden	Edad	Género	Tiempo (seg)	Orden	Edad	Género	Tiempo (seg)
130	32	H	9631	7082	38	H	13851
265	39	H	10209	7093	32	M	13854
314	39	H	10351	7933	50	H	14057
490	36	H	10641	7966	43	H	14066
547	34	H	10723	8011	25	H	14078
708	28	H	10905	8027	39	H	14082
834	42	H	11061	8042	31	H	14086
944	46	H	11188	8186	37	H	14121
1084	32	H	11337	8225	46	H	14128
1086	34	H	11338	8609	23	M	14216
1132	41	H	11382	8707	30	M	14235
1593	36	H	11738	8823	24	H	14256
1625	50	H	11761	9451	29	H	14375
1735	36	H	11830	9630	30	H	14402
1792	40	H	11874	10130	36	H	14512
1826	33	H	11897	10191	40	H	14528
2052	29	M	12047	10556	51	H	14617
2108	28	H	12077	10585	51	H	14623
2167	40	H	12115	10643	51	H	14632
2505	30	M	12289	10821	30	H	14677
2550	28	H	12312	10910	38	H	14698
3344	44	H	12639	10979	59	H	14720
3376	45	H	12652	10982	28	M	14721
4115	45	H	12940	11091	49	H	14752
4252	54	H	12986	11413	55	H	14836
4459	33	H	13063	11699	53	H	14919
4945	49	H	13217	11769	53	H	14935
5269	45	H	13315	11792	40	H	14942
5286	40	H	13322	11869	38	H	14964
5559	26	H	13408	11896	35	H	14971
6169	23	M	13593	11997	54	H	14996
6235	21	H	13615	12019	21	H	15002
6552	50	M	13704	12160	33	M	15036
6618	33	H	13722	12306	58	M	15077
6904	38	H	13802	12683	43	H	15167
6996	40	H	13829	12845	33	H	15210

(continúa)

Conjunto de datos 8: Finalistas de la carrera de maratón de la ciudad de Nueva York (*continuación*)

Orden	Edad	Género	Tiempo (seg)	Orden	Edad	Género	Tiempo (seg)
12942	35	H	15232	21013	38	H	17396
13226	31	H	15309	21017	47	H	17397
13262	38	H	15318	21524	34	H	17563
13297	28	M	15326	21787	37	M	17636
13434	30	M	15357	22009	37	H	17711
13597	23	M	15402	22042	31	M	17726
14391	40	H	15608	22258	29	M	17799
14633	43	H	15671	22285	49	H	17807
14909	43	H	15741	22638	31	H	17918
15282	29	H	15825	22993	52	H	18041
16030	34	M	16013	23092	38	H	18080
16324	30	H	16090	24018	30	M	18469
16723	65	H	16194	24283	31	M	18580
16840	50	H	16229	24290	40	H	18583
17104	37	M	16297	24417	50	M	18647
17298	30	M	16352	24466	29	H	18677
17436	32	H	16389	24649	21	H	18784
17483	19	M	16401	24845	53	H	18906
17487	42	H	16402	25262	41	H	19164
17694	33	H	16461	25287	50	M	19177
18132	42	H	16582	25956	45	H	19669
18765	51	H	16752	26471	27	M	20084
18783	54	M	16758	26545	32	H	20164
18825	32	M	16771	26637	53	H	20269
18897	34	M	16792	27035	42	M	20675
19002	31	H	16812	27046	45	H	20698
19210	50	M	16871	27133	39	H	20808
19264	60	H	16886	27152	31	H	20841
19278	49	H	16889	27196	68	M	20891
19649	51	M	16991	27277	51	H	20970
19789	45	H	17034	27800	51	H	21649
20425	40	M	17211	27955	31	M	21911
20558	30	H	17245	27995	25	M	21983
20562	25	H	17246	28062	25	H	22087
20580	32	H	17252	28085	61	H	22146
20592	34	H	17257	28578	31	H	23545
20605	42	M	17260	28779	32	H	24384
20700	34	M	17286	28986	47	M	25399
20826	52	H	17327	29045	61	M	25898

Conjunto de datos 9: Osos (osos salvajes anestesiados)

EDAD es en meses, MES es el mes de la medición (1 = enero), SEXO está codificado con 1 = macho y 2 = hembra, CABEZAL es la longitud de la cabeza (pulgadas), CABEZAA es la anchura de la cabeza (pulgadas), CUELLO es la circunferencia del cuello (en 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: Nombres de los archivos de STATDISK y de los archivos de texto: BAGE, BMNTH, BSEX, BHDLN, BHDWD, BNECK, BLEN, BCHST, BWGHT.
Minitab: El nombre de la hoja de cálculo es BEARS.MTW.
Excel: El nombre del libro de trabajo es BEARS.XLS.
TI-83 Plus: El nombre de la App es BEARS y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

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 10: Temperaturas reales y pronosticadas

Las temperaturas son en grados Fahrenheit y las cantidades de precipitación son en pulgadas. Todas las mediciones se registraron cerca de la casa del autor.

STATDISK Nombres de los archivos de STATDISK y de los archivos de texto: ACTHI, ACTLO, PHI1, PLO1, PHB, PLO3, PHI5, PLO5, y PREC.
Minitab: El nombre de la hoja de cálculo es WEATHER.MTW.
Excel: El nombre del libro de trabajo es WEATHER.XLS.
TI-83 Plus: El nombre de la App es WEATHER y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Fecha	Real máxima	Real mínima	Pronóstico 1 día máxima	Pronóstico 1 día mínima	Pronóstico 3 días máxima	Pronóstico 3 días mínima	Pronóstico 5 días máxima	Pronóstico 5 días mínima	Precip. (pulg.)
Ene. 1	30	1	28	13	30	18	28	16	0
Ene. 2	25	−5	29	13	26	17	27	16	0
Ene. 3	31	−5	32	14	30	13	28	20	0
Ene. 4	33	23	29	13	32	19	30	22	0
Ene. 5	29	9	30	19	35	26	26	15	0
Ene. 6	36	14	35	23	36	24	35	24	0.26
Ene. 7	36	12	35	21	38	25	34	23	0
Ene. 8	37	18	32	18	35	22	34	22	0.01
Ene. 9	32	26	27	17	33	18	33	21	0.21
Ene. 10	28	13	25	16	34	21	35	24	0.02
Ene. 11	43	7	41	22	37	26	38	26	0
Ene. 12	37	10	30	7	37	20	37	28	0
Ene. 13	36	6	33	20	31	14	36	16	0
Ene. 14	37	10	40	27	44	35	36	22	0
Ene. 15	34	29	34	24	38	29	45	26	0.02
Ene. 16	41	33	38	24	39	25	36	22	0.05
Ene. 17	40	36	33	15	37	21	33	21	0
Ene. 18	33	18	35	28	37	20	34	25	0
Ene. 19	35	32	40	25	39	32	36	21	0.01
Ene. 20	33	24	27	15	28	20	33	22	0.02
Ene. 21	31	19	27	10	30	16	31	18	0.21
Ene. 22	33	1	30	15	31	18	30	15	0.08
Ene. 23	35	0	37	19	40	23	38	20	0
Ene. 24	38	6	38	18	40	24	39	22	0
Ene. 25	37	26	29	14	31	18	33	17	0.01
Ene. 26	31	5	36	23	32	24	36	23	0
Ene. 27	38	20	34	16	36	24	37	21	0.01
Ene. 28	35	24	30	14	36	18	35	24	0
Ene. 29	33	9	36	28	39	25	40	26	0
Ene. 30	39	26	41	31	42	35	36	29	0
Ene. 31	46	32	42	26	42	30	40	18	0

Conjunto de datos 11: Precipitación pluvial (en pulgadas) en Boston durante un año

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: RNMON, RNTUE, RNWED, RNTHU, RNFRI, RNSAT, RNSUN.
Minitab: El nombre de la hoja de cálculo es BOSTRAIN.MTW.
Excel: El nombre del libro de trabajo es BOSTRAIN.XLS.
TI-83 Plus: El nombre de la App es BOSTRAIN y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Lun	Mar	Miér	Jue	Vie	Sáb	Dom
0	0	0	0.04	0.04	0	0.05
0	0	0	0.06	0.03	0.1	0
0	0	0	0.71	0	0	0
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
0.01	0	0	0	0.3	0.05	0
0	0	0.01	0	0	0	0
0	0.01	0.01	0.16	0	0	0.09
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
1.41	0.65	0.31	0	0	0.54	0
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
0.92	0.36	0.02	0.09	0.27	0	0
0.01	0	0.06	0	0	0	0.27
0.01	0	0	0	0	0	0.01
0	0	0	0	0	0	0
0	0	0	0	0.71	0	0
0	0	0.27	0.08	0	0	0.33
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 12: Temperatura, lluvia y conductividad en los Everglades

Las temperaturas están en grados Celsius y se midieron en el fondo. La conductividad es conductancia específica y tuvo una correlación muy alta con la salinidad. Todas las mediciones provienen de la estación hidrológica Garfield Bight en los Everglades de Florida. Los datos son de Kevin Kotun y el National Park Service.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: EVTMP, EVCON, EVRN.
Minitab: El nombre de la hoja de cálculo es EVERGLADE.MTW.
Excel: El nombre del libro de trabajo es EVERGLADE.XLS.
TI-83 Plus: El nombre de la App es EVERGLADE y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Temp	Conductividad	Precipitación pluvial (pulgadas)	Temp	Conductividad	Precipitación pluvial (pulgadas)
27.6	57.8	0.10	29.2	30.2	0.07
29.1	57.8	0.17	28.2	33.5	0.03
29.4	57.1	0.65	29.1	40.5	0.00
28.5	57.0	0.00	29.9	42.4	0.01
28.6	57.3	0.00	29.9	46.7	0.00
28.0	58.4	0.00	30.6	46.7	0.00
27.9	59.2	0.65	30.6	46.5	0.00
29.0	57.7	0.00	30.9	45.6	0.00
30.6	56.8	0.67	30.0	47.1	0.00
31.2	56.8	0.03	30.7	48.1	0.00
30.7	55.2	1.72	31.9	50.5	0.00
28.0	53.6	0.00	31.5	51.2	0.02
28.3	52.0	0.84	31.2	50.4	0.00
30.1	51.9	0.00	30.9	49.9	0.94
31.3	49.8	0.00	30.6	49.0	0.00
31.0	49.8	0.06	30.1	48.5	0.38
30.8	51.7	0.50	31.1	51.3	0.05
28.5	48.6	1.50	31.5	52.1	0.34
25.9	44.3	1.40	31.8	52.4	0.02
28.5	43.2	0.00	32.0	51.0	0.00
31.9	41.5	0.00	32.6	52.2	0.34
31.3	40.6	0.18	32.9	50.3	0.02
29.4	35.9	2.77	32.7	48.5	0.00
30.0	33.8	0.04	33.5	49.7	0.00
30.1	32.8	0.00	33.8	49.9	0.09
28.8	30.5	1.11	33.7	48.5	0.00
29.5	32.7	0.00	33.6	48.3	0.00
30.5	32.1	0.04	32.3	49.0	0.00
29.2	30.3	1.72	31.6	49.9	0.00
28.8	28.1	0.00	32.0	51.0	0.00
30.1	29.3	0.05			

Conjunto de datos 13: Géiser Old Faithful

Las duraciones son en segundos, los intervalos de tiempo se midieron en minutos hasta la erupción siguiente, y las alturas de las erupciones están en pies. Los datos son cortesía del National Park Service y del geólogo investigador Rick Hutchinson.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: OFDTN, OFINT, OFHT.
Minitab: El nombre de la hoja de cálculo es OLDFAITH.MTW.
Excel: El nombre del libro de trabajo es OLDFAITH.XLS.
TI-83 Plus: El nombre de la App es OLDFAITH y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Duración	Intervalo	Altura
240	86	140
237	86	154
122	62	140
267	104	140
113	62	160
258	95	140
232	79	150
105	62	150
276	94	160
248	79	155
243	86	125
241	85	136
214	86	140
114	58	155
272	89	130
227	79	125
237	83	125
238	82	139
203	84	125
270	82	140
218	78	140
226	91	135
250	89	141
245	79	140
120	57	139
267	100	110
103	62	140
270	87	135
241	70	140
239	88	135
233	82	140
238	83	139
102	56	100
271	81	105
127	74	130
275	102	135
140	61	131
264	83	135
134	73	153
268	97	155
124	67	140
270	90	150
249	84	153
237	82	120
235	81	138
228	78	135
265	89	145
120	69	130
275	98	136
241	79	150

Conjunto de datos 14: Libros de Clancy, Rowling y Tolstoi

Cada renglón de datos representa una página seleccionada al azar.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: CLWDS, CLCHR, CLFRE, CLFKG, RWWDS, RWCHR, RWFRE, RWFKG, TLWDS, TLCHR, TLFRE, TLFKG.
Minitab: Los nombres de las hojas de cálculo son CLANCY.MTW, ROWLING. MTW, TOLSTOY.MTW.
Excel: Los nombres de los libros de trabajo son CLANCY.XLS, ROWLING.XLS, TOLSTOY.XLS.
TI-83 Plus: Los nombres de las Apps son CLANCY, ROWLING, TOLSTOY, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Tom Clancy: *El oso y el dragón*

Palabras/oración	Caracteres/palabra	Facilidad de lectura de Flesch	Nivel de Flesch-Kincaid
15.0	4.8	58.2	8.8
9.8	4.5	73.4	5.4
8.1	4.6	73.1	5.0
13.5	4.5	64.4	7.6
24.0	4.0	72.7	9.0
9.8	4.0	89.2	3.2
33.0	4.6	43.9	12.0
9.4	4.5	76.3	4.9
8.3	4.4	76.4	4.6
11.3	4.4	78.9	5.0
11.4	4.3	69.4	6.4
12.4	4.3	72.9	6.1

J. K. Rowling: *Harry Potter y la piedra filosofal*

Palabras/oración	Caracteres/palabra	Facilidad de lectura de Flesch	Nivel de Flesch-Kincaid
15.7	4.1	85.3	5.2
9.0	4.2	84.3	3.7
16.3	4.2	79.5	6.1
14.5	4.4	82.5	4.9
9.7	4.3	80.2	4.4
7.4	4.2	84.6	3.2
14.0	4.5	79.2	5.6
16.1	4.5	70.9	6.9
13.9	4.3	78.6	5.7
12.5	4.0	86.2	4.1
17.2	4.4	74.0	6.7
11.5	4.3	83.7	4.4

León Tolstoi: *La guerra y la paz*

Palabras/oración	Caracteres/palabra	Facilidad de lectura de Flesch	Nivel de Flesch-Kincaid
20.6	4.3	69.4	8.6
28.0	4.5	64.2	9.8
12.0	4.5	71.4	6.1
11.5	4.5	71.6	5.9
17.4	4.5	68.5	7.7
19.7	4.8	51.9	10.9
20.3	4.3	72.2	8.2
17.8	4.2	74.4	7.2
22.1	4.7	52.8	11.0
31.4	4.3	58.4	11.5
18.3	4.4	65.4	8.4
11.7	4.5	73.6	5.9

Conjunto de datos 15: Edades de los polizones del *Queen Mary*

Los datos son de la Cunard Steamship Co., Ltd.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: WEST, EAST.
Minitab: El nombre de la hoja de cálculo es STOWAWAY.MTW.
Excel: El nombre del libro de trabajo es STOWAWAY.XLS.
TI-83 Plus: El nombre de la App es STOWAWAY, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Costa oeste

41	24	32	26	39	45	24	21	22	21
40	18	33	33	19	31	16	16	23	19
16	20	18	22	26	22	38	42	25	21
29	24	18	17	24	18	19	30	18	24
31	30	48	29	34	25	23	41	16	17
15	19	18	66	27	43				

Costa este

24	24	34	15	19	22	18	20	20	17
17	20	18	23	37	15	25	28	21	15
48	18	12	15	23	25	22	21	30	19
20	20	35	19	38	26	19	20	19	41
31	20	19	18	42	25	19	47	19	22
20	23	24	37	23	30	32	28	32	48
27	31	22	34	26	20	22	15	19	20
18	26	36	31	35					

Conjunto de datos 16: Cereal

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: CRCST, CRCAL, CRFAT, CRSGR, CRCHO, CRSOD, CRPRO, CRSHL.
Minitab: El nombre de la hoja de cálculo es CEREAL.MTW.
Excel: El nombre del libro de trabajo es CEREAL.XLS.
TI-83 Plus: El nombre de la App es CEREAL, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Cereal	Costo ($) por 100 gramos de cereal	Calorías por gramo de cereal	Gramos de grasa por gramo de cereal	Gramos de azúcar por gramo de cereal	Colesterol por gramo de cereal	Sodio (mg) por gramo de cereal	Proteínas (g) por gramo de cereal	Posición del anaquel
Cheerios	0.67	3.7	0.07	0.03	0	9.3	0.10	1
Harmony	0.82	3.6	0.02	0.24	0	6.4	0.09	3
Smart Start	0.78	3.6	0.01	0.30	0	6.6	0.06	4
Cocoa Puffs	1.03	4.0	0.03	0.47	0	5.7	0.03	2
Lucky Charms	0.83	4.0	0.03	0.43	0	7.0	0.07	2
Corn Flakes	0.55	3.6	0.00	0.07	0	7.1	0.07	1
Fruit Loops	0.68	3.8	0.03	0.47	0	4.7	0.03	2
Wheaties	0.78	3.7	0.03	0.13	0	7.3	0.10	1
Cap'n Crunch	0.73	4.1	0.06	0.44	0	7.4	0.04	1
Frosted Flakes	0.65	3.9	0.00	0.39	0	4.8	0.03	1
Apple Jacks	0.81	3.9	0.02	0.48	0	4.5	0.03	2
Bran Flakes	0.70	3.3	0.02	0.17	0	7.0	0.10	4
Special K	0.78	3.5	0.00	0.13	0	7.1	0.23	1
Rice Krispies	0.95	3.6	0.00	0.09	0	9.7	0.06	4
Corn Pops	0.84	3.9	0.00	0.45	0	3.9	0.03	2
Trix	0.94	4.0	0.03	0.43	0	6.3	0.03	2

Conjunto de datos 17: Pesos y volúmenes de bebidas de cola

Los pesos están en libras y los volúmenes en onzas.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: CRGWT, CRGVL, CDTWT, CDTVL, PRGWT, PRGVL, PDTWT, PDTVL.
Minitab: El nombre de la hoja de cálculo es COLA.MTW.
Excel: El nombre del libro de trabajo es COLA.XLS.
TI-83 Plus: El nombre de la App es COLA, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

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 18: Diamantes

El precio está en dólares. El fondo es 100 veces la proporción de la altura al diámetro. La mesa es el tamaño de la superficie plana superior (el fondo y la mesa determinan el "corte"). Los índices de color están en una escala estándar, con 1 = sin color y los números crecientes indican más amarillo. En la escala de claridad, 1 = impecable y 6 indica inclusiones visibles a simple vista.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: PRICE, CARAT, DEPTH, TABLE, COLOR, CLRTY.
Minitab: El nombre de la hoja de cálculo es DIAMONDS.MTW.
Excel: El nombre del libro de trabajo es DIAMONDS.XLS.
TI-83 Plus: El nombre de la App es DIAMONDS, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Precio	Quilates	Fondo	Mesa	Color	Claridad
6958	1.00	60.5	65	3	4
5885	1.00	59.2	65	5	4
6333	1.01	62.3	55	4	4
4299	1.01	64.4	62	5	5
9589	1.02	63.9	58	2	3
6921	1.04	60.0	61	4	4
4426	1.04	62.0	62	5	5
6885	1.07	63.6	61	4	3
5826	1.07	61.6	62	5	5
3670	1.11	60.4	60	9	4
7176	1.12	60.2	65	2	3
7497	1.16	59.5	60	5	3
5170	1.20	62.6	61	6	4
5547	1.23	59.2	65	7	4
18596	1.25	61.2	61	1	2
7521	1.29	59.6	59	6	2
7260	1.50	61.1	65	6	4
8139	1.51	63.0	60	6	4
12196	1.67	58.7	64	3	5
14998	1.72	58.5	61	4	3
9736	1.76	57.9	62	8	2
9859	1.80	59.6	63	5	5
12398	1.88	62.9	62	6	2
25322	2.03	60.1	62	2	3
11008	2.03	62.0	63	8	3
38794	2.06	58.2	63	2	2
66780	3.00	63.3	62	1	3
46769	4.01	57.1	51	3	4
28800	4.01	63.0	63	6	5
28868	4.05	59.3	60	7	4

Conjunto de datos 19: Pesos de una muestra de dulces M&M clásicos

Los pesos están en gramos.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: RED, ORNG, YLLW, BROWN, BLUE, GREEN.
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 Plus: El nombre de la App es MM, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Rojo	Naranja	Amarillo	Café	Azul	Verde
0.870	0.903	0.906	0.932	0.838	0.911
0.933	0.920	0.978	0.860	0.875	1.002
0.952	0.861	0.926	0.919	0.870	0.902
0.908	1.009	0.868	0.914	0.956	0.930
0.911	0.971	0.876	0.914	0.968	0.949
0.908	0.898	0.968	0.904		0.890
0.913	0.942	0.921	0.930		0.902
0.983	0.897	0.893	0.871		
0.920		0.939	1.033		
0.936		0.886	0.955		
0.891		0.924	0.876		
0.924		0.910	0.856		
0.874		0.877	0.866		
0.908		0.879	0.858		
0.924		0.941	0.988		
0.897		0.879	0.936		
0.912		0.940	0.930		
0.888		0.960	0.923		
0.872		0.989	0.867		
0.898		0.900	0.965		
0.882		0.917	0.902		
		0.911	0.928		
		0.892	0.900		
		0.886	0.889		
		0.949	0.875		
		0.934	0.909		
			0.976		
			0.921		
			0.898		
			0.897		
			0.902		
			0.920		
			0.909		

Conjunto de datos 20: Cargas axiales de latas de aluminio

Las cargas axiales están medidas en libras.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: CN109, CN111.
Minitab: El nombre de la hoja de cálculo es CANS.MTW.
Excel: El nombre del libro de trabajo es CANS.XLS.
TI-83 Plus: El nombre de la App es CANS, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Muestra	Latas de aluminio de 0.0109 pulgadas Carga (libras)							Muestra	Latas de aluminio de 0.0111 pulgadas 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 21: Películas

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: MVBUD, MVGRS, MVLEN, MVRAT.
Minitab: El nombre de la hoja de cálculo es MOVIES.MTW.
Excel: El nombre del libro de trabajo es MOVIES.XLS.
TI-83 Plus: El nombre de la App es MOVIES, y los nombres de los archivos son los mismos que para los archivos de texto de STATDISK.

Título	Año	Clasificación	Presupuesto ($) en millones	Ganancias ($) en millones	Duración en minutos	Tasa de audiencia
Alien	1986	Adultos	18.5	81.843	137	8.2
Armagedón	1998	Con supervisión-13	140	194.125	144	6.7
Mejor, imposible	1997	Con supervisión-13	50	147.54	138	8.1
Corazón valiente	1995	Adultos	72	75.6	177	8.3
Chasing Amy	1997	Adultos	0.25	12.006	105	7.9
Contacto	1997	Con supervisión	90	100.853	153	8.3
El pico de Dante	1997	Con supervisión-13	104	67.155	112	6.7
Impacto profundo	1998	Con supervisión-13	75	140.424	120	6.4
Decisión ejecutiva	1996	Adultos	55	68.75	129	7.3
Forrest Gump	1994	Con supervisión-13	55	329.691	142	7.7
Ghost, la sombra del amor	1990	Con supervisión-13	22	217.631	128	7.1
Lo que el viento se llevó	1939	General	3.9	198.571	222	8.0
Good Will Hunting	1997	Adultos	10	138.339	126	8.5
Vaselina	1978	Con supervisión	6	181.28	110	7.3
Halloween	1978	Adultos	0.325	47	93	7.7
Hard Rain	1998	Adultos	70	19.819	95	5.2
Sé lo que hicieron el verano pasado	1997	Adultos	17	72.219	100	6.5
El día de la Independencia	1996	Con supervisión-13	75	306.124	142	6.6
Indiana Jones y la última cruzada	1989	Con supervisión-13	39	197.171	127	7.8
Tiburón	1975	Con supervisión	12	260	124	7.8
Hombres de negro	1997	Con supervisión-13	90	250.147	98	7.4
Multiplicity	1996	Con supervisión-13	45	20.1	117	6.8
Pulp Fiction	1994	Adultos	8	107.93	154	8.3
Los cazadores del arca perdida	1981	Con supervisión	20	242.374	115	8.3
Salvando al soldado Ryan	1998	Adultos	70	178.091	170	9.1
La lista de Schindler	1993	Adultos	25	96.067	197	8.6
Scream	1996	Adultos	15	103.001	111	7.7
Velocidad máxima 2	1997	Con supervisión-13	110	48.068	121	4.3
Terminator	1984	Adultos	6.4	36.9	108	7.7
El presidente	1995	Con supervisión-13	62	65	114	7.6
El quinto elemento	1997	Con supervisión-13	90	63.54	126	7.8
El juego	1997	Adultos	50	48.265	128	7.6
El hombre de la máscara de hierro	1998	Con supervisión-13	35	56.876	132	6.5
Titanic	1997	Con supervisión-13	200	600.743	195	8.4
Mentiras verdaderas	1994	Adultos	100	146.261	144	7.2
Volcán	1997	Con supervisión-13	90	47.474	102	5.8

Conjunto de datos 22: Automóviles

CIUDAD es el consumo de combustible en ciudad en millas/galón, CARR es el consumo de combustible en carretera en millas/galón, PESO es el peso del automóvil en libras, CILINDROS es el número de cilindros, DESPLAZAMIENTO es el desplazamiento del motor en litros, MAN/AUT indica transmisión manual o automática, GIN es la cantidad emitida de gases invernadero (en toneladas/año), y OXN es la cantidad de emisiones de NO_x en el escape (en libras/año).

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: CRCTY, CRHWY, CRWT, CRCYL, CRDSP, CRGHG, CRNOX.
Minitab: El nombre de la hoja de cálculo es CARS.MTW.
Excel: El nombre del libro de trabajo es CARS.XLS.
TI-83 Plus: El nombre de la App es CARS, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Automóvil	Ciudad	Carr	Peso	Cilindros	Desplazamiento	MAN/AUT	GIN	OXN
Chev. Camaro	19	30	3545	6	3.8	M	12	34.4
Chev. Cavalier	23	31	2795	4	2.2	A	10	25.1
Dodge Neón	23	32	2600	4	2	A	10	25.1
Ford Taurus	19	27	3515	6	3	A	12	25.1
Honda Accord	23	30	3245	4	2.3	A	11	25.1
Lincoln Cont.	17	24	3930	8	4.6	A	14	25.1
Mercury Mystique	20	29	3115	6	2.5	A	12	34.4
Mitsubishi Eclipse	22	33	3235	4	2	M	10	25.1
Olds. Aurora	17	26	3995	8	4	A	13	34.4
Pontiac Grand Am	22	30	3115	4	2.4	A	11	25.1
Toyota Camry	23	32	3240	4	2.2	M	10	25.1
Cadillac DeVille	17	26	4020	8	4.6	A	13	34.4
Chev. Corvette	18	28	3220	8	5.7	M	12	34.4
Chrysler Sebring	19	27	3175	6	2.5	A	12	25.1
Ford Mustang	20	29	3450	6	3.8	M	12	34.4
BMW 3-Series	19	27	3225	6	2.8	A	12	34.4
Ford Crown Victoria	17	24	3985	8	4.6	A	14	25.1
Honda Civic	32	37	2440	4	1.6	M	8	25.1
Mazda Protege	29	34	2500	4	1.6	A	9	25.1
Hyundai Accent	28	37	2290	4	1.5	A	9	34.4

Conjunto de datos 23: Pesos de basura desechada en una semana

Los pesos están en libras. TAMAÑO es el tamaño del hogar. Datos proporcionados por Masakuza Tani, el Garbedad Project, University of Arizona.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: HHSIZ, METAL, PAPER, PLAS, GLASS, FOOD, YARD, TEXT, OTHER, TOTAL.
Minitab: El nombre de la hoja de cálculo es GARBAGE.MTW.
Excel: El nombre del libro de trabajo es GARBAGE.XLS.
TI-83 Plus: El nombre de la App es GARBAGE, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Hogar	Tamaño	Metal	Papel	Plas.	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

(continúa)

Conjunto de datos 23: Pesos de basura desechada en una semana (*continuación*)

Hogar	Tamaño	Metal	Papel	Plas.	Vidrio	Comida	Patio	Textos	Otros	Total
36	2	3.61	7.72	0.92	2.43	2.07	0.68	0.63	0.00	18.06
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 24: Casas vendidas en el condado Dutchess

STATDISK: **Nombres de los archivos de STATDISK y de los archivos de texto: HMSP, HMLST, HMLA, HMRMS, HMBRS, HMBTH, HMAGE, HMACR, HMTAX.**
Minitab: **El nombre de la hoja de cálculo es HOMES.MTW.**
Excel: **El nombre del libro de trabajo es HOMES.XLS.**
TI-83 Plus: **El nombre de la App es HOMES, y los nombres de los archivos son los mismos que para los archivos de texto STATDISK.**

Precio de venta (miles)	Precio de lista (miles)	Área habitable (cientos de pies cuadrados)	Habitaciones	Recámaras	Baños	Antigüedad (años)	Acres	Impuestos (dólares)
142.0	160	28	10	5	3	60	0.28	3167
175.0	180	18	8	4	1	12	0.43	4033
129.0	132	13	6	3	1	41	0.33	1471
138.0	140	17	7	3	1	22	0.46	3204
232.0	240	25	8	4	3	5	2.05	3613
135.0	140	18	7	4	3	9	0.57	3028
150.0	160	20	8	4	3	18	4.00	3131
207.0	225	22	8	4	2	16	2.22	5158
271.0	285	30	10	5	2	30	0.53	5702
89.0	90	10	5	3	1	43	0.30	2054
153.0	157	22	8	3	3	18	0.38	4127
86.5	90	16	7	3	1	50	0.65	1445
234.0	238	25	8	4	2	2	1.61	2087
105.5	116	20	8	4	1	13	0.22	2818
175.0	180	22	8	4	2	15	2.06	3917
165.0	170	17	8	4	2	33	0.46	2220
166.0	170	23	9	4	2	37	0.27	3498
136.0	140	19	7	3	1	22	0.63	3607
148.0	160	17	7	3	2	13	0.36	3648
151.0	153	19	8	4	2	24	0.34	3561
180.0	190	24	9	4	2	10	1.55	4681
293.0	305	26	8	4	3	6	0.46	7088
167.0	170	20	9	4	2	46	0.46	3482
190.0	193	22	9	5	2	37	0.48	3920
184.0	190	21	9	5	2	27	1.30	4162
157.0	165	20	8	4	2	7	0.30	3785
110.0	115	16	8	4	1	26	0.29	3103
135.0	145	18	7	4	1	35	0.43	3363
567.0	625	64	11	4	4	4	0.85	12192
180.0	185	20	8	4	2	11	1.00	3831
183.0	188	17	7	3	2	16	3.00	3564
185.0	193	20	9	3	2	56	6.49	3765
152.0	155	17	8	4	1	33	0.70	3361
148.0	153	13	6	3	2	22	0.39	3950

(continúa)

Conjunto de datos 24: Casas vendidas en el condado Dutchess (*continuación*)

Precio de venta (miles)	Precio de lista (miles)	Área habitable (cientos de pies cuadrados)	Habitaciones	Recámaras	Baños	Antigüedad (años)	Acres	Impuestos (dólares)
152.0	159	15	7	3	1	25	0.59	3055
146.0	150	16	7	3	1	31	0.36	2950
170.0	190	24	10	3	2	33	0.57	3346
127.0	130	20	8	4	1	65	0.40	3334
265.0	270	36	10	6	3	33	1.20	5853
157.0	163	18	8	4	2	12	1.13	3982
128.0	135	17	9	4	1	25	0.52	3374
110.0	120	15	8	4	2	11	0.59	3119
123.0	130	18	8	4	2	43	0.39	3268
212.0	230	39	12	5	3	202	4.29	3648
145.0	145	18	8	4	2	44	0.22	2783
129.0	135	10	6	3	1	15	1.00	2438
143.0	145	21	7	4	2	10	1.20	3529
247.0	252	29	9	4	2	4	1.25	4626
111.0	120	15	8	3	1	97	1.11	3205
133.0	145	26	7	3	1	42	0.36	3059

Conjunto de datos 25: Misceláneos: DJ, ventas de automóviles, muertes en vehículos motorizados, homicidios, manchas de sol y Super Bowl

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: DJIA, CRSLS, MVDTH, MURDR, SNSPT, SUPER.
Minitab: El nombre de la hoja de cálculo es MISC.MTW.
Excel: El nombre del libro de trabajo es MISC.XLS.
TI-83 Plus: El nombre de la App es MISC, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Año	DJ alto	Ventas de automóviles en EUA (miles)	Muertes en vehículos motorizados en EUA	Homicidios y homicidios sin agravantes en EUA	Número de manchas de sol	Puntos en el Súper Bowl
1980	1000	8979	53172	23040	154.6	50
1981	1024	8536	51385	22520	140.5	37
1982	1071	7982	45779	21010	115.9	57
1983	1287	9182	44452	19310	66.6	44
1984	1287	10390	46263	18690	45.9	47
1985	1553	11042	45901	18980	17.9	54
1986	1956	11460	47865	20610	13.4	56
1987	2722	10277	48290	20100	29.2	59
1988	2184	10530	49078	20680	100.2	36
1989	2791	9773	47575	21500	157.6	65
1990	3000	9300	46814	23440	142.6	39
1991	3169	8175	43536	24700	145.7	61
1992	3413	8213	40982	23760	94.3	69
1993	3794	8518	41893	24530	54.6	43
1994	3978	8991	42524	23330	29.9	75
1995	5216	8635	43363	21610	17.5	44
1996	6561	8527	43649	19650	8.6	56
1997	8259	8272	43458	18210	21.5	55
1998	9374	8142	43501	16970	64.3	53
1999	11568	8698	41300	15522	93.3	39
2000	11401	8847	43000	15517	119.6	41

Conjunto de datos 26: Lotería estatal de Nueva York

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: LOT1, LOT2, LOT3, LOT4, LOT5, LOT6, WIN1, WIN2, WIN3, WIN4.
Minitab: El nombre de la hoja de cálculo es LOTTO.MTW.
Excel: El nombre del libro de trabajo es LOTTO.XLS.
TI-83 Plus: El nombre de la App es LOTTO, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

Lotería del estado de Nueva York						NYS Win 4			
19	22	26	38	44	48	9	2	5	4
4	6	16	24	37	49	7	7	5	4
22	31	35	38	41	48	0	1	7	5
4	11	22	31	34	35	7	3	7	6
9	15	19	23	24	51	5	5	7	1
3	8	21	28	30	45	5	2	6	4
1	2	15	32	33	48	1	5	4	3
18	29	30	32	38	43	9	3	5	0
1	8	13	35	44	46	0	6	2	7
11	21	25	32	37	49	0	7	2	7
6	8	9	33	34	40	8	9	1	9
11	14	20	25	31	33	7	0	0	9
6	11	25	30	42	49	6	6	6	2
19	32	33	41	50	51	0	0	1	5
8	13	24	42	43	47	1	6	6	0
12	13	16	25	27	31	6	1	9	3
3	23	26	36	40	45	7	1	5	6
11	18	20	24	25	41	2	7	5	9
2	10	17	19	42	43	4	4	1	0
5	18	20	23	46	49	7	2	8	6
12	13	17	31	32	35	5	7	4	5
5	23	26	32	45	46	5	9	3	3
8	12	27	39	40	50	5	7	7	6
6	21	41	43	50	51	2	4	0	4
18	19	21	23	38	49	0	8	7	2
13	14	32	39	44	51	4	3	5	7
17	19	21	22	31	35	4	0	4	7
6	12	19	41	47	49	9	6	1	5
5	15	38	41	42	50	9	2	9	5
3	4	6	14	24	46	0	6	4	7
6	28	29	46	47	51	5	4	6	9
8	9	29	30	33	50	3	0	6	0
3	6	22	26	41	45	3	7	4	7
2	15	33	36	38	46	1	9	1	6
10	16	36	37	46	51	9	0	9	8
8	10	13	23	33	45	2	6	7	6
20	23	26	39	48	50	5	2	2	9
12	22	31	33	43	50	6	8	6	8
22	30	31	40	45	49	8	7	4	7
9	23	25	27	37	38	2	4	0	7

Conjunto de datos 27: Resultados del solitario

Resultados del juego de solitario de Microsoft (reglas de Las Vegas de “sacar 3” con apuesta de $52 y reembolso de $5 por naipe. Las cantidades están en dólares y representan la ganancia o pérdida netas del jugador).

STATDISK: Nombre del archivo de STATDISK y del archivo de texto: SOLTR.
Minitab: El nombre de la hoja de cálculo es SOLITAIRE.MTW.
Excel: El nombre del libro de trabajo es SOLITAIRE.XLS.
TI-83 Plus: El nombre de la App es SOLITAIRE, y el nombre del archivo es SOLTR.

−42	−17	−17	−37	−12	−22	−32	−12	−27	−27
−27	−17	−47	−47	−12	−2	−22	−27	−37	−27
−12	−37	−2	23	−17	−42	23	−42	−32	53
−27	−17	18	−2	−37	−32	−2	−7	23	−7
−7	−42	−42	−27	38	−27	−42	−47	−32	−32
−27	28	3	−7	−47	13	−42	3	−17	−27
−37	−27	23	−47	−47	−27	−27	−42	−17	−37
−32	−22	−12	208	−27	3	−17	−42	−32	−22
−52	−12	−47	−47	−22	−42	−12	−12	−32	−37
13	208	−42	−7	−47	−47	−42	−7	−32	−47
−42	−27	23	−52	−37	−42	−32	−42	−52	−17
−37	−47	−42	−47	−47	−37	−17	−17	−52	33
−17	18	−52	208	−47	−47	−37	−42	108	−2
−47	−42	23	18	−37	−7	−27	43	−27	−12
8	−27	8	−27	−2	−47	−12	18	−47	−17
−32	−52	−32	28	−2	−12	−27	208	−7	−37
−22	−12	−42	−7	−12	−27	−22	−27	−42	−42
−7	−12	−27	−27	−22	−12	8	−27	−32	3
8	−47	−37	−17	8	−42	−7	−47	−47	13
−7	−17	−22	−12	208	−2	−37	−32	−32	−27
−47	−12	53	−42	−32	−47	3	−42	−42	−37
−12	−27	−37	−37	−42	208	−22	−17	−32	−37
18	3	−42	−22	−22	−42	−42	−27	−32	−17
−17	−32	−2	−2	−2	−27	−27	−22	−22	−52
−32	−17	3	58	−22	−37	−42	−27	−52	−7
−32	−37	−12	−42	−52	−47	−37	208	−47	−27
−22	−12	−42	−32	−27	−37	−12	−37	−32	−7
−47	−42	−47	−37	−37	−37	−52	73	−32	53
−12	−27	−2	−17	−37	−37	−52	−22	−17	8
−32	−42	−2	−22	−32	−37	−47	68	−32	−17
−2	−12	−7	−37	−42	−42	−22	−37	−47	−32
−22	−47	−42	−47	−27	−47	208	−17	−17	−7
−37	−37	−32	−17	−32	−37	−37	−32	−37	−12
−47	−22	−27	−32	−42	−32	−37	−7	−22	−52
−22	−27	3	−47	−27	−42	208	−22	−7	−32
−27	−42	−22	−37	−52	−27	−42	−17	−37	18
−2	3	−32	−37	−42	18	−12	−17	43	−17
−42	−12	−22	−37	−17	18	−32	8	−12	208
−37	−32	−32	−2	−52	−42	−37	−27	−37	−27
8	−52	−32	208	13	−22	48	−27	−37	−42
−32	−37	−47	−7	−22	−17	−52	−2	−37	−37
−17	−42	−22	98	−37	−37	−22	13	−22	3
−47	33	−37	−47	23	−12	−17	−22	23	8
3	−32	−32	−17	3	−37	−27	−22	−7	18
−47	−37	−32	−12	−12	−42	−2	−27	−32	−37
−22	−32	−7	−12	3	−52	−22	−42	8	−7
−47	−2	−27	−7	−32	−17	−37	−17	−32	−42
−7	−42	−32	−52	−42	−27	208	−32	−32	−42
−42	−52	208	−37	−7	−27	−2	−37	23	−37
−12	−22	−7	−2	−32	13	8	−47	−37	−32

Conjunto de datos 28: Pesos de sobres de azúcar Dominó

Los pesos están en gramos.

STATDISK: Nombre del archivo de STATDISK y del archivo de texto: SUGAR.
Minitab: El nombre de la hoja de cálculo es SUGAR.MTW.
Excel: El nombre del libro de trabajo es SUGAR.XLS.
TI-83 Plus: El nombre de la App es SUGARWT, y el nombre del archivo es SUGAR.

3.647	3.638	3.635	3.645	3.521	3.617	3.666
3.588	3.545	3.590	3.621	3.532	3.511	3.516
3.531	3.678	3.643	3.583	3.723	3.673	3.588
3.600	3.611	3.580	3.667	3.506	3.632	3.450
3.660	3.569	3.573	3.526	3.494	3.601	3.604
3.407	3.522	3.598	3.585	3.577	3.522	3.464
3.604	3.508	3.718	3.635	3.643	3.507	3.687
3.582	3.622	3.654	3.482	3.494	3.475	3.492
3.542	3.625	3.688	3.468	3.639	3.582	3.491
3.535	3.548	3.671	3.665	3.726	3.576	3.725

Conjunto de datos 29: Pesos de monedas de 25 centavos de dólar

Los pesos están en gramos.

STATDISK: Nombre del archivo de STATDISK y del archivo de texto: QRTRS.
Minitab: El nombre de la hoja de cálculo es QUARTERS.MTW.
Excel: El nombre del libro de trabajo es QUARTERS.XLS.
TI-83 Plus: El nombre de la App es QUARTERS, y el nombre del archivo es QRTRS.

5.60	5.63	5.58	5.56	5.66	5.58	5.57	5.59	5.67	5.61
5.84	5.73	5.53	5.58	5.52	5.65	5.57	5.71	5.59	5.53
5.63	5.68	5.62	5.60	5.53	5.58	5.60	5.58	5.59	5.66
5.73	5.59	5.63	5.66	5.67	5.60	5.74	5.57	5.62	5.73
5.60	5.60	5.57	5.71	5.62	5.72	5.57	5.70	5.60	5.49

Conjunto de datos 30: Distancias de jonrones

Las distancias de jonrones para Mark McGwire (1998), Sammy Sosa (1998) y Barry Bonds (2001) están dadas en pies.

STATDISK: Nombres de los archivos de STATDISK y de los archivos de texto: MCGWR, SOSA, BONDS.
Minitab: El nombre de la hoja de cálculo es HOMERUNS.MTW.
Excel: El nombre del libro de trabajo es HOMERUNS.XLS.
TI-83 Plus: El nombre de la App es HOMERUNS, y los nombres de los archivos son los mismos que para los archivos de texto y de STATDISK.

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							

Apéndice C: TI-83 Plus

CLEAR

To **CLEAR** data in list L_1:

L_1
STAT [4:ClrList] 2nd 1 ENTER

ENTER

To **ENTER** data in list L_1:

STAT [1:Edit] ENTER value, press ENTER,...

Quit
When all data have been ENTERed, press 2nd Mode

STATS

To get **STATISTICS** for data in list L_1:

L_1
STAT [CALC] [1:1-Var Stats] 2nd 1 ENTER

Notes: Sx is the sample standard deviation s.
Q_1 and Q_3 may be different from textbook.

GRAPH

To get **HISTOGRAM** or **BOXPLOT** for data in L_1:

STAT PLOT
1. 2nd Y= ENTER ENTER
2. Select "Type" (for boxplot, middle of second row).
3. ZOOM [9: ZoomStat]

FREQ. DIST.

To get **STATISTICS FROM A FREQUENCY DISTRIBUTION:**

L_1 L_2
1. Clear L_1 and L_2: STAT [4:ClrList] 2nd 1 , 2nd 2 ENTER
2. ENTER the data in L_1 and L_2: ENTER CLASS MIDPOINTS IN L_1.
ENTER FREQUENCIES IN L_2.
3. To get the statistics:

L_1 L_2
STAT [CALC] [1:1-VarStats] 2nd 1 , 2nd 2 ENTER

BINOM.

To find **BINOMIAL PROBABILITIES:**

L_1 L_2
1. Clear L_1 and L_2: STAT [4:ClrList] 2nd 1 , 2nd 2 ENTER

DISTR number of trials prob.
2. 2nd Vars [0:binompdf(] n , p ENTER

L_2
STO→ 2nd 2 ENTER

3. Now use STAT Edit to view the probabilities in list L_2 and to ENTER the x-values (such as 0, 1, 2, ...) in L_1.
4. You can get the mean μ and the standard deviation σ with

L_1 L_2
STAT [CALC] [1:1-Var Stats] 2nd 1 , 2nd 2 ENTER

NORMAL

NORMAL DISTRIBUTION:

$\boxed{\text{2nd}}$ $\boxed{\text{Vars}}$ (DISTR) to get Area → 2: normalcdf (lower score, upper score, μ, σ)

to get Score → 3: invNorm (LEFT area, μ, σ)

Total area to *LEFT* of score.

CONF INT

To construct CONFIDENCE INTERVALS:

MEAN
- σ known: $\boxed{\text{STAT}}$ [TESTS] [7:ZInterval] $\boxed{\text{ENTER}}$
- σ not known: $\boxed{\text{STAT}}$ [TESTS] [8:TInterval] $\boxed{\text{ENTER}}$

Proportion: $\boxed{\text{STAT}}$ [TESTS] [A:1-PropZInt] $\boxed{\text{ENTER}}$

HYP TEST

HYPOTHESES TESTING

MEAN
- σ known: $\boxed{\text{STAT}}$ [TESTS] [1:Z-Test] $\boxed{\text{ENTER}}$
- σ not known: $\boxed{\text{STAT}}$ [TESTS] [2:T-Test] $\boxed{\text{ENTER}}$

PROPORTION: $\boxed{\text{STAT}}$ [TESTS] [5:1-PropZInt} $\boxed{\text{ENTER}}$

St. dev. or variance: You're on your own: $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ and Table A-4

CORR REG

CORRELATION and REGRESSION

1. Enter **PAIRED** data in lists L_1 and L_2.
2. $\boxed{\text{STAT}}$ [TESTS] [E:LinRegTTest] Choose Freq. 1 and $\neq$ 0
3. Interpret: Correlation: If P-value $\leq \alpha$, there IS a significant linear correlation.
   Regression: Get equation $y = a + bx$
   fill in values

CONTIN. TABLE

CONTINGENCY TABLE

1. Enter Table as a matrix: $\boxed{\text{2nd}}$ $\boxed{x^{-1}}$ (MATRIX) [EDIT] $\boxed{\text{ENTER}}$
   then press $\boxed{\text{2nd}}$ $\boxed{\text{MODE}}$ (QUIT) when done.
2. $\boxed{\text{STAT}}$ [TESTS] [C:χ^2-Test]

Apéndice D: Glosario

Alfa (α) Símbolo empleado para representar la probabilidad de un error tipo I. *Vea* nivel de significancia.

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 de varianza de un solo factor *Vea* análisis de varianza de un factor.

Análisis de varianza Método para analizar la varianza de población que permite hacer pruebas de hipótesis acerca de medias de poblaciones.

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 *Vea* análisis de varianza.

Aproximación clásica a la probabilidad Aproximación en la que la probabilidad de un suceso se determina al dividir el número de maneras en que éste puede suceder, entre el número total de resultados posibles.

Aproximación de la probabilidad por frecuencia relativa Valor de probabilidad estimado con base en observaciones reales.

Beta (β) Símbolo empleado para representar la probabilidad de un error tipo II.

Bimodal Que 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 $(\overline{x}, \overline{y})$, determinado a partir de una colección de datos bivariados.

CM (error) Cuadrado medio del error; se usa en el análisis de varianza.

CM (total) Cuadrado medio de la variación total; se usa en el análisis de varianza.

CM (tratamiento) Cuadrado medio de tratamientos; 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 *Vea* 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 *Vea* 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) Cociente de la desviación estándar con respecto a 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 ocurre 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 a fin de lograr y mantener un estado de control estadístico y mejorar la capacidad del proceso.

Control estadístico dentro de una muestra *Vea* proceso estadísticamente estable.

Corrección por continuidad Ajuste que se hace cuando una variable aleatoria discreta se aproxima con una variable aleatoria continua (sección 5-6).

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 de orden en cuatro grupos, con aproximadamente el 25% de los puntajes en cada grupo.

Curva de densidad Gráfica de una distribución de probabilidad continua.

Datos Información o números que describen alguna característica.

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 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 acomodados 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 que se han reunido 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 o 1 posibilidad o 2 posibilidades, etcétera.

Datos distantes Valores poco comunes, en el sentido de que están muy lejos de la mayoría de los datos.

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 puntaje respecto a la media, dividida entre el número de puntajes.

Desviación estándar Medida de variación igual a la raíz cuadrada de la varianza.

Desviación explicada Para un par de valores de una colección 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 puntaje respecto a la media, dividida entre el número de puntajes.

Desviación no explicada Para un par de valores de una colección 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 una colección de datos bivariados.

Diagrama de árbol Representación gráfica de los diferentes resultados posibles en un suceso compuesto.

Diagrama de cuadro y bigotes *Vea* 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 equiparados según características similares.

Diseño rigurosamente controlado Diseño experimental en el que se obliga a que todos los factores sean constantes, a fin de eliminar los efectos de factores ajenos.

Diseño totalmente aleatorizado Procedimiento de un experimento en el que cada elemento tiene la misma posibilidad de pertenecer a las diferentes categorías o tratamientos.

Distribución chi cuadrada Una distribución de probabilidad continua (se presenta en la sección 6-5).

Distribución de frecuencias Lista de valores de datos (individualmente o por grupos de intervalos) junto con sus correspondientes frecuencias (o conteos).

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 Colección de valores de una variable aleatoria, junto con sus correspondientes probabilidades.

Distribución *F* Distribución de probabilidad continua, que se introduce en la sección 8-5.

Distribución muestral de medias muestrales Distribución de las medias muestrales que se obtiene al seleccionar repetidamente muestras del mismo 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 5-1 de la sección 5-1.

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 *t* Distribución normal que suele estar asociada con datos muestrales de una población con una desviación estándar desconocida.

Distribución *t* de Student *Vea* 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 ($x_1, x_2, \ldots x_k$).

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.

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 distribución Distribución normal con una media igual a 0 y una desviación estándar igual a 1.

Error estándar de estimado Medida de la dispersión de puntos de muestra 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 *Vea* 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 cuanto ésta es falsa.

Errores no 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 Colección de métodos para planear experimentos, para obtener, organizar, resumir, presentar, analizar e interpretar datos, y sacar conclusiones con base en esos datos.

Estadística descriptiva Métodos empleados para resumir las características clave de los 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 de muestra 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 no predispuesto Estadístico de muestra que tiende a acercarse al parámetro de población para cuya estimación se usa.

Estudio ciego Procedimiento utilizado en experimentos en los que el sujeto no sabe si está recibiendo un tratamiento o un placebo.

Estudio cohorte Estudio de sujetos en grupos identificados que comparten factores comunes (denominados *cohortes*), en el que los datos se reunirán en el futuro.

Estudio de control de caso Estudio en el que se reúnen datos del pasado (a través del examen de registros, entrevistas y otros).

Estudio doble ciego Procedimiento utilizado en un experimento, en el que el sujeto no sabe si está recibiendo un tratamiento o un placebo, y el experimentador tampoco lo sabe.

Estudio longitudinal Estudio de sujetos en grupos identificados que comparten factores comunes (llamados *cohortes*), donde 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 *cohortes*), 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).

Estudio transversal Estudio en el que los datos se observan, miden y reúnen en un punto del tiempo.

Experimento La aplicación de un tratamiento, seguida por la 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 (vea la fórmula 4-5 de la sección 4-3).

Fractiles Números que dividen los datos en partes de el mismo tamaño aproximadamente.

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 contingencia 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 modo 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 dado; 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 circular Representación gráfica de datos en forma de círculo que contiene divisiones radiales.

Gráfica de control Cualquiera de varios tipos de gráficas (capítulo 13), que representa alguna característica de un proceso, para 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 Representación gráfica de datos (x, y) apareados, 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 de un proceso.

Gráfica de tallo y hojas Método para clasificar y acomodar datos a modo de 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.

Gráfica $\overline{x}$ Gráfica de control que se usa para vigilar la media de un proceso.

Grupo control En un experimento, grupo de sujetos a quienes no se les da tratamiento.

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 el 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 los factores 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 significan algo.

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 de 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 se ajusta mejor a una colección de puntos que representan datos muestrales apareados.

Marca de clase Mitad de la suma de los valores máximo y mínimo.

Margen de error Máxima diferencia probable (con probabilidad $1 - \alpha$) entre el estadístico de muestra observado y el verdadero valor del parámetro de población.

Media La suma de un conjunto de puntajes, dividida entre el número de puntajes.

Media aritmética Suma de un conjunto de puntajes dividida entre el número de puntajes; normalmente se denomina media.

Media ponderada Media de una colección de puntajes a los que se han asignado diferentes grados de importancia.

Mediana Valor que está a la mitad de un conjunto de puntajes 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 comprobación de hipótesis Método para probar hipótesis, que se basa en una comparación del estadístico de prueba con los valores críticos.

Método tradicional de comprobación de hipótesis Método de comprobación de hipótesis que se basa en una comparación del estadístico de prueba y los valores críticos.

Moda Puntaje que ocurre con mayor frecuencia.

Modelo matemático Función matemática que se "ajusta" o 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 posibilidad de ser escogido.

Muestra aleatoria simple Muestra de cierto tamaño seleccionada de modo que toda posible muestra del mismo tamaño tenga la misma posibilidad de ser elegida.

Muestra autoseleccionada Muestra en la que los sujetos deciden por sí mismos ser incluidos; también se llama muestra de respuesta voluntaria.

Muestra de respuesta voluntaria Vea autoseleccionada.

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 Relación entre dos muestras, de modo que cada valor de una muestra está apareado con un valor correspondiente de la otra muestra.

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 son asequibles.

Muestreo estratificado Muestreo en el que se sacan muestras de cada estrato (clase).

Muestreo por racimos Tipo de muestreo en el que se divide el área de población en secciones (o racimos) y luego se seleccionan en forma aleatoria algunas de esas secciones; después se eligen *todos* los miembros de las secciones escogidas.

Muestreo sistemático Muestreo en el que se selecciona cada k-ésimo elemento.

Multimodal Que 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, rótulos 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.

Pendiente Medida de la inclinación de una línea recta.

Percentil Los 99 valores que dividen datos de orden en 100 grupos, con aproximadamente el 1% de los puntajes 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 de la probabilidad de que un suceso ocurra a que no ocurra; suele expresarse como la proporción de dos enteros sin factores comunes.

Posibilidades en contra Razón de la probabilidad de que un suceso no ocurra a 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 del suceso.

Posibilidades reales en contra El cociente $P(\overline{A})/P(A)$, generalmente expresado en la forma $a{:}b$ (“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 de ganar Razón de la ganancia neta (si se gana), en relación con lo apostado.

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 todas 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 una colección 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 de muestra, 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 afirmación 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 contingencia, 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 rachas Método no paramétrico que sirve para detectar aleatoriedad.

Prueba de rangos con signo de Wilcoxon Prueba de hipótesis no paramétrica que se utiliza para comparar dos muestras dependientes.

Prueba de significancia *Vea* prueba de hipótesis.

Prueba de signo Prueba de hipótesis no paramétrica que sirve para comparar muestras de dos poblaciones.

Prueba de suma de rangos ordenados de Wilcoxon Prueba de hipótesis no paramétrica que se utiliza para comparar dos muestras independientes.

Prueba *H* *Vea* prueba 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. *Vea* 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 abajo de la media; también se llama puntuación z.

Puntuación *z* Número de desviaciones estándar que un valor dado está por arriba o por abajo de la media.

Racha Secuencia de datos que presentan la misma característica; se usan 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 de un conjunto de muestra 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 tienen significado 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 (poco común) Si bajo un supuesto dado, la probabilidad de un resultado específico observado es en extremo pequeña, se concluye que posiblemente el supuesto no sea correcto.

Regla empírica Regla que usa la desviación estándar para proporcionar información sobre datos que tienen una distribución normal (sección 2-5).

Regla factorial Regla que afirma que *n* cosas distintas se pueden acomodar de *n*! maneras distintas.

Regla práctica del rango Regla que dice que el rango de un conjunto de datos abarca aproximadamente cuatro desviaciones estándar (4*s*).

Regresión múltiple Estudio de relaciones lineales entre tres o más variables.

Regresión por pasos Proceso de usar 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 cifras Puntaje mínimo, puntaje máximo, mediana, y el primer y tercer cuartiles de un conjunto de datos.

SC (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 (total) Medida de la variación total (alrededor de $\overline{\overline{x}}$) en todos los datos muestrales combinados; se usa en el análisis de varianza.

SC (tratamiento) Medida de la variación entre las medias muestrales; se usa en el análisis de varianza.

Selección aleatoria Selección de elementos muestrales de modo que todos los elementos disponibles para ser seleccionados tienen la misma posibilidad 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 trazando una línea vertical por la mitad.

Simulación Proceso que se comporta de forma similar a algún experimento, de modo 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 independientes Sucesos para los cuales la ocurrencia de cualquiera de los sucesos no afecta las probabilidades de ocurrencia de los demás.

Sucesos mutuamente excluyentes Sucesos que no pueden ocurrir simultáneamente.

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 bidireccional.

Tabla de dos factores *Vea* 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 proporcionar información acerca de la distribución de los datos.

Teorema del límite central Teorema que afirma que las medias muestrales tienden a estar distribuidas normalmente, con una media μ y una desviación estándar $\sigma/\sqrt{n}$.

Tratamiento Propiedad o característica que permite distinguir entre las diferentes poblaciones entre sí; 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 *Vea* 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 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 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 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 ella misma en el estudio.

Variables predictoras 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 producir todos los bienes o servicios exactamente de la misma forma todo el tiempo.

Variación asignable Tipo de variación en un proceso, que es el resultado de causas que pueden identificarse.

Variación debida al error *Vea* variación dentro de las muestras.

Variación debida al tratamiento *Vea* varianza entre muestras.

Variación dentro de las muestras En 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 E: Bibliografía

***Un asterisco indica un libro recomendado para leer. Otros libros sólo se recomiendan como referencias.**

Andrews D. y A. Herzberg. 1985. *Data: A Collection of Problems from Many Fields for the Student and Research Worker.* Nueva York: Springer.

Bell, M. 2003. *The TI-83 Plus Companion to Accompany Elementary Statistics.* 9a. ed. Reading, Mass.: Addison-Wesley.

Bennett, D. 1998. *Randomness.* Cambridge: Harvard University Press.

*Best, J. 2001. *Damned Lies and Statistics.* Berkeley: University of California Press.

Beyer, W. 1991. *CRC Standard Probability and Statistics Tables and Formulae.* Boca Raton, Fla.: CRC Press.

*Campbell, S. 1974. *Flaws and Fallacies in Statistical Thinking.* Englewood Cliffs, N.J.: Prentice-Hall.

*Crossen, C. 1994. *Tainted Truth: The Manipulation of Fact in America.* Nueva York: Simon & Schuster.

Devore, J. y R. Peck. 1997. *Statistics: The Exploration and Analysis of Data.* 3a. ed. y St. Paul, Minn.: West Publishing.

*Fairley, W. y F. Mosteller. 1977. *Statistics and Public Policy.* Reading, Mass.: Addison-Wesley.

Fisher, R. 1966. *The Design of Experiments.* 8a. ed. Nueva York: Hafner.

*Freedman, D., R. Pisani, R. Purves y A. Adhikari. 1991. *Statistics.* 2a. ed. Nueva York: Norton.

*Gonick, L. y W. Smith. 1993. *The Cartoon Guide to Statistics.* Nueva York: HarperCollins.

Halsey, J. y E. Reda. 2003. *Excel Student Laboratory Manual and Workbook.* Reading, Mass.: Addison-Wesley.

*Heyde, C. y E. Seneta(eds.). 2001. *Statisticians of the Centuries.* Nueva York: Springer-Verlag.

Hoaglin, D., F. Mosteller y J. Tukey, eds. 1983. *Understanding Robust and Exploratory Data Analysis.* Nueva York: Wiley.

*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.

*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. 2003. *Student Solutions Manual to Accompany Elementary Statistics.* 9a. ed. Reading, Mass.: Addison-Wesley.

*Moore, D. 1997. *Statistics: Concepts and Controversies.* 4a. ed. San Francisco: Freeman.

Mosteller, F., R. Rourke y G. Thomas, Jr. 1970. *Probability with Statistical Applications.* 2a. ed. Reading, Mass.: Addison-Wesley.

Ott, L. y W. Mendenhall. 1994. *Understanding Statistics.* 6a. ed. Boston: Duxbury Press.

Owen, D. 1962. *Handbook of Statistical Tables.* Reading, Mass.: Addison-Wesley.

*Paulos, J. 1988. *Innumeracy: Mathematical Illiteracy and Its Consequences.* Nueva York: Hill and Wang.

Peck, R. 2003. *SPSS Student Laboratory Manual and Workbook.* Reading, Mass.: Addison-Wesley.

*Reichard, R. 1974. *The Figure Finaglers.* Nueva York: McGraw-Hill.

*Reichmann, W. 1962. *Use and Abuse of Statistics.* Nueva York: Oxford University Press.

*Rossman, A. 1996. *Workshop Statistics: Discovery with Data.* Nueva York: Springer.

Ryan, T., B. Joiner y B. Ryan. 1995. *MINITAB Handbook.* 3a. ed. Boston: Duxbury.

*Salsburg, D. 2000. *The Lady Tasting Tea: How Statistics Revolutionized the Twentieth Century.* Nueva York: W. H. Freeman.

Schaeffer, R., M. Gnanadesikan, A. Watkins y J. Witmer. 1996. *Activity-Based Statistics: Student Guide.* Nueva York: Springer.

Schmid, C. 1983. *Statistical Graphics.* Nueva York: Wiley.

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.

Smith, G. 1995. *Statistical Process Control and Quality Improvement.* 2a. ed. Columbus: Merrill.

*Stigler, S. 1986. *The History of Statistics.* Cambridge, Mass.: Harvard University Press.

*Tanur, J.(ed.). 1989. *Statistics: A Guide to the Unknown.* 3a. ed. Belmont, Calif.: Wadsworth.

Triola, M. 2003. *Minitab Student Laboratory Manual and Workbook.* 9a. ed. Reading, Mass.: Addison-Wesley.

Triola, M. 2003. *STATDISK 9.0 Student Laboratory Manual and Workbook.* 9a. ed. Reading, Mass.: Addison-Wesley.

Triola, M. y L. Franklin. 1994. *Business Statistics.* Reading, Mass.: Addison-Wesley.

*Tufte, E. 1983. *The Visual Display of Quantitative Information.* Cheshire, Conn.: Graphics Press.

Tukey, J. 1977. *Exploratory Data Analysis.* Reading, Mass.: Addison-Wesley.

Utts, J. 1996. *Seeing Through Statistics.* Belmont, Calif.: Wadsworth.

Apéndice F: Soluciones de los ejercicios impares (y de todos los ejercicios de repaso y de los ejercicios de repaso acumulativo)

Sección 1-2

1. Parámetro.
3. Estadístico.
5. Continuo.
7. Discreto.
9. De razón.
11. De intervalo.
13. Ordinal.
15. De razón.
17. Muestra: los 10 adultos seleccionados; población, todos adultos; no es representativa.
19. Muestra: los 1059 adultos seleccionados; población: todos los adultos; representativa.
21. Sin punto de partida natural, las temperaturas están a un nivel de medición de intervalo; razones tales como "dos veces" carecen de significado.
23. Ordinal o de intervalo son respuestas aceptables, aunque ordinal es más sensato, porque las diferencias entre los valores no tienden a ser significativas. Por ejemplo, la diferencia entre un alimento con calificación de 1 y un alimento con calificación de 2 no es la misma que existe entre un alimento con calificación de 9 y otro con calificación de 10.

Sección 1-3

1. A los conductores de camiones con frecuencia las condiciones les obligan a comer en restaurantes de comida rápida, por lo que tienen dietas con contenidos más altos de grasa. Probablemente sea la dieta de comida rápida lo que causa un mayor peso y no los camiones por sí mismos. Evite hablar de causalidad y mejor diga que conducir camiones es una actividad que se asocia con un mayor peso.
3. Una posible alternativa: Sí hay discriminación racial, de modo que la policía del condado de Orange tiende a detener e infraccionar a más individuos de minorías que a personas blancas.
5. Ya que el estudio fue financiado por una compañía de dulces y la Chocolate Manufacturers Association, existe una posibilidad real de que tal hecho motivara a los investigadores, de alguna manera, a obtener resultados favorables para el consumo de chocolate.
7. No, ella utilizó una muestra de respuesta voluntaria.
9. A las personas sin teléfono o cuyos números no aparecen en el directorio se les excluyó.
11. Los motociclistas muertos.
13. No. A cada uno de los 29 cigarros se le da el mismo peso, pero algunos cigarros se consumen en mayores cantidades que otros. Además, hay cigarros a los que no se les incluyó en el conjunto de datos 5.
15. Los resultados no serían buenos ya que usted estaría muestreando únicamente a las personas que se sometieron a la dieta a una edad relativamente joven. Muchas personas que nacieron después de 1945 aún están vivas.
17. a) 68%
    b) 0.352
    c) 855
    d) 48.6%
19. a) 540
    b) 5%
21. El 62% del 8% de 1875 es únicamente 93.
23. Todos los porcentajes de éxitos deben ser múltiplos de cinco. Los porcentajes dados no pueden ser correctos.
25. La respuesta varía.

Sección 1-4

1. Experimento.
3. Estudio observacional.
5. Retrospectivo.
7. Transeccional.
9. De conveniencia.
11. Aleatorio.
13. Por racimos.
15. Sistemático.
17. Estratificado.
19. Por racimos.
21. Sí; sí.
23. No; no.
25. Sí; no.
27. Las respuestas varían.
29. No, no todos los votantes tienen la misma posibilidad de ser elegidos. Los votantes de los estados menos poblados tienen mayores posibilidades de ser seleccionados.
31. Pedir a los conductores que utilicen teléfonos celulares podría ponerlos en una situación peligrosa. La población de conductores que no tienen teléfonos celulares diferirían fundamentalmente de la población de conductores que poseen teléfonos celulares. La magnitud del uso del teléfono celular variaría considerablemente, de manera que los efectos del uso de un teléfono celular no quedarían claros. Los usuarios de teléfonos celulares saben que forman parte del grupo de tratamiento y quizá se comportarían de forma diferente, además de tender a culpar al teléfono celular por los problemas al conducir o por los choques.

Capítulo 1 Ejercicios de repaso

1. No, puesto que se trata de una muestra de respuesta voluntaria podría no ser 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) El valor más grande, ya que representa a los accionistas que lograrían el control de la compañía.
   f) La muestra de respuesta voluntaria tiende a sesgarse.

5. a) Sistemático; representativo.
   b) De conveniencia; no representativo.
   c) Por racimos; no representativo.
   d) Aleatorio; representativo.
   e) Estratificado; no representativo.
6. a) Diseñe el experimento de manera que los sujetos no sepan sí están utilizando Sleepeze o un placebo; también diséñelo de forma que quienes observan y evalúan a los sujetos no sepan cuáles sujetos están utilizando Sleepeze y cuáles un placebo.
   b) El estudio ciego ayudará a distinguir entre la eficacia del Sleepeze y el efecto placebo, ya que los sujetos y los evaluadores tienden a creer que la mejoría sucede sólo porque se está aplicando un tratamiento.
   c) Los sujetos se asignan a diferentes grupos a través de un proceso de *selección aleatoria*.
   d) Los sujetos se *eligen cuidadosamente* para los diferentes grupos, de manera que los grupos se conforman de forma similar en los aspectos que son importantes.
   e) La replicación se utiliza cuando el experimento se repite. Es importante tener una muestra de sujetos que sea lo suficientemente grande para conocer la verdadera naturaleza de cualquier efecto. Asimismo, lo es para no confundirnos con un comportamiento errático o con muestras que son muy pequeñas.

Capítulo 1 Ejercicios de repaso acumulativo

1. 163.85.
2. −0.64516129.
3. −6.6423420.
4. 216.09.
5. 4.3588989.
6. 18.647867.
7. 0.47667832.
8. 0.89735239.
9. 0.0000000000072744916.
10. 4,389,046,500,000.
11. 282,429,540,000.
12. 0.000000000058207661.

Capítulo 2 Respuestas

Sección 2-2

1. Anchura de clase: 10. Marcas de clase: 94.5, 104.5, 114.5, 124.5, 134.5, 144.5, 154.5.
   Fronteras de clase: 89.5, 99.5, 109.5, 119.5, 129.5, 139.5, 149.5, 159.5.
3. Anchura de clase: 200. Marcas de clase: 99.5, 299.5, 499.5, 699.5, 899.5, 1099.5, 1299.5.
   Fronteras de clase: −0.5, 199.5, 399.5, 599.5, 799.5, 999.5, 1199.5, 1399.5.

5.

Presión sanguínea sistólica de hombres	Frecuencia relativa
90–99	2.5%
100–109	10.0%
110–119	42.5%
120–129	30.0%
130–139	12.5%
140–149	0.0%
150–159	2.5%

7.

Colesterol de hombres	Frecuencia relativa
0–199	32.5%
200–399	27.5%
400–599	12.5%
600–799	20.0%
800–999	5.0%
1000–1199	0.0%
1200–1399	2.5%

9.

Presión sanguínea sistólica de hombres	Frecuencia acumulativa
menor que 100	1
menor que 110	5
menor que 120	22
menor que 130	34
menor que 140	39
menor que 150	39
menor que 160	40

11.

Colesterol de hombres	Frecuencia acumulativa
menor que 200	13
menor que 400	24
menor que 600	29
menor que 800	37
menor que 1000	39
menor que 1200	39
menor que 1400	40

13. Cambie el encabezado de "frecuencia" por "frecuencia relativa" e ingrese las siguientes frecuencias relativas: 13.5%, 15.5%, 21.0%, 20.0%, 14.0% y 16.0%. Las frecuencias relativas parecen variar de alguna manera. (Con el uso de métodos que se describen posteriormente en el libro, las diferencias no son significativas).

15.

Peso (b)	Frecuencia
0–49	6
50–99	10
100–149	10
150–199	7
200–249	8
250–299	2
300–349	4
350–399	3
400–449	3
450–499	0
500–549	1

17. Las circunferencias de las mujeres parecen ser ligeramente inferiores, pero la diferencia no parece ser significativa.

Circunferencia (cm)	Hombres	Mujeres
34.0–35.9	2	1
36.0–37.9	0	3
38.0–39.9	5	14
40.0–41.9	29	27
42.0–43.9	14	5

19. Las corredoras mujeres parecen ser algunos años menores.

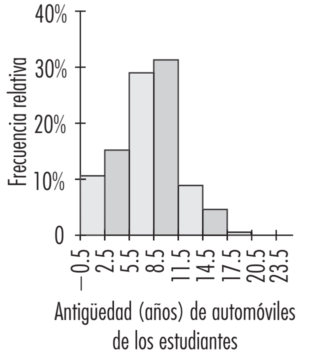
Edad	Hombre	Mujer
19–28	9.9%	20.5%
29–38	38.7%	46.2%
39–48	27.9%	10.3%
49–58	19.8%	17.9%
59–68	3.6%	5.1%

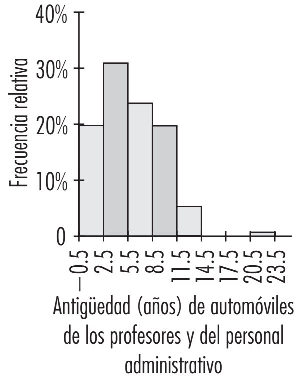
21. Un dato distante puede afectar en forma drástica la tabla de frecuencias.

Peso (lb)	Con dato distante	Sin dato distante
200–219	6	6
220–239	5	5
240–259	12	12
260–279	36	36
280–299	87	87
300–319	28	28
320–339	0	
340–359	0	
360–379	0	
380–399	0	
400–419	0	
420–439	0	
440–459	0	
460–479	0	
480–499	0	
500–519	1	

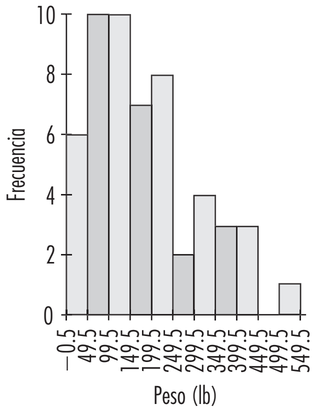
Sección 2-3

1. 26 años.
3. 71%.
5. 40%; 200.
7. La distribución de los automóviles de los profesores y el personal administrativo se carga ligeramente más hacia la izquierda, de manera que sus automóviles son un poco más nuevos.

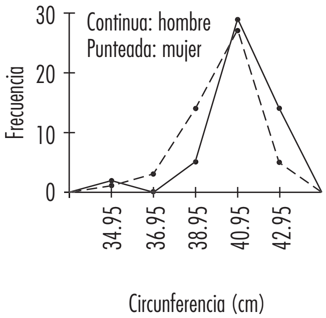




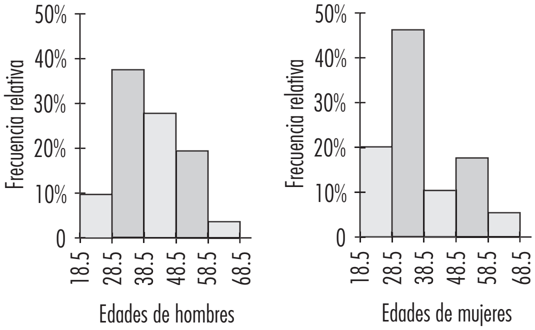
9. 183 libras



11. No parece haber una diferencia significativa.

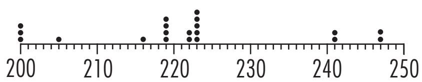


13. Las edades de los hombres parecen tener una distribución que se carga más hacia la derecha, de manera que tienden a ser ligeramente mayores.



15. 200, 200, 200, 205, 216, 219, 219, 219, 219, 222, 222, 223, 223, 223, 223, 223, 241, 241, 247, 247.

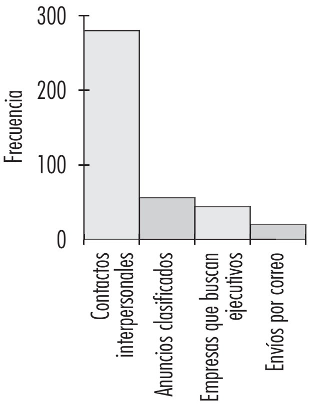
17.



19.

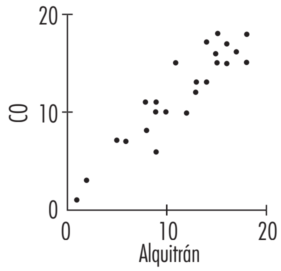
3	67
4	00134
4	667889
5	023334
5	788999
6	0111223334444
6	557789
7	01222234
7	57

21. Los contactos interpersonales parecen ser la forma más eficaz de obtener un empleo.

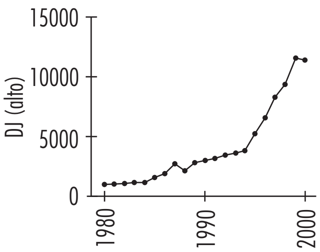


23.

25. Conforme aumenta la cantidad de alquitrán de cigarros, la cantidad de monóxido de carbono también se incrementa.



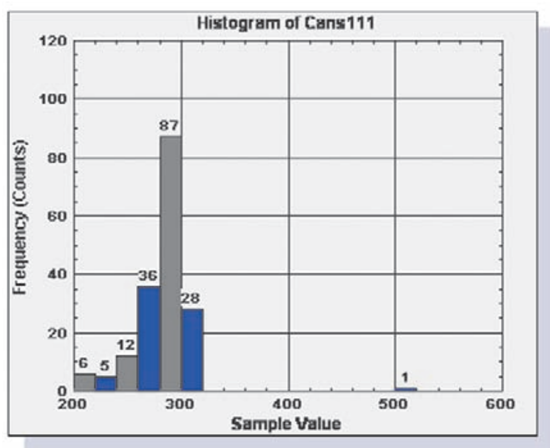
27. Parece haber una tendencia creciente, lo que sugiere que el mercado bursátil es una buena inversión.



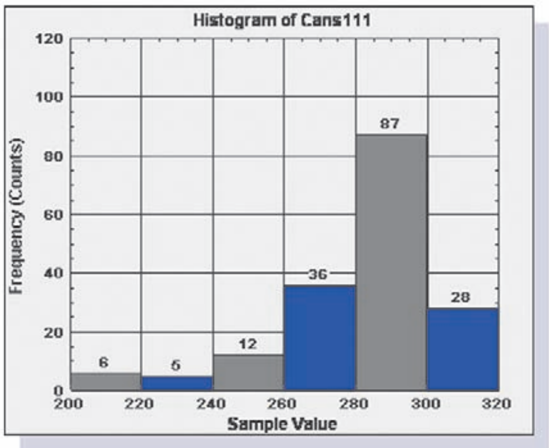
29. 10,000/422,000: 2.4%.

31. 13,000 (desde 37,000 hasta 24,000).

33. a)



b)



c) El dato distante llega a tener un efecto importante en el histograma. El uso de una anchura de clase mayor que la utilizada en los incisos *a*) y *b*) escondería la verdadera naturaleza de la distribución.

Sección 2-4

1. $\bar{x} = 157.8$ seg; mediana $= 88.0$ seg; moda $= 0$ seg; mitad del intervalo $= 274.0$ seg.
   Sí, los niños no deben ser influidos por la exposición al consumo de tabaco.
3. $\bar{x} = 0.295$ g; mediana $= 0.345$ g; moda: 0.13 g, 0.43 g, 0.47g; mitad del intervalo $= 0.255$ g.
   No necesariamente. Hay otros cereales que no se incluyen y los estadounidenses consumirían mucho más de algunas otras marcas.
5. $\bar{x} = 0.187$ g; mediana $= 0.170$; moda: 0.16, 0.17; mitad del intervalo $=$ 0.205.
   Sí.
7. $\bar{x} = 18.3$; mediana $= 18.0$; moda $= 17$; mitad del intervalo $= 18.0$.
   Los resultados son muy consistentes; por lo tanto, la media debe ser un buen estimado.
9. Jefferson Valley: $\bar{x} = 7.15$ min; mediana $= 7.20$ min; moda $= 7.7$ min; mitad del intervalo $= 7.10$ min.
   Providence: los mismos resultados que el Jefferson Valley.
   Aunque las medidas de tendencia central son las mismas, los tiempos del Providence varían mucho más que los tiempos del Jefferson Valley.
11. McDonald´s: $\bar{x} = 186.3$ seg; mediana $= 184.0$ seg; moda $=$ ninguna; mitad del intervalo $= 189.5$ seg.
    Jack in the Box: $\bar{x} = 262.5$ seg; mediana $= 262.5$ seg; moda $=$ 109 seg; mitad del intervalo $= 277.5$ seg.
    McDonald´s parece ser significativamente más rápido.
13. Hombres: $\bar{x} = 41.10$ cm; mediana $= 41.10$ cm.
    Mujeres: $\bar{x} = 40.05$ cm; mediana $= 40.20$ cm.
    Sí parece haber una pequeña diferencia.
15. Jueves: $\bar{x} = 0.069$ pulgadas; mediana $= 0.000$ pulgadas.
    Domingo: $\bar{x} = 0.068$ pulgadas; mediana $= 0.000$ pulgadas.
    No parece haber una diferencia importante.
17. 74.4 min.
19. 46.8 mi/h; la media es significativamente más alta que el límite de 30 mi/h.
21. a) 182.9 lb.
    b) 171.0 lb.
    c) 159.2 lb.
    Los resultados difieren en grandes cantidades, lo que sugiere que la media del conjunto original de pesos es afectada fuertemente por valores extremos.
23. a) 52.
    b) $n - 1$.
25. 84.5.
27. 48.0 mi/h.
29. 62.9 volts.

Sección 2-5

1. Rango $= 548.0$ seg; $s^2 = 46308.2$ seg^2; $s = 215.2$ seg; varían ampliamente.
3. Rango $= 0.450$ g; $s^2 = 0.028$ g^2; $s = 0.168$ g.
5. Rango $= 0.170$; $s^2 = 0.003$ g^2; $s = 0.051$.
   No, la intención es disminuir todos los valores individuales, lo cual daría como resultado una media menor.
7. Rango $= 6.0$; $s^2 = 2.5$; $s = 1.6$.
   Las medidas de variación son valores bajos.
9. Jefferson Valley: rango $= 1.20$ min; $s^2 = 0.23$ min^2; $s = 0.48$ min.
   Providence: rango $= 5.80$ min; $s^2 = 3.32$ min^2; $s = 1.82$ min.
11. McDonald´s: rango $= 195.0$ seg; $s^2 = 4081.7$ seg^2; $s = 63.9$ seg.
    Jack in the Box: rango $= 407.0$ seg; $s^2 = 16644.3$ seg^2; $s = 129.0$ seg.
13. Hombres: 1.50 cm; mujeres: 1.64 cm; la diferencia no parece ser muy grande.
15. Jueves: 0.167 pulgadas; domingo: 0.200 pulgadas.
17. 14.7 min.
19. 4.1 mi/h.
21. Aproximadamente 12 años (con base en un mínimo de 23 años y un máximo de 70 años).
23. Mínimo: 31.30 cm; máximo: 46.42 cm; sí.
25. a) 68%.
    b) 99.7%.
27. El porcentaje es de al menos el 75%.
29. Calorías: 5.9%; azúcar 56.9%. El contenido de azúcar tiene una variación mucho mayor cuando se compara con las calorías.
31. Todos los valores son iguales.
33. La baterías Everlast son mejores, porque son más consistentes y predecibles.
35. Sección 1: rango $= 19.0$; $s = 5.7$.
    Sección 2: rango $= 17.0$; $s = 6.7$.
    Los rangos sugieren que la sección 2 tiene menor variación, pero las desviaciones estándar sugieren que la sección 1 tiene menor variación.
37. 1.44.
39. 15.8.

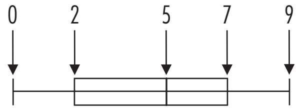
41. a) 6.
    b) 6.
    c) 3.0.
    d) $n - 1$.
    e) No, la media de las varianzas muestrales (6) es igual a la varianza poblacional (6), pero la media de las desviaciones estándar poblacionales (1.9) no es igual a la media de la desviación estándar poblacional (2.4).

Sección 2-6

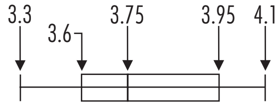
1. a) 60.
   b) 3.75.
   c) 3.75.
   d) Poco común.
3. a) –3.21.
   b) 5.71.
   c) 0.26.
5. 2.56; poco común.
7. 4.52; sí; el paciente está enfermo.
9. La prueba psicológica, ya que $z = -0.50$ es mayor que $z = -2.00$.
11. –3.56; sí.
13. 43.
15. 15.
17. 46.
19. 251.5.
21. 121.
23. 0.
25. 25.
27. 65.
29. 415.5.
31. 117.5.
33. 98.
35. 254.
37. La puntuación z permanece igual.
39. a) Uniforme.
    b) Con forma de campana.
    c) La forma de la distribución permanece igual.
41. a) 165.
    b) 169.
    c) 279.5.
    d) Sí; sí.
    e) No; no.
43. a) P_{10}, P_{50}, P_{80}.
    b) 10, 46, 107.5, 130.5, 170, 209, 239.5, 265.5, 289.5.
    c) 46, 130.5, 209, 265.5.

Sección 2-7

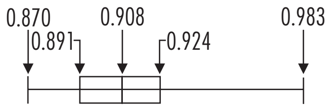
1. 0, 2, 5, 7, 9. Las separaciones en la gráfica de cuadro son aproximadamente las mismas, lo que indica que los valores son casi probables.



3. 3.3, 3.6, 3.75, 3.95, 4.1. No, el consumo de cereal no se distribuye de manera uniforme entre las marcas; por lo tanto, se deben utilizar valores ponderados.

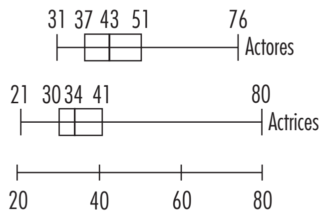


5. 0.870, 0.891, 0.908, 0.924, 0.983; sí

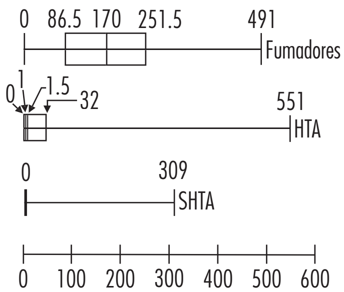


7. 0, 0, 1.5, 39, 414. Sesgada.

9. Actores: 31, 37, 43, 51, 76.
   Actrices: 21, 30, 34, 41, 80.
   Las actrices parecen ser más jóvenes.



11. Fumadores: 0, 86.5, 170, 251.5, 491. HTA: 0, 1, 1.5, 32, 551. SHTA: 0, 0, 0, 0, 309. Las diferencias son significativas e indican incrementos de cotinina ante la exposición al consumo del tabaco.



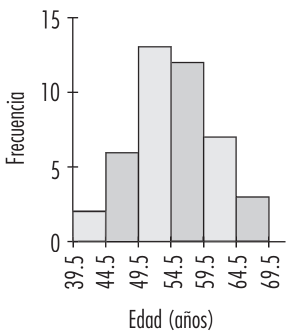
13. RIC = 165.
    Datos ligeramente distantes: valores x tales que $-408.5 \leq x < -161$ o $499 < x \leq 746.5$.
    Datos extremadamente distantes: valores x tales que $x < -408.5$ o $x > 746.5$.
    No hay datos ligeramente distantes o datos extremadamente distantes.

Capítulo 2 Ejercicios de repaso

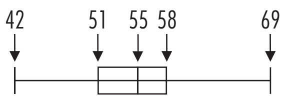
1. a) 54.8 años.
   b) 55.0 años.
   c) 51 años, 54 años.
   d) 55.5 años.
   e) 27.0 años.
   f) 6.2 años.
   g) 38.7 años^2.
   h) 51 años.
   i) 58 años.
   j) 47 años.
2. a) –1.90.
   b) No, porque la puntuación *z* está a dos desviaciones estándar de la media.
   c) 42, 68, 69.
   d) Sí; sí.
3.

Edad	Frecuencia
40–44	2
45–49	6
50–54	13
55–59	12
60–64	7
65–69	3

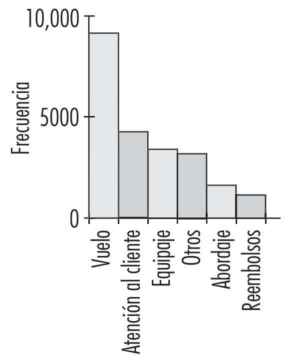
4. Con forma de campana.



5. 42, 51, 55, 58, 69



6. a) El porcentaje es 68%.
   b) El porcentaje es 95%.
7. La puntuación de 19 es mejor, porque $z = -0.20$ es mayor que $z = -0.67$.
8. a) Las respuestas varían, pero 7 u 8 años es razonable.
   b) 5 años (con base en un mínimo de 0 años y un máximo de 20 años).
9. a) 140 min.
   b) 15 min.
   c) 225 min^2.
10.



Capítulo 2 Ejercicios de repaso acumulativo

1. a) $\bar{x} = 20.5$ seg; mediana = 27.0 seg; moda = 20 seg; mitad del intervalo = 42.0 seg.
   b) $s = 142.2$ seg; $s^2 = 20216.4$ seg^2; rango = 566.0 seg.
   c) Los tiempos exactos originales son continuos, pero los datos parece que se redondearon a valores discretos.
   d) De razón.
2. a) La moda, porque las otras medidas de tendencia central requieren cálculos que no es posible (o no es conveniente) hacer con datos a un nivel de medición nominal.
   b) De conveniencia.
   c) Por racimos.
   d) Desviaciones estándar; disminuirse.
3. No, los 50 valores deben ponderarse, utilizando las poblaciones estatales como pesos.

Capítulo 3 Respuestas

Sección 3-2

1. a) 0.5.
   b) 0.20.
   c) 0.
3. -1, 2, 5/3, $\sqrt{2}$.
5. a) 3/8.
   b) 3/8.
   c) 1/8.
7. 0.153; sí.
9. a) 1/17 o 0.0588.
   b) No.
11. a) 0.0501.
    b) No.
13. a) 0.0154 (no 0.0156).
    b) Sí.
15. a) 1/365.
    b) Sí.
    c) Él ya lo sabía.
    d) 0.

17. a) 1/365.
    b) 31/365.
    c) 1.
19. 0.130.
21. a) 77/500 o 0.154.
    b) 13/500 o 0.026.
23. a) niño niño, niño niña, niña niño, niña niña.
    b) 1/4.
    c) 1/2.
25. a) $21.
    b) 21:2.
    c) 14:1.
    d) $30.
27. Como la probabilidad de mostrar mejoría con un fármaco ineficaz es tan baja (0.04), parece que el fármaco es eficaz.
29. 5/8.
31. a) 4/1461.
    b) 400/146,097.
33. 1/4.

Sección 3-3

1. a) No.
   b) No.
   c) Sí.
3. a) 0.95.
   b) 0.782.
5. 5/7 o 0.714.
7. 364/365 o 0.997.
9. 0.239.
11. 0.341.
13. 0.600.
15. 0.490.
17. 0.140.
19. 0.870.
21. 0.5.
23. 0.290.
25. a) Son mutuamente excluyentes.
    b) No son mutuamente excluyentes.
27. $P(A \text{ o } B) = P(A) + P(B) - 2P(A \text{ y } B)$

Sección 3-4

1. a) Independiente.
   b) Independiente.
   c) Dependiente.
3. 1/12.
5. a) 9/49.
   b) 1/7.
7. a) 0.288.
   b) 0.288.
   c) Aunque los resultados son ligeramente diferentes, son iguales cuando se redondea a tres decimales.
   d) Muestreo sin reemplazo para evitar la duplicación.
9. a) 1/1024.
   b) No, porque existen otras formas de pasar.
11. a) 1/133225 o 0.00000751.
    b) 1/365.
13. 0.694.
15. 1/1024; sí, porque la probabilidad de obtener 10 niñas por azar es muy baja.
17. 1/64.
19. 0.739 (o 0.738, si se asume dependencia); no.
21. 0.702.
23. 0.736.
25. a) 0.992.
    b) 0.973.
    c) 0.431.
27. 0.0192.

Sección 3-5

1. Ninguno de los estudiantes es del grupo sanguíneo A.
3. Al menos una de la devolución de impuestos es correcta.
5. 0.97; no.
7. 31/32; sí.
9. 0.410.
11. 0.5; no.
13. 11/14; realice otra prueba.
15. 0.999999; sí, porque la posibilidad de estar despierto se incrementa desde 0.99 hasta 0.999999.
17. 0.271.
19. 0.897.
21. 0.0793.
23. 1/12; 35.
25. a)

	Positivo	Negativo
Infectado con VIH	285	15
No infectado con VIH	4985	94,715

    b) 0.0541.
27. 1/3.

Sección 3-6

1. V, V, F, F, V.
3. Bueno, bueno, defectuoso, bueno, bueno.
5. Con número impar = niña: 17/20 o 0.85. El resultado se encuentra razonablemente cerca de 0.813.
7. De los 20 renglones, hay al menos un 0 en siete renglones, de manera que la probabilidad que se estima es 7/20 o 0.35, que está razonablemente cerca del resultado correcto de 0.410.

9. Aproximadamente 0.813.
11. Aproximadamente 0.410.
13. Debe cambiar: P(ganar) = 2/3; no cambiar: P(ganar) = 1/3.
15. No; no.

Sección 3-7

1. 720.
3. 600.
5. 300.
7. 2,598,960.
9. 1/13,983,816.
11. 1/45,057,474.
13. 1/35,960, parece que se seleccionó a los empleados más grandes.
15. 1/3,776,965,920.
17. 1/5005; sí.
19. 4; 40,320.
21. 10.
23. 720; satire; 1/720.
25. 1/125,000.
27. a) 256.
    b) 70.
    c) 70/256 = 0.273.
29. 144.
31. 1/41,416,353.
33. 2,095,681,645,538 (más de 2 billones).
35. a) Calculadora: 3.0414093×10^{64}; aproximación: 3.0363452×10^{64}
    b) 615.

Capítulo 3 Ejercicios de repaso

1. 0.2.
2. 0.32.
3. 0.35.
4. 0.83.
5. 0.638.
6. 0.100.
7. 15/32 o 0.469.
8. 15/80 = 3/16 o 0.188.
9. a) 0.248.
   b) 0.0615.
   c) 0.575.
10. 0.0777.
11. 1/4096; sí.
12. a) 1/120.
    b) 720.
13. a) 9/19.
    b) 10:9.
    c) \$5.
14. 0.000000531; no.
15. 0.979.
16. a) 1/20,358,520.
    b) 1/142,506.
    c) 1/76,275,360.

Capítulo 3 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) 63.6 pulgadas.
   b) 1/4.
   c) 3/4.
   d) 1/16.
   e) 5/16.

Capítulo 4 Respuestas

Sección 4-2

1. a) Continua.
   b) Discreta.
   c) Continua.
   d) Discreta.
   e) Discreta.
3. $\mu = 1.5$, $\sigma = 0.9$.
5. No es una distribución de probabilidad, porque $\Sigma P(x) = 0.94 \neq 1$.
7. $\mu = 0.7$, $\sigma = 0.9$.
9. $\mu = 5.8$, $\sigma = 1.1$; no.
11. −7.07 ¢; 1.4 ¢.
13. a) Vive: −\$250 (una pérdida); muere: \$99,750 (una ganancia).
    b) −\$100.
    c) \$150.
    d) El valor negativo que se espera es un precio relativamente bajo para asegurar la tranquilidad financiera de sus herederos.
15. a) 10,000.
    b) 0.0001.
    c) \$2787.50.
    d) −22.12¢.
    e) Pick 4, porque −22.12 ¢ es mayor que −22.5¢.
17. a) 0.122.
    b) 0.212.
    c) El inciso b). La ocurrencia de nueve niñas entre 14 sería muy poco común si la probabilidad de nueve niñas o más es muy baja (tal como menor que 0.05).
    d) No, ya que la probabilidad de nueve niñas o más no es muy baja (0.212). El resultado de nueve niñas o más con facilidad sucedería por el azar.

19. a) 0.029.
    b) Sí, porque la probabilidad de 11 niñas o más es muy baja (0.029). El resultado de 11 niñas o más no sucedería con facilidad por el azar.
21. Como la probabilidad de adivinar correctamente ocho respuestas o más es de 0.395, dicho resultado ocurriría fácilmente; por lo tanto, no hay evidencia de que Bob tenga poderes especiales.
23. Los bonos A son mejores porque el valor que se espera es de $49.40, que es mayor que el valor que se espera de $26 para los bonos B. Ella debe seleccionar los bonos A porque el valor que se espera es positivo, lo que indica una posible ganancia.
25. $\mu = 0.6$, $\sigma = 0.6$.
27. a) 3.
    b) $\sqrt{2}$.
    c) $\mu = 10.5$, $\sigma = 5.8$.

Sección 4-3

1. No es binomial; más de dos resultados; no hay un número fijo de ensayos.
3. No es binomial: más de dos resultados.
5. Binomial.
7. No es binomial; más de dos resultados.
9. a) 0.128.
   b) IIC, ICI, CII; 0.128 para cada una.
   c) 0.384.
11. 0.980.
13. 0.171.
15. 0+.
17. 0.278.
19. 0.208.
21. 0.4711; no.
23. 0.9925 (o 0.9924); sí.
25. 0.0833.
27. a) 0+ (o 0.00000980).
    b) 0+ (o 0.00000985).
    c) Probablemente están siendo blanco de las auditorías.
29. 0.0874; no.
31. a) 0.107.
    b) 0.893.
    c) 0.375 (o 0.376).
    d) No, porque con una tasa del 20% la probabilidad de al menos uno es alta (es mayor que 0.05).
33. 0.000201; sí.
35. P(nueve niñas o más) = 0.073, de modo que nueve niñas ocurrirían fácilmente por el azar. No hay evidencia suficiente para concluir que la técnica de selección del género sea eficaz.
37. 0.0524.
39. 0.000535.

Sección 4-4

1. $\mu = 80.0$, $\sigma = 8.0$, mínimo = 64.0, máximo = 96.0.
3. $\mu = 1488.0$, $\sigma = 19.3$, mínimo = 1449.4, máximo = 1526.6.
5. a) $\mu = 5.0$, $\sigma = 1.6$.
   b) No, porque 7 está dentro de dos desviaciones estándar de la media.
7. a) $\mu = 2.6$, $\sigma = 1.6$.
   b) No, porque 0 triunfos están dentro de dos desviaciones estándar de la media.
9. a) Las probabilidades de 0, 1, 2, 3, . . . , 15 son 0+, 0+, 0.003, 0.014, . . . , 0+ (de la tabla A-1).
   b) $\mu = 7.5$, $\sigma = 1.9$.
   c) No, porque 10 está dentro de dos desviaciones estándar de la media. Además, P(10 niñas o más) = 0.151, lo que demuestra que es fácil obtener 10 o más niñas por el azar.
11. a) $\mu = 27.2$, $\sigma = 5.1$.
    b) Sí, parece que el programa de entrenamiento tuvo efecto.
13. a) $\mu = 142.8$, $\sigma = 11.9$.
    b) No, 135 no es poco común, porque está dentro de dos desviaciones estándar de la media.
    c) Con base en los resultados, los teléfonos celulares no constituyen un riesgo para la salud que incremente la posibilidad de tener cáncer cerebral o del sistema nervioso.
15. a) 901.
    b) $\mu = 506.0$, $\sigma = 15.9$.
    c) Sí, porque 901 está a más de dos desviaciones estándar por arriba de la media.
17. a) Sí (con base en el histograma de probabilidad).
    b) La probabilidad es de 0.95.
    c) La probabilidad es de 0.997.
    d) Al menos 75% de dichos grupos de 100 tendrán entre 40 y 60 niñas.

Sección 4-5

1. 0.180.
3. 0.0399.
5. a) 62.2.
   b) 0.0155 (0.0156 empleando una media redondeada).
7. a) 0.497.
   b) 0.348.
   c) 0.122.
   d) 0.0284.
   e) 0.00497.
   Las frecuencias que se esperan de 139, 97, 34, 8 y 1.4 se comparan razonablemente bien con las frecuencias reales, de manera que la distribución de Poisson proporciona buenos resultados.
9. a) 0.00518 (si se utiliza la binomial: 0.00483).
   b) 0.995.
   c) 0.570.
   d) 0.430.
11. 4.82×10^{-64} es tan pequeño que, para propósitos prácticos, consideraremos que es cero.

Capítulo 4 Ejercicios de repaso

1. a) Una variable aleatoria es aquella que tiene un solo valor numérico (que se determina por el azar) para cada resultado de algún procedimiento.
   b) Una distribución de probabilidad da la probabilidad de cada valor de la variable aleatoria.
   c) Sí, porque cada valor de probabilidad está entre 0 y 1 y la suma de las probabilidades es 1.
   d) 4.2 días.
   e) 2.1 días.
   f) No, porque la probabilidad de 0.08 indica que es fácil obtener 0 días por el azar.
2. a) 3.0.
   b) 3.0.
   c) 1.6.
   d) 0.103.
   e) Sí, porque la probabilidad de 0 televisores es de 0.0388, lo que indica que es muy poco probable que ningún televisor esté sintonizando *West Wing*.
3. a) 0.026.
   b) 0.992 (o 0.994).
   c) $\mu = 8.0$, $\sigma = 1.3$.
   d) No, porque 6 está dentro de dos desviaciones estándar de la media.
4. a) 0.00361.
   b) Esta compañía parece ser muy diferente, porque el suceso de al menos cuatro despidos es muy poco probable, con una probabilidad de 0.00361.
5. a) 7/365.
   b) 0.981.
   c) 0.0188.
   d) 0.0002.
   e) No, porque el suceso es muy poco común.

Capítulo 4 Ejercicios de repaso acumulativo

1. a) $\bar{x} = 1.8$, $s = 2.6$
   b)

x	Frecuencias relativas
0	64.4%
1	4.1%
2	1.4%
3	0.0%
4	4.1%
5	15.1%
6	4.1%
7	4.1%
8	1.4%
9	1.4%

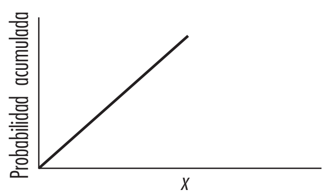
   c) $\mu = 4.5$, $\sigma = 2.9$
   d) La cantidad excesiva de ceros sugiere que las distancias se estimaron, no se midieron. Los dígitos parece que no se seleccionaron aleatoriamente.

2. a) 0.2.
   b) $\mu = 0.2$, $\sigma = 0.4$.
   c) 0.182.
   d) Sí, porque 3 está a más de dos desviaciones estándar por arriba de la media.
   e) 0.0001.

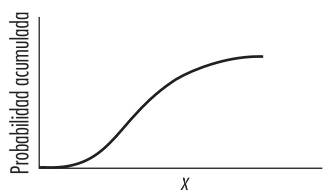
Capítulo 5 Respuestas

Sección 5-2

1. 0.15.
3. 0.15.
5. 1/3.
7. 1/2.
9. 0.4013.
11. 0.5987.
13. 0.0099.
15. 0.9901.
17. 0.2417.
19. 0.1359.
21. 0.8959.
23. 0.6984.
25. 0.0001.
27. 0.5.
29. 68.26%.
31. 99.74%.
33. 0.9500.
35. 0.9950.
37. 1.28.
39. -1.645.
41. a) 68.26%.
    b) 95%.
    c) 99.74%.
    d) 81.85%.
    e) 4.56%.
43. a) 1.23.
    b) 1.50.
    c) 1.52.
    d) -2.42.
    e) -0.13.
45. a)



    b)



Sección 5-3

1. 0.8413.
3. 0.4972.
5. 87.4.
7. 115.6.
9. a) 0.0001; sí.
   b) 99.2°.
11. a) 69.15%.
    b) 1049; si se selecciona el 40% de las mejores calificaciones del grupo de solicitantes, nadie sabría si lo aceptaron o rechazaron, sino hasta después de obtener las calificaciones de todos los solicitantes.
13. a) 0.0018.
    b) 5.6 años.
15. a) 25% de coincidencia.
    b) 0.8895; bastante.
17. 0.52%.
19. 0.1222; 12.22%; sí, todas están muy por arriba de la media.
21. a) Las puntuaciones *z* son números reales que carecen de unidad de medición.
    b) $\mu = 0$; $\sigma = 1$; la distribución es normal.
    c) $\mu = 64.9$ kg, $\sigma = 13.2$ kg, la distribución es normal.
23. a) 75; 5.
    b) No, la conversión también debería tomar en cuenta la variación.
    c) 31.4, 27.6, 22.4, 18.6.
    d) El inciso c), porque la variación se incluye en la conversión.
25. a) 1087; 22.9.
    b) 26.0.

Sección 5-4

1. No, por la variabilidad de muestreo, las proporciones muestrales variarán de forma natural de la proporción poblacional verdadera, incluso si el muestreo se hace con un procedimiento perfectamente válido.
3. No, el histograma representa la forma de la distribución de una muestra, pero la distribución de muestreo incluye a todas las muestras posibles del mismo tamaño, tales como todas las medias calculadas a partir de todas las muestras posibles de 106 personas.
5. 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)

Media	10.0	8.0	7.5	8.0	6.0	5.5	7.5	5.5	5.0
Probabilidad	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9

   c) 7.0.
   d) Sí; sí.
7. a) Medias: 85.0, 82.0, 83.5, 79.0, 81.5, 82.0, 79.0, 80.5, 76.0, 78.5, 83.5, 80.5, 82.0, 77.5, 80.0, 79.0, 76.0, 77.5, 73.0, 75.5, 81.5, 78.5, 80.0, 75.5, 78.0.
   b) La probabilidad de cada media es de 1/25. La distribución muestral consiste en las 25 medias muestrales que se aparearon con la probabilidad de 1/25.
   c) 79.4.
   d) Sí; sí.
9. a) 0, 0.5, 0.5, 0.5, 0.5, 1, 1, 1, 0.5, 1, 1, 1, 0.5, 1, 1, 1.
   b) La distribución muestral consiste en las 16 proporciones que se aparearon con la probabilidad de 1/16.
   c) 0.75.
   d) Sí; sí.
11. a) La respuesta varía.
    b) La respuesta varía, pero debe ser alguna de éstas: 0, 0.2, 0.4, 0.6, 0.8, 1.
    c) Un estadístico.
    d) No; no.
    e) Debe ser 10/13 o 0.769.
13. a) 59.4; 4.6.
    b) 59.4; 3.1.
    c) 59.4; 1.9.
    d) Sí. Cada distribución muestral tiene una media de 59.4, que es la media de la población.
    e) Conforme el tamaño de la muestra aumenta, la variación de la distribución muestral de las medias de muestras disminuye.
15. Medianas: 2.5; medias: 2.7. Las medias muestrales nuevamente coinciden con la media poblacional, pero las medianas no. La mediana no es un buen estadístico para estimar la media poblacional.

Sección 5-5

1. a) 0.4325.
   b) 0.1515.
3. a) 0.0677.
   b) 0.5055.
5. a) 0.9808.
   b) Si la población original tiene una distribución normal, el teorema del límite central proporciona buenos resultados para cualquier tamaño de muestra.
7. a) 0.5302.
   b) 0.7323.
   c) El inciso a), porque los asientos los ocuparán mujeres individuales, no grupos de mujeres.
9. a) 0.0119.
   b) No; sí.
11. a) 0.0001.
    b) No, pero a los consumidores no se les engaña, ya que las latas se llenan de más, no de menos.
13. a) 0.0051.
    b) Sí.
15. a) 0.1170.
    b) No, porque la probabilidad de 0.1170 indica que es fácil obtener una media tal como 0.882 g, suponiendo que no se cambian las cantidades de nicotina.
17. 0.0069; el nivel es aceptable.
19. 2979 lb.
21. a) 0.9750.
    b) 1329 lb.

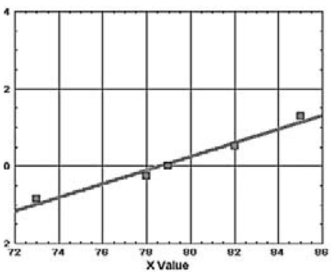
23. 0.0240. Concluiríamos que el generador de números aleatorios es defectuoso si obtuviéramos una media muestral que difiera de 0.500, de tal manera que haya una probabilidad muy baja de obtener una media muestral "al menos tan extrema" como el valor de la media muestral que ya se tiene. Con un tamaño muestral de 100, no hay una media muestral, entre 0.499 y 0.501, que cumpla ese criterio; por lo tanto, no debemos concluir que el generador de números aleatorios es defectuoso.

Sección 5-6

1. El área a la derecha de 15.5.
3. El área a la izquierda de 99.5.
5. El área a la izquierda de 4.5.
7. El área entre 7.5 y 10.5.
9. Tabla: 0.122; aproximación normal: 0.1218.
11. Tabla: 0.549; la aproximación normal no es posible.
13. 0.1357; no.
15. 0.0287; no.
17. 0.2676; no.
19. 0.7389; no, no es muy confiable.
21. 0.0708; sí.
23. 0.0080; sí.
25. 0.6368; es posible que el grupo sea suficiente, pero la probabilidad debe ser mucho más alta. Sería mejor incrementar el grupo de voluntarios.
27. 0.0026; sí.
29. 6; 0.4602.
31. a) 0.821.
    b) 0.9993.
    c) 0.0000165.
    d) 0.552.

Sección 5-7

1. No es normal.
3. No es normal.
5. No es normal.
7. Es normal.
9. No es normal.
11. Es normal.
13. Las estaturas parecen ser normales, aunque los niveles de colesterol no parecen ser normales. Los niveles de colesterol se afectan mucho por la dieta, por lo cual la dieta variaría en tantas formas que no produce resultados que se distribuyan normalmente.
15. −1.28, −0.52, 0, 0.52, 1.28; normal.



17. No; la transformación a puntuaciones *z* implica restar una constante y dividir entre una constante, de modo que la gráfica de los puntos (*x*, *z*) será siempre una línea recta, sin importar la naturaleza de la distribución.

Capítulo 5 Ejercicios de repaso

1. a) 0.0222.
   b) 0.2847.
   c) 0.6720.
   d) 254.6.
2. a) 0.69% de 900 = 6.21 bebés.
   b) 2405 g.
   c) 0.0119.
   d) 0.9553.
3. 0.1020; no; suponiendo que la tasa correcta sea del 25%, hay una alta probabilidad (0.1020) de que 19 descendientes o menos tengan ojos azules. Puesto que el suceso que se observa ocurriría fácilmente por el azar, no existe evidencia en contra de la tasa del 25%.
4. a) 0.9626.
   b) 63.3 pulgadas, 74.7 pulgadas.
   c) 0.9979.
5. a) 0.5.
   b) 1.
   c) 0.
   d) 0.25.
6. a) Distribución normal.
   b) 51.2 lb.
   c) Distribución normal.
7. Aproximación normal: 0.0436; valor exacto: 0.0355. Como la probabilidad de obtener sólo dos mujeres por el azar es tan baja, parece que la compañía está discriminando con base en el género.
8. Sí. El histograma se aproxima burdamente a una forma de campana y la gráfica cuantilar normal contiene puntos que se aproximan razonablemente al patrón de una línea recta. Además, no hay datos distantes.

Capítulo 5 Ejercicios de repaso acumulativo

1. a) 63.0 mm.
   b) 64.5 mm.
   c) 66 mm.
   d) 4.2 mm.
   e) −0.95.
   f) 75%.
   g) 82.89%.
   h) De razón.
   i) Continuo.
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 daría como resultado errores demasiado grandes.
   d) 5.0.
   e) 2.1.
   f) No; 8 está dentro de dos desviaciones estándar de la media y dentro del rango de valores que ocurrirían fácilmente por el azar.

Capítulo 6 Respuestas

Sección 6-2

1. 2.575.
3. 2.33.
5. $p = 0.250 \pm 0.030$.
7. $p = 0.654 \pm 0.050$.
9. $\hat{p} = 0.464$; $E = 0.020$.
11. $\hat{p} = 0.655$; $E = 0.023$.
13. 0.0300.
15. 0.0405.
17. $0.708 < p < 0.792$.
19. $0.0887 < p < 0.124$.
21. 461.
23. 232.
25. a) Se tiene una confianza del 95% de que los límites de 0.0489 y 0.0531 contienen la proporción poblacional.
    b) Sí, cerca del 5% de los hombres entre 18 y 20 años conduce en estado de ebriedad.
    c) 5.31%.
27. a) 29%.
    b) $25.4\% < p < 32.6\%$.
    c) 32.6%.
29. a) $22.6\% < p < 29.8\%$.
    b) No, los límites del intervalo de confianza incluyen el 25%.
31. a) $0.134\% < p < 6.20\%$ utilizando $x = 7$, $n = 221$.
    b) El ziac no parece causar mareo como reacción adversa.
33. 4145.
35. a) 473.
    b) 982.
    c) Puesto que se basan en una muestra de respuesta voluntaria, los resultados no se validarían.
37. a) $0.0355 < p < 0.139$.
    b) 373.
    c) Sí.
39. a) $0.0267\% < p < 0.0376\%$.
    b) No, porque el 0.0340% está dentro del intervalo de confianza.
41. a) $1.07\% < p < 8.68\%$.
    b) $70.1\% < p < 75.3\%$.
    c) Sí; si utilizar ropa color naranja no tuviera efecto alguno, esperaríamos que el porcentaje de cazadores con ropa naranja heridos estuviera entre el 70.1% y el 75.3%, pero es mucho menor.
43. $13.0\% < p < 29.0\%$; sí.
45. $x = 419$ resulta en el intervalo de confianza (0.471, 0.539) y $x = 426$ en el intervalo de confianza (0.480, 0.548). No difieren en cantidades importantes.
47. $p > 0.818$; 81.8%.
49. $0.894 < p < 1.006$; los límites del intervalo de confianza superior exceden 1; utilice un límite superior de 1.
51. 602.

Sección 6-3

1. 2.33.
3. 2.05.
5. Sí.
7. Sí.
9. \$2419.62; $\$92{,}580 < \mu < \$97{,}420$.
11. 0.823 seg; $4.42 \text{ seg} < \mu < 6.06$ seg.
13. 62.
15. 250.
17. 318.1.
19. $\mu = 318.10 \pm 56.01$.
21. $30.0°C < \mu < 30.8°C$; es poco realista conocer σ.
23. $141.4 < \mu < 203.6$; es poco realista conocer σ.
25. 217.
27. 601.
29. 80,770; no; incremente el margen de error.
31. El rango es 40, de modo que se estima que σ es $40/4 = 10$ por medio de la regla práctica del intervalo, y el tamaño de la muestra es 97. La desviación estándar muestral es $s = 11.3$, que resulta en un tamaño de muestra de 123. Es probable que el tamaño muestral de 123 sea mejor porque s es un mejor estimado de σ que rango/4.
33. $105 < \mu < 115$.

Sección 6-4

1. $t_{\alpha/2} = 2.776$.
3. No se aplica la distribución normal ni la distribución t.
5. $t_{\alpha/2} = 1.662$.
7. $t_{\alpha/2} = 2.33$.
9. 60; $436 < \mu < 556$.
11. $112.84 < \mu < 121.56$; hay una confianza del 95% de que el intervalo de 112.84 a 121.56 contiene el valor verdadero de la media poblacional μ.
13. $\$16{,}142 < \mu < \$36{,}312$; hay una confianza del 95% de que el intervalo de \$16,142 a \$36,312 contiene el valor verdadero de la media poblacional μ.
15. a) $-2.248° < \mu < 1.410°$
    b) El intervalo de confianza incluye los 0°. La aseveración no parece ser válida, porque una media de 0° no representa diferencia alguna entre las temperaturas altas reales y las temperaturas del pronóstico para tres días, en tanto que el intervalo de confianza sí incluye los 0°, lo que indica que 0° es un valor muy probable de la diferencia.
17. $0.075 < \mu < 0.168$; no; es imposible que se satisfaga el requisito, pero también es muy posible que la media no sea menor que 0.165 gramos/milla.

19. a) $164 < \mu < 186$.
    b) $111 < \mu < 137$.
    c) 186.
    d) Puesto que una conclusión definitiva sobre la igualdad de medias no debe basarse en el traslape de intervalos de confianza, la siguiente es una conclusión tentativa: los intervalos de confianza no se traslapan en lo absoluto, lo que sugiere que es probable que las dos medias poblacionales sean significativamente diferentes, en tanto que la frecuencia cardiaca media de quienes palean la nieve a mano parece ser mayor que la frecuencia cardiaca media de quienes utilizan el aparato eléctrico para retirar la nieve.
21. Intervalo de confianza del 95% para 4000 a.C.: $125.7 < \mu < 131.6$. Intervalo de confianza del 95% para 150 d.C.: $130.1 < \mu < 136.5$. Puesto que una conclusión definitiva sobre la igualdad de medias no debe basarse en el traslape de intervalos de confianza, la siguiente es una conclusión tentativa: los dos intervalos de confianza se traslapan, entonces es posible que las dos medias poblacionales sean iguales y no concluiríamos que el tamaño de las cabezas parezca cambiar.
23. a) 0.82217 lb $< \mu <$ 0.82603 lb.
    b) 0.78238 lb $< \mu <$ 0.78533 lb.
    c) Puesto que una conclusión definitiva sobre la igualdad de medias no debe basarse en el traslape de intervalos de confianza, la siguiente es una conclusión tentativa: las latas de Pepsi dietética parecen tener un peso medio significativamente menor que el peso medio de latas de Pepsi clásica, quizá por el contenido de azúcar.
25. $-12.244 < \mu < 29.613$; el intervalo de confianza es muy diferente con el dato distante. Los límites del intervalo de confianza son muy sensibles a los datos distantes. Los datos distantes deben examinarse cuidadosamente y descartarse si se descubre que constituyen errores.
27. a) E se multiplica por 5/9.
    b) $\frac{5}{9}(a - 32), \frac{5}{9}(b - 32)$
    c) Sí.

Sección 6-5

1. 6.262, 27.488.
3. 51.172, 116.321.
5. \$9388 $< \sigma <$ \$18,030.
7. 2.06 seg $< \sigma <$ 3.20 seg.
9. 191.
11. 133,448; no.
13. \$11,244 $< \sigma <$ \$26,950; tenemos una confianza del 95% de que los límites de \$11,244 y \$26,950 contienen el valor verdadero de la desviación estándar poblacional σ.
15. $1.195 < \sigma < 4.695$; sí, es probable que el intervalo de confianza sea un estimado pobre, ya que el valor de 5.40 parece ser un dato distante, lo que sugiere que el supuesto de una población con distribución normal no es correcto.
17. a) $10 < \sigma < 27$.
    b) $12 < \sigma < 33$.
    c) Puesto que las conclusiones definitivas sobre la igualdad de desviaciones estándar no deben basarse en el traslape de intervalos de confianza, la siguiente es una conclusión tentativa: la variación no parece ser significativamente diferente.
19. a) 0.33 min $< \sigma <$ 0.87 min.
    b) 1.25 min $< \sigma <$ 3.33 min.
    c) Puesto que las conclusiones definitivas sobre la igualdad de desviaciones estándar no deben basarse en el traslape de intervalos de confianza, la siguiente es una conclusión tentativa: la variación parece ser significativamente menor con una sola fila. Una sola fila parece ser mejor.
21. a) 98%.
    b) 27.0.

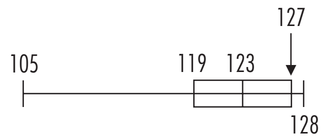
Capítulo 6 Ejercicios de repaso

1. a) 9.00%.
   b) 7.40% $< p <$ 10.6%.
   c) 2653.
2. a) 5.47 años $< \mu <$ 8.55 años.
   b) 2.92 años $< \sigma <$ 5.20 años.
   c) 1484.
   d) No; la muestra no sería representativa de la población de todos los propietarios de automóviles.
3. a) 50.4%.
   b) 45.7% $< p <$ 55.1%.
   c) No, tal vez quienes responden están tratando de impresionar a los encuestadores o quizá sus recuerdos tienen la tendencia a indicar que votaron por el ganador.
4. a) $4.94 < \mu < 8.06$.
   b) $4.33 < \mu < 5.82$.
   c) $7.16 < \mu < 9.71$.
   d) Puesto que las conclusiones definitivas sobre la igualdad de medias no deben basarse en el traslape de intervalos de confianza, la siguiente es una conclusión tentativa: Tolstoi tiene una media significativamente más alta, de manera que su trabajo es más difícil de leer que los de Clancy o Rowling.
5. 65.
6. $0.83 < \sigma < 1.99$.
7. 2944.
8. 221.

Capítulo 6 Ejercicios de repaso acumulativo

1. a) 121.0 lb.
   b) 123.0 lb.
   c) 119 lb, 128 lb.
   d) 116.5 lb.
   e) 23.0 lb.
   f) 56.8 lb^2.
   g) 7.5 lb.
   h) 119.0 lb.
   i) 123.0 lb.
   j) 127.0 lb.
   k) De razón.

l)



m) 112.6 lb $< \mu <$ 129.4 lb.
n) 4.5 lb $< \sigma <$ 18.4 lb.
o) 95.
p) Los pesos individuales de las supermodelos no parecen ser muy diferentes de los pesos de mujeres que se seleccionaron al azar, ya que todos están dentro de 1.31 desviaciones estándar de la media de 143 lb. Sin embargo, cuando se consideran como grupo su media es significativamente menor que la media de 143 lb [véase el inciso (*m*)].

2. a) 0.0089.
   b) $0.260 < p < 0.390$.
   c) Puesto que los límites del intervalo de confianza no contienen 0.25, es poco probable que el experto esté en lo correcto.
3. a) 39.0%.
   b) $36.1\% < p < 41.9\%$.
   c) Sí, porque todo el intervalo de confianza está por debajo del 50%.
   d) El tamaño muestral que se requiere depende del intervalo de confianza y de la proporción muestral, no del tamaño muestral.

Capítulo 7 Respuestas

Sección 7-2

1. No hay evidencia suficiente para sustentar la aseveración de que el método de selección del género es eficaz.
3. Sí, parece haber evidencia suficiente para sustentar la aseveración de que a la mayoría de los adultos estadounidenses les gusta la pizza.
5. H_0: $\mu = \$50{,}000$. H_1: $\mu > \$50{,}000$.
7. H_0: $p = 0$. H_1: $p > 0.5$.
9. H_0: $\sigma = 2.8$. H_1: $\sigma < 2.8$.
11. H_0: $\mu = 12$. H_1: $\mu < 12$.
13. $z = \pm 1.96$.
15. $z = 2.33$.
17. $z = \pm 1.645$.
19. $z = -2.05$.
21. -13.45.
23. 1.85.
25. 0.2912.
27. 0.0512.
29. 0.0244.
31. 0.4412.
33. Hay evidencia suficiente para sustentar la aseveración de que la proporción de mujeres casadas es mayor que 0.5.
35. No existe evidencia suficiente para sustentar la aseveración de que la proporción de accidentes fatales de aviación comercial difiere de 0.038.
37. Tipo I: Concluir que hay evidencia suficiente para sustentar que $p > 0.5$, cuando en realidad $p = 0.5$. Tipo II: No rechazar que $p = 0.5$ (y, por lo tanto, no sustentar que $p > 0.5$) cuando en realidad $p > 0.5$.
39. Tipo I: Concluir que hay evidencia suficiente para sustentar que $p \neq 0.038$ cuando en realidad $p = 0.038$. Tipo II: No rechazar que $p = 0.038$ (y, por lo tanto, no sustentar que $p \neq 0.038$) cuando en realidad $p \neq 0.038$.
41. Valor $P = 0.9999$. Con una hipótesis alternativa de que $p > 0.5$, es imposible que un estadístico muestral de 0.27 caiga en la región crítica. Ninguna proporción muestral menor que 0.5 sustentaría la aseveración de que $p > 0.5$.
43. 0.01, porque este valor más bajo de *P* correspondería a datos muestrales que serían los que sustentarían más la aseveración de que la tasa de defectos es más baja.
45. No hay valores críticos finitos correspondientes a $\alpha = 0$; por lo tanto, es imposible tener un valor $P \leq 0$. Con $\alpha = 0$, la hipótesis nula nunca se rechazará.

Sección 7-3

1. a) $z = -0.12$.
   b) $z = \pm 1.96$.
   c) 0.9044.
   d) No hay evidencia suficiente para justificar el rechazo de la aseveración de que los chícharos con flores verdes sucedan con una tasa del 25%.
   e) No, una prueba de hipótesis no puede utilizarse para probar que una proporción es igual a algún valor aseverado.
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 rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que menos del 62% de las novias que se comprometieron gastan menos de $750. Si las respuestas las hubiesen enviado lectores voluntarios, se trataría de una muestra de respuesta voluntaria y los resultados de la prueba de hipótesis se invalidarían.
5. 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. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que más del 15% de los hogares estadounidenses utilizan el correo electrónico. La conclusión no se vale hoy porque las características de la población (uso del correo electrónico) están cambiando rápidamente con el tiempo.
7. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 0.58$. Valor crítico: $z = 1.28$. Valor *P*: 0.2810. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la proporción es mayor que 0.5.
9. 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. Rechace H_0: $p = 0.01$. Hay evidencia suficiente para justificar el rechazo de la aseveración de que el 1% de las ventas tienen sobreprecios. Puesto que el 1.62% de los artículos muestreados tienen sobreprecios, parece que la tasa de error es peor con los verificadores de precios, no mejor.
11. H_0: $p = 0.61$. H_1: $p > 0.61$. Estadístico de prueba: $z = 1.60$. Valor crítico: $z = 1.645$. Valor *P*: 0.0548. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la tasa de *strikes* de Morrison es mayor del 61%.
13. 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. No rechace H_0. No hay suficiente evidencia para sustentar la aseveración de que la tasa difiere de 0.0340%. Los usuarios de teléfonos celulares no deben preocuparse por el cáncer cerebral o del sistema nervioso.

15. H_0: $p = 0.27$. H_1: $p < 0.27$. Estadístico de prueba: $z = -5.46$. Valor crítico: $z = -2.33$. Valor P: 0.0001. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la tasa de tabaquismo entre los individuos con cuatro años en la universidad es menor al 27%.
17. H_0: $p = 0.75$. H_1: $p > 0.75$. Estadístico de prueba: $z = 8.26$. Valor crítico: $z = 2.33$. Valor P: 0.0000, redondeado a cuatro decimales. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que más de tres cuartas partes de los accidentes de aviación provocan muertes.
19. H_0: $p = 0.10$. H_1: $p \neq 0.10$. Estadístico de prueba: $z = -1.67$. Valores críticos, con un nivel de significancia de 0.05: $z = \pm 1.96$. Valor P: 0.0950. No rechace H_0. No existe evidencia suficiente para justificar el rechazo de la aseveración de que el 10% de los dulces M&M son azules.
21. a) H_0: $p = 0.10$. H_1: $p \neq 0.10$. Estadístico de prueba: $z = 2.00$. Valores críticos: $z = \pm 1.96$. Rechace H_0. Hay 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. Existe evidencia suficiente para justificar 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 se incluye dentro del intervalo de confianza, no rechace H_0: $p = 0.10$. No hay 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.
23. Aseveración original: $p \leq c$; H_0: $p = c$; H_1: $p > c$. Podemos o no rechazar la aseveración original. Los datos muestrales no “sustentan” la aseveración de que los niños que viven cerca de líneas eléctricas de alta tensión no tienen mayores posibilidades de padecer leucemia que otros niños.
25. H_0: $p = 0.10$. H_1: $p \neq 0.10$. Estadístico de prueba: $z = -2.36$. Valores críticos: $z = \pm 2.575$. Valor P: 0.0182. No rechace H_0. Aun cuando se obtuvieron los dulces azules, no hay evidencia suficiente para justificar el rechazo de la aseveración de que el 10% de los dulces son azules.
27. 47% no es un resultado posible porque, con 20 ratones, las únicas tasas posibles de éxito son 0%, 5%, 10%, . . . , 100%.

Sección 7-4

1. Sí.
3. No.
5. $z = 1.18$, Valor P: 0.1190; Valor crítico: $z = 1.645$. No existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 118.
7. $z = 0.89$, Valor P: 0.3734; Valores críticos: $z = \pm 2.575$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 5.00 segundos.
9. H_0: $\mu = 30.0$. H_1: $\mu > 30.0$. Estadístico de prueba: $z = 1.84$. Valor P: 0.0329. (Valor crítico: $z = 1.645$). Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 30.0°C.
11. H_0: $\mu = 200.0$. H_1: $\mu \neq 200.0$. Estadístico de prueba: $z = -1.46$. Valor P: 0.1442. (Valores críticos: $z = \pm 2.575$). No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 200.0.
13. H_0: $\mu = 0.9085$. H_1: $\mu \neq 0.9085$. Estadístico de prueba: $z = 1.68$. Valor P: 0.093. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la media difiere de 0.9085 g.
15. H_0: $\mu = 0$. H_1: $\mu \neq 0$. Estadístico de prueba: $z = -0.63$. Valor P: 0.5288. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que la media es diferente de 0. Tales resultados sugieren que las altas temperaturas que se pronosticaron para tres días son bastante precisas, ya que no parecen ser diferentes de las altas temperaturas reales por una cantidad significativa.
17. a) No es probable que se conozca σ.
    b) 2.10.
    c) No, de manera que el supuesto de que $\sigma = 0.62$ es un supuesto seguro, en el sentido de que si σ en realidad es algún valor diferente de 0.62, es muy poco probable que el resultado de la prueba de hipótesis se afecte.
19. a) 0.6178.
    b) 0.0868.

Sección 7-5

1. t de Student.
3. Normal.
5. Entre 0.005 y 0.01.
7. Menor que 0.01.
9. $t = 0.745$. El valor P es mayor que 0.10. Valor crítico: $t = 1.729$. No hay evidencia suficiente para sustentar la aseveración de que la media es mayor que 118.
11. $t = 0.900$; eel valor P es mayor que 0.20. Valores críticos: $t = \pm 2.639$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 5.00 seg.
13. H_0: $\mu = 4$. H_1: $\mu > 4$. Estadístico de prueba: $t = 3.188$. El valor P es menor que 0.005. Valor crítico: $t = 1.796$. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 4.
15. H_0: $\mu = 0$. H_1: $\mu \neq 0$. Estadístico de prueba: $t = -0.63$. El valor P es mayor que 0.20. Valores críticos: $t = \pm 2.042$. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que la media es diferente de 0°. Con base en estos resultados, parece que las altas temperaturas que se pronosticaron para tres días son bastante precisas.
17. H_0: $\mu = 0$. H_1: $\mu \neq 0$. Estadístico de prueba: $t = 4.010$. El valor P es menor que 0.01. Valores críticos: $t = \pm 2.704$. Rechace H_0. Existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 0 seg. Los relojes no parecen ser bastante precisos.
19. H_0: $\mu = 69.5$. H_1: $\mu > 69.5$. Estadístico de prueba: $t = 2.652$. El valor P se encuentra entre 0.005 y 0.01. Valor crítico: $t = 1.691$. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la media es mayor que 69.5 años. Sin embargo, los hombres no se vuelven directores de orquesta sino hasta que tienen al menos 25 años, en tanto que la expectativa de vida de dichos hombres es naturalmente mayor que la expectativa de vida de los hombres al nacer.
21. H_0: $\mu = \$1000$. H_1: $\mu < \$1000$. Estadístico de prueba: $t = -1.83$. Valor P: 0.071. Con un nivel de significancia de 0.05, no rechace H_0 y concluya que no existe evidencia suficiente para sustentar la aseveración de que la media es menor que $1000.

23. H_0: $\mu = 3.39$. H_1: $\mu \neq 3.39$. Estadístico de prueba: $t = 1.734$. Valor P: 0.1034. No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 3.39 kg (suponiendo un nivel de significancia de 0.05). No existe evidencia suficiente para concluir que el complemento de vitaminas tiene un efecto en el peso al momento de nacer.
25. H_0: $\mu = 1.5$. H_1: $\mu > 1.5$. Estadístico de prueba: $t = 0.049$. El valor P es mayor que 0.10. Valor crítico: $t = 2.015$. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que la media es mayor que 1.5 μg/m^3. El supuesto de una distribución normal es cuestionable, porque 5.40 parece ser un dato distante.
27. H_0: $\mu = 11$. H_1: $\mu < 11$. Estadístico de prueba: $t = -0.095$. Los valores críticos dependen del nivel de significancia, pero el estadístico de prueba no caerá en la región crítica en ninguna opción razonable. El valor P es mayor que 0.10. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la media es menor que 11 segundos. Puesto que los datos se toman de Juegos Olímpicos consecutivos, la media poblacional cambia conforme los atletas se vuelven más rápidos. No es posible concluir que los tiempos futuros serán de alrededor de 10.5 segundos.
29. H_0: $\mu = 6$. H_1: $\mu > 6$. Estadístico de prueba: $t = 0.707$. El valor P es mayor que 0.10. Valor crítico: $t = 1.796$ (suponiendo que $\alpha = 0.05$). No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 6.
31. H_0: $\mu = 12$. H_1: $\mu > 12$. Estadístico de prueba: $t = 10.166$. El valor P es menor que 0.005. Valor crítico: $t = 2.441$ (aproximadamente). Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la media es mayor que 12 oz. El método de producción se ajustaría para que haya menor desperdicio.
33. El valor P se convierte en 0.070. El estadístico de prueba no cambia. La aseveración de que la media difiere de 420 h no se rechaza al nivel de significancia de 0.05.
35. El estadístico de prueba cambia a $t = 0.992$ y el valor P a 0.182. Un dato distante puede cambiar el estadístico de prueba y el valor P de manera sustancial. Aunque la conclusión aquí no cambia, en otros casos sí sería posible que lo hiciera.
37. 0.10.

Sección 7-6

1. Estadístico de prueba: $\chi^2 = 8.444$. Valores críticos: $\chi^2 = 8.907$, 32.852. Valor P: entre 0.02 y 0.05. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que $\sigma \neq 15$.
3. Estadístico de prueba: $\chi^2 = 10.440$. Valor crítico: $\chi^2 = 14.257$. Valor P: menor que 0.005. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que $\sigma < 50$.
5. H_0: $\sigma = 0.04$. H_1: $\sigma > 0.04$. Estadístico de prueba: $\chi^2 = 2342.438$. Valor crítico: $\chi^2 = 63.691$ (aproximadamente). (Valor P: menor que 0.005.) Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que los pesos de los M&M de cacahuate varían más que los pesos de los M&M clásicos.
7. H_0: $\sigma = 43.7$. H_1: $\sigma \neq 43.7$. Estadístico de prueba: $\chi^2 = 114.586$. Valores críticos: $\chi^2 = 57.153$, 106.629. (Valor P: entre 0.01 y 0.02.) Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la desviación estándar difiere de 43.7 pies. Puesto que la desviación estándar muestral es mayor que en el pasado, parece que el nuevo método de producción es peor que entonces.
9. H_0: $\sigma = 6.2$. H_1: $\sigma < 6.2$. Estadístico de prueba: $\chi^2 = 9.016$. Valor crítico: $\chi^2 = 13.848$. (Valor P: menor que 0.005). Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que una sola fila corresponde a una menor variación. Dichos resultados no implican necesariamente que una sola fila dé como resultado un menor tiempo de espera.
11. 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.) Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que los pesos de las supermodelos varían menos que los pesos de las mujeres en general.
13. H_0: $\sigma = 0.10$. H_1: $\sigma < 0.10$. Estadístico de prueba: $\chi^2 = 28.750$. El valor crítico de χ^2 se encuentra entre 18.493 y 26.509. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que los volúmenes tienen una desviación estándar menor que 0.10 oz.
15. H_0: $\sigma = 28.7$. H_1: $\sigma \neq 28.7$. Estadístico de prueba: $\chi^2 = 32.818$. Valores críticos: $\chi^2 = 24.433$, 59.342 (aproximadamente). (Valor P: mayor que 0.20.) No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la desviación estándar es 28.7 lb.
17. Utilice la interpolación y estime el valor crítico como $\chi^2 = 22.501$ para obtener $s = 0.08$ oz.
19. a) Valores estimados: 74.216, 129.565; valores de la tabla A-5: 74.222, 129.561.
    b) 117.093, 184.690.
21. a) La desviación estándar será menor.
    b) No se satisface el requisito de una población que se distribuye normalmente.

Capítulo 7 Ejercicios de repaso

1. a) No; se trata de una muestra de respuesta voluntaria, de manera que los resultados no necesariamente se aplican a la población de adultos estadounidenses.
   b) No; aunque parece haber una pérdida de peso estadísticamente significativa, la cantidad promedio de peso que se pierde es tan pequeña que el fármaco no resulta práctico.
   c) 0.001, porque este valor P corresponde a resultados que proporcionan el mayor sustento para la eficacia de la cura.
   d) No hay evidencia suficiente para sustentar la aseveración de que la media es mayor que 12 oz.
   e) Rechazar una hipótesis nula verdadera.
2. a) H_1: $\mu <$ \$10,000; distribución t de Student.
   b) H_1: $\sigma > 1.8$ seg; distribución chi cuadrada.
   c) H_1: $p > 0.5$; distribución normal.
   d) H_1: $\mu \neq 100$; distribución normal.
3. a) H_0: $\mu = 100$. H_1: $\mu \neq 100$. Estadístico de prueba: $z = -0.75$. Valor P: 0.4532. (Valores críticos: $z = \pm 1.645$). No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 100.

b) H_0: $\mu = 100$. H_1: $\mu \neq 100$. Estadístico de prueba: $t = -0.694$. Valor P: mayor que 0.20. (Valores críticos: $t = \pm 1.676$, aproximadamente). No rechace H_0. No existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 100.
c) H_0: $\sigma = 15$. H_1: $\sigma \neq 15$. Estadístico de prueba: $\chi^2 = 57.861$. Valores críticos: $\chi^2 = 34.764, 67.505$. (Valor P: mayor que 0.20). No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la desviación estándar es igual a 15.
d) Sí.

4. 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 rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que menos de la mitad de todos los ejecutivos identifican que el error más común en una entrevista es "no conocer o conocer poco la empresa".
5. H_0: $\mu = 5.670$ g. H_1: $\mu \neq 5.670$ g. Estadístico de prueba: $t = -4.991$. El valor P es menor que 0.01. Valores críticos: $t = \pm 2.678$, aproximadamente. Rechace H_0. Hay evidencia suficiente para justificar el rechazo de la aseveración de que el peso medio es 5.670 g. Una posible explicación es que las monedas de 25¢ perdieron peso al ser manipuladas en su circulación.
6. H_0: $\mu = 0.9085$ g. H_1: $\mu < 0.9085$ g. Estadístico de prueba: $t = -0.277$. El valor P es mayor que 0.10. Valor crítico: $t = -2.132$. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la media es menor que 0.9085 g. El peso aseverado parece ser el correcto, de acuerdo con lo impreso en la envoltura.
7. H_0: $p = 0.10$. H_1: $p < 0.10$. Estadístico de prueba: $z = -1.17$. Valor crítico: $z = -1.645$. Valor P: 0.1210. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que menos del 10% de los viajes incluyen una visita a un parque temático.
8. H_0: $p = 0.43$. H_1: $p \neq 0.43$. Estadístico de prueba: $z = 3.70$. Valores críticos: $z = \pm 2.05$. Valor P: 0.0002. Rechace H_0. Existe evidencia suficiente para justificar el rechazo de la aseveración de que el porcentaje de votantes que votaron por el candidato ganador es igual al 43%.
9. H_0: $\mu = 12$ oz. H_1: $\mu < 12$ oz. Estadístico de prueba: $t = -4.741$. El valor P es menor que 0.005. Valor crítico: $t = -1.714$ (suponiendo que $\alpha = 0.05$). Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la media es menor que 12 oz. El argumento de Windsor no es válido.
10. H_0: $p = 0.10$. H_1: $p < 0.10$. Estadístico de prueba: $z = -2.36$. Valor crítico: $z = -2.33$. Valor P: 0.0091. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que el verdadero porcentaje es menor que el 10%. La frase "casi 1 de cada 10" no se justifica.
11. H_0: $\sigma = 0.15$. H_1: $\sigma < 0.15$. Estadístico de prueba: $\chi^2 = 44.800$. Valor crítico: $\chi^2 = 51.739$. (Valor P: entre 0.005 y 0.01). Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la variación es menor con la nueva máquina. Se debe comprar la nueva máquina.
12. H_0: $\mu = 3.5$ g. H_1: $\mu \neq 3.5$ g. Estadístico de prueba: $t = 9.720$. El valor P es menor que 0.01. Valores críticos: $t = \pm 1.994$ (aproximadamente, suponiendo que $\alpha = 0.05$). Rechace H_0. Existe evidencia suficiente para justificar el rechazo de la aseveración de que la media es igual a 3.5 g. Parece que los paquetes contienen más azúcar de lo que se indica en la etiqueta.

Capítulo 7 Ejercicios de repaso acumulativo

1. a) 0.0793 ng/m^3.
b) 0.044 ng/m^3.
c) 0.0694 ng/m^3.
d) 0.0048.
e) 0.158 ng/m^3.
f) $0.0259 < \mu < 0.1326$.
g) H_0: $\mu = 0.16$. H_1: $\mu < 0.16$. Estadístico de prueba: $t = -3.491$. El valor P es menor que 0.005. Valor crítico: $t = -1.860$. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la media es menor que 0.16 ng/m^3.
h) Sí, los datos se listan en orden y parece haber una tendencia de valores decrecientes. La población va cambiando con el paso el tiempo.
2. a) 0.4840.
b) 0.0266 (de 0.4840^5).
c) 0.4681.
d) 634.
3. a) 6.3.
b) 2.2.
c) Binom: 0.0034; normal: 0.0019.
d) Con base en el bajo valor de probabilidad del inciso *c*), rechace H_0: $P = 0.25$. Hay evidencia suficiente para rechazar la aseveración de que el sujeto trató de adivinar.
e) 423.

Capítulo 8 Respuestas

Sección 8-2

1. 30.
3. 85.
5. a) 0.417.
b) 2.17.
c) ± 1.96.
d) 0.0300.
7. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 2.17$. Valor P: 0.0150. Valor crítico: $z = 1.645$. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que es mayor la proporción de empleados que la proporción de jefes que dijeron que vigilar el correo electrónico atenta, de manera grave, contra la ética.
9. $0.00216 < p_1 - p_2 < 0.00623$; parece que la actividad física corresponde a una tasa menor de enfermedad coronaria cardiaca.
11. H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$. Estadístico de prueba: $z = -0.21$. Valor P: 0.8336. Valores críticos: $z = \pm 1.96$. No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la tasa de revocación es la misma para ambos años.
13. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = -12.39$. Valor P: 0.0001. Valor crítico de $\alpha = 0.05$: $z = -1.645$. Rechace H_0. Existe evidencia muestral suficiente para sustentar la aseveración.
15. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = 1.07$. Valor P: 0.1423. Valor crítico: $z = 1.645$. No rechace H_0. No. Con base en la evidencia disponible, posponga cualquier acción.

17. Con un estadístico de prueba de $z = 4.94$ y un valor P de 0.000, rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la tasa de sentencias de culpabilidad para las esposas es menor que la tasa de sentencias de culpabilidad para los esposos.
19. H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$. Estadístico de prueba: $z = -2.01$. Valor P: 0.0444. Valores críticos: $z = \pm 1.96$. Rechace H_0. Parece haber una diferencia significativa. Puesto que la tasa de fallas de autozone es menor, parece ser la mejor opción.
21. $-0.135 < p_1 - p_2 < 0.0742$ (con $x_1 = 49$ y $x_2 = 70$); no parece existir una diferencia por género.
23. $-0.0144 < p_1 - p_2 < 0.0086$; sí.
25. a) H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$. Estadístico de prueba: $z = -3.06$. Valores críticos: $z = \pm 2.575$ (suponiendo un nivel de significancia de 0.01). Valor P: 0.0022. Rechace H_0: $p_1 = p_2$. Hay evidencia suficiente para sustentar la aseveración de que los dos porcentajes poblacionales son diferentes.
    b) $-0.0823 < p_1 - p_2 < -0.00713$; porque los límites del intervalo de confianza no incluyen al 0, parece haber una diferencia significativa (aunque sería mejor utilizar una prueba de hipótesis para la hipótesis nula $p_1 = p_2$).
27. H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$. Estadístico de prueba: $z = 4.46$. Valor P: 0.0002. Valores críticos: $z = \pm 2.575$. Rechace H_0. Existe evidencia suficiente para justificar el rechazo de la aseveración de que la tasa de rechazos de las ciudades centrales es igual a la tasa de rechazos de otras áreas.
29. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = -0.60$. Valor P: 0.2743. No hay suficiente evidencia para sustentar la aseveración de que la proporción de películas infantiles que muestran consumo de alcohol es menor que la proporción que muestra consumo de tabaco. Los resultados no se aplican al conjunto de datos 7 porque las muestras no son independientes.
31. a) $0.0227 < p_1 - p_2 < 0.217$; ya que los límites del intervalo de confianza no contienen el 0, parece que puede rechazarse $p_1 = p_2$.
    b) $0.491 < p_1 < 0.629$; $0.371 < p_2 < 0.509$; ya que los intervalos de confianza sí se traslapan, 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: $z = \pm 1.96$. Rechace H_0. Hay evidencia suficiente para rechazar $p_1 = p_2$.
    d) Rechace $p_1 = p_2$. El menos efectivo: utilizar el traslape entre intervalos de confianza individuales.
33. El estadístico de prueba cambia a $z = 2.03$ y el intervalo de confianza del 90% cambia a 0.00231 y 0.0277, de forma que ahora hay evidencia suficiente para sustentar la aseveración dada.
35. a) Estadístico de prueba: $z = 1.48$. Valor P: 0.1388. Valores críticos: $z = \pm 1.96$. No rechace H_0: $p_1 = p_2$.
    b) Estadístico de prueba: $z = 1.63$. Valor P: 0.1032. Valores críticos: $z = \pm 1.96$. No rechace H_0: $p_2 = p_3$.
    c) Estadístico de prueba: $z = 3.09$. Valor P: 0.0020. Valores críticos: $z = \pm 1.96$. Rechace H_0: $p_1 = p_3$.
    d) No.
37. a) No, porque las condiciones $np \geq 5$ y $nq \geq 5$ no se satisfacen en ambas muestras.
    b) Con 144 personas en el grupo placebo, 1.8% no es un resultado posible.

Sección 8-3

1. Muestras independientes.
3. Datos apareados.
5. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. Estadístico de prueba: $t = 2.790$. Valor crítico: $t = 2.660$. Valor $P < 0.01$. (Si se utiliza la TI-83: gl $= 122$, valor $P = 0.003$). Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la población de consumidores frecuentes de marihuana tiene una media menor que la de los consumidores ligeros. Los consumidores frecuentes de marihuana deberían preocuparse por el deterioro de sus capacidades mentales.
7. $-0.65 < \mu_1 - \mu_2 < 3.03$ (TI-83: gl $= 69$ y $-0.61 < \mu_1 - \mu_2 < 2.99$). Como el intervalo de confianza sí contiene el cero, no tendremos que concluir que las dos medias poblacionales son diferentes. Parece que el tratamiento no es eficaz, por lo que la paroxetina no debería prescribirse.
9. 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: 0.461, aproximadamente. (Si se usa la TI-83: gl $= 34$, valor $P = 0.448$). No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que los magnetos son eficaces para reducir el dolor. Se valdría el argumento de que los magnetos son eficaces si los tamaños muestrales fueran más grandes.
11. a) H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 22.098$. Valores críticos: $t = \pm 2.728$. Valor $P < 0.01$. (Si se usa la TI-83: gl $= 56$, valor $P = 0.000$). Rechace H_0. Hay evidencia suficiente para justificar el rechazo de la aseveración de que la Coca Cola clásica y la Coca Cola dietética tienen el mismo peso medio. La diferencia tal vez se debe al azúcar que contiene la Coca Cola clásica pero no la Coca Cola dietética.
    b) $0.02808 < \mu_1 - \mu_2 < 0.03598$ (TI-83: gl $= 56$ y $0.02817 < \mu_1 - \mu_2 < 0.03589$).
13. $-0.01 < \mu_1 - \mu_2 < 0.23$; puesto que este intervalo de confianza contiene al cero, sí parece haber una diferencia significativa entre las dos medias poblacionales, de manera que no parece que los trastornos obsesivo-compulsivos tengan una base biológica. (Con una calculadora TI-83, gl $= 18$ y $0.01 < \mu_1 - \mu_2 < 0.21$, que no contiene el cero, lo que sugiere que existe una diferencia significativa, de tal forma que los trastornos obsesivo-compulsivos parecen tener una base biológica. Se trata de un caso poco común en el cual el estimado simple y conservador de gl conduce a una conclusión diferente que la fórmula 8.1 más precisa).
15. $1.46 < \mu_1 - \mu_2 < 3.52$ (TI-83: gl $= 25$ y $1.47 < \mu_1 - \mu_2 < 3.51$.) Puesto que el intervalo de confianza no contiene al cero, parece haber una diferencia significativa entre las dos medias poblacionales. Aquellos que consumieron alcohol cometieron significativamente más errores.
17. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. Estadístico de prueba: $t = 2.879$. Valor crítico: $t = 2.429$. Valor $P = 0.006$, aproximadamente. (Con una TI-83: gl $= 77$, valor $P = 0.003$). Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la población con estrés tiene una media menor que la población sin estrés. Sin embargo, no es posible concluir que el estrés *disminuye* la memoria.

19. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 1.130$. Valores críticos: $t = \pm 1.983$. Valor $P = 0.261$. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia significativa entre las dos medias poblacionales.
21. 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 $P = 0.000$. (Utilizando TI-83: gl $= 26$, valor $P = 0.0000$). Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la cantidad media de alquitrán en cigarros largos con filtro es menor que la cantidad media de alquitrán en cigarros largos sin filtro.
23. Hombres: $n_1 = 40$, $\bar{x}_1 = 25.9975$, $s_1 = 3.4307$. Mujeres: $n_2 = 40$, $\bar{x}_2 = 25.7400$, $s_2 = 6.1656$. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 0.231$. Valores críticos: $t = \pm 2.024$ (suponiendo un nivel de significancia de 0.05). Valor $P = 0.842$, aproximadamente. (Utilizando TI-83: gl $= 61$, valor $P = 0.818$). No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que la media del IMC de los hombres es igual a la media del IMC de las mujeres.
25. $-0.62 < \mu_1 - \mu_2 < 3.00$ (TI-83: $-0.60 < \mu_1 - \mu_2 < 2.98$). Los resultados no cambiaron mucho.
27. Estos nuevos resultados son muy cercanos a los que se obtuvieron en el ejercicio 9: valor crítico: $t = 1.686$; valor P: 0.460, aproximadamente. (Utilizando TI-83: Valor $P = 0.448$). Los demás resultados son iguales.
29. a) El estadístico de prueba cambia sustancialmente de $t = 1.130$ a $t = 1.508$, pero no es un cambio suficiente para provocar un cambio en la conclusión.
    b) El numerador del estadístico de prueba sí se incrementa sustancialmente, puesto que las medias muestrales presentan una diferencia mucho mayor, pero el denominador también se incrementa de forma importante, por el aumento en la varianza de la primera muestra.
31. a) 50/3
    b) 2/3
    c) $50/3 + 2/3 = 52/3$.
    d) El rango de los valores de *x-y* es igual al rango de los valores de *x* más el rango de los valores de *y*.
33. gl $= 18$ (en lugar de 9), los valores críticos se convierten en $t = \pm 2.878$ (en lugar de ± 3.250) y los límites del intervalo de confianza cambian a 0.007 y 0.213, en tanto que el valor P es menor que 0.01 (en lugar de 0.01 y 0.02). Utilizando la fórmula 8-1, el intervalo de confianza es un poco más angosto, el valor crítico es un poco menor y el valor P es un poco menor. Con gl $= 9$ no parece que los trastornos obsesivo-compulsivos tengan una base biológica; con gl$= 18$, de la fórmula 8-1, sí parece que los trastornos obsesivo-compulsivos tienen una base biológica. Es más conservador el uso del menor de $n_1 - 1$ y $n_2 - 1$ (que el uso de la fórmula 8-1), en el sentido de que los datos muestrales necesitan ser más extremos para considerarse significativos, como se aprecia en las distintas conclusiones.

Sección 8-4

1. a) -0.2
   b) 2.8
   c) $t = -0.161$
   d) ± 2.776
3. $-3.6 < \mu_d < 3.2$
5. a) H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -0.831$. Valores críticos: $t = \pm 2.201$. Valor P: 0.440, aproximadamente. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas reportadas y las estaturas medidas.
   b) $-1.7 < \mu_d < 0.8$; puesto que los límites del intervalo de confianza contienen al 0, no existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas que se reportaron y las estaturas medidas.
7. a) H_0: $\mu_d = 0$. H_1: $\mu_d < 0$. Estadístico de prueba: $t = -1.718$. Valor crítico: $t = -1.833$. Valor P: 0.062. No rechace H_0. No hay evidencia suficiente para concluir que el curso de preparación sea eficaz para elevar las calificaciones.
   b) $-25.5 < \mu_d < 3.5$; tenemos una confianza del 95% de que el intervalo que va de -25.5 a 3.5 en realidad contiene la verdadera diferencia de la media poblacional.
9. 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: $t = 1.895$. Valor P: 0.007. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que las mediciones sensoriales son más bajas después de la hipnosis.
   c) Sí.
11. a) H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -1.690$. Valores críticos $t = \pm 2.228$. Valor P: 0.120. No rechace H_0. No hay 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.
13. a) H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -0.41$. Valor P: 0.691. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que el astemizole surta un efecto. No tome astemizole para el mareo que causa el movimiento.
    b) 0.3455; no hay evidencia suficiente para sustentar la aseveración de que el astemizole previene el mareo que causa el movimiento.
15. H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -0.501$. Valores críticos: $t = \pm 2.201$. Valor P: 0.626. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre los pesos que se reportan y los pesos que se miden de hombres cuyas edades fluctúan entre 12 y 16 años.
17. a) $-1.40 < \mu_d < -0.17$.
    b) H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -2.840$. Valores críticos: $t = \pm 2.228$. Valor $P < 0.02$. Rechace H_0. Hay evidencia suficiente para justificar el rechazo de la aseveración de que la diferencia media es 0. Las temperaturas corporales de la mañana y de la noche no parecen ser aproximadamente las mismas.

19. a) H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -2.966$. Valores críticos: $t = \pm 2.042$. Valor P: 0.006. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las bajas temperaturas reales y las bajas temperaturas que se pronosticaron cinco días antes.
    b) $-10.1 < \mu_d < -1.9$.
    c) Con el conjunto de datos más grande, consistente en 31 datos apareados, existe evidencia suficiente para concluir que hay una diferencia significativa entre las bajas temperaturas reales y las bajas temperaturas que se pronosticaron cinco días antes.
21. a) Sí.
    b) La prueba de hipótesis no afecta. Los límites del intervalo de confianza cambiarán de la escala Fahrenheit a los valores equivalentes de la escala Celsius.
23. a) Estadístico de prueba: $t = 1.861$. Valor crítico: $t = 1.833$. Valor P: 0.045. Rechace H_0. Hay evidencia suficiente para sustentar $\mu_d > 0$.
    b) Estadístico de prueba: $t = 1.627$, Valor crítico: $t = 1.833$. Valor P: 0.072. No rechace H_0. No existe evidencia suficiente para sustentar $\mu_1 > \mu_2$.
    c) Sí, la conclusión se afecta por la prueba que se utiliza.

Sección 8-5

1. H_0: $\sigma_1^2 = \sigma_2^2$. H_1: $\sigma_1^2 \neq \sigma_2^2$. Estadístico de prueba; $F = 2.2500$. Valor crítico superior: $F = 2.1540$. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que las poblaciones de tratamiento y placebo tienen diferentes varianzas.
3. 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. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que la reducción del dolor en el grupo de tratamiento que se simula varía más que la reducción del dolor del grupo de tratamiento con magnetos.
5. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 2.9228$. El valor crítico superior de F se encuentra entre 1.8752 y 2.0739. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que las poblaciones tienen distintas desviaciones estándar.
7. H_0: $\sigma_1^2 = \sigma_2^2$. H_1: $\sigma_1^2 > \sigma_2^2$. Estadístico de prueba: $F = 3.7539$. Valor crítico: $F = 3.4445$. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que los cigarros largos con filtro tienen cantidades de nicotina que varían más que las cantidades de nicotina de los cigarros largos sin filtro.
9. H_0: $\sigma_1^2 = \sigma_2^2$. H_1: $\sigma_1^2 > \sigma_2^2$. Estadístico de prueba: $F = 1.0110$. El valor crítico de F es menor que 1.3519 (suponiendo un nivel de significancia de 0.05). (Aunque la conclusión no está clara a partir del estadístico de prueba y del valor crítico, los valores de las desviaciones estándar [3.67 y 3.65] sugieren que la diferencia no es significativa. Con una calculadora TI-83 Plus, resulta un valor P de 0.4745). No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que la antigüedad de los automóviles de los profesores varíe más que la antigüedad de los automóviles de los estudiantes.
11. a) Estadístico de prueba: $F = 2.1722$. El valor crítico superior de F se encuentra entre 1.6668 y 1.8752. Rechace H_0. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las cantidades de lluvia del miércoles y el sábado tienen la misma desviación estándar.
    b) Puesto que incluyen muchos ceros como valores menores, ni las cantidades de lluvia de los miércoles ni las cantidades de lluvia de los sábados se distribuyen normalmente.
    c) Puesto que las poblaciones no parecen distribuirse normalmente, la conclusión del inciso *a*) no es necesariamente válida. Los métodos de la sección 8-5 no se aplican.
13. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 1.2478$. Valor P: 0.6852. No rechace H_0. No existe evidencia suficiente para justificar el rechazo de la aseveración de que los dos grupos muestrales provienen de poblaciones con desviaciones estándar iguales. Sí.
15. H_0: $\sigma_1^2 = \sigma_2^2$. H_1: $\sigma_1^2 \neq \sigma_2^2$. Estadístico de prueba: $F = 2.8176$. El valor crítico superior de F se encuentra entre 3.5257 y 3.4296. No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que las dos muestras provienen de poblaciones con la misma variación.
17. El estadístico de prueba cambia de $F = 1.5824$ a $F = 1.0000$. No rechace H_0. No existe evidencia suficiente para justificar el rechazo de la aseveración de que las poblaciones tienen la misma desviación estándar. La conclusión cambia. El dato distante sí produce un efecto sumamente importante en los resultados.
19. a) $F_I = 0.2484$, $F_D = 4.0260$.
    b) $F_I = 0.2315$, $F_D = 5.5234$.
    c) $F_I = 0.1810$, $F_D = 4.3197$.

Capítulo 8 Ejercicios de repaso

1. 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. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración. Parece que los pacientes quirúrgicos deben mantenerse tibios como rutina.
   b) 90%.
   c) $-0.205 < p_1 - p_2 < -0.0543$.
   d) No, las conclusiones serían distintas.
2. a) H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = -1.532$. Valores críticos: $t = \pm 2.228$. Valor P: 0.164, aproximadamente. No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencia.
   b) $-2.7 < \mu_d < 0.5$.
   c) No, no existe una diferencia significativa.
3. a) $-27.80 < \mu_1 - \mu_2 < 271.04$; TI-83 utiliza gl $= 17.7$ para obtener los límites de -17.32 y 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: 0.106, aproximadamente. No rechace H_0. No hay evidencia suficiente para justificar el rechazo de la aseveración de ninguna diferencia.
   c) No.
4. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 1.2922$. Valor crítico superior: $F = 4.0260$. No rechace H_0. No existe evidencia suficiente para sustentar la aseveración de que las dos poblaciones tienen distintas cantidades de variación.
5. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = -3.500$. Valores críticos: $t = \pm 2.365$. Valor P: 0.010. Rechace H_0. Hay evidencia suficiente para justificar el rechazo de la aseveración de que las medias son iguales. Los filtros parecen ser eficaces para reducir el monóxido de carbono.

6. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. Estadístico de prueba: $t = 2.169$. Valores críticos: $t = \pm 1.968$, aproximadamente. El valor P se encuentra entre 0.01 y 0.025. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que el complemento de zinc se asocia con mayores pesos al momento de nacer.
7. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 2.41$. Valor crítico: $z = 1.645$. Valor P: 0.0080. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración.
8. a) H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 2.301$. Valores críticos: $t = \pm 2.262$ (suponiendo un nivel de significancia de 0.05). Valor P: 0.046. Rechace H_0 (suponiendo un nivel de significancia de 0.05). Hay evidencia suficiente para concluir que existe una diferencia entre los pesos antes del entrenamiento y los pesos después del entrenamiento.
   b) $0.0 < \mu_d < 4{,}0$.

Capítulo 8 Ejercicios de repaso acumulativo

1. a) 0.0707
   b) 0.369
   c) 0.104
   d) 0.0540
   e) H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = -2.52$. Valor crítico: $z = -1.645$. Valor P: 0.0059. Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que el porcentaje de mujeres infraccionadas por ir a exceso de velocidad es menor que el porcentaje de hombres.
2. Debe haber un error, puesto que las tasas del 13.7% y 10.6% no son posibles con tamaños muestrales de
3. a) $0.0254 < p < 0.0536$ (usando $x = 29$).
   b) $0.0103 < p < 0.0311$ (usando $x = 15$).
   c) $0.00133 < p_1 - p_2 < 0.0363$.
   d) método (iii).
4. a) H_0: $p = 0.5$. H_1: $p < 0.5$. Estadístico de prueba: $z = -5.88$. Valor crítico: $z = -1.645$ (suponiendo un nivel de significancia de 0.05). Valor P: 0.0001. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que la proporción de mujeres es menor que 0.5.
   b) $\bar{x} = 17{,}198.3$ seg; mediana $= 16{,}792$ seg; $s = 3107.2$; la distribución es aproximadamente normal; sin datos distantes.
   c) H_0: $\mu = 18{,}000$ seg. H_1: $\mu < 18{,}000$ seg. Estadístico de prueba: $t = -1.611$. Valor crítico: $t = -1.686$. Valor P: 0.059. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que las mujeres tienen un tiempo medio de carrera menor a cinco horas.
   d) H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 3.101$. Valores críticos: $t = \pm 2.024$. Valor P: 0.004. Rechace H_0. Existe evidencia suficiente para sustentar la aseveración de que el tiempo medio de los hombres es diferente del tiempo medio de las mujeres.
   e) El uso de las proporciones muestrales de 39/150 y 111/150 resulta en el valor que se agrupa de $\bar{p} = 150/300$, que supone que el tamaño muestral total es de 300 en lugar de 150. La muestra de 150 valores proviene de una población, no de dos poblaciones.

Capítulo 9 Respuestas

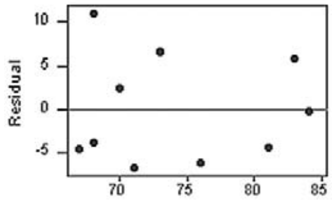
Sección 9-2

1. a) Sí, puesto que el valor absoluto del estadístico de prueba excede los valores críticos $r = \pm 0.707$.
   b) 0.986.
3. a) No, puesto que el valor absoluto del estadístico de prueba no excede los valores críticos $r = \pm 0.444$ (aproximadamente).
   b) 0.0177.
5. El diagrama de dispersión sugiere que hay una correlación, pero no es lineal. Con $r = 0$ y valores críticos de $r = \pm 0.878$ (para un nivel de significancia de 0.05), no hay una correlación lineal significativa.
7. a) Parece que 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 significativa.
   c) $r = 0$. Valores críticos: $r = \pm 0.666$ (para un nivel de significancia de 0.05). No parece haber una correlación lineal significativa.
   d) El efecto de un solo par de valores puede ser muy grande y cambiar la conclusión.
9. $r = -0.118$. Valores críticos: $r = \pm 0.707$. No existe una correlación lineal significativa. Más bajo: Susan Lucci. Más alto: Kelsey Grammer.
11. $r = 0.658$. Valores críticos: $r = \pm 0.532$. Hay una correlación lineal significativa. Otro aspecto es la precisión de las mediciones, las cuales parecen variar con amplitud. Se realizaría un estudio para determinar si la presión sanguínea de los sujetos en realidad varía considerablemente o si las mediciones son erróneas, por otros factores.
13. $r = 0.262$. Valores críticos: $r = \pm 0.576$. No existe una correlación lineal significativa. Ya que se reportaron los conteos de cigarros, tal vez los sujetos dieron valores incorrectos. Quizá los sujetos estuvieron expuestos a niveles variables de humo de cigarros.
15. $r = 0.359$. Valores críticos: $r = \pm 0.497$. No hay una correlación lineal significativa.
17. $r = 0.482$. Valores críticos: $r = \pm 0.312$. Existe una correlación lineal significativa.
19. a) $r = 0.997$. Valores críticos: $r = \pm 0.279$. Hay una correlación lineal significativa.
    b) $r = 0.899$. Valores críticos: $r = \pm 0.279$. Existe una correlación lineal significativa. Hay una correlación entre la cantidad de impuestos y el valor de la casa.
21. a) $r = 0.574$. Valores críticos: $r = \pm 0.361$, aproximadamente. Existe una correlación lineal significativa. No, puede haber una correlación alta, aunque las temperaturas que se pronosticaron sean muy imprecisas.
    b) $r = 0.685$. Valores críticos: $r = \pm 0.361$, aproximadamente. Existe una correlación lineal significativa. No, puede existir una correlación alta, aun cuando las temperaturas que se pronosticaron sean muy imprecisas.
    c) Las temperaturas que se pronosticaron a un día son mejores, puesto que tienen una correlación más alta con las temperaturas reales. Sin embargo, una correlación alta no implica que las temperaturas que se pronosticaron sean precisas.

23. a) $r = 0.870$. Valores críticos: $r = \pm 0.279$. Existe una correlación lineal significativa.
    b) $r = -0.010$. Valores críticos: $r = \pm 0.279$. No hay una correlación lineal significativa.
    c) Duración, puesto que tiene una correlación lineal significativa con el intervalo.
25. Con un coeficiente de correlación lineal muy cercano a 0, no parece haber una correlación, pero la conclusión sugiere que hay una correlación.
27. Aunque no existe una *correlación lineal*, las variables tal vez se relacionen de alguna otra manera *no lineal*.
29. $r = 0.819$ (aproximadamente). Valores críticos: $r = \pm 0.553$. Hay una correlación lineal significativa.
31. a) ± 0.279.
    b) ± 0.191.
    c) -0.378.
    d) 0.549.
    e) 0.658.
33. $0.386 < \rho < 0.753$.

Sección 9-3

1. a) 18.00.
   b) 5.00.
3. 401 lb.
5. $\hat{y} = 2 + 0x$ (o $\hat{y} = 2$).
7. 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 número llega a afectar de manera importante la ecuación de regresión.
9. $\hat{y} = 6.76 - 0.0111x$; 6.5 millones. El valor predicho de 6.5 millones se aleja mucho del valor real de 24 millones.
11. $\hat{y} = -14.4 + 0.769x$; 79.
13. $\hat{y} = 139 + 2.48x$; 175.2.
15. $\hat{y} = 3.68 + 3.78x$; 3.76.
17. $\hat{y} = 21.9 + 0.0160x$; 29.9.
19. a) $\hat{y} = 7.26 + 0.914x$; $190,060.
    b) $\hat{y} = 380 + 19.5x$; $8180.
21. a) $\hat{y} = 13.8 + 0.611x$; 31°.
    b) $\hat{y} = 13.8 + 0.634x$; 32°.
    c) Inciso *b*), puesto que el coeficiente de correlación es más alto.
23. a) $\hat{y} = 41.9 + 0.179x$; 79 min.
    b) $\hat{y} = 81.9 - 0.009x$; 81 min.
    c) El inciso *a*), puesto que hay una correlación lineal significativa entre la duración y los intervalos, pero no entre la estatura y los intervalos.
25. Sí; no. El punto se encuentra muy alejado del resto, pero no tiene un efecto muy importante en la recta de regresión.
27. $\hat{y} = -182 + 0.000351x$; $\hat{y} = -182 + 0.351x$. La pendiente se multiplica por 1000 y el intercepto *y* no cambia. Si cada valor de *y* se divide entre 1000, tanto la pendiente como el intercepto *y* se dividen entre 1000.
29. La ecuación $\hat{y} = -49.9 + 27.2x$ es mejor, ya que tiene $r = 0.997$, que es más alto que $r = 0.963$ para $\hat{y} = -103.2 + 134.9 \ln x$.
31. No.



Sección 9-4

1. 0.64; 64%.
3. 0.253: 25.3%.
5. 0.961; sí.
7. 1.3.
9. a) 287.37026.
   b) 166.62974.
   c) 454.
   d) 0.63297415.
   e) 4.8789597.
11. a) 3696.9263.
    b) 1690.5830.
    c) 5387.5093.
    d) 0.68620324.
    e) 11.869369.
13. a) 116 lb.
    b) $103.7 \text{ lb} < y < 128.8 \text{ lb}$.
15. a) 44 pies.
    b) $6.2 \text{ pies} < y < 81.4 \text{ pies}$.
17. $54.2 < y < 107.0$.
19. $71.7 < y < 112.2$.
21. $-170.8 < \beta_0 < -54.6$; $1.5 < \beta_1 < 3.1$.
23. a) $(n - 2)s_e^2$.
    b) $\dfrac{r^2 \cdot (\textit{variación inexplicable})}{1 - r^2}$
    c) $r = -0.949$.

Sección 9-5

1. $\hat{y} = -272 - 0.870x_1 + 0.554x_2 + 12.2x_3$.
3. Sí, puesto que el valor *P* es 0.000 y el valor de R^2 ajustado es 0.924.
5. El consumo de combustible en carretera, porque tiene la R^2 ajustada más alta.
7. La ecuación de regresión con el consumo de combustible en carretera y el peso tiene la R^2 más alta de 0.861, pero puede ser posible argumentar a favor de utilizar sólo la variable independiente del consumo de combustible en carretera, puesto que su R^2 ajustada es de 0.853, que es ligeramente menor.

9. a) $\hat{y} = 21.6 + 0.690x$.
   b) $\hat{y} = 45.7 + 0.293x$.
   c) $\hat{y} = 9.80 + 0.658x_1 + 0.200x_2$; (donde x_1 = estatura de la madre).
   d) El inciso *c*), puesto que la R^2 ajustada es mayor.
   e) No, puesto que el valor más alto de la R^2 ajustada es de tan sólo 0.366, que no es muy alto.
11. a) $\hat{y} = 3.68+3.78x$.
   b) $\hat{y} = 3.46 + 1.01x$
   c) $\hat{y} = 3.40 + 3.20x_1, + 0.982x_2$ (donde x_1 = grasa).
   d) El inciso *c*), puesto que la R^2 ajustada es la más alta.
   e) Sí, pero los valores predichos no serán necesariamente muy precisos.
13. $\hat{y} = 0.154 + 0.0651x$, donde x es la cantidad de alquitrán. Aun cuando la ecuación de regresión múltiple con el alquitrán y la nicotina, como variables independientes, tiene la R^2 ajustada más alta de 0.928, la R^2 ajustada que incluye únicamente el alquitrán como variable independiente tiene una R^2 ajustada de 0.921, que es muy cercana. Con valores de R^2 tan cercanos, es mejor seleccionar la ecuación con una variable independiente que con dos.
15. $\hat{y} = 7.26 + 0.914x_1$ (donde x_1 representa el precio de lista); si se utilizan más variables, la R^2 ajustada se llega a incrementar de 0.995 a 0.996, aunque el pequeño incremento en la R^2 ajustada no justifica la inclusión de variables adicionales.

Sección 9-6

1. Cuadrático: $y = 2x^2 - 12x + 18$.
3. Exponencial: $y = 3^x$.
5. Cuadrático: $y = 0.0516657x^2 + 1.50881\,x + 18.6857$, donde x se codifica como 1 para 1980, 2 para 1981 y así sucesivamente. Valor predicho: 77.
7. Cuadrático: $y = 1.21445x^2 + 42.4084x + 371.958$, donde x se codifica como 1 para 1990, 2 para 1991 y así sucesivamente. Valor predicho: 1361.
9. a) Exponencial: $y = 2^{\frac{2}{3}(x-1)}$ [o $y = (0.629961)(1.587401)^x$ para un valor inicial de 1, que se duplica cada 1.5 años].
   b) Exponencial: $y = (2.32040)(1.36587)^x$.
   c) La ley de Moore no parece estar funcionando muy bien.
11. a) 189.1.
   b) 0.9979.
   c) El valor de R^2 cuadrática de 0.9992 es mayor que $R^2 = 0.9979$ para el modelo logístico, en tanto que la suma de cuadrados de los residuales es menor para el modelo cuadrático (73.2) que para el modelo logístico (189.1).

Capítulo 9 Ejercicios de repaso

1. $r = -0.069$. Valores críticos: $r = \pm 0.707$ (suponiendo un nivel de significancia de 0.05). No hay una correlación lineal significativa. El nivel de CAS no parece relacionarse con la edad de la persona a quien se realizó la prueba.
2. a) $r = 0.828$. Valores críticos: $r = \pm 0.811$ (suponiendo un nivel de significancia de 0.05). Existe una correlación lineal significativa.
   b) La ecuación de regresión de $\hat{y} = -0.347 + 0.149x$ (donde x representa la cuenta) indica que la cantidad de propina predicha es 35 centavos menor que el 15% de la cuenta, que es aproximadamente el 15% de la cuenta.
3. a) $r = 0.338$. Valores críticos: $r = \pm 0.632$; no hay una correlación lineal significativa.
   b) 11%.
   c) $\hat{y} = -0.488 + 0.611x$.
   d) 0.347 pintas por persona a la semana.
4. a) $r = 0.116$. Valores críticos: $r = \pm 0.632$; correlación lineal no significativa.
   b) 1.3%.
   c) $\hat{y} = 0.0657 + 0.000792x$.
   d) 0.347 pintas por persona a la semana.
5. a) $r = 0.777$. Valores críticos: $r = \pm 0.632$. Existe una correlación lineal significativa.
   b) 60%.
   c) $\hat{y} = 0.193 + 0.00293x$.
   d) 0.286 pintas por persona a la semana.
6. $\hat{y} = -0.0526 + 0.747x_1 - 0.00220x_2 + 0.00303x_3$; $R^2 = 0.726$; R^2 ajustada $= 0.589$; valor $P = 0.040$. Puesto que el valor total de P de 0.040 es menor que 0.05, la ecuación se llega a utilizar para predecir el consumo de helado. Si se utilizan los datos consumo/temperatura, la R^2 ajustada es 0.554. Aunque la R^2 ajustada es ligeramente mayor si se utilizan las tres variables, el ligero incremento en la R^2 ajustada no justifica la inclusión de variables adicionales; por lo tanto, la mejor ecuación de regresión parece resultar del uso de la temperatura como única variable independiente.

Capítulo 9 Ejercicios de repaso acumulativo

1. a) $r = -0.884$. Valores críticos: $r = \pm 0.576$ (suponiendo un nivel de significancia de 0.05). Existe una correlación lineal significativa.
   b) $\hat{y} = 95.3 - 3.46x$.
   c) Es posible realizar la prueba de dos medias poblacionales iguales, pero la prueba no tendría sentido, ya que las dos variables miden la facilidad de lectura por medio de criterios diferentes y con distintas escalas.
   d) $61.16 < \mu < 71.14$.
2. a) $\bar{x} = 99.1$, $s = 8.5$.
   b) $\bar{x} = 102.8$, $s = 8.7$.
   c) No, pero una mejor comparación implicaría tratar los datos como datos que se aparearon, en lugar de dos muestras independientes.
   d) H_0: $\mu = 100$. H_1: $\mu \neq 100$. Estadístico de prueba: $r = 0.546$. Valores críticos: $t = \pm 2.069$. El valor P es mayor que 0.20. No rechace H_0. No hay evidencia suficiente para sustentar la aseveración de que la puntuación media del CI de gemelos que se criaron aparte sea diferente del CI medio de 100.
   e) Sí. $r = 0.702$ y los valores críticos son $r = \pm 0.576$ (suponiendo un nivel de significancia de 0.05). Existe una correlación lineal significativa.

Capítulo 10 Respuestas

Sección 10-2

1. a) H_0: $p_1 = p_2 = p_3 = p_4$.
   b) 8, 8, 8, 8.
   c) $\chi^2 = 4.750$.
   d) $\chi^2 = 7.815$.
   e) No existe evidencia suficiente para justificar el rechazo de la aseveración de que las cuatro categorías son igualmente probables.

3. a) gl = 37; entonces, $\chi^2 = 51.805$ (aproximadamente).
   b) $0.10 <$ valor $P < 0.90$.
   c) No hay evidencia suficiente para justificar el rechazo de la aseveración de que las ranuras de la ruleta son igualmente probables.
5. Estadístico de prueba: $\chi^2 = 5.860$. Valor crítico: $\chi^2 = 11.071$. No existe evidencia suficiente para sustentar la aseveración de que los resultados no sean igualmente probables. El dado que se cargó no parece comportarse de manera diferente que un dado que se balanceó.
7. Estadístico de prueba: $\chi^2 = 9.233$. Valor crítico: $\chi^2 = 12.592$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que los accidentes suceden con la misma frecuencia durante los diferentes días.
9. Estadístico de prueba: $\chi^2 = 10.653$. Valor crítico: $\chi^2 = 9.488$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para justificar el rechazo de la aseveración de que los accidentes suceden en proporciones iguales los cinco días hábiles.
11. Estadístico de prueba: $\chi^2 = 16.333$. Valor crítico: $\chi^2 = 14.067$ (suponiendo un nivel de significancia de 0.05). Hay evidencia suficiente para sustentar la aseveración de que las probabilidades de ganar en las distintas posiciones que se asignaron no son iguales.
13. Estadístico de prueba: $\chi^2 = 18.500$. Valor crítico: $\chi^2 = 16.919$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que los dígitos suceden con la misma frecuencia. La conclusión cambia con un nivel de significancia de 0.01. El proceso de selección tendría que cambiarse de inmediato si hubiera una fuerte evidencia que sugiriera que los dígitos no son igualmente probables.
15. Estadístico de prueba: $\chi^2 = 5.950$. Valor crítico: $\chi^2 = 11.071$. No existe evidencia suficiente para justificar el rechazo de la distribución que asevera Mars, Inc.
17. Estadístico de prueba: $\chi^2 = 4.200$. Valor crítico: $\chi^2 = 16.919$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que los dígitos se distribuyen de manera uniforme.
19. Estadístico de prueba: $\chi^2 = 14.421$. Valor crítico: $\chi^2 = 15.507$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que los dígitos provienen de una población de dígitos líder que se ajustan a la ley de Benford.
21. El estadístico de prueba cambia de 4.600 a 76.638, de forma que el dato distante surte un efecto muy importante.
23. a) Valor crítico es $\chi^2 = 3.841$ y el estadístico de prueba es
$$\chi^2 = \frac{\left(f_1 - \dfrac{f_1 + f_2}{2}\right)^2}{\dfrac{f_1 + f_2}{2}} + \frac{\left(f_2 - \dfrac{f_1 + f_2}{2}\right)^2}{\dfrac{f_1 + f_2}{2}}$$
$$= \frac{(f_1 - f_2)^2}{f_1 + f_2}$$
   b) Valores críticos: el valor crítico χ^2 es 3.841 y aproximadamente igual al cuadrado de $z = 1.96$.
25. a) 0.0853, 0.2968, 0.3759, 0.1567, 0.0853.
   b) 17.06, 59.36, 75.18, 31.34, 17.06.
   c) Estadístico de prueba: $\chi^2 = 60.154$. Valor crítico: $\chi^2 = 13.277$. Rechace H_0: Las puntuaciones de CI provienen de una población que se distribuye normalmente, con la media y la desviación estándar dadas. Hay evidencia suficiente para justificar el rechazo de la aseveración de que las puntuaciones de CI se seleccionaron al azar de una población que se distribuye normalmente, con una media de 100 y una desviación estándar de 15.

Sección 10-3

1. Estadístico de prueba: $\chi^2 = 0.413$. Valor P: 0.521. No existe evidencia suficiente para justificar el rechazo de la aseveración de que a las personas las detiene la policía independientemente de su raza y su grupo. No existe evidencia suficiente para sustentar una aseveración de discriminación racial.
3. Estadístico de prueba: $\chi^2 = 4.698$. Valor crítico: $\chi^2 = 3.841$. Hay evidencia suficiente para justificar el rechazo de independencia entre respuesta y el hecho de que el sujeto sea un trabajador o un jefe de alto nivel. La conclusión cambia si se utiliza un nivel de significancia de 0.01.
5. Estadístico de prueba: $\chi^2 = 51.458$. Valor crítico: $\chi^2 = 6.635$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las proporciones de las respuestas de acuerdo/en desacuerdo son iguales para los sujetos que entrevistaron hombres y los sujetos que entrevistaron mujeres.
7. Estadístico de prueba: $\chi^2 = 63.908$. Valor crítico: $\chi^2 = 3.841$. Hay evidencia suficiente para justificar el rechazo de la aseveración de que el género es independiente del miedo a volar.
9. Estadístico de prueba: $\chi^2 = 3.062$. Valor crítico: $\chi^2 = 5.991$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que el éxito es independiente del método que se emplee. La evidencia no sugiere que alguno de los métodos sea significativamente mejor que los demás.
11. Estadístico de prueba: $\chi^2 = 65.524$. Valor crítico: $\chi^2 = 7.815$ (suponiendo un nivel de significancia de 0.05). Hay evidencia suficiente para justificar el rechazo de la aseveración de que la ocupación es independiente de si la causa de la muerte fue un homicidio. Los cajeros parecen ser los más vulnerables al homicidio.
13. Estadístico de prueba: $\chi^2 = 20.271$. Valor crítico: $\chi^2 = 15.086$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que la cooperación del sujeto es independiente de la categoría de edad.
15. Estadístico de prueba: $\chi^2 = 119.330$. Valor crítico: $\chi^2 = 5.991$. Hay 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.
17. Estadístico de prueba: $\chi^2 = 42.557$. Valor crítico: $\chi^2 = 3.841$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que la sentencia es independiente de la declaración de inocencia. Los resultados motivan que los acusados que son culpables se declaren inocentes.
19. Estadístico de prueba: $\chi^2 = 1.199$. Valor crítico: $\chi^2 = 7.815$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que tener un dolor de cabeza es independiente de la cantidad de atorvastatin que se utilice como tratamiento.
21. Sin la corrección de Yates: $\chi^2 = 0.413$. Con la 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.

Capítulo 10 Ejercicios de repaso

1. $\chi^2 = 16.747$. Valor crítico: $\chi^2 = 9.488$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las llamadas se distribuyen de manera uniforme durante los días laborales de la semana.
2. $\chi^2 = 6.780$. Valor crítico: $\chi^2 = 12.592$. No hay evidencia suficiente para sustentar la teoría de que suceden más muertes por arma de fuego durante los fines de semana.
3. $\chi^2 = 49.731$. Valor crítico: $\chi^2 = 11.071$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para sustentar la aseveración de que el tipo de crimen se relaciona con el hecho de que el criminal sea bebedor o abstemio.
4. $\chi^2 = 5.297$. Valor crítico: $\chi^2 = 3.841$. Hay evidencia suficiente para justificar el rechazo de la aseveración de que si un recién nacido es dado de alta antes o después es independiente del hecho de que el recién nacido se hospitalice durante la semana posterior. La conclusión cambia si el nivel de significancia se modifica a 0.01.

Capítulo 10 Ejercicios de repaso acumulativo

1. $\bar{x} = 80.9$; mediana: 81.0; rango: 28.0; $s^2 = 75.4$; $s = 8.6$; resumen de los cinco números: 66, 76.0, 81.0, 86.5, 94.
2. a) 0.272.
   b) 0.468.
   c) 0.614.
   d) 0.282.
3. Tabla de contingencia; véase la sección 10-3. Estadístico de prueba: $\chi^2 = 0.055$. Valor crítico: $\chi^2 = 7.815$ (suponiendo un nivel de significancia de 0.05). No existe evidencia suficiente para justificar el rechazo de la aseveración de que hombres y mujeres eligen las distintas respuestas en las mismas proporciones.
4. Utilice la correlación; véase la sección 9-2. Estadístico de prueba: $r = 0.978$. Valores críticos: $r = \pm 0.950$ (suponiendo un nivel de significancia de 0.05). Hay evidencia suficiente para sustentar la aseveración de que existe una relación entre la memoria y las puntuaciones de razonamiento.
5. Utilice la prueba para datos que se aparean; véase la sección 8-4. $\bar{d} = -10.25$; $s_d = 1.5$. Estadístico de prueba: $t = -13.667$. Valor crítico: $t = -2.353$ (suponiendo un nivel de significancia de 0.05). Rechace H_0. Hay evidencia suficiente para sustentar la aseveración de que la sesión de entrenamiento es eficaz para aumentar las puntuaciones.
6. Pruebe la diferencia entre dos muestras independientes: véase la sección 8-3. Estadístico de prueba: $t = -2.014$. Valores críticos: $t = \pm 3.182$ (suponiendo un nivel de significancia de 0.05). No rechace H_0: $\mu_1 = \mu_2$. No existe evidencia suficiente para justificar el rechazo de la aseveración de que hombres y mujeres tienen la misma puntuación media.

Capítulo 11 Respuestas

Sección 11-2

1. a) $\mu_1 = \mu_2 = \mu_3$.
   b) Al menos una de las tres medias es diferente de las demás.
   c) $F = 8.98$.
   d) $F = 3.3158$, aproximadamente.
   e) 0.001.
   f) Hay evidencia suficiente para justificar el rechazo de la aseveración de que los tres autores tienen la misma puntuación media del nivel de Flesch-Kincaid.
3. a) $\mu_1 = \mu_2 = \mu_3$.
   b) Al menos una de las tres medias es diferente de las demás.
   c) $F = 0.1887$.
   d) 3.0804.
   e) 0.8283.
   f) No.
5. Estadístico de prueba: $F = 0.9922$. Valor crítico: $F = 3.2389$. Valor P: 0.4216. No rechace 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.
7. Estadístico de prueba: $F = 4.0497$. Valor crítico: $F = 3.4028$. Valor P: 0.0305. Rechace H_0: $\mu_1 = \mu_2 = \mu_3$. Hay evidencia suficiente para sustentar la aseveración de que la anchura media no es la misma para las distintas épocas.
9. Estadístico de prueba: $F = 0.5083$. Valor crítico: $F = 2.2899$ (aproximadamente). Valor P: 0.7694. No rechace 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 las poblaciones de M&M de diferentes colores tienen la misma media.
11. Estadístico de prueba: $F = 9.0646$. Valor crítico: $F = 3.8056$. Valor P: 0.0034. Rechace H_0: $\mu_1 = \mu_2 = \mu_3$. Hay evidencia suficiente para justificar el rechazo de la aseveración de que la media de las cantidades de azúcar en los diferentes anaqueles son las mismas. El anaquel 2 parece tener una media mucho más grande, lo cual sustentaría la aseveración de que los cereales con mayor contenido de azúcar fueron colocados en los anaqueles que se encuentran al nivel de los ojos de los niños.
13. a) 10.
    b) 0.599.
    c) 0.95.
    d) Análisis de varianza.

Sección 11-3

1. "Dos factores" se refiere a la inclusión de dos factores diferentes, que son propiedades o características que se utilizan para distinguir distintas poblaciones entre sí. "Análisis de varianza" se refiere al método empleado, que se basa en dos estimados diferentes de la supuesta varianza poblacional común.
3. Si existe una interacción entre factores, no debemos considerar los efectos de alguno de los factores sin considerar los del otro.
5. Estadístico de prueba: $F = 1.28$. Valor P: 0.313. No rechace la hipótesis nula de ninguna interacción. No parece haber un efecto significativo de la interacción entre el lugar y la edad.
7. Estadístico de prueba: $F = 249.85$. Valor P: 0.000. Rechace la hipótesis nula de que la edad no tiene un efecto en la cantidad de DDT. Existe evidencia suficiente para sustentar la aseveración de que la edad surte un efecto en la cantidad de DDT.
9. Estadístico de prueba: $F = 5.03$. Valor P: 0.031. Rechace la hipótesis nula de que el género no ejerce un efecto en las calificaciones del SAT. Hay evidencia suficiente para sustentar la aseveración de que el género produce un efecto en las calificaciones del SAT.

11. Estadístico de prueba: $F = 3.87$. Valor P: 0.000. Rechace la hipótesis nula de que la selección del sujeto no tiene un efecto en la puntuación de la prueba de audición. Existe evidencia suficiente para sustentar la aseveración de que la selección del sujeto produce un efecto en 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 hay un efecto de interacción significativo. Para el género, el estadístico de prueba es $F = 0.09$ y el valor P es 0.762, por lo que no hay un efecto significativo a partir del género. Para la edad, el estadístico de prueba es $F = 0.36$ y el valor P es 0.701; por lo tanto, no hay un efecto significativo a partir 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 dato distante puede afectar de manera importante y a modificar los resultados y las conclusiones.

Capítulo 11 Ejercicios de repaso

1. Estadístico de prueba: $F = 46.90$. Valor P: 0.000. Rechace H_0: $\mu_1 = \mu_2 = \mu_3$. Existe evidencia suficiente para justificar el rechazo de la aseveración de medias poblacionales iguales.
2. Estadístico de prueba: $F = 9.4827$. Valor crítico: $F = 3.0984$. Rechace H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$. Hay evidencia suficiente para sustentar la aseveración de precios medios de venta diferentes.
3. Estadístico de prueba: $F = 0.19$. Valor P: 0.832. No rechace la hipótesis nula de ninguna interacción. No parece haber un efecto significativo de la interacción entre el género y el área de estudios.
4. Estadístico de prueba: $F = 0.78$. Valor P: 0.395. No rechace la hipótesis nula de que el género no produce efecto en las calificaciones del SAT. No hay evidencia suficiente para sustentar la aseveración de que la longitud que se estima se afecte por el género.
5. Estadístico de prueba: $F = 0.13$. Valor P: 0.876. No rechace la hipótesis nula de que el área de estudios no produce un efecto en las calificaciones del SAT. No existe evidencia suficiente para sustentar la aseveración de que la longitud que se estima se afecte por el área de estudios.
6. a) Estadístico de prueba: $F = 1.00$. Valor P: 0.423. No hay evidencia suficiente para sustentar la aseveración de que las cantidades de gases invernadero que se emiten se afecten por el tipo de transmisión.
   b) Estadístico de prueba: $F = 7.00$. Valor P: 0.125. No existe evidencia suficiente para sustentar la aseveración de que las cantidades de gases invernadero estén afectadas por el número de cilindros.
   c) Quizá 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 dichas aseveraciones.

Capítulo 11 Ejercicios de repaso acumulativo

1. a) 0.100 in.
   b) 0.263 in.
   c) 0.00, 0.00, 0.00, 0.010, 1.41.
   d) 0.92 in, 1.41 in.
   e) La respuesta varía, dependiendo del número de clases que se emplee, pero el histograma tiene que describir una distribución sesgada hacia la derecha.
   f) No, puesto que los datos no parecen provenir de una población con distribución normal.
   g) 19/52 o 0.365.
2. a) 960.5, 980.0, 1046.0; no.
   b) 914.5, 1010.5, 1008.5; no.
   c) 174.6, 239.6, 226.8; no.
   d) Estadístico de prueba: $t = -0.294$. Valores críticos: $t = \pm 2.093$ (suponiendo un nivel de significancia de 0.05). No rechace H_0: $\mu_1 = \mu_2$.
   e) $878.8 < \mu < 1042.2$.
   f) Estadístico de prueba: $F = 0.8647$. Valor P: 0.4266. No rechace H_0: $\mu_1 = \mu_2 = \mu_3$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que las tres poblaciones tienen la misma calificación media en el SAT.
3. a) 0.3372.
   b) 0.0455.
   c) 1/8.

Capítulo 12 Respuestas

Sección 12-2

1. El estadístico de prueba de $x = 5$ no es menor que o igual al valor crítico de 3. No existe evidencia suficiente para justificar el rechazo de la aseveración de ninguna diferencia.
3. El estadístico de prueba $z = -0.95$ no es menor que o igual al valor crítico de -1.96. No hay evidencia suficiente para justificar el rechazo de la aseveración de ninguna diferencia.
5. El estadístico de prueba $x = 5$ no es menor que o igual al valor crítico de 2. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas que se reportaron y las estaturas que se midieron.
7. El estadístico de prueba $x = 1$ es menor que 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.
9. Convierta $x = 301$ en el estadístico de prueba $z = -12.60$. Valor crítico: $z = -1.645$ (suponiendo un nivel de significancia de 0.05). Hay evidencia suficiente para sustentar la aseveración de que la mayoría de las personas votaron, asegura, en la elección.
11. Convierta $x = 1$ en el estadístico de prueba $z = -5.32$. Valor crítico: $z = -1.645$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para sustentar la aseveración de que las latas de Coca Cola tienen volúmenes con una mediana mayor que 12 oz.
13. (En lugar de una prueba de cola derecha para determinar si $x = 37$ es suficientemente grande para ser significativa, use una prueba de cola izquierda para saber si $x = 13$ es suficientemente pequeña para ser significativa). Convierta $x = 13$ en el estadístico de prueba $z = -3.25$. Valor crítico: $z = -1.645$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para sustentar la aseveración de que la mediana es mayor que 77 min.
15. Primer método: $z = -1.90$; rechace H_0.
    Segundo método: $z = -1.73$; rechace H_0.
    Tercer método: $z = 0$; no rechace H_0.

17. Convierta $x = 18$ en el estadístico de prueba $z = -2.31$. Valor crítico: $z = -2.33$. No hay evidencia suficiente para sustentar una acusación de discriminación 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, de manera que existe evidencia suficiente para sustentar una acusación de discriminación por género. Si se utiliza la aproximación normal, el estadístico de prueba se encuentra apenas fuera de la región crítica; con la distribución binomial, el estadístico de prueba se encuentra apenas dentro de la región crítica.

Sección 12.3

1. Estadístico de prueba: $T = 1$. Valor crítico: $T = 2$. Rechace la hipótesis nula de que ambas muestras provienen de la misma distribución poblacional.
3. Estadístico de prueba: $T = 34$. Valor crítico: $T = 14$. No rechace la hipótesis nula de que ambas muestras provienen de la misma distribución poblacional.
5. Estadístico de prueba: $T = 0$. Valor crítico: $T = 8$. Rechace la hipótesis nula de que ambas muestras provienen de la misma distribución poblacional. Existe evidencia suficiente para justificar el rechazo de la aseveración de que no hay diferencia.
7. Estadístico de prueba: $T = 178$. Valor crítico: $T = 117$ (suponiendo un nivel de significancia de 0.05). No rechace la hipótesis nula de que ambas muestras provienen de la misma distribución poblacional. No parece haber una diferencia.
9. Convierta $T = 661$ en el estadístico de prueba $z = -5.67$. Valores críticos: $z = \pm 1.96$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que los adultos saludables tienen una temperatura media corporal que es igual a 98.6°F.

Sección 12-4

1. $R_1 = 120.5$, $R_2 = 155.5$, $\mu_R = 132$, $\sigma_R = 16.248$. Estadístico de prueba: $z = -0.71$. Valores críticos: $z = \pm 1.96$. No rechace la hipótesis nula de que las poblaciones tienen la misma distribución.
3. $\mu_R = 150$, $\sigma_R = 17.321$, $R = 96.5$, $z = -3.09$. Estadístico de prueba: $z = -3.09$. Valores críticos: $z = \pm 2.575$. Hay evidencia suficiente para justificar el rechazo de la aseveración de que las dos muestras provienen de poblaciones idénticas.
5. $\mu_R = 525$, $\sigma_R = 37.417$, $R = 437$, $z = -2.35$. Estadístico de prueba; $z = -2.35$. Valores críticos: $z = \pm 1.96$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las dos muestras provienen de poblaciones idénticas.
7. $\mu_R = 150$, $\sigma_R = 17.321$, $R = 86.5$, $z = -3.67$. Estadístico de prueba: $z = -3.67$. Valores críticos: $z = \pm 1.96$. Rechace la hipótesis nula de que las muestras de Rowling y Tolstoi provienen de poblaciones con la misma distribución.
9. $\mu_R = 3696$, $\sigma_R = 214.94$, $R = 3861$, $z = 0.77$. Estadístico de prueba: $z = 0.77$. Valores críticos: $z = \pm 1.96$. No rechace la hipótesis nula de que las dos poblaciones de edades tienen la misma distribución.
11. $z = -3.67$; el estadístico de prueba es el mismo número con signo opuesto.

Sección 12-5

1. Sí. El valor P de 0.747 indica que no rechazamos la hipótesis nula de que las tres categorías de edad tienen poblaciones idénticas.
3. Estadístico de prueba: $H = 1.1914$. Valor crítico: $\chi^2 = 7.815$. No hay evidencia suficiente para sustentar la aseveración de que los traumatismos craneales de las cuatro categorías de peso no son todos iguales. Los datos no proporcionan evidencia suficiente para concluir que los automóviles que pesan más son más seguros en un choque.
5. Estadístico de prueba: $H = 6.631$. Valor crítico: $\chi^2 = 5.991$. Existe evidencia suficiente para justificar el rechazo de la aseveración de que las tres muestras provienen de poblaciones idénticas.
7. Estadístico de prueba: $H = 2.075$. Valor crítico: $\chi^2 = 11.071$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que los pesos sean iguales en cada una de las seis poblaciones de colores diferentes.
9. a) El estadístico de prueba H no cambia.
   b) El estadístico de prueba H no cambia.
   c) El valor del estadístico de prueba no cambia mucho (puesto que se emplea el rango de orden en lugar de la magnitud del dato distante).
11. 14.840 (usando $T = 6, 6, 24$); no.

Sección 12-6

1. a) $r_s = 1$ y parece haber una correlación entre x y y.
   b) $r_s = -1$ y parece existir una correlación entre x y y.
   c) $r_s = 0$ y no parece haber una correlación entre x y y.
3. $r_s = 0.855$. Valores críticos: $r_s = \pm 0.648$. Correlación significativa. Parece existir una correlación entre el salario y el estrés.
5. $r_s = 0.103$. Valores críticos: $r_s = \pm 0.648$. Correlación no significativa. No parece haber una correlación entre los rangos de orden institucional y de graduados de escuelas de negocios.
7. $r_s = 0.557$. Valores críticos: $r_s = \pm 0.683$. Correlación no significativa. No parece existir una correlación entre la estatura y el peso.
9. $r_s = 0.506$. Valores críticos: $r_s = \pm 0.507$. Correlación no significativa. No parece haber una correlación entre las cantidades de grasa y el conteo de calorías.
11. a) $r_s = 0.918$. Valores críticos; $r_s = \pm 0.370$. Correlación significativa. Parece que existe una correlación entre el alquitrán y la nicotina.
    b) $r_s = 0.739$. Valores críticos: $r_s = \pm 0.370$. Correlación significativa. Parece que hay una correlación entre el monóxido de carbono y la nicotina.
    c) El alquitrán, puesto que tiene una mayor correlación con la nicotina.
13. a) ± 0.707.
    b) ± 0.514.
    c) ± 0.361.
    d) ± 0.463.
    e) ± 0.834.

Sección 12-7

1. $n_1 = 11$, $n_2 = 7$, $G = 3$, valores críticos: 5, 14; rechace aleatoriedad.
3. $n_1 = 10$, $n_2 = 10$, $G = 10$, valores críticos: 6, 16; no rechace aleatoriedad.
5. $n_1 = 12$, $n_2 = 8$, $G = 10$, valores críticos: 6, 16; no rechace aleatoriedad.
7. $n_1 = 19$, $n_2 = 13$, $G = 6$, valores críticos: 10, 23; rechace aleatoriedad.
9. $n_1 = 18$, $n_2 = 14$, $G = 15$, valores críticos: 10, 23; no rechace aleatoriedad.
11. $n_1 = 10$, $n_2 = 10$, $G = 2$, valores críticos: 6, 16; rechace aleatoriedad. Puesto que la tendencia es ascendente, el mercado bursátil parece ser un buen medio de inversión.
13. $n_1 = 49$, $n_2 = 51$, $G = 43$, $\mu_G = 50.98$, $\sigma_G = 4.9727$. Estadístico de prueba: $z = -1.60$. Valores críticos: $z = \pm 1.96$. No rechace aleatoriedad.
15. $n_1 = 111$, $n_2 = 39$, $G = 54$, $\mu_G = 58.720$, $\sigma_G = 4.6875$. Estadístico de prueba: $z = -1.01$. Valores críticos: $z = \pm 1.96$. No rechace aleatoriedad. No existe evidencia suficiente para sustentar la aseveración de que los hombres corredores tienden a llegar a la meta antes que las mujeres corredoras.
17. El mínimo es 2; el máximo es 4. Los valores críticos de 1 y 6 nunca pueden suceder, por lo cual la hipótesis nula de aleatoriedad nunca puede ser rechazada.

Capítulo 12 Ejercicios de repaso

1. El estadístico de prueba $x = 1$ no es menor que o igual al valor crítico de 1. No existe evidencia suficiente para justificar el rechazo de la aseveración de que el curso no surte efecto alguno.
2. Estadístico de prueba: $T = 9.5$. Valor crítico: $T = 6$. No hay evidencia suficiente para justificar el rechazo de la aseveración de que el curso no produce efecto alguno.
3. Prueba del signo: Convierta $x = 22$ en el estadístico de prueba $z = -2.58$. Valor crítico: $z = -2.33$. Existe evidencia suficiente para sustentar la aseveración de un sesgo a favor de los hombres.
4. Prueba de la suma de rangos de Wilcoxon: $\mu_R = 162$, $\sigma_R = 19.442$, $H = 89.5$, $z = -3.73$. Estadístico de prueba: $z = -3.73$. Valores críticos: $z = \pm 1.96$. Rechace la hipótesis nula de que las dos muestras provienen de poblaciones idénticas. Hay evidencia suficiente para justificar el rechazo de la aseveración de que los bebedores de cerveza y los bebedores de licor tienen los mismos niveles de CAS.
5. Correlación de rangos: $r_s = -0.796$. Valores críticos: $r_s = \pm 0.648$. Correlación significativa. Parece que hay una correlación entre el peso y el consumo de combustible en carretera.
6. 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: $z = \pm 1.96$. No rechace aleatoriedad. Parece que los números pares e impares suceden aleatoriamente.
7. Prueba de Kruskal-Wallis: estadístico de prueba: $H = 4.234$. Valor crítico: $\chi^2 = 7.815$. No existe evidencia suficiente para sustentar la aseveración de que las mediciones del daño no son las mismas para las cuatro categorías. No hay evidencia suficiente para sustentar la aseveración de que los automóviles que más pesan son más seguros.
8. Correlación de rangos: $r_s = 0.190$. Valores críticos: $r_s = \pm 0.738$. No existe evidencia suficiente para sustentar la aseveración de una correlación entre el rendimiento y el precio. Compre las cintas de menor precio.

Capítulo 12 Ejercicios de repaso acumulativo

1. a) $n_1 = 20$, $n_2 = 7$, $G = 15$, valores críticos: 6, 16; no rechace aleatoriedad.
   b) Estadístico de prueba: $z = -2.50$. Valores críticos: $z = \pm 1{,}96$. Rechace la hipótesis nula de que la proporción de mujeres es igual a 0.5.
   c) Convierta $x = 7$ en el estadístico de prueba $z = -2.31$. Valores críticos: $z = \pm 1.96$. Existe evidencia suficiente para sustentar la aseveración de que la proporción de mujeres difiere de 0.5.
   d) $0.0940 < p < 0.425$.
   e) La secuencia parece estar en un orden aleatorio, pero los sujetos parecen sesgarse con respecto a las mujeres. Se debe realizar una mayor investigación para determinar si la población tiene una proporción de mujeres menor que 0.5.
2. a) Estadístico de prueba: $r = -0.515$. Valores críticos: $r = \pm 0.707$. Correlación lineal no significativa.
   b) Estadístico de prueba: $r_s = -0.463$. Valores críticos: $r_s = \pm 0.738$. Correlación no significativa.
   c) El estadístico de prueba $x = 3$ no es menor que o igual al valor crítico de 0. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas de los candidatos ganadores y los perdedores.
   d) Estadístico de prueba: $T = 10$. Valor crítico: $T = 2$. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas de los candidatos ganadores y los perdedores.
   e) Estadístico de prueba: $t = 0.851$. Valores críticos: $t = \pm 2.365$. No rechace H_0: $\mu_1 = \mu_2$. No existe evidencia suficiente para sustentar la aseveración de que hay una diferencia entre las estaturas de los candidatos ganadores y los perdedores.
   f) No existe evidencia suficiente para concluir que las estaturas de los candidatos ganadores y las estaturas de los candidatos perdedores se relacionan y no hay evidencia suficiente para concluir que existe una diferencia entre las estaturas de los candidatos ganadores y las estaturas de los candidatos perdedores.

Capítulo 13

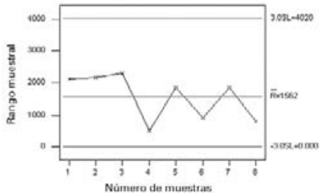
Sección 13-2

1. a) Los datos de proceso son aquellos que se ordenan de acuerdo con alguna secuencia de tiempo.
   b) Un proceso se encuentra fuera de control estadístico si tiene una variación distinta a la variación natural y patrones, ciclos o puntos poco comunes.
   c) Existe un patrón, una tendencia o un ciclo que obviamente no es aleatorio, o un punto que se encuentra fuera de los límites de control superior e inferior, o ocho puntos consecutivos, todos por arriba o por debajo de la línea central.

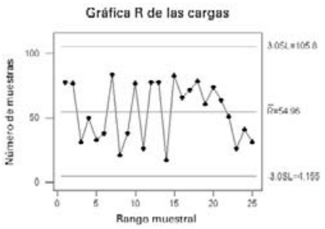
d) La variación aleatoria se debe al azar, pero la variación asignable resulta de causas que no es posible identificar.

e) Una gráfica R muestra el patrón de rangos muestrales y se utiliza para determinar si la variación se encuentra bajo control estadístico, mientras que una gráfica $\overline{x}$ indica el patrón de medias muestrales y se emplea para determinar si la media de un proceso está bajo control estadístico.

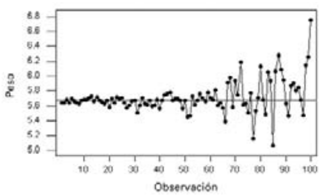
3. La variación del proceso parece estar bajo control estadístico.



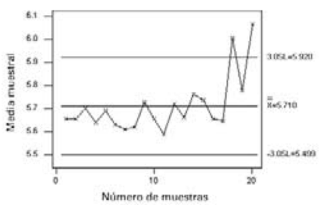
5. La variación del proceso parece estar bajo control estadístico.



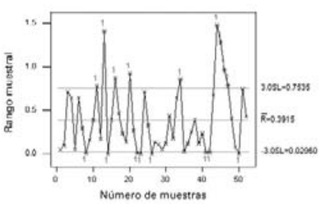
7. Existe un patrón de variación en incremento; por lo tanto, el proceso está fuera de control estadístico. La variación en incremento resultará en más y más defectos.



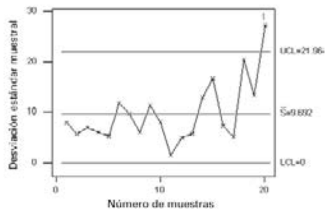
9. Hay un patrón de variación en incremento, puntos que caen más allá del límite de control superior y ocho puntos consecutivos por debajo de la línea central; por lo tanto, la media del proceso está fuera de control estadístico. Este proceso necesita una acción correctiva.



11. La variación del proceso al parecer está fuera de control estadístico. Existen puntos que caen más allá de los límites de control.

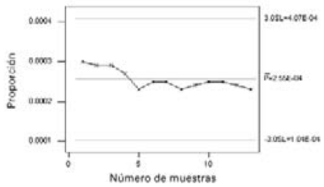


13. La variación del proceso al parecer está fuera de control estadístico. Existe un punto más allá del límite de control superior y parece haber una tendencia creciente.

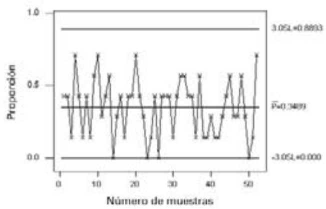


Sección 13-3

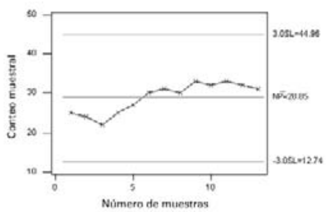
1. El proceso al parecer está bajo control estadístico.
3. El proceso al parecer está fuera de control estadístico, ya que hay un patrón de una tendencia creciente y un punto que cae más allá del límite de control superior.
5. El proceso está fuera de control estadístico, puesto que hay una tendencia descendente y ocho (o más) puntos consecutivos que caen por debajo de la línea central. Dicha tendencia descendente es buena, por lo que deberían identificarse sus causas para continuar.



7. El proceso tal vez sea estadísticamente estable.

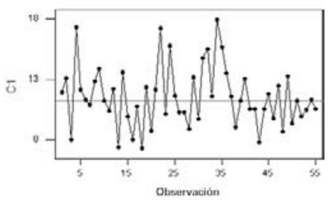


9. Excepto por la escala que se utiliza, las gráficas son idénticas.

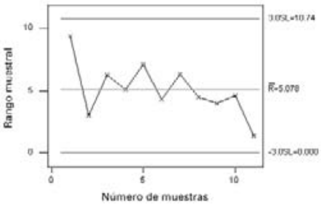


Capítulo 13 Ejercicios de repaso

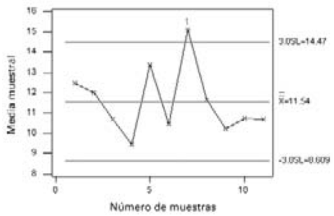
1. El proceso al parecer está bajo control estadístico.



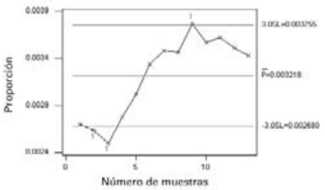
2. La variación del proceso al parecer está bajo control estadístico.



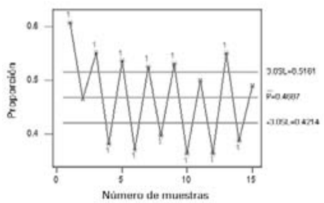
3. Puesto que hay un punto que cae más allá del límite de control superior, la media del proceso no está bajo control estadístico.



4. El proceso está fuera de control, puesto que hay un cambio ascendente y puntos que se encuentran más allá de los límites de control.

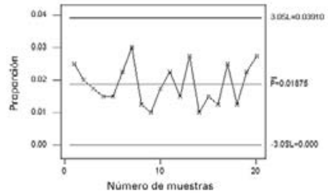


5. El proceso está fuera de control, puesto que hay puntos que se encuentran más allá de los límites de control. Además, existe un patrón cíclico.

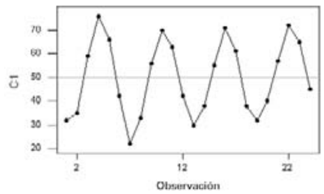


Capítulo 13 Ejercicios de repaso acumulativo

1. a) El proceso al parecer está bajo control estadístico.



   b) $0.0158 < p < 0.0217$.
   c) Estadístico de prueba: $z = 7.87$. Valor crítico: $z = 1.645$. Existe evidencia suficiente para sustentar la aseveración de que la tasa de defectos es mayor que 1%.
2. a) 1/256.
   b) 1/256.
   c) 1/128.
3. La gráfica de rachas muestra ciclos muy claros, así que el proceso no es estadísticamente estable.



4. a) $r = -0.484$. Valores críticos: $r = \pm 0.396$ (aproximadamente, suponiendo un nivel de significancia de 0.05). Hay evidencia suficiente para sustentar la aseveración de una correlación lineal significativa entre la temperatura y el consumo de energía.
   b) $\hat{y} = 4278 - 23.9x$.
   c) 2844 kilowatts/hora.

Créditos

Fotografías

Las siguientes fotos tienen copyright PhotoDisc, Inc.

Capítulo 1
p. 5,
p. 6,
p. 13,
p. 17,

Capítulo 2
p. 61,
p. 66,
p. 75,

Capítulo 3
p. 126,
p. 127,
p. 150,
p. 153,
p. 159,
p. 163,
p. 165,
p. 166,

Capítulo 4
p. 188,
p. 190,
p. 198,
p. 218,

Capítulo 5
p. 229,
p. 273,

Capítulo 6
p. 300,
p. 320,

Capítulo 7
p. 371,
p. 374,
p. 393,

Capítulo 8
p. 443,
p. 445,
p. 457,
p. 458,
p. 468,
p. 480,

Capítulo 9
p. 498,
p. 520,
p. 544,
p. 546,

Capítulo 10
p. 586,

Capítulo 12
p. 659,

Ilustraciones

Las siguientes ilustraciones laterales fueron elaboradas por Bob Giuliani.

Capítulo 1
p. 16,
p. 22,

Capítulo 2
p. 42,
p. 52,
p. 60,
p. 95,

Capítulo 3
p. 128,
p. 136,
p. 157,

Capítulo 4
p. 201,
p. 202,

Capítulo 6
p. 319,
p. 331,
p. 332,

Capítulo 7
p. 409,
p. 422,

Capítulo 8
p. 469,

Capítulo 9
p. 522,

Capítulo 10
p. 583,

Capítulo 11
p. 607,

Capítulo 13
p. 699,
p. 711,

Las siguientes ilustraciones laterales fueron elaboradas por James Bryant.

Capítulo 3
p. 140,
p. 144,
p. 151,
p. 164,

Capítulo 5
p. 243,
p. 260,

Capítulo 6
p. 302,

Capítulo 9
p. 501,

Capítulo 12
p. 644,

Capítulo 13
p. 710,

Las siguientes ilustraciones fueron elaboradas por Darwen Hennings.

Capítulo 5
p. 242,

Tufte Illustration

p. 53,

Índice

Tabla de símbolos

$\overline{A}$	Complemento del suceso A
H_0	Hipótesis nula
H_1	Hipótesis alternativa
α	Alfa; probabilidad de un error tipo I o el área de la región crítica
β	Beta; probabilidad de un error tipo II
r	Coeficiente de correlación lineal muestral
ρ	Ro; coeficiente de correlación lineal poblacional
r^2	Coeficiente de determinación
R^2	Coeficiente de determinación múltiple
r_s	Coeficiente de correlación de rangos de Spearman
b_1	Estimado puntual de la pendiente de la recta de regresión
b_0	Estimado puntual del intercepto y de la recta de regresión
$\hat{y}$	Valor predicho de y
d	Diferencia entre dos valores apareados
$\overline{d}$	Media de las diferencias d calculada a partir de los datos muestrales apareados
s_d	Desviación estándar de las diferencias d calculada a partir de los datos muestrales apareados
s_e	Error estándar de un estimado
T	Suma de rangos; utilizada en la prueba de rangos con signo de Wilcoxon
H	Estadístico de prueba de Kruskal-Wallis
R	Suma de rangos para una muestra; utilizada en la prueba de la suma de rangos de Wilcoxon
μ_R	Rango medio esperado; utilizado en la prueba de la suma de rangos de Wilcoxon
σ_R	Desviación estándar de rangos esperada; utilizada en la prueba de la suma de rangos de Wilcoxon
G	Número de rachas en la prueba de rachas para detectar aleatoriedad
μ_G	Media esperada del número de rachas; utilizado en la prueba de rachas para detectar aleatoriedad
σ_G	Desviación estándar esperada para el número de rachas; utilizada en la prueba de rachas para detectar aleatoriedad
$\mu_{\bar{x}}$	Media poblacional de todas las medias muestrales posibles.
$\sigma_{\bar{x}}$	Desviación estándar de la población de todas las medias muestrales posibles.
E	Margen de error del estimado de un parámetro poblacional, o valor esperado
Q_1, Q_2, Q_3	Cuartiles
$D_1, D_2, \ldots, D_9$	Deciles
$P_1, P_2, \ldots, P_{99}$	Percentiles
x	Valor de datos

Tabla de símbolos

f	Frecuencia con la que ocurre un valor
Σ	Sigma mayúscula; sumatoria
$\boldsymbol{\sigma} x$	Suma de los valores
Σx^2	Suma de los cuadrados de los valores
$(\Sigma x)^2$	Cuadrado de la suma de todos los valores
$\boldsymbol{\sigma} xy$	Suma de los productos de cada valor x multiplicado por el valor y correspondiente
n	Número de valores en una muestra
$n!$	n factorial
N	Número de valores en una población finita; también se utiliza como el tamaño de todas las muestras combinadas
k	Número de muestras o poblaciones o categorías
$\bar{x}$	Media de los valores en una muestra
$\bar{R}$	Media de los rangos muestrales
μ	Mu; media de todos los valores en una población
s	Desviación estándar de un conjunto de valores muestrales
σ	Sigma minúscula; desviación estándar de todas las variables en una población
s^2	Varianza de un conjunto de datos muestrales
σ^2	Varianza de todos los valores en una población
z	Puntuación estándar
$z_{\alpha/2}$	Valor crítico de z
t	Distribución t
$t_{\alpha/2}$	Valor crítico de t
df	Número de grados de libertad
F	Distribución F
χ^2	Distribución chi cuadrada
χ^2_R	Valor crítico de cola derecha de chi cuadrada
χ^2_L	Valor crítico de cola izquierda de chi cuadrada
p	Probabilidad de un suceso o la proporción poblacional
q	Probabilidad de una proporción, igual a $1 - p$
$\hat{p}$	Proporción muestral
$\hat{q}$	Proporción muestral igual a $1 - \hat{p}$
$\bar{p}$	Proporción obtenida por agrupar dos muestras
$\bar{q}$	Proporción de probabilidad igual a $1 - \bar{p}$
$P(A)$	Probabilidad de un suceso A
$P(A'' B)$	Probabilidad de un suceso A, suponiendo que el suceso B ya ocurrió
${}_nP_r$	Número de permutaciones de n elementos seleccionando r elementos a la vez
${}_nC_r$	Número de combinaciones de n elementos seleccionando r elementos a la vez