Mario F. Triola

ESTADÍSTICA

DECIMOSEGUNDA EDICIÓN

Pearson

12A EDICIÓN

ESTADÍSTICA

12A EDICIÓN

ESTADÍSTICA

MARIO F. TRIOLA

Con la colaboración especial de

Laura Lossi,

Broward College

TRADUCCIÓN

Jesús Elmer Murrieta Murrieta

Maestro en Investigación de Operaciones

Tecnológico de Monterrey

REVISIÓN TÉCNICA

Gerardo Montes Sifuentes

Universidad Regiomontana

Instituto de Especialización de Ejecutivos – Campus Monterrey

Alberto de la Rosa Elizalde

Facultad de Contaduría y Administración

Facultad de Ciencias

Universidad Nacional Autónoma de México

Julio Sergio Acosta Rodríguez

Facultad de Contaduría y Administración

Universidad Nacional Autónoma de México

Pearson

Datos de catalogación bibliográfica

MARIO F. TRIOLA
Estadística
Decimosegunda edición
Pearson Educación de México, S.A. de C.V., 2018
ISBN: 978-607-32-4378-0
Área: Matemáticas
Formato: 21 × 27 cm Páginas: 784

Estadística

Authorized translation from the English Language edition entitled *Elementary Statistics, 13th Edition*, by *Mario F. Triola*, published by Pearson Education, Inc., Copyright © 2018. All rights reserved. ISBN 9780134462455

Traducción autorizada de la edición en idioma inglés titulada *Elementary Statistics, 13th Edition*, por *Mario F. Triola*, publicada por Pearson Education, Inc., Copyright © 2018. Todos los derechos reservados.

Edición en español
Director general: Sergio Fonseca ■ **Director de innovación y servicios educativos:** Alan David Palau ■ **Gerente de contenidos y servicios editoriales:** Jorge Luis Íñiguez ■ **Coordinador de desarrollo de contenidos:** Lilia Moreno ■ **Editora especialista en contenidos de aprendizaje:** Rosa Díaz Sandoval ■ **Coordinador de arte y diseño:** Mónica Galván ■ **Editor de desarrollo:** Bernardino Gutiérrez Hernández ■ **Traductor:** Jesús Elmer Murrieta Murrieta ■ **Corrector de estilo:** César Romero ■ **Gestor de arte y diseño:** José Hernández Garduño ■ **Lector de pruebas:** Felipe Martínez ■ **Composición y diagramación:** Servicios Editoriales 6Ns.

Esta edición en español es la única autorizada.

Contacto: soporte@pearson.com

Decimosegunda edición, 2018

ISBN LIBRO IMPRESO: 978-607-32-4378-0
ISBN LIBRO E-BOOK: 978-607-32-4377-3

Impreso en México. *Printed in Mexico*.

1 2 3 4 5 6 7 8 9 0 – 21 20 19 18

D.R. © 2018 por Pearson Educación de México, S.A. de C.V.
Avenida Antonio Dovalí Jaime núm. 70
Torre B, Piso 6, Colonia Zedec, Ed. Plaza Santa Fe
Delegación Álvaro Obregón, México, Ciudad de México, C. P. 01210

www.pearsonenespañol.com

Pearson

Pearson Hispanoamérica
Argentina ■ Belice ■ Bolivia ■ Chile ■ Colombia ■ Costa Rica ■ Cuba ■ República Dominicana ■ Ecuador ■ El Salvador ■ Guatemala ■ Honduras ■ México ■ Nicaragua ■ Panamá ■ Paraguay ■ Perú ■ Uruguay ■ Venezuela

AGRADECIMIENTOS A LA EDICIÓN EN ESPAÑOL

Pearson Educación agradece a los centros de estudio y profesores usuarios de esta obra por su apoyo y retroalimentación, elemento fundamental para el logro de esta nueva edición de *Estadística*.

MÉXICO

Instituto Tecnológico de Culiacán
José Luis Cázarez Contreras

Instituto Tecnológico y de Estudios Superiores de Monterrey
Campus Ciudad de México
Carlos A. Díaz Tufinio
Francisco Javier Hernández Moreno

Universidad Autónoma de San Luis Potosí
Marco Antonio Villa Cerda

Universidad Nacional Autónoma de México
Facultad de Contaduría y Administración
Francisco A. Piña Salazar

Universidad Nacional Autónoma de México
Facultad de Estudios Superiores Cuautitlán
Jorge Altamira Ibarra

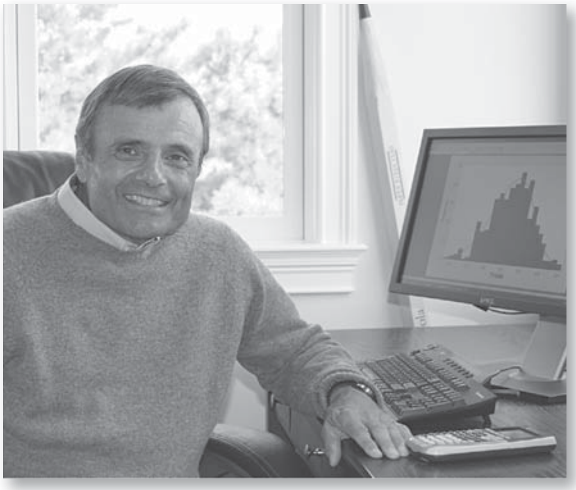
Universidad de Monterrey
Roberto Hernández Ramírez

Costa Rica

Instituto Tecnológico de Costa Rica
Francisco Navarro Henríquez

Para Ginny
Marc, Dushana y Marisa
Scott, Anna, Siena y Kaia

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*, quinta edición, *Elementary Statistics Using Excel*, sexta edición y *Elementary Statistics Using the TI-83/84 Plus Calculator*, cuarta edición; también es coautor de los libros *Biostatistics for the Biological and Health Sciences*, segunda edición, *Statistical Reasoning for Everyday Life*, quinta edición, y *Business Statistics*. En la actualidad existe una edición internacional de *Estadística* que ha sido traducida a varios idiomas. Marty diseñó el software estadístico Statdisk original y ha escrito diversos manuales y libros de trabajo para educación en estadística con apoyos tecnológicos. Asimismo, ha sido orador en muchas conferencias y universidades. Su trabajo de consultoría incluye el diseño de máquinas tragamonedas para casinos y de cañas de pescar; ha trabajado con abogados en la determinación de probabilidades en casos de demandas de paternidad, en la identificación de desigualdades salariales entre géneros, en el análisis de datos de demandas por malas prácticas médicas y en el análisis de resultados de elecciones en disputa. También ha utilizado métodos estadísticos para analizar encuestas de escuelas de medicina y los resultados de una encuesta para la oficina de movilidad de la Ciudad de Nueva York. Por otro lado, Marty también ha fungido como testigo experto en la Suprema Corte del estado de Nueva York. La Text and Academic Authors Association otorgó a Mario F. Triola el premio Texty a la Excelencia por su trabajo en el libro *Estadística*.

CONTENIDO

PREFACIO

La estadística permea casi todos los aspectos de nuestras vidas. Desde los sondeos de opinión hasta las pruebas clínicas en medicina, los automóviles autoconducidos, los drones y la seguridad biométrica, la estadística influye y da forma al mundo que nos rodea. El presente libro, *Estadística,* forja la relación entre la estadística y nuestro mundo mediante el uso extensivo de una amplia variedad de aplicaciones reales que dan vida a la teoría y a los métodos presentados.

Objetivos de esta nueva edición

- Fomentar el crecimiento personal de los estudiantes a través del pensamiento crítico, el uso de la tecnología, el trabajo en equipo y el desarrollo de sus habilidades de comunicación.
- Incorporar los mejores y más novedosos métodos utilizados por los estadísticos profesionales.
- Incorporar características que aborden todas las recomendaciones incluidas en las *Guidelines for Assesment and Instruction in Statistics Education* (GAISE) según lo recomendado por la American Statistical Association.
- Proporcionar una gran cantidad de nuevos e interesantes datos, conjuntos de ejemplos y series de ejercicios; como los que involucran seguridad biométrica, ciberseguridad, drones y velocidades de datos en teléfonos inteligentes.
- Mejorar la enseñanza y el aprendizaje con el mejor y más amplio conjunto de complementos y recursos digitales.

Público / requisitos necesarios

El libro *Estadística* se escribió para estudiantes de cualquier carrera. Aun cuando el uso del álgebra es mínimo se recomienda que los estudiantes hayan cursado al menos una materia de álgebra elemental o que aprendan los componentes básicos del álgebra a través de un curso integrado. En muchos casos se incluyen teorías subyacentes, pero este libro no enfatiza el rigor matemático que es más adecuado para carreras en matemáticas.

Características distintivas

Se ha tenido mucho cuidado de asegurar que cada capítulo ayude a los estudiantes a comprender los conceptos presentados. Las siguientes características se diseñaron para lograr el objetivo de la comprensión conceptual.

Datos reales

Cientos de horas se han dedicado a encontrar datos que sean reales, significativos e interesantes para los estudiantes. 94% de los ejemplos y 92% de los ejercicios se basan en datos reales. Algunos ejercicios se refieren a los 32 conjuntos de datos listados en el apéndice B, y 12 de esos conjuntos son nuevos en esta edición. Los ejercicios que requieren el uso de los conjuntos de datos del apéndice B se ubican hacia el final de cada serie de ejercicios y están marcados con un icono especial .

En todo el libro se presentan conjuntos de datos reales para proporcionar aplicaciones estadísticas relevantes e interesantes del mundo real, incluyendo seguridad biométrica, automóviles autoconducidos, velocidades de datos en teléfonos inteligentes y uso de drones para entrega de productos. El apéndice B incluye descripciones de los 32 conjuntos de datos que se pueden descargar desde el sitio del autor, www.TriolaStats.com; también pueden consultarse los enlaces de descarga en el sitio web de este libro: www.pearsonenespañol.com/triola.

El sitio del autor, incluye conjuntos de datos descargables en formatos compatibles con tecnologías como Excel, Minitab, JMP, SPSS y las calculadoras TI-83/84 Plus. Los conjuntos de datos también se incluyen en el software gratuito Statdisk, disponible en el sitio del autor.

Legibilidad

Se ha tenido gran cuidado, entusiasmo y pasión para crear un libro legible, comprensible, interesante y relevante. Los estudiantes de cualquier carrera pueden estar seguros de que encontrarán aplicaciones relacionadas con su trabajo futuro.

Material adicional

Este libro de texto cuenta con el apoyo de material adicional en inglés para instructores que adopten el libro en sus cursos y para sus estudiantes. En el sitio de este libro puede consultar cómo obtener esos recursos, que van desde el manual de soluciones y presentaciones de clase para los instructores hasta el conjunto de datos del apéndice B, manuales y software estadístico para los estudiantes que lleven este libro como texto en sus cursos.

Además, el autor ha puesto al alcance de los usuarios de sus libros el sitio www.TriolaStats.com, que se actualiza continuamente para proporcionar los más recientes recursos digitales para la serie Estadística de Triola, entre los que se encuentran:

- Statdisk: Un paquete de software estadístico robusto y gratuito diseñado especialmente para este libro.
- Conjuntos de datos descargables del apéndice B en una variedad de formatos tecnológicos.
- Complementos descargables del libro de texto, incluyendo la sección 4-5 *Probabilidades mediante simulaciones, un glosario de términos estadísticos, así como fórmulas y tablas*.
- Videos de enseñanza en línea creados específicamente para esta edición que proporcionan instrucciones paso a paso en el uso de la tecnología.
- El blog de Triola que destaca aplicaciones actuales de la estadística, la estadística en las noticias y recursos en línea.
- Vínculo de contacto que proporciona acceso con un solo clic, para que los profesores y estudiantes hagan preguntas y comentarios al autor, Marty Triola.

Características de los capítulos

Características al inicio de los capítulos

- Los capítulos inician con un ***Problema del capítulo*** que utiliza datos reales y da sentido al material del capítulo.
- Los ***Objetivos del capítulo*** proporcionan un resumen de las metas de aprendizaje para cada sección del capítulo.

Ejercicios Muchos ejercicios requieren la *interpretación* de los resultados. Se ha tenido gran cuidado en asegurar su utilidad, relevancia y precisión. Los ejercicios se organizan en orden creciente de dificultad y se dividen en dos grupos: (1) *Habilidades y conceptos básicos* y (2) *Más allá de lo básico*. Los ejercicios que están *Más allá de lo básico* abordan conceptos más difíciles o requieren una formación matemática más sólida. En algunos casos, estos ejercicios introducen un nuevo concepto.

Características al final de los capítulos

- El ***Examen rápido del capítulo*** proporciona 10 preguntas que requieren respuestas breves.

- Los ***Ejercicios de repaso*** ofrecen prácticas sobre los conceptos y procedimientos presentados en el capítulo.
- Los ***Ejercicios de repaso acumulado*** refuerzan el material estudiado previamente.
- El ***Proyecto de tecnología*** ofrece una actividad que puede ser utilizada con una variedad de tecnologías.
- ***De los datos a la decisión*** es un problema concluyente que requiere pensamiento crítico y redacción.
- Las ***Actividades en equipo*** fomentan el aprendizaje activo en grupos.

Otras características

Ensayos al margen Existen 106 ensayos al margen diseñados para resaltar temas del mundo real y fomentar el interés de los estudiantes. También hay muchos artículos del tipo *En cifras* que describen brevemente números o estadísticas interesantes.

Diagramas de flujo El texto incluye diagramas de flujo que simplifican y aclaran conceptos y procedimientos más complejos.

Fórmulas y tablas En el apéndice A y al final del libro se presentan varias tablas útiles en estadística. Además de estas tablas, existen guías resumidas de referencia rápida que se pueden obtener en el sitio del autor o a través de los vínculos que se encuentran en el sitio web de este libro.

Integración tecnológica

Al igual que en la edición anterior, a lo largo del libro se presentan muchas pantallas de tecnología, y algunos ejercicios se basan en los resultados mostrados en ellas. Cuando resulta apropiado, las secciones terminan con una nueva subsección llamada **Centro de tecnología** que incluye videos específicos de la tecnología e instrucciones detalladas para Statdisk, Minitab®, Excel®, StatCrunch o una calculadora TI-83/84 Plus®. (En este texto se utiliza "TI-83/84 Plus" para identificar una calculadora TI-83 Plus o TI-84 Plus). Las características al final de los capítulos incluyen un *Proyecto de tecnología*.

El paquete de software estadístico Statdisk se diseñó específicamente para este texto y contiene todos los conjuntos de datos del apéndice B. Statdisk es gratuito para los usuarios de este libro y se puede descargar en www.Statdisk.org. Consulte el sitio web del libro para mayor información sobre los recursos adicionales y para indicaciones de cómo comprar StatCrunch.

Cambios en esta edición

Características nuevas

Los ***Objetivos del capítulo*** proporcionan un resumen de las metas de aprendizaje para cada sección del capítulo.

Su turno: Muchos ejemplos incluyen una nueva característica llamada "su turno" que guía a los estudiantes hacia un ejercicio relevante para que puedan aplicar inmediatamente lo que acaban de aprender en el ejemplo.

Centro de tecnología: Instrucciones mejoradas sobre tecnología, con el apoyo de videos creados personalmente por el autor y contenido descargable desde su sitio web.

Videos de tecnología: Los nuevos videos sobre tecnología, dirigidos por el autor, proporcionan detalles paso a paso para procedimientos estadísticos clave con Excel, calculadoras TI-83/84 y Statdisk.

Conjuntos de datos más grandes: Algunos de los conjuntos de datos del apéndice B son mucho más grandes que en ediciones anteriores. Ya no resulta práctico imprimir todos los conjuntos de datos en este libro, por lo que se describen en el apéndice B y se pueden descargar directamente del sitio del autor o puede consultar los enlaces de descarga en el sitio web del libro.

Nuevo contenido: Los nuevos ejemplos, ejercicios y problemas de capítulo proporcionan aplicaciones estadísticas relevantes e interesantes del mundo real, incluyendo seguridad biométrica, automóviles autoconducidos, velocidades de datos en teléfonos inteligentes y uso de drones para la entrega de productos.

	Número	Nuevo en esta edición	Uso de datos reales
Ejercicios	1756	81% (1427)	92% (1618)
Ejemplos	211	73% (153)	94% (198)
Problemas de capítulo	14	93% (13)	100% (14)

Cambios en la organización

Nuevos objetivos de capítulo: Ahora todos los capítulos comienzan con una lista de metas de aprendizaje clave para ese capítulo. Los *Objetivos del capítulo* reemplazan la antigua sección numerada de *Repaso y panorama general*. La primera sección numerada de cada capítulo cubre ahora un tema importante relativo al capítulo.

Nueva subsección 1-3, Parte 2: *Datos grandes y datos faltantes: Demasiado y no suficientes*

Nueva sección 2-4: *Diagramas de dispersión, correlación y regresión*

La edición anterior contenía diagramas de dispersión en el capítulo 2, pero esta nueva sección incluye los diagramas de dispersión en la parte 1, el coeficiente de correlación lineal r en la parte 2 y la regresión lineal en la parte 3. Estas adiciones están destinadas a facilitar enormemente la cobertura para aquellos profesores que prefieren una cobertura temprana de los conceptos de correlación y regresión. El capítulo 10 incluye estos temas, analizados con mucho mayor detalle.

Nueva subsección 4-3*, *Parte 3: *Teorema de Bayes*

Nueva sección 7-4: *Bootstrapping: Uso de la tecnología para realizar estimaciones*

Secciones combinadas:

- **4-2: Regla de la suma y regla de la multiplicación**
  Combina las secciones 4-3 (*Regla de la suma*) y 4-4 (*Regla de la multiplicación: fundamentos*) de la edición anterior.
- **5-2: Distribuciones binomiales de probabilidad**
  Combina las secciones 5-3 (*Distribuciones binomiales de probabilidad*) y 5-4 (*Parámetros para distribuciones binomiales*) de la edición anterior.

Secciones eliminadas:
La sección 15-2 (*Proyectos*) ahora es un inserto en la edición del profesor y complementa el primer conjunto de *Actividades de cooperación en equipo* del capítulo 1. Las secciones 15-3 (*Procedimientos*) y 15-4 (*Perspectivas*) se han eliminado.

Terminología modificada

Significativo: Las referencias a los resultados “inusuales”, en la edición anterior, se describen ahora en términos de “significativamente bajo” o “significativamente alto”, de modo que el vínculo con la prueba de hipótesis se refuerce aún más.

Regla de conteo de la multiplicación: Las referencias en la sección 4-4 (*Conteo*) a la “regla fundamental del conteo” ahora usan la “regla de conteo de la multiplicación” para que su nombre haga una mejor sugerencia de cómo se aplica.

Plan de estudios flexible

La organización de este libro refleja las preferencias de la mayoría de los profesores de estadística, pero hay dos variaciones comunes:

- ***Cobertura temprana de la correlación y la regresión:*** Algunos profesores prefieren cubrir los fundamentos de la correlación y la regresión a inicios del curso. La sección 2-4 ahora incluye conceptos básicos de diagramas de dispersión, correlación y regresión sin el uso de fórmulas y sin tanta profundidad como en las secciones 10-1 (*Correlación*) y 10-2 (*Regresión*).

- ***Probabilidad al mínimo:*** Algunos profesores prefieren una amplia cobertura de la probabilidad, mientras que otros optan por incluir sólo conceptos básicos. Quienes prefieren una cobertura mínima pueden incluir la sección 4-1 y omitir las secciones restantes del capítulo 4, puesto que no son esenciales para los capítulos siguientes. Quienes prefieren cubrir los fundamentos de probabilidad junto con los fundamentos de las reglas de la suma y la multiplicación deben ver la sección 4-2.

Directrices para la evaluación y enseñanza de la estadística Este libro refleja las recomendaciones de la American Statistical Association y sus *Guidelines for Assessment and Instruction in Statistics Educaction* (GAISE). Esas directrices sugieren los siguientes objetivos y estrategias.

1. ***Énfasis en el conocimiento estadístico y el desarrollo del pensamiento crítico:*** Cada ejercicio de sección inicia con ejercicios de *Conocimiento estadístico y pensamiento crítico*. Muchos de los ejercicios del libro están diseñados para fomentar el pensamiento estadístico en lugar del uso ciego de procedimientos mecánicos.
2. ***Utilización de datos reales:*** 94% de los ejemplos y 92% de los ejercicios usan datos reales.
3. ***Realce de la comprensión conceptual más que del mero conocimiento de los procedimientos:*** En vez de buscar respuestas numéricas sencillas, la mayoría de los ejercicios y ejemplos implican la comprensión conceptual a través de preguntas que alientan las interpretaciones prácticas de los resultados. Además, cada capítulo incluye un proyecto del tipo *De los datos a la decisión*.
4. ***Fomento del aprendizaje activo en el aula:*** Cada capítulo termina con varias actividades de cooperación en equipo.
5. ***Uso de la tecnología para desarrollar la comprensión conceptual y analizar los datos:*** Se incluyen pantallas de software a lo largo del libro. En especial, las subsecciones *Centro de tecnología* incluyen instrucciones sobre cómo usar el software. Cada capítulo incluye un *Proyecto de tecnología*. Cuando hay discrepancias entre las respuestas basadas en tablas y las respuestas basadas en la tecnología consulte el apéndice D, el cual proporciona *ambas* respuestas. El autor ofrece en su sitio un software gratuito específico para este texto (Statdisk), conjuntos de datos formateados para varias tecnologías y videos de enseñanza de las tecnologías.
6. ***Uso de evaluaciones para mejorar y evaluar el aprendizaje de los estudiantes:*** Las herramientas de evaluación incluyen una abundancia de ejercicios por sección, *exámenes rápidos del capítulo*, *ejercicios de repaso del capítulo*, *proyectos de tecnología*, *proyectos de los datos a la decisión* y *actividades en equipo*.

Reconocimientos

Estoy muy agradecido con los miles de profesores y estudiantes de estadística que han contribuido al éxito de este libro. Agradezco a los revisores por sus sugerencias a esta nueva edición: Eric Gorenstein, Bunker Hill Community College; Rhonda Hatcher, Texas Christian University; Ladorian Latin, Franklin University; Joseph Pick, Palm Beach State College; y Lisa Whitaker, Keiser University. Un agradecimiento especial a Laura Iossi del Broward College por su extenso trabajo de revisión y contribución a esta edición.

Entre los revisores más recientes están Raid W. Amin, University of West Florida; Robert Black, United States Air Force Academy; James Bryan, Merced College; Donald Burd, Monroe College; Keith Carroll, Benedictine University; Monte Cheney, Central Oregon Community College; Christopher Donnelly, Macomb Community College; Billy Edwards, University of

Tennessee—Chattanooga; Marcos Enriquez, Moorpark College; Angela Everett, Chattanooga State Technical Community College; Joe Franko, Mount San Antonio College; Rob Fusco, Broward College; Sanford Geraci, Broward College; Laura Heath, Palm Beach State College; Richard Herbst, Montgomery County Community College; Richard Hertz; Diane Hollister, Reading Area Community College; Michael Huber, George Jahn, Palm Beach State College; Gary King, Ozarks Technical Community College; Kate Kozak, Coconino Community College; Dan Kumpf, Ventura College; Mickey Levendusky, Pima County Community College; Mitch Levy, Broward College; Tristan Londre, Blue River Community College; Alma Lopez, South Plains College; Kim McHale, Heartland Community College; Carla Monticelli, Camden County Community College; Ken Mulzet, Florida State College en Jacksonville; Julia Norton, California State University Hayward; Michael Oriolo, Herkimer Community College; Jeanne Osborne, Middlesex Community College; Ali Saadat, University of California—Riverside; Radha Sankaran, Passaic County Community College; Steve Schwager, Cornell University; Pradipta Seal, Boston University; Kelly Smitch, Brevard College; Sandra Spain, Thomas Nelson Community College; Ellen G. Stutes, Louisiana State University, Eunice; Sharon Testone, Onondaga Community College; Chris Vertullo, Marist College; Dave Wallach, University of Findlay; Cheng Wang, Nova Southeastern University; Barbara Ward, Belmont University; Richard Weil, Brown College; Gail Wiltse, St. John River Community College; Claire Wladis, Borough of Manhattan Community College; Rick Woodmansee, Sacramento City College; Yong Zeng, University of Missouri en Kansas City; Jim Zimmer, Chattanooga State Technical Community College; Cathleen Zucco-Teveloff, Rowan University; Mark Z. Zuiker, Minnesota State University, Mankato.

Esta nueva edición de *Estadística* es en realidad un esfuerzo de equipo, y me considero afortunado de trabajar con el dedicado y comprometido equipo de Pearson. Agradezco a Suzy Bainbridge, Justin Billing, Deirdre Lynch, Peggy McMahon, Vicki Dreyfus, Christine O'Brien, Joe Vetere, y Rose Keman de Cenveo Publisher Services.

Agradezco especialmente a Marc Triola, M. D., de la New York University School of Medicine, por su excelente trabajo en la creación de nueva edición del software Statdisk. Doy las gracias a Scott Triola por su gran ayuda durante todo el proceso de producción de esta edición.

Agradezco a las siguientes personas por su ayuda en la revisión de la precisión del texto y de las respuestas en esta edición: James Lapp, Paul Lorczak y Dirk Tempelaar.

M. F. T.
Madison, Connecticut
Septiembre 2016

Recursos tecnológicos (en inglés)

Los siguientes recursos se pueden encontrar en el sitio web del autor o bien puede consultar los enlaces de descarga en el sitio web del libro.

- Conjuntos de datos del apéndice B formateados para Minitab, SPSS, SAS, Excel, JMP y como archivos de texto. Además, estos conjuntos de datos están disponibles como APP y programas suplementarios para las calculadoras TI-83/84 Plus.
- Instrucciones descargables para el software estadístico Statdisk. Las nuevas características incluyen la capacidad de utilizar directamente listas de datos en vez de requerir el uso de sus estadísticas resumidas.
- Conjuntos de datos adicionales, *Probabilidades a través de simulaciones*, *Teorema de Bayes*, un índice de aplicaciones y una tabla de símbolos.

Se han ampliado y actualizado los **Recursos en video** y ahora dan soporte a la mayor parte de las secciones de este libro con muchos temas presentados por el autor. Los videos, en inglés, apoyan tanto a los estudiantes como a los profesores mediante clases frente al grupo, reforzando los fundamentos estadísticos a través de la tecnología y aplicando conceptos como:

- **Sección de videos de clase frente al grupo**
- **Videos de los ejercicios de repaso del capítulo** que guían a los estudiantes a través de los ejercicios y les ayudan a entender los conceptos clave del capítulo.
- **¡Nuevo! Tutoriales de tecnología en video.** Estos videos cortos y novedosos enseñan cómo usar Excel, StatDisk y la calculadora gráfica TI para resolver los ejercicios.
- **Videos StatTalk: 24 videos conceptuales que ayudan a comprender a profundidad la estadística.** El divertido estadístico Andrew Vickers toma las calles de Brooklyn, NY, para demostrar conceptos estadísticos importantes mediante historias interesantes y eventos de la vida real. Estos divertidos y atractivos videos le ayudarán a comprender los conceptos estadísticos. Estos recursos están disponibles junto con una guía de instrucciones y preguntas de evaluación para los profesores que lleven este libro en sus cursos.

Los videos contienen la opción de subtítulos en inglés y español.

Los siguientes recursos están disponibles para su compra desde Estados Unidos:

Minitab® 17 y **Minitab Express**™ facilitan el aprendizaje de la estadística y proporcionan a los estudiantes un conjunto de habilidades requeridas en la fuerza de trabajo actual. El paquete de software Minitab® con materiales educativos garantiza que los estudiantes tengan acceso al software que requerirán en el aula, en el campus y en casa. La disposición de las versiones más actuales de Minitab 17 y Minitab Express asegura que los estudiantes puedan utilizar el software durante todo su curso.

ISBN 13: 978-0-13-445640-9
ISBN 10: 0-13-445640-8 (sólo con tarjeta de acceso).

JMP Student Edition, Versión 12 es una versión simplificada y fácil de usar del novedoso software estadístico JMP del SAS Institute, Inc., y está disponible para su utilización conjunta con este texto.

ISBN-13: 978-0-13-467979-2
ISBN-10: 0-13-467979-2

StatCrunch es un software de estadística online que se puede adquirir en https://www.statcrunch.com/.

1

INTRODUCCIÓN A LA ESTADÍSTICA

PROBLEMA DEL CAPÍTULO

Pregunta de encuesta: ¿Prefiere leer un libro *impreso* o un libro *electrónico*?

Las encuestas proveen datos que nos permiten mejorar los productos o servicios. Las encuestas guían a los candidatos políticos, modelan las prácticas comerciales, influyen en los medios sociales y afectan muchos aspectos de nuestras vidas. Las encuestas nos permiten percibir con claridad las opiniones y los puntos de vista de los demás. Consideremos una encuesta de *USA Today* en la que se preguntó a los encuestados si preferían leer un libro impreso o uno electrónico. De 281 sujetos, 65% prefirió un libro impreso y 35% prefirió un libro electrónico. La figura 1-1 de la página siguiente presenta gráficas con estos resultados.

Los resultados de la encuesta sugieren que, marcadamente, las personas prefieren leer libros impresos a leer libros electrónicos. Las gráficas de la figura 1-1 representan de manera visual los resultados de la encuesta y apoyan una afirmación de que la gente prefiere los libros impresos a los libros electrónicos por un amplio margen. Uno de los objetivos más importantes de este libro es fomentar el uso del pensamiento crítico para que tales resultados no sean aceptados de forma irreflexiva. Podríamos cuestionar si los resultados de la encuesta son válidos. ¿Quién llevó a cabo la encuesta? ¿Cómo fueron seleccionados los encuestados? ¿Las gráficas de la figura 1-1 representan los resultados con fidelidad o de alguna manera son engañosas?

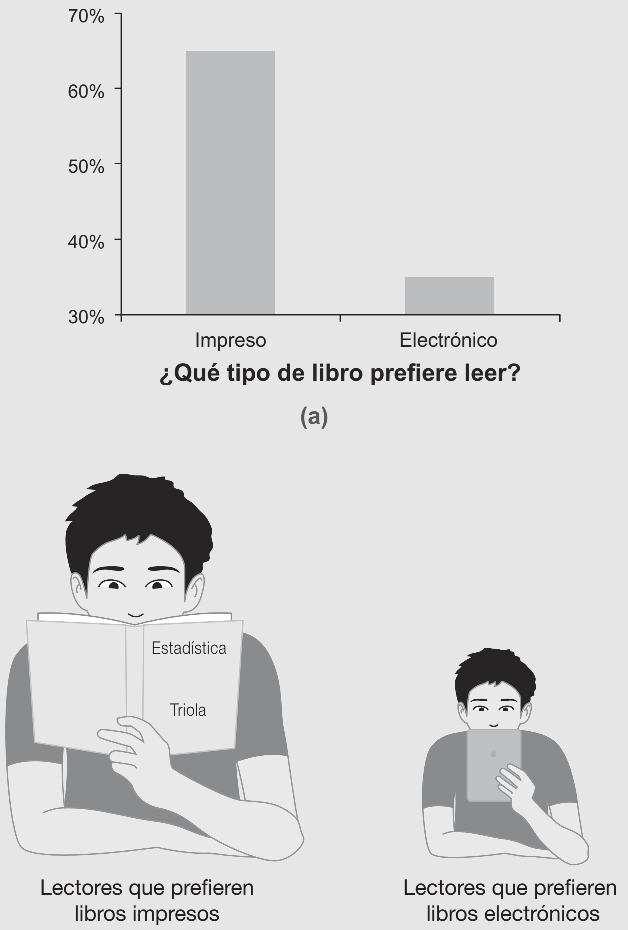
Los resultados de la encuesta presentan defectos relevantes que suelen ser frecuentes, por lo que resulta especialmente importante reconocerlos. A continuación se presentan descripciones breves de cada uno de ellos:

Defecto 1: Gráficas engañosas La gráfica de barras de la figura 1-1(a) es muy engañosa. Mediante el uso de una escala vertical que no empieza en cero, la diferencia entre los dos porcentajes se exagera. La figura 1-1(a) hace parecer que alrededor de ocho veces más personas eligen un libro impreso sobre un libro electrónico, aunque la proporción real es aproximadamente 2:1, no 8:1 (las proporciones en las respuestas son 65 y 35%).

La ilustración de la figura 1-1(b) también es engañosa. Una vez más, la diferencia entre las proporciones reales de respuesta de 65% para los libros impresos y 35% para los libros electrónicos se observa muy distorsionada. La gráfica de imagen (o "pictograma") de la figura 1-1(b) hace parecer que las personas prefieren los libros impresos a los electrónicos en una proporción de aproximadamente 4:1, en vez de en la proporción correcta de 65:35, o aproximadamente 2:1. (Los objetos con área o volumen pueden distorsionar las percepciones porque es posible dibujarlos desproporcionadamente mayores o menores de lo que indican los datos).

Las gráficas engañosas se analizan con más detalle en la sección 2-3, pero aquí se observa que las ilustraciones de la figura 1-1 exageran la preferencia por los libros impresos.

Defecto 2: Mal método de muestreo Las respuestas de la encuesta antes mencionada provienen de un sondeo de *USA Today* entre sus usuarios de Internet. La pregunta de la encuesta se publicó en su sitio web y sus usuarios decidieron responder. Este es un ejemplo de una muestra de respuesta voluntaria en la que los encuestados deciden participar. Con una muestra de respuesta voluntaria, suele ocurrir que aquellos con un fuerte interés en el tema son los más propensos a participar, por lo que los resultados son muy cuestionables. En este caso, es razonable sospechar que los usuarios de Internet podrían preferir, más bien, libros electrónicos en una mayor proporción que la población general. Cuando se usan datos muestrales para conocer algo sobre una población, es extremadamente importante obtener



(b)

FIGURA 1-1 Resultados de la encuesta

información que sea representativa de la población de la que se extrae. A medida que avance en este capítulo y se analicen los tipos de datos y métodos de muestreo, será importante concentrarse en los siguientes conceptos clave:

- **Los datos muestrales deben recopilarse de una manera apropiada, por ejemplo a través de un proceso de selección aleatoria.**
- **Si los datos muestrales no se recopilan de manera apropiada, pueden ser tan inútiles que ningún tipo de tratamiento estadístico pueda salvarlos.**

Sería fácil aceptar los resultados de la encuesta anterior y proceder ciegamente con los cálculos y análisis estadísticos, pero no se tendrían en cuenta los dos defectos críticos descritos previamente. Entonces sería probable sacar conclusiones erróneas y engañosas. En vez de esto, es necesario desarrollar habilidades de pensamiento estadístico y crítico para entender por qué la encuesta es esencialmente defectuosa.

OBJETIVOS DEL CAPÍTULO

El concepto más importante presentado en este capítulo es el siguiente: cuando se utilizan métodos estadísticos con datos muestrales para obtener conclusiones sobre una población, es esencial recopilar los datos de muestra en forma apropiada. Los objetivos del capítulo son:

1-1 Pensamiento estadístico y crítico

- Analizar los datos muestrales en relación con el contexto, la fuente y el método de muestreo.
- Entender la diferencia entre la significancia estadística y la significancia práctica.
- Definir e identificar una *muestra de respuesta voluntaria* y entender que las conclusiones estadísticas basadas en los datos de una muestra de este tipo por lo general no son válidas.

1-2 Tipos de datos

- Distinguir entre un *parámetro* y un *dato estadístico*.
- Distinguir entre *datos cuantitativos* y *datos categóricos* (o *cualitativos* o *de atributo*).
- Distinguir entre *datos discretos* y *datos continuos*.
- Determinar si los cálculos estadísticos básicos son apropiados para un conjunto de datos determinado.

1-3 Recopilación de datos muestrales

- Definir e identificar una *muestra aleatoria simple*.
- Comprender la importancia de los métodos de muestreo correctos y la importancia del buen diseño de experimentos.

1-1 Pensamiento estadístico y crítico

Concepto clave En esta sección comenzamos con algunas definiciones básicas y luego presentaremos una visión general del proceso implicado en la realización de un estudio estadístico. Este proceso consiste en "preparar, analizar y concluir". La "preparación" abarca definir el contexto, la fuente de datos y el método de muestreo. En los capítulos siguientes elaboraremos gráficas adecuadas, exploraremos los datos y llevaremos a cabo los cálculos requeridos para el método estadístico que se esté utilizando. También obtendremos conclusiones determinando si los resultados tienen significancia estadística y significancia práctica.

El pensamiento estadístico involucra pensamiento crítico y capacidad de dar sentido a los resultados. El pensamiento estadístico exige mucho más que hacer cálculos complicados. A través de numerosos ejemplos, ejercicios y análisis, este texto le ayudará a desarrollar las habilidades de pensamiento estadístico que son tan importantes en el mundo actual.

En cifras

78%: El porcentaje de estudiantes de veterinaria que son mujeres, según *The Herald* en Glasgow, Escocia.

Comenzamos con algunas definiciones básicas.

DEFINICIONES

Los **datos** son el conjunto de observaciones como mediciones, géneros o respuestas de encuestas.

Estadística es la ciencia que se encarga de planear estudios y experimentos, obtener datos y luego organizar, resumir, presentar, analizar e interpretar esos datos para obtener conclusiones basadas en ellos.

Población es el conjunto completo de todos los individuos, las cosas o los eventos sobre los que se quiere investigar con respecto a una particularidad dada. A la población le correspondería la colección completa de datos –casi siempre imposible de elaborar por su tamaño u otras condiciones– sobre los cuales se harán inferencias.

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

Una **muestra** es un *subconjunto* de miembros seleccionados de una población.

Debido a que las poblaciones suelen ser muy grandes, un objetivo común del uso de la estadística es obtener datos de una muestra y luego utilizarlos para sacar una conclusión acerca de la población.

EJEMPLO 1 **Detectores domésticos de monóxido de carbono**

En el artículo de revista "Tasas de falla en los detectores de residuos de monóxido de carbono en Estados Unidos" (*Residential Carbon Monoxide Detector Failure Rates in the United States*, de Ryan y Arnold, *American Journal of Public Health*, vol. 101, no. 10), se afirmó que hay 38 millones de detectores de monóxido de carbono instalados en Estados Unidos. Cuando 30 de ellos fueron seleccionados al azar y probados, se encontró que 12 no dieron la alarma en condiciones peligrosas de monóxido de carbono. En este caso, la población y la muestra fueron:

Población: Los 38 millones de detectores de monóxido de carbono que hay en Estados Unidos

Muestra: Los 30 detectores de monóxido de carbono seleccionados y probados

El objetivo era utilizar los datos muestrales como base para llegar a una conclusión acerca de la población de todos los detectores de monóxido de carbono, y los métodos estadísticos son útiles para extraer tales conclusiones.

SU TURNO Resuelva el inciso (a) del ejercicio 2 "Reportado contra medido".

Ahora describiremos el proceso involucrado en un estudio estadístico. Vea en la figura 1-2 un resumen de este proceso y observe que el enfoque está puesto en el pensamiento crítico, no en los cálculos matemáticos. Gracias a los maravillosos desarrollos en tecnología, tenemos potentes herramientas que llevan a cabo el procesamiento numérico de manera eficiente para que podamos centrarnos en la comprensión y la interpretación de los resultados.

Preparar

Contexto La figura 1-2 sugiere que comencemos nuestra preparación considerando el *contexto* de los datos, así que iniciemos con el contexto de los datos de la tabla 1-1, que presenta el número de embarcaciones recreativas registradas en Florida (en decenas de miles) y la cantidad de decesos de manatíes resultantes de sus encuentros con barcos, en Florida, durante varios años recientes. El formato de la tabla 1-1 sugiere la siguiente meta: determinar si hay una relación entre el número de embarcaciones y las muertes de manatíes ocasionados

TABLA 1-1 Embarcaciones recreativas y muertes de manatíes en encuentros con barcos

Embarcaciones recreativas (decenas de miles)	99	99	97	95	90	90	87	90	90
Muertes de manatíes	92	73	90	97	83	88	81	73	68

por barcos. Este objetivo sugiere una hipótesis razonable: a medida que aumenta el número de barcos, aumenta el número de muertes de manatíes.

Fuente de los datos El segundo paso en nuestra preparación es considerar la fuente (como se indica en la figura 1-2). Los datos en la tabla 1-1 provienen del Departamento de Seguridad en Carreteras y Vehículos Motorizados de Florida y el Instituto de Investigación Marina de Florida. Las fuentes ciertamente parecen ser respetables.

Método de muestreo La figura 1-2 sugiere que concluyamos nuestra preparación considerando el método de muestreo. Los datos de la tabla 1-1 se obtuvieron de registros oficiales del gobierno que, en este caso, son confiables. El método de muestreo parece ser sólido.

Los métodos de muestreo y el uso de la aleatorización se analizará en la sección 1-3, pero por ahora enfatizamos que un método de muestreo adecuado es absolutamente esencial para obtener buenos resultados en un estudio estadístico. En general, es mala práctica utilizar muestras de respuesta voluntaria (o auto-seleccionadas), aunque su uso es común.

Preparar

1. **Contexto**
   - ¿Qué representan los datos?
   - ¿Cuál es la meta del estudio?
2. **Origen de los datos**
   - ¿Los datos provienen de una fuente con un interés especial, tal que haya presión para obtener resultados favorables a la fuente?
3. **Método de muestreo**
   - ¿Los datos se recolectaron de manera imparcial o en forma sesgada (como un procedimiento en el cual los encuestados se ofrecen voluntariamente para participar)?

↓

Analizar

1. **Graficar los datos**
2. **Explorar los datos**
   - ¿Hay datos atípicos (números muy alejados de casi todos los otros datos)?
   - ¿Qué estadísticas importantes resumen los datos (como la media y la desviación estándar descritas en el capítulo 3)?
   - ¿Cómo se distribuyen los datos?
   - ¿Hay datos faltantes?
   - ¿Muchos sujetos seleccionados se negaron a responder?
3. **Aplicar métodos estadísticos**
   - Usar la tecnología para obtener resultados.

↓

Concluir

1. **Significancia**
   - ¿Los resultados tienen significancia estadística?
   - ¿Los resultados tienen significancia práctica?

FIGURA 1-2 Pensamiento estadístico y crítico

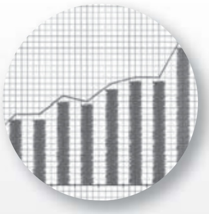
Sesgo de supervivencia

En la Segunda Guerra Mundial, el estadístico Abraham Wald salvó muchas vidas con su trabajo en el Panel de Matemáticas Aplicadas. Los líderes militares le preguntaron al panel cómo podrían mejorar las probabilidades de que los bombarderos regresaran a sus bases después de llevar a cabo sus misiones. La idea era agregar blindaje; registraron los puntos en el fuselaje donde había orificios dañinos. Razonaron que el blindaje se debía colocar en los sitios donde se habían registrado más agujeros, pero Wald dijo que la estrategia sería un gran error. Argumentó que el blindaje se debía colocar en los lugares donde los bombarderos que regresaban no mostraban daños. Su razonamiento era el siguiente: los bombarderos que volvían con daños eran sobrevivientes; es decir, el daño sufrido no había sido tal que los hubiera destruido. Los sitios en los aviones que no sufrían daños eran los más vulnerables y los bombarderos que sufrieron daños en esas áreas vulnerables no pudieron regresar. Los líderes militares habrían cometido un gran error con el sesgo de supervivencia al estudiar los aviones que sobrevivieron en lugar de pensar en los aviones que no lo hicieron.

En cifras

17%: El porcentaje de hombres estadounidenses de entre 20 y 40 años de edad, con una estatura mayor a 2.13 m, que juegan al baloncesto en la NBA.

Origen de "estadística"



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

DEFINICIÓN

Una **muestra de respuesta voluntaria** (o **muestra auto-seleccionada**) es aquella en la que los propios encuestados deciden si serán incluidos.

Los siguientes tipos de encuestas son ejemplos comunes de muestras de respuesta voluntaria. Por su propia naturaleza, todas son esencialmente defectuosas debido a que no debemos obtener conclusiones sobre una población con base en muestras con una fuerte posibilidad de sesgo:

- Encuestas por Internet, donde las personas en línea pueden decidir si responden o no.
- Encuestas por correo, donde las personas pueden decidir si responden o no.
- Encuestas telefónicas, en las que se pide mediante anuncios de periódico, radio o televisión que las personas llamen voluntariamente a un número especial para registrar su opinión.

El problema del capítulo involucra una encuesta de *USA Today* con una muestra de respuesta voluntaria. Vea también el siguiente ejemplo.

EJEMPLO 2 **Muestra de respuesta voluntaria**

El programa de televisión *Nightline* transmitido por ABC pidió a los espectadores que llamaran para dar su opinión sobre si la sede de las Naciones Unidas debería permanecer en Estados Unidos. 67% de los 186,000 que decidieron llamar dijo que las Naciones Unidas debían trasladarse fuera de Estados Unidos. En otra encuesta independiente, se seleccionaron al azar 500 participantes, y 38% de este grupo quería que las Naciones Unidas se mudaran fuera de Estados Unidos. Las dos encuestas produjeron resultados notoriamente diferentes. A pesar de que la encuesta de *Nightline* involucró 186,000 encuestados voluntarios, es probable que la encuesta más pequeña (de 500 encuestados elegidos al azar) haya proporcionado mejores resultados debido a un mejor método de muestreo.

SU TURNO Resuelva el ejercicio 1 "Información médica en línea".

Analizar

La figura 1-2 indica que después de completar nuestra preparación, considerando el contexto, la fuente y el método de muestreo debemos comenzar a analizar los datos.

Graficar y explorar Un análisis debe comenzar con las gráficas y las exploraciones adecuadas de los datos. Las gráficas se estudian en el capítulo 2 y las estadísticas importantes se analizan en el capítulo 3.

Aplicar métodos estadísticos En capítulos posteriores se describen métodos estadísticos importantes, pero con frecuencia la aplicación de tales métodos se realiza usando calculadoras y/o paquetes de software estadístico. Un buen análisis estadístico no requiere grandes habilidades computacionales, sino el uso del sentido común y una cuidadosa atención a los métodos estadísticos.

Sacar conclusiones

La figura 1-2 muestra que el paso final en nuestro proceso estadístico implica conclusiones, y debemos desarrollar la capacidad de distinguir entre la significancia estadística y la significancia práctica.

Significancia estadística La *significancia estadística* en un estudio se logra cuando obtenemos un resultado que es muy improbable que ocurra por casualidad. Un criterio común es que se logra la significancia estadística si la probabilidad de que ocurra un evento por casualidad es 5% o menos.

- Obtener 98 niñas en 100 nacimientos aleatorios *es* estadísticamente significativo porque no es probable que tal resultado extremo resulte del azar.
- Obtener 52 niñas en 100 nacimientos *no es* estadísticamente significativo porque ese evento podría ocurrir fácilmente con el azar.

Significancia práctica Es posible que algún tratamiento o hallazgo sea efectivo, pero el sentido común podría sugerir que éste no constituye una diferencia suficiente para justificar su uso o para ser práctico, como lo ilustra el ejemplo 3.

EJEMPLO 3 Significancia estadística contra significancia práctica

En cierta ocasión, ProCare Industries suministró un producto llamado Gender Choice que supuestamente aumentaba la probabilidad de que una pareja tuviera un bebé con el sexo que deseaban. En ausencia de cualquier evidencia de su eficacia, el producto fue prohibido por la Administración de Alimentos y Medicamentos (FDA) como un "engaño claro al consumidor". Pero supongamos que el producto fue probado con 10,000 parejas que querían tener niñas, y los resultados fueron 5,200 niñas nacidas. Este resultado es estadísticamente significativo porque la probabilidad de que ocurra por casualidad es de sólo 0.003%, por lo que el azar no parece una explicación factible. Esa tasa de 52% de niñas es estadísticamente significativa, pero carece de significado práctico porque 52% está sólo ligeramente por encima del 50%. Las parejas no querrían gastar tiempo y dinero para aumentar la probabilidad de tener una niña de 50% a 52%. (*Nota:* En realidad, la probabilidad de que un bebé sea una niña es de aproximadamente 48.8%, no de 50%).

SU TURNO Resuelva el ejercicio 15 "Selección de género".

Análisis de datos: errores potenciales

A continuación se presentan algunos elementos adicionales que podrían causar problemas al analizar datos.

Conclusiones engañosas Al formular una conclusión basada en un análisis estadístico, debemos hacer declaraciones que sean claras incluso para aquellos que no comprenden las estadísticas y su terminología. Debemos evitar cuidadosamente aquellas que el análisis estadístico no justifique. Por ejemplo, más adelante en este libro presentamos el concepto de una correlación, o asociación entre dos variables, como el número de embarcaciones recreativas registradas y el número de muertes de manatíes por encuentros con barcos. Un análisis estadístico podría justificar la afirmación de que hay una correlación entre el número de embarcaciones y el número de muertes de manatíes, pero no una declaración en el sentido de que un aumento en el número de embarcaciones provoca un aumento en el número de muertes de manatíes. Esta afirmación sobre la causalidad puede ser justificada por la evidencia física, no por el análisis estadístico.

Correlación no implica causalidad.

Datos muestrales reportados en vez de medidos Cuando recopile datos de personas, es mejor que usted mismo tome las medidas en lugar de pedir a los sujetos que *reporten* los resultados. Pregunte a las personas cuánto pesan y es probable que obtenga sus pesos *deseados*, no sus pesos reales. La gente tiende a redondear, por lo general hacia abajo, a veces *muy* hacia abajo. Al preguntarle a alguien con un peso de 85 kg podría responder que pesa 72 kg. Los pesos precisos se recopilan mediante una báscula, no preguntándoselos a las personas.

Sesgo de publicación

Existe un "sesgo de publicación" en las revistas científicas, que es la tendencia a publicar resultados positivos (como demostrar que algún tratamiento es eficaz) con mucha mayor frecuencia que resultados negativos (como demostrar que cierto tratamiento no tiene efecto alguno). En el artículo "Registro de pruebas clínicas" (*Registering Clinical Trials* en *Journal of the American Medical Association*, vol. 290, núm. 4), los autores Kay Dickerson y Drummond Rennie afirman que "no saber quién realizó tal o cual acción (en este caso, un ensayo clínico) resulta en la pérdida y distorsión de la evidencia, el desperdicio y la duplicación de ensayos, la incapacidad de planeación por parte de las agencias patrocinadoras y un sistema caótico del que sólo ciertos patrocinadores se pueden beneficiar, lo cual invariablemente va en contra de los intereses de quienes se ofrecieron a participar en los ensayos y de los pacientes en general". Los autores del artículo apoyan un proceso donde *todos* los ensayos clínicos queden registrados en un sistema central, de modo que los futuros investigadores tengan acceso a todos los estudios previos, no sólo a los estudios publicados.

La estadística es sexy

CareerCast.com es un sitio web de empleos, y sus organizadores analizaron las profesiones utilizando cinco criterios: ambiente, ingresos, perspectivas de empleo, demandas físicas y estrés. Con base en ese estudio, a continuación se presentan los 10 mejores trabajos: (1) matemático, (2) actuario, (3) *estadístico* (énfasis del autor), (4) biólogo, (5) ingeniero de software, (6) analista de sistemas informáticos, (7) historiador, (8) sociólogo, (9) diseñador industrial, (10) contador. Los leñadores están en la parte inferior de la lista con un salario muy bajo, un trabajo peligroso y malas perspectivas de empleo.

El reportero Steve Lohr escribió el artículo "Para el graduado de hoy, sólo una palabra: estadística" en el *New York Times*. En ese artículo citó al economista en jefe de Google que dijo "el empleo sexy en los próximos 10 años será el de estadístico y no estoy bromeando".

Preguntas sesgadas Si las preguntas de la encuesta no se redactan cuidadosamente, los resultados de un estudio pueden ser engañosos. Las preguntas de la encuesta pueden estar "sesgadas" o formuladas intencionalmente para obtener una respuesta deseada. A continuación se muestran las proporciones reales de respuestas "sí" para dos redacciones de una misma pregunta:

97% sí: "¿Debe el presidente tener poder de veto para así evitar los despilfarros?"

57% sí: "¿Debe el presidente tener poder de veto, o no?"

Orden de las preguntas En ocasiones las preguntas de las encuestas se sesgan involuntariamente por factores como el orden de los elementos que se están considerando. Vea las dos preguntas siguientes de una encuesta realizada en Alemania, junto con las muy diferentes proporciones de respuestas:

"¿Diría usted que el tráfico contribuye más o menos a la contaminación del aire que la industria?" (El 45% culpó al tráfico y el 27% culpó a la industria).

"¿Diría usted que la industria contribuye más o menos a la contaminación del aire que el tráfico?" (El 24% culpó al tráfico, el 57% culpó a la industria).

Además del orden de los elementos dentro de una pregunta, como se acaba de ilustrar, el orden de las preguntas también podría afectar las respuestas.

Sin respuesta Una pregunta *sin respuesta* ocurre cuando alguien se niega a responder a la encuesta o no está disponible. Cuando se le hacen preguntas a las personas, algunas de ellas se niegan a responder. La tasa de rechazo ha estado creciendo en los años recientes, en parte debido a que muchos vendedores por teléfono persistentes tratan de vender bienes o servicios comenzando con un tono que suena como una encuesta de opinión. En *Lies, Damn Lies, and Statistics*, el autor Michael Wheeler hace esta importante observación:

Es probable que las personas que se niegan a hablar con los encuestadores sean diferentes de las que no lo hacen. Algunas pueden tener miedo de los extraños y otras ser celosas de su privacidad, pero su negativa a hablar demuestra que su visión del mundo que les rodea es muy diferente a la de las personas que dejan entrar a los encuestadores en sus hogares.

Porcentajes Algunos estudios citan porcentajes engañosos o poco claros. Tenga en cuenta que el 100% de alguna cantidad es *toda* ella, pero si hay referencias a porcentajes que exceden el 100%, a menudo no están justificadas. Un anuncio de The Club, un dispositivo utilizado para desalentar los robos de automóviles, establecía que "The Club reduce las probabilidades de robo de automóviles en un 400%". Si The Club eliminara todos los robos de automóviles, reduciría las probabilidades de robo de autos en un 100%, por lo que la cifra de 400% es engañosa y no tiene sentido.

La siguiente lista identifica algunos principios clave a aplicar cuando se trata de porcentajes. Estos principios usan el concepto básico de que % o "porcentaje" significa en realidad "dividido por 100". El primer principio se usa con frecuencia en este libro.

Porcentaje de: Para encontrar el porcentaje de una cantidad, reemplace el símbolo % por una división entre 100 y luego interprete "de" como una multiplicación. Este ejemplo muestra que el 6% de 1200 es 72:

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

Decimal → Porcentaje: Para convertir de un decimal a un porcentaje, multiplique por 100%. Este ejemplo muestra que 0.25 es equivalente a 25%:

$$0.25 \rightarrow 0.25 \times 100\% = 25\%$$

Fracción → Porcentaje: *Para convertir de una fracción a un porcentaje*, divida el numerador entre el denominador para obtener un número decimal equivalente; después multiplique por 100%. Este ejemplo muestra que la fracción 3/4 es equivalente a 75%:

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

Porcentaje → Decimal: Para convertir de un porcentaje a un número decimal, reemplace el símbolo % por una división por 100. Este ejemplo muestra que el 85% es equivalente a 0.85:

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

1-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Información médica en línea *USA Today* publicó la siguiente pregunta en su sitio web: "¿Con qué frecuencia buscas información médica en línea?" De los 1072 usuarios de Internet que optaron por responder, 38% respondió "frecuentemente". ¿Qué término se utiliza para describir este tipo de encuesta en la que las personas encuestadas son aquellas que decidieron responder? ¿Qué hay de malo en este tipo de método de muestreo?

2. Reportado contra medido En una encuesta a 1046 adultos realizada por Bradley Corporation, se preguntó con qué frecuencia se lavaban las manos después de usar un baño público; 70% de ellos respondió "siempre".

a. Identifique la muestra y la población.

b. ¿Por qué se obtendrían mejores resultados al observar el lavado de manos en vez de preguntar acerca de él?

3. Significancia estadística contra significancia práctica Al probar un nuevo tratamiento, ¿cuál es la diferencia entre la significancia estadística y la significancia práctica? ¿Puede un tratamiento tener significancia estadística, pero no significancia práctica?

4. Correlación Un estudio mostró que durante un período reciente de 11 años hubo una fuerte correlación (o asociación) entre el número de personas que se ahogaron en piscinas y las cantidades de energía generadas por las centrales nucleares (con base en datos de Los Centros para el Control y la Prevención de Enfermedades y el Departamento de Energía de Estados Unidos). ¿Esto implica que el aumento de energía de las centrales nucleares es la causa de más muertes en las piscinas? ¿Por qué sí o por qué no?

Considere la fuente. ***En los ejercicios 5 a 8, determine si la fuente dada tiene el potencial de crear un sesgo en un estudio estadístico.***

5. Comité de Médicos para la Medicina Responsable El Comité de Médicos para la Medicina Responsable tiende a oponerse al uso de carne y productos lácteos en nuestra dieta, tal organización ha recibido cientos de miles de dólares en apoyos de la Sociedad Protectora de Animales.

6. Arsénico en el arroz Las cantidades de arsénico en muestras de arroz cultivado en Texas fueron medidas por la Administración de Alimentos y Medicamentos en Estados Unidos (FDA, por sus siglas en inglés).

7. Tamaño del cerebro Un conjunto de datos del apéndice B incluye volúmenes cerebrales de 10 pares de gemelos monocigóticos (idénticos). Los datos fueron recopilados por investigadores de la Harvard University, el Massachusetts General Hospital, el Dartmouth College y la University of California en Davis.

8. Chocolate Un artículo en la *Journal of Nutrition* (Vol. 130, núm. 8) señaló que el chocolate es rico en flavonoides. El artículo señala que "el consumo regular de alimentos ricos en flavonoides puede reducir el riesgo de enfermedades coronarias". El estudio recibió financiamiento de Mars. Inc., la compañía de dulces, y de la Asociación de Fabricantes de Chocolate.

Método de muestreo. ***En los ejercicios 9 a 12, determine si el método de muestreo parece ser correcto o defectuoso.***

9. Plantas eléctricas nucleares En una encuesta de 1368 sujetos, la siguiente pregunta fue publicada en el sitio web de *USA Today*: "En su opinión, ¿las plantas nucleares son seguras?". Los sujetos de la encuesta fueron usuarios de Internet que optaron por responder a la pregunta publicada en la edición electrónica de *USA Today*.

10. Ensayos clínicos Los investigadores de la Yale University realizan una amplia variedad de ensayos clínicos utilizando sujetos voluntarios después de leer anuncios que solicitan voluntarios remunerados.

11. Pagos con tarjeta de crédito En una encuesta de AARP, Inc. con 1019 adultos seleccionados al azar, a cada uno se le preguntó qué cantidad de su deuda con tarjetas de crédito paga mensualmente.

12. Uso de teléfonos inteligentes En una encuesta sobre la propiedad de teléfonos inteligentes, el Pew Research Center seleccionó aleatoriamente a 1006 adultos en Estados Unidos.

Significancia estadística y significancia práctica. ***En los ejercicios 13 a 16, determine si los resultados parecen tener significancia estadística y/o significancia práctica.***

13. Programa de dieta y ejercicio En un estudio del programa de dieta y ejercicio de Kingman, 40 sujetos perdieron en promedio 22 libras de peso. Hay aproximadamente un 1% de probabilidad de obtener esos resultados con un programa que en realidad no produce ningún efecto.

14. MCAT El Examen de Admisión a la Facultad de Medicina (MCAT, por sus siglas en inglés) se utiliza comúnmente como parte del proceso de toma de decisiones para determinar qué estudiantes deben aceptarse en las escuelas de medicina. Para probar la eficacia del curso de preparación Siena MCAT, 16 estudiantes realizan el examen MCAT, luego completan el curso propedéutico y después presentan de nuevo el examen MCAT; con el resultado de que la puntuación promedio (media) para este grupo aumenta de 25 a 30. Existe una probabilidad de 0.3% de obtener esos resultados por casualidad. ¿El curso parece ser efectivo?

15. Selección de género En un estudio sobre el método Gender Aide para la selección de género, utilizado para aumentar la probabilidad de que un bebé nazca niña, 2000 usuarios del método tuvieron 980 niños y 1020 niñas. Hay alrededor de 19% de probabilidad de conseguir tal cantidad de niñas si el método no tuviera ningún efecto.

16. Puntuaciones de IQ La mayoría de las personas tienen puntuaciones de IQ (coeficiente intelectual) entre 70 y 130. Por $39.99, usted puede comprar un programa para PC o Mac, desarrollado por HighIQPro, que alega aumentar su puntuación de inteligencia entre 10 y 20 puntos. El programa pretende ser "el único software comprobado para el aumento del IQ en el mercado del entrenamiento del cerebro", pero el autor de este texto no pudo encontrar ningún dato que apoye tal afirmación, por lo que supondremos que se obtuvieron los siguientes resultados: En un estudio de 12 usuarios del programa, el aumento promedio en la puntuación del IQ fue de 3 puntos. Hay un 25% de probabilidad de obtener tales resultados si el programa no tuviera ningún efecto.

En los ejercicios 17 a 20, consulte la muestra de temperaturas corporales (grados Fahrenheit) en la siguiente tabla. (Las temperaturas corporales provienen de un conjunto de datos del apéndice B).

	Sujeto				
	1	2	3	4	5
8 AM	97.0	98.5	97.6	97.7	98.7
12 AM	97.6	97.8	98.0	98.4	98.4

17. Contexto de los datos Consulte la tabla de temperaturas corporales. ¿Hay algún modo significativo en el que cada temperatura corporal registrada a las 8 AM se corresponda con la temperatura de las 12 AM?

18. Fuente Las temperaturas corporales de la lista fueron obtenidas de los doctores Steven Wasserman, Philip Mackowiak y Myron Levine, investigadores de la University of Maryland. ¿Es probable que la fuente de los datos esté sesgada?

19. Conclusión Dadas las temperaturas corporales de la tabla, ¿qué problema se puede abordar mediante la realización de un análisis estadístico de los datos?

20. Conclusión Si analizamos las temperaturas corporales listadas con métodos estadísticos adecuados, concluimos que al encontrar las diferencias entre las temperaturas corporales de las 8 AM y las 12 AM, existe un 64% de probabilidad de que éstas puedan explicarse mediante resultados aleatorios obtenidos de poblaciones que tienen las mismas temperaturas corporales a las 8 AM y las l2 AM. ¿Qué deberíamos concluir sobre la significancia estadística de esas diferencias?

En los ejercicios 21 a 24, consulte los datos de la tabla siguiente. Las entradas son conteos de glóbulos blancos (1000 células/μl) y de glóbulos rojos (millones de células/μl) en varones examinados como parte de un gran estudio de salud realizado por el Centro Nacional de Estadísticas de la Salud. Los datos se encuentran listados de modo que el primer sujeto tiene un recuento de glóbulos blancos de 8.7 y un recuento de glóbulos rojos de 4.91, y así sucesivamente.

	Sujeto				
	1	2	3	4	5
Blancos	8.7	5.9	7.3	6.2	5.9
Rojos	4.91	5.59	4.44	4.80	5.17

21. Contexto Dada la forma en que se listan los datos y considerando sus unidades, ¿tiene sentido usar la diferencia entre cada recuento de glóbulos blancos y de glóbulos rojos correspondiente? ¿Por qué sí o por qué no?

22. Análisis Dado el contexto de los datos de la tabla, ¿qué problema se puede abordar realizando un análisis estadístico de las mediciones?

23. Fuente de los datos Si se considera la fuente de los datos, ¿parece que tal fuente está sesgada de alguna manera?

24. Conclusión Si analizamos los datos muestrales y concluimos que existe una correlación entre el recuento de glóbulos blancos y el recuento de glóbulos rojos, ¿se puede concluir que un mayor recuento de glóbulos blancos es la causa de mayores recuentos de glóbulos rojos?

¿Qué es incorrecto? ***En los ejercicios 25 a 28, identifique lo que es incorrecto.***

25. Papas En una encuesta patrocinada por la Comisión de la Papa de Idaho, se pidió a 1000 adultos que seleccionaran su vegetal favorito, y la opción preferida fueron las papas, que fueron escogidas por 26% de los encuestados.

26. Agua saludable En una encuesta en línea de *USA Today*, 951 usuarios de Internet optaron por responder, y el 57% dijo que prefiere beber agua embotellada en vez de agua del grifo.

27. Motocicletas y crema agria En los últimos años, ha habido una fuerte correlación entre el consumo *per cápita* de crema agria y el número de motociclistas muertos en accidentes sin colisión. Por lo tanto, el consumo de crema agria causa muertes en motocicleta.

28. Fumadores El fabricante de cigarrillos electrónicos V2 Cigs patrocinó una encuesta que mostró que el 55% de los fumadores encuestados dicen sentirse excluidos “a veces”, “a menudo” o “siempre”.

Porcentajes. ***En los ejercicios 29 a 36, responda las preguntas relacionadas con porcentajes.***

29. Vestimenta de trabajo En una encuesta realizada por Opinion Research Corporation, se pidió a 1000 adultos que identificaran "lo que es inapropiado en el lugar de trabajo". De los 1000 sujetos, 70% dijo que las minifaldas no eran apropiadas en el lugar de trabajo.

a. ¿Cuál es el 70% de 1000?

b. De los 1000 encuestados, 550 dijeron que los pantalones cortos son inaceptables en el lugar de trabajo. ¿Qué porcentaje de encuestados dijo que los pantalones cortos son inaceptables en el lugar de trabajo?

30. Verificación de solicitantes de empleo En un estudio realizado por la Sociedad para la Gestión de Recursos Humanos, se encuestó a 347 profesionales de los recursos humanos. De los encuestados, 73% dijo que sus empresas verifican los antecedentes penales de todos los solicitantes de empleo.

a. ¿Cuál es el valor exacto del 73% de los 347 sujetos encuestados?

b. ¿Podría el resultado del inciso (a) ser el número real de sujetos encuestados que dijeron que sus compañías verifican los antecedentes penales de todos los solicitantes de empleo? ¿Por qué sí o por qué no?

c. ¿Cuál es el número real de sujetos encuestados que dijeron que su empresa verifica los antecedentes penales de todos los solicitantes de empleo?

d. Suponga que 112 de los sujetos encuestados son mujeres, ¿qué porcentaje de los encuestados son mujeres?

31. Propuestas de matrimonio En una encuesta realizada por TheKnot.com, se preguntó a 1165 mujeres comprometidas o casadas sobre la importancia de arrodillarse al hacer una propuesta de matrimonio. Entre las 1165 encuestadas, 48% dijo que ese gesto es esencial.

a. ¿Cuál es el valor exacto del 48% de las 1165 encuestadas?

b. ¿Podría el resultado del inciso (a) ser el número real de sujetos encuestados que dijeron que arrodillarse es esencial? ¿Por qué sí o por qué no?

c. ¿Cuál es el número real de encuestadas que dijeron que arrodillarse es esencial?

d. De las 1165 encuestadas, 93 dijeron que hincar una rodilla es cursi y anticuado. ¿Qué porcentaje de encuestadas afirmó que hincar una rodilla es cursi y anticuado?

32. Chillax *USA Today* reportó los resultados de una encuesta de Research Now para Keurig en la que se preguntó a 1458 hombres y 1543 mujeres: "En una semana típica, ¿con qué frecuencia puedes relajarte?"

a. Entre las mujeres, 19% respondió "en muy pocas ocasiones". ¿Cuál es el valor exacto del 19% del número de mujeres encuestadas?

b. ¿Podría el resultado del inciso (a) ser el número real de mujeres que respondieron "en muy pocas ocasiones"? ¿Por qué sí o por qué no?

c. ¿Cuál es el número real de mujeres que respondieron "en muy pocas ocasiones"?

d. Entre los hombres encuestados, 219 respondieron "en muy pocas ocasiones". ¿Cuál es el porcentaje de hombres que respondieron de esta forma?

e. Considere la pregunta formulada a los sujetos. ¿Es esa pregunta clara e inequívoca para que todos los encuestados interpreten la pregunta de la misma manera? ¿Cómo se podría mejorar la encuesta?

33. Porcentajes en publicidad Un anuncio de las billeteras Big Skinny incluye la afirmación de que uno de sus productos "reduce el tamaño de su billetera llena entre 50 y 200%". ¿Qué hay de erróneo en esta afirmación?

34. Porcentajes en publicidad Continental Airlines publicó anuncios que afirman que el equipaje perdido es "un aspecto en el que hemos mejorado 100% en los últimos seis meses". ¿Qué hay de erróneo en esta afirmación?

35. Porcentajes en publicidad Un editorial del *New York Times* criticó el título de una gráfica que afirmaba que un enjuague dental "reduce la placa en los dientes en más del 300%". ¿Qué es incorrecto en esta afirmación?

36. Porcentajes en negociaciones Cuando el autor de este texto estaba negociando un contrato para el profesorado y la administración en una universidad, un decano presentó el argumento de que si el profesorado recibía un aumento de 4% y los administrativos un incremento de 4%, se tendría un aumento de 8%, lo cual nunca sería aprobado. ¿Qué tiene de erróneo este argumento?

1-1 Más allá de lo básico

37. ¿Qué tiene de erróneo este panorama? El *Newport Chronicle* realizó una encuesta pidiendo a los lectores que respondieran la siguiente pregunta: "¿Apoyan el desarrollo de armas atómicas que podrían matar a millones de personas inocentes?" Se reportó que 20 lectores respondieron y que 87% dijeron "no", mientras que 13% dijeron "sí". Identifique cuatro fallas importantes en esta encuesta.

38. Falsificación de datos Un investigador del Centro de Investigación del Cáncer Sloan-Kettering fue criticado una vez por falsificar datos. Entre sus datos se obtuvieron cifras de 6 grupos de 20 ratones cada uno. Se dieron los siguientes valores para el porcentaje de éxito en cada grupo: 53%, 58%, 63%, 46%, 48%, 67%. ¿Qué hay de erróneo en estos valores?

1-2 Tipos de datos

Concepto clave Un uso importante de la estadística es recopilar y utilizar datos muestrales para obtener conclusiones sobre las poblaciones. Debemos conocer y entender los significados de los términos *dato estadístico* y *parámetro*, como se definen a continuación. En esta sección describimos algunos tipos de datos. El tipo de datos es uno de los factores clave que determinan el método estadístico a utilizar en un análisis.

En la Parte 1 de esta sección describimos los fundamentos de los tipos de datos, y después, en la Parte 2, consideraremos "los grandes datos" y los datos faltantes.

PARTE 1 Tipos básicos de datos

Parámetro / Dato estadístico

DEFINICIONES

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

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

SUGERENCIA La asociación entre "parámetro" y "población", por un lado, y "dato estadístico" y "muestra", por el otro, nos ayuda a recordar los significados de estos términos.

Obtener datos estadísticos es la finalidad principal de la estadística; en la sección 1-1 se definió la estadística como la ciencia de la planificación de estudios y experimentos, la obtención de datos, la organización, el resumen, la presentación, el análisis y la interpretación de esos datos para después obtener conclusiones basadas en ellos. A partir de las definiciones de parámetro y dato estadístico es posible determinar cuál de las dos definiciones se aplica considerando el contexto en el cual se usa el término. En el ejemplo siguiente se utiliza el significado de dato estadístico tal y como se indica en la página anterior.

EJEMPLO 1 **Parámetro / Dato estadístico**

Hay 17,246,372 estudiantes de preparatoria en Estados Unidos. En un estudio de 8505 estudiantes estadounidenses de preparatoria de 16 años de edad o más, 44.5% de ellos dijeron que enviaron mensajes de texto al conducir al menos una vez durante los 30 días anteriores (con base en datos de “Envío de mensajes de texto al manejar y otros comportamientos de riesgo entre estudiantes de bachillerato en Estados Unidos” (*“Texting While Driving and Other Risq Motor Behavior US High School Students”*, de Olsen, Shults y Eaton, *Pediatrics*, vol. 131, núm. 6).

1. **Parámetro:** El tamaño de la población de 17,246,372 estudiantes de preparatoria es un parámetro, porque es la población total de estudiantes de preparatoria en Estados Unidos. Si de alguna manera supiéramos el porcentaje de todos los 17,246,372 estudiantes de preparatoria que reportan haber enviado mensajes de texto mientras conducían, ese porcentaje también sería un parámetro.
2. **Dato estadístico:** El tamaño de la muestra de 8505 estudiantes de preparatoria encuestados es un dato estadístico, porque se basa en una muestra, no en la población entera de todos los estudiantes de preparatoria de Estados Unidos. El valor del 44.5% es otro dato estadístico, porque también se basa en la muestra, no en toda la población.

SU TURNO Resuelva el ejercicio 1 “Parámetro y dato estadístico”.

Cuantitativo / Categórico

Algunos datos son números que representan conteos o mediciones (como una puntuación de IQ de 135), mientras que otros son atributos (como el color de ojos verde o marrón) que no son recuentos ni mediciones. Los términos datos cuantitativos y datos categóricos distinguen entre estos tipos de datos.

DEFINICIONES

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

Los **datos categóricos** (o **cualitativos** o **de atributo**) consisten en nombres o etiquetas (no números que representan conteos o mediciones).

PRECAUCIÓN En ocasiones los datos categóricos se codifican con números que reemplazan los nombres. Aunque tales números pueden parecer cuantitativos, en realidad son datos categóricos. Vea la tercera parte del ejemplo 2 que se presenta más adelante.

Incluya las unidades de medida Al usar datos cuantitativos es importante utilizar las unidades de medida apropiadas, como dólares, horas, pies o metros. Debemos observar cuidadosamente la información dada sobre las unidades de medida, como “todas las cantidades están en *miles de dólares*” o “todas las unidades se dan en *kilogramos*”. No tomar en cuenta tales unidades de medida puede ser muy costoso. La Administración Nacional de Aeronáutica y del Espacio (NASA) perdió su Orbitador de Marte, con un costo de 125 millones de dólares, cuando éste se estrelló debido a que el software de control poseía datos de aceleración en unidades *inglesas*, pero se había supuesto incorrectamente que estaban en unidades *métricas*.

En cifras

7 mil millones: La población mundial superada a principios de 2012, sólo 13 años después de haber pasado la barrera de los 6 mil millones.

EJEMPLO 2 **Cuantitativo / Categórico**

1. **Datos cuantitativos:** Las edades (en años) de los sujetos inscritos en un ensayo clínico.
2. **Datos categóricos como etiquetas:** Los géneros (masculino/femenino) de los sujetos inscritos en un ensayo clínico
3. **Datos categóricos como números:** Los números de identificación 1, 2, 3, ..., 25 asignados aleatoriamente a los 25 sujetos en un ensayo clínico. Estos números son sustitutos de los nombres; no miden ni cuentan nada, por lo que son datos categóricos.

SU TURNO Resuelva el ejercicio 2 “Datos cuantitativos/categóricos”.

Discreto / Continuo

Los datos cuantitativos se pueden describir con mayor detalle al distinguir entre los tipos *discreto* y *continuo*.

DEFINICIONES

Los **datos discretos** resultan cuando los valores de los datos son cuantitativos y el número de valores es finito –cuando “puede contarse”– (si hay una infinidad de valores, la colección de valores puede contarse si es posible contabilizarlos individualmente, como el número de lanzamientos de una moneda antes de obtener una cruz).

Los **datos continuos (numéricos)** resultan de una cantidad infinita de valores cuantitativos posibles, en los que la colección de valores no puede contarse. Es decir, es imposible contar los elementos individuales porque al menos algunos de ellos están en una escala continua, como las longitudes de las distancias entre 0 y 12 cm.

PRECAUCIÓN El concepto de datos que pueden contarse desempeña un papel clave en las definiciones anteriores, pero no es un concepto particularmente fácil de entender. Los datos continuos se pueden medir, pero no se cuentan. Si usted selecciona un valor particular entre una serie de datos continuos, no hay un valor que sea necesariamente el “siguiente” en los datos. Vea el ejemplo 3.

Datos continuos

Datos discretos

Datos estadísticos engañosos en el periodismo

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

EJEMPLO 3 **Discreto / Continuo**

1. **Datos discretos del tipo finito:** Cada uno de varios médicos planea contar el número de exámenes físicos que realice durante toda la próxima semana. Los datos son discretos porque son números finitos, como 27 y 46, que resultan de un proceso de conteo.

2. **Datos discretos del tipo infinito:** Los empleados de un casino planean lanzar un dado hasta que aparezca el número 5, y contar el número de lanzamientos requeridos para obtener ese resultado. Es posible que los lanzamientos pudieran durar para siempre sin obtener nunca un 5, pero el número de lanzamientos se puede contar, aunque el conteo dure por siempre. Por lo tanto, la colección del número de lanzamientos es contabilizable.

3. **Datos continuos:** Cuando a un paciente típico le extraen sangre como parte de un examen de rutina, el volumen de sangre extraída está entre 0 y 50 mL. Hay una cantidad infinita de valores entre 0 y 50 mL. Debido a que es imposible contar el número de valores posibles en una escala continua, estas cantidades son datos continuos.

SU TURNO Resuelva el ejercicio 3 "Datos discretos/continuos".

Niveles de medición

Otra forma común de clasificar datos es usar cuatro niveles de medición: nominal, ordinal, de intervalo y de razón, que se definen a continuación (también vea en la tabla 1-2 las descripciones breves de los cuatro niveles de medición). Cuando se aplica la estadística a problemas reales, el nivel de medición de los datos ayuda a decidir qué procedimiento debe utilizarse. A lo largo de este libro, habrá referencias a estos niveles de medición, pero la cuestión relevante en este caso tiene qué ver con el sentido común: *No hacer cálculos y no utilizar métodos estadísticos que no sean apropiados para los datos*. Por ejemplo, no tendría sentido calcular un promedio (media) de los números de Seguro Social, porque esos números son datos que se utilizan como identificación, y no representan mediciones o conteos de nada.

TABLA 1-2 Niveles de medición

Nivel de medición	Descripción breve	Ejemplo
De razón	Hay un punto de inicio cero natural y las proporciones tienen sentido.	Alturas, longitudes, distancias, volúmenes
De intervalo	Las diferencias son significativas, pero no hay un punto de inicio cero natural y las proporciones no tienen sentido.	Temperaturas corporales en grados Fahrenheit o Celsius
Ordinal	Los datos pueden ponerse en orden, pero no se pueden encontrar diferencias o éstas carecen de significado.	Clasificación de universidades en el *U.S. News & World Report*
Nominal	Sólo categorías. Los datos no se pueden poner en orden.	Colores de los ojos

DEFINICIÓN

El **nivel nominal de medición** se caracteriza por datos que consisten únicamente en nombres, etiquetas o categorías. Los datos no se pueden organizar en algún orden (por ejemplo, de bajo a alto).

EJEMPLO 4 **Nivel nominal**

A continuación se presentan ejemplos de datos muestrales con el nivel de medición nominal.

1. **Sí/No/Indeciso:** Respuestas de encuesta del tipo *sí*, *no* e *indeciso*
2. **Respuestas de encuesta codificadas:** Para una pregunta de una encuesta, los encuestados tienen una serie de respuestas posibles, las cuales se codifican de la siguiente manera: "Estoy de acuerdo" se codifica como 1; "No estoy de acuerdo" como 2; "No me importa" es 3; "Me niego a responder", 4; y "Váyase y deje de molestarme", 5. Los números 1, 2, 3, 4 y 5 no miden ni cuentan nada.

SU TURNO Resuelva el ejercicio 22 "Encuesta de salida".

Debido a que los datos nominales carecen de cualquier orden o significancia numérica, no deben usarse en cálculos. En ocasiones se asignan números como 1, 2, 3 y 4 a las diferentes categorías (especialmente cuando los datos se codifican para computadoras), pero estos números no tienen significado computacional real y cualquier promedio (media) calculado a partir de ellos carece de sentido y posiblemente sea engañoso.

DEFINICIÓN

Los datos están en el **nivel de medición ordinal** si pueden colocarse en cierto orden, pero las diferencias (obtenidas por sustracción) entre los valores de los datos no se pueden determinar o carecen de significado.

EJEMPLO 5 **Nivel ordinal**

A continuación se presenta un ejemplo de datos muestrales con el nivel de medición ordinal.

Calificaciones del curso: Un profesor universitario asigna calificaciones de A, B, C, D o F. Estas calificaciones se pueden poner en orden, pero no es posible determinar las diferencias entre las calificaciones. Por ejemplo, se sabe que A es mayor que B (por lo que hay un ordenamiento), pero no se puede restar B de A (por lo que no es posible encontrar la diferencia).

SU TURNO Resuelva el ejercicio 21 "Clasificaciones de universidades".

Los datos ordinales proporcionan información sobre comparaciones relativas, pero no de *magnitudes* de las diferencias. Por lo general, los datos ordinales no deben usarse para cálculos como un promedio (media), pero a veces esta directriz puede ignorarse (como cuando se usan calificaciones con letras para calcular un promedio general).

DEFINICIÓN

Los datos están en un **nivel de medición de intervalo** si se pueden poner en orden y es posible encontrar diferencias significativas entre los valores de los datos. *Los datos en este nivel no tienen un punto de inicio cero natural en el que no hay ninguna cantidad presente*.

Medición de la desobediencia

¿De qué manera se recolectan datos sobre algo que parece que no es medible, como el nivel de desobediencia de las personas? El psicólogo Stanley Milgram diseñó el siguiente experimento. Un investigador enseñó a un sujeto voluntario a operar un tablero de control que administraba "descargas eléctricas" cada vez más dolorosas a una tercera persona. En realidad no se aplicaban tales descargas, y la tercera persona era un actor. El voluntario iniciaba con 15 volts y recibía la instrucción de incrementar las descargas en 15 volts cada vez. El nivel de desobediencia era el punto donde el sujeto se negaba a incrementar el voltaje. Fue sorprendente que dos terceras partes de los sujetos obedecieron las órdenes, aun cuando el actor gritaba y fingía sufrir un ataque cardiaco.

Seis grados de separación

Los psicólogos sociales, historiadores, científicos políticos y especialistas en comunicaciones están interesados en "El problema del mundo pequeño": dadas dos personas cualesquiera en el mundo, ¿cuántos vínculos intermedios son necesarios para conectar a ambas? En las décadas de 1950 y 1960, el psicólogo social Stanley Milgram realizó un experimento en el que los sujetos trataron de ponerse en contacto con otras personas objetivo mediante el envío de una carpeta de información a un conocido que pensaban que estaría más cerca de dicho objetivo. De las 160 cadenas iniciadas, sólo 44 se completaron, por lo que la tasa de fracaso fue 73%. Entre los éxitos, el número de conocidos intermedios varió de 2 a 10, con una mediana de 6 (de ahí los "seis grados de separación"). El experimento ha sido criticado por su alta tasa de fracaso y su inclusión desproporcionada de sujetos con ingresos por encima de la media. Un estudio más reciente conducido por el investigador de Microsoft, Eric Horvitz y el profesor asistente de Stanford, Jure Leskovec involucró 30 mil millones de mensajes instantáneos y 240 millones de personas. Este estudio encontró que para los mensajes instantáneos utilizados por Microsoft, la longitud media de un vínculo entre dos individuos es 6.6, lo que sugiere "siete grados de separación". En la actualidad, se continúa trabajando en este importante e interesante campo.

EJEMPLO 6 Nivel de intervalo

Los siguientes ejemplos ilustran el nivel de medición de intervalo.

1. **Temperaturas:** Las temperaturas corporales de 98.2°F y 98.8°F son ejemplos de datos con el nivel de medición de intervalo. Tales valores pueden ordenarse, y es posible determinar su diferencia de 0.6°F. Sin embargo, no hay un punto de inicio cero natural. El valor de 0°F puede parecer un punto de inicio, pero es arbitrario y no representa la ausencia total de calor.
2. **Años:** Los años 1492 y 1776 se pueden poner en orden, y es posible encontrar la diferencia de 284 años, la cual es significativa. Sin embargo, el tiempo no comenzó en el año 0, por lo que ese año es arbitrario en vez de ser un punto de inicio cero natural que represente "la ausencia total de tiempo".

SU TURNO Resuelva el ejercicio 25 "Béisbol".

DEFINICIÓN

Los datos tienen un **nivel de medición de razón** si se pueden poner en orden, es posible encontrar diferencias significativas, y hay un punto de inicio cero natural (donde cero indica que no hay ninguna cantidad presente). Para los datos con este nivel, las diferencias y las razones son significativas.

EJEMPLO 7 Nivel de razón

Los siguientes son ejemplos de datos con nivel de medición de razón. Observe la presencia del valor cero natural, así como el uso de razones significativas del tipo "doble" y "triple".

1. **Estaturas de estudiantes:** Estaturas de 180 cm y 90 cm para estudiantes de preparatoria y preescolar (0 cm representa la inexistencia de altura y 180 cm es el *doble* de 90 cm).
2. **Tiempos de clase:** Tiempos de 50 min y 100 min para una clase de estadística (0 min representa la inexistencia de tiempo de clase y 100 minutos es el *doble* de 50 minutos).

SU TURNO Resuelva el ejercicio 24 "Tiempos de servicio en la comida rápida"

SUGERENCIA La distinción entre los niveles de medición de intervalo y de razón puede ser un poco difícil. A continuación se describen dos herramientas para ayudar con esa distinción:

1. **Prueba de razón** Enfocarse en el término "razón" y saber que el término "doble" describe la relación de un valor que es el doble de otro valor. Para distinguir entre los niveles de medición de intervalo y de razón, utilice una "prueba de razón" formulando la siguiente pregunta: ¿El uso del término "doble" tiene sentido? El "doble" tiene sentido para los datos en el nivel de medición de razón, pero no para los datos en el nivel de medición de intervalo.
2. **Cero verdadero** Para que las razones tengan sentido, debe haber un valor "cero verdadero", donde este valor indica que no hay ninguna cantidad presente y que no es simplemente un valor arbitrario en una escala. La temperatura de 0°F es arbitraria y no indica la inexistencia de calor, por lo que las temperaturas en la escala Fahrenheit están en el nivel de medición de intervalo, no en el nivel de razón.

EJEMPLO 8 **Distinción entre el nivel de razón y el nivel de intervalo**

En cada uno de los siguientes casos, determine si los datos están en el nivel de medición de razón o de intervalo:

a. Tiempos (minutos) que requieren los estudiantes para completar un examen de estadística.

b. Temperatura corporal (Celsius) de los estudiantes de estadística.

SOLUCIÓN

a. Aplique la “prueba de razón” descrita en la sugerencia previa. Si un estudiante completa el examen en 40 minutos y otro lo termina en 20, ¿tiene sentido decir que el primer estudiante usó el doble de tiempo? ¡Sí! Así que los tiempos están en el nivel de medición de razón. También se podría aplicar la prueba del “cero verdadero”. Un tiempo de 0 minutos representa “nada de tiempo”, por lo que el valor de 0 es un cero verdadero que indica que no se usó ningún tiempo.

b. Aplique la “prueba de razón” descrita en la sugerencia previa. Si un estudiante tiene una temperatura corporal de 40°C y otro de 20, ¿tiene sentido decir que el primer estudiante está dos veces más caliente que el segundo? ¡No! Así que las temperaturas corporales no están en el nivel de medición de razón. Debido a que la diferencia entre 40°C y 20°C es igual a la diferencia entre 90°C y 70°C, las diferencias son significativas, pero como las razones no tienen sentido, las temperaturas corporales están en el nivel de medición de intervalo. Además, la temperatura de 0°C no representa “inexistencia de calor” por lo que el valor de 0 no es un cero verdadero.

Grandes datos en lugar de un ensayo clínico

Nicholas Tatonetti de la Columbia University buscó bases de datos de la Administración de Alimentos y Medicinas sobre las reacciones adversas que presentaron pacientes debido a diferentes combinaciones de medicamentos. Descubrió que el medicamento Paxil (paroxetina), para la depresión, y el fármaco pravastatina, para el colesterol alto, interactuaban para crear aumentos en los niveles de glucosa (azúcar en la sangre). Cuando los pacientes los tomaron por separado, ninguno de los medicamentos aumentó los niveles de glucosa, a diferencia de cuando los tomaron juntos. Este hallazgo resultó de una búsqueda general en bases de datos de las interacciones de muchos pares de fármacos, no de un ensayo clínico con la participación de pacientes que usaban Paxil y pravastatina.

PARTE 2 Grandes datos y datos faltantes: demasiado e insuficiente

Al trabajar con datos, podemos encontrar algunos conjuntos de datos que son excesivamente grandes y otros conjuntos con elementos individuales que hacen falta. Aquí, en la parte 2, se analizan brevemente ambos casos.

Grandes datos

Algunos hablan de ellos como de héroes, mientras que otros los consideran traidores, pero Edward Snowden usó su empleo en la Agencia Nacional de Seguridad (NSA, por sus siglas en inglés) para revelar documentos secretos sustanciales que llevaron a entender que la NSA estaba vigilando a los ciudadanos de Estados Unidos, así como a los líderes mundiales, por teléfono e Internet. La NSA recolectaba enormes cantidades de datos que se analizaban con el fin de prevenir el terrorismo. El control de las comunicaciones telefónicas y por Internet es posible gracias a la tecnología moderna. La NSA ahora puede recolectar enormes cantidades datos, y tales conjuntos han llevado al nacimiento de la ciencia de los datos. No hay un acuerdo universal sobre las siguientes definiciones, y es posible encontrar versiones diferentes en otros lugares.

Estadística para las citas en línea

Los cuatro fundadores del sitio de citas en línea OkCupid son matemáticos que utilizan métodos estadísticos para analizar los resultados de su sitio web. El director ejecutivo de OkCupid ha dicho: "No somos psicólogos, somos matemáticos" (en "¿Buscando pareja? Un sitio le sugiere que se informe" *"Looking for a Date", "A Site Suggest You Check the Data"*, de Jenna Wortham, *New York Times*). El sitio web de OkCupid es único por su uso de métodos estadísticos para encontrar personas coincidentes de manera efectiva.

Mediante el análisis de las fotos y respuestas de 7000 usuarios, los analistas de OkCupid descubrieron que al crear una foto de perfil, los hombres no deben mirar directamente a la cámara, y no deben sonreír. Para las mujeres, una apariencia interesante produce mejores resultados que la apariencia sexy; también descubrieron que la brevedad en el primer mensaje publicado es positiva; la longitud ideal del primer mensaje es de 40 palabras, aproximadamente lo que una persona común puede escribir en un minuto.

DEFINICIONES

Los **grandes datos** se refieren a conjuntos de datos tan grandes y tan complejos que su análisis está fuera del alcance de las herramientas de software tradicionales. El análisis de los grandes datos puede requerir que el software funcione simultáneamente en paralelo en muchas computadoras.

La **ciencia de los datos** implica la aplicación de la estadística, informática e ingeniería de software, junto con otros campos relevantes (como la sociología o las finanzas).

Ejemplos de magnitudes de los conjuntos de datos A partir de la definición anterior de grandes datos, puede verse que no hay un número fijo que sirva como límite exacto para determinar si un conjunto de datos puede calificarse como "grandes datos". Pero puede decirse que los grandes datos suelen involucrar cantidades de datos como las siguientes:

- Terabytes (10^{12} o 1,000,000,000,000 bytes) de datos
- Petabytes (10^{15} bytes) de datos
- Exabytes (10^{18} bytes) de datos
- Zettabytes (10^{21} bytes) de datos
- Yottabytes (10^{24} bytes) de datos

Ejemplos de aplicaciones de los grandes datos. Los siguientes son algunos ejemplos que involucran grandes datos:

- Google proporciona mapas de tráfico en vivo mediante el registro y el análisis de datos GPS (sistema de posicionamiento global) recopilados en los teléfonos inteligentes de las personas que viajan en sus automóviles.
- Intentos de pronosticar epidemias de gripe analizando las búsquedas en Internet de los síntomas de esa enfermedad.
- La búsqueda del Sloan Digital Sky comenzó en el año 2000, y rápidamente recogió más datos astronómicos que en toda la historia de la humanidad. Ahora posee más de 140 terabytes de datos astronómicos.
- Walmart tiene una base de datos de ventas con más de 2.5 petabytes (2,500,000,000,000,000 bytes) de datos. Para sus ventas en línea, Walmart desarrolló el motor de búsqueda Polaris que incrementó las ventas entre 10% y 15%, por un valor de miles de millones de dólares.
- Amazon monitorea y rastrea 1400 millones de artículos de su tienda, que se distribuyen a través de cientos de centros de entrega en todo el mundo.

Ejemplos de empleos. De acuerdo con Analytic Talent, hay 6000 empresas que contratan a científicos de datos, aquí se muestran algunos ejemplos de ofertas de trabajo:

- Facebook: Científico de datos
- IBM: Científico de datos
- PayPal: Científico de datos
- The College Board: Programador SAS/Científico de datos
- Netflix: Ingeniero/Científico de datos en jefe

Estadística en ciencias de la información El científico de datos moderno tiene una sólida formación en estadística y sistemas computacionales, así como conocimientos especializados en campos que se extienden más allá de la estadística. El científico de datos moderno podría ser experto en el software Hadoop, que utiliza el procesamiento en paralelo en muchas computadoras para analizar grandes cantidades de datos. También podría tener una sólida formación en algún otro campo como la psicología, biología, medicina, química o economía. Debido a la amplia gama de disciplinas requeridas, un proyecto de ciencias de la información podría incluir un equipo de individuos con experiencia en campos diversos. Un curso introductorio de estadística es un muy buen primer paso para convertirse en un científico de datos.

Datos faltantes

Cuando se recopilan datos muestrales, es bastante común encontrar que faltan algunos valores. Si estos datos faltantes se desprecian, en ocasiones es posible obtener resultados engañosos. Si además se comete el error de omitir algunos valores muestrales al escribirlos manualmente en un programa de software estadístico, es poco probable que los valores faltantes afecten de manera seria los resultados. Sin embargo, si una encuesta incluye muchas entradas de salario faltantes porque las personas con ingresos muy bajos son reacias a revelar sus ingresos, los valores bajos faltantes tendrán el importante efecto de que los salarios estimados sean más altos de lo que realmente son.

Para obtener un ejemplo de datos faltantes, consulte la siguiente tabla. Falta la temperatura corporal para el Sujeto 2 a las 12 AM del Día 2. (La siguiente tabla incluye las tres primeras filas de datos del Conjunto de datos 3 "Temperaturas corporales" del apéndice B).

Temperaturas corporales (en grados Fahrenheit) de adultos saludables

				Temperatura el Día 1		Temperatura el Día 2	
Sujeto	**Edad**	**Sexo**	**Fuma**	**8 AM**	**12 AM**	**8 AM**	**12 AM**
1	22	M	Y	98.0	98.0	98.0	98.6
2	23	M	Y	97.0	97.6	97.4	----
3	22	M	Y	98.6	98.8	97.8	98.6

Hay diferentes categorías de datos faltantes. Vea las siguientes definiciones.

DEFINICIÓN

Un valor de datos **falta completamente al azar** si la probabilidad de su inexistencia es independiente de su valor o del de cualquiera de los otros valores en el conjunto de datos. Es decir, es tan probable que falte cualquier valor como cualquier otro valor de los datos.

(*NOTA:* Un análisis más completo de los datos faltantes distingue entre los datos que *faltan completamente al azar* y los que *faltan al azar*; lo que significa que la probabilidad de que un valor falte es independiente de su valor después de estar controlado por otra variable. No hay necesidad de conocer tal distinción en este libro).

Ejemplo de datos faltantes—Al azar: Mientras utiliza un teclado para introducir manualmente las edades de los encuestados, un capturista se distrae con su compañero cantando "Daydream Believer" y comete el error de no ingresar la edad de 37 años. Este valor de datos falta completamente al azar.

DEFINICIÓN

Un valor de datos **falta de manera no aleatoria** si el valor faltante se relaciona con la razón de su inexistencia.

Ejemplo de datos faltantes-no aleatorios. Una pregunta de encuesta pide a cada encuestado introducir sus ingresos anuales, pero los encuestados con ingresos muy bajos omiten esta pregunta porque les resulta vergonzoso responder.

¿Resultados sesgados? Con base en las dos definiciones y ejemplos de la página anterior, tiene sentido concluir que si se omiten los datos *faltantes completamente al azar*, es poco probable que los valores restantes estén sesgados y deben obtenerse buenos resultados. Sin embargo, si se desprecian los datos que *faltan de manera no aleatoria*, es muy posible que los valores restantes estén sesgados y que los resultados sean engañosos.

Corrección de datos faltantes. Existen diferentes métodos para tratar los datos faltantes.

1. **Eliminación de casos:** Un método muy común para tratar los datos faltantes es eliminar todos los sujetos con valores faltantes.
   - Si los datos son faltantes completamente al azar, es poco probable que los valores restantes estén sesgados y es posible obtener buenos resultados, pero con un tamaño de muestra más pequeño.
   - Si los datos faltantes no son aleatorios, la eliminación de los sujetos con valores faltantes puede dar como resultado un sesgo entre los valores restantes, por lo que los resultados pueden ser engañosos.
2. **Imputación de valores faltantes:** Los valores de datos faltantes se "imputan" al sustituir valores en su lugar. Existen diferentes métodos para determinar los valores de reemplazo, como usar la media de los valores restantes o usar un valor seleccionado aleatoriamente de otros casos similares; también es posible utilizar un método basado en el análisis de regresión (esto tendrá más sentido después de estudiar el capítulo 10).

En este libro no se realiza mucho trabajo con datos faltantes, pero es importante entender lo siguiente:

> **Cuando analice datos muestrales con valores faltantes, trate de determinar por qué están ausentes, y después decida si tiene sentido tratar los valores restantes como representativos de la población. Si parece que faltan valores de manera no aleatoria (es decir, sus valores se relacionan con las razones por las que están ausentes), sepa que los datos restantes pueden estar sesgados y cualquier conclusión basada en tales valores puede resultar engañosa.**

1-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Parámetro y dato estadístico En una encuesta de Harris Interactive aplicada a 2276 adultos en Estados Unidos, se encontró que 33% de los encuestados nunca viajan en líneas aéreas comerciales. Identifique la población y la muestra. ¿Es el valor de 33% un dato estadístico o un parámetro?

2. Datos cuantitativos/categóricos Identifique cada uno de los siguientes casos como datos cuantitativos o categóricos.

a. El número de plaquetas en el Conjunto de datos 1 "Datos corporales" del apéndice B.

b. Las marcas de cigarrillos en el Conjunto de datos 13 "Contenido de los cigarrillos" del apéndice B.

c. Los colores de los caramelos M&M en el Conjunto de datos 27 "Pesos de los M&M" en el apéndice B.

d. Los pesos de los caramelos M&M en el Conjunto de datos 27 "Pesos de los M&M" en el apéndice B.

3. Datos discretos/continuos ¿Cuál de los siguientes casos describe datos discretos?

a. El número de personas encuestadas en cada uno de los próximos años para las Encuestas Nacionales de Exámenes de Salud y Nutrición.

b. Las longitudes exactas de los pies (medidas en cm) de una muestra aleatoria de estudiantes de estadística.

c. Los tiempos exactos en que conductores seleccionados al azar envían mensajes de texto mientras conducen durante los últimos 7 días.

4. Encuesta de salud En una encuesta de 1020 adultos en Estados Unidos, 44% dijo que se lavan las manos después de viajar en transporte público (con base en datos de KRC Research).

a. Identifique la muestra y la población.

b. ¿Es el valor de 44% un dato estadístico o un parámetro?

c. ¿Cuál es el nivel de medición del valor de 44%? (nominal, ordinal, de intervalo o de razón).

d. ¿El número de sujetos en estas encuestas es discreto o continuo?

En los ejercicios 5 a 12, identifique si el valor dado es un estadístico o un parámetro.

5. Vuelos a tiempo En un estudio de los vuelos de American Airlines del Aeropuerto JFK en Nueva York a LAX en Los Ángeles, se seleccionan 48 vuelos al azar y el tiempo promedio (medio) de llegada es con 8.9 minutos de retraso.

6. CHIS Una reciente encuesta de las Entrevistas de Salud de California (CHIS, por sus siglas en inglés) incluyó a 2799 adolescentes residentes en California.

7. Viviendas De acuerdo con la Oficina de Censos, el número total de viviendas en Estados Unidos es de 132,802,859.

8. Muertes en el incendio de Triangle Un desastre mortal en Estados Unidos fue el incendio en la fábrica de Triangle Shirtwaist en la ciudad de Nueva York. En ese incendio falleció una población de 146 trabajadores.

9. Peso al nacer En un estudio de 400 bebés nacidos en cuatro hospitales del Estado de Nueva York, se encontró que la media (promedio) del peso al nacer fue de 3152.0 gramos.

10. Género de nacimiento En el mismo estudio citado en el ejercicio anterior, 51% de los bebés fueron niñas.

11. *Titanic* Se realizó un estudio de los 2223 pasajeros que estaban a bordo del *Titanic* cuando se hundió.

12. Tabla periódica La media (el promedio) del peso atómico de todos los elementos en la tabla periódica es de 134.355 unidades de masa atómica unificadas.

En los ejercicios 13 a 20, determine si los datos provienen de un conjunto de datos discretos o continuos.

13. Estudiante de primer año 15 En un estudio del aumento de peso de los estudiantes universitarios en su primer año, los investigadores registran las cantidades de peso que estudiantes seleccionados al azar aumentaron (como en el Conjunto de datos 6 "Estudiantes de primer año 15" en el apéndice B).

14. CHIS Entre los sujetos de la encuesta de Entrevistas de Salud en California (CHIS, por sus siglas en inglés), se seleccionaron varios sujetos al azar y se registraron sus estaturas.

15. McDonald's En un estudio de los tiempos de servicio en una ventanilla de autoservicio de McDonald's, se registra el número de automóviles atendidos cada hora durante varios días.

16. Asistencia a la Cámara El secretario de la Cámara de Representantes de Estados Unidos registra el número de representantes presentes en cada sesión.

17. Corvettes Un gerente de turno registra las cantidades de Corvettes fabricados durante cada día de producción.

18. Técnicas de criminalística El estudio de la relación entre las longitudes de los pies y las estaturas, cuando la huella es evidencia en una escena del crimen, permite la estimación de la estatura del sospechoso; un investigador registra la longitud exacta de los pies a partir de una gran muestra aleatoria de sujetos.

19. Teléfonos inteligentes Los estudiantes de una clase de estadística registran el tiempo exacto que usan secretamente sus teléfonos inteligentes durante la clase.

20. Muertes por mensajes de texto El Instituto para la Seguridad en Carreteras recopila datos que consisten en el número de muertes en vehículos de motor causadas por conducir mientras se envían mensajes de texto.

En los ejercicios 21 a 28, determine cuál de los cuatro niveles de medición (nominal, ordinal, de intervalo o de razón) es el más apropiado.

21. Clasificaciones de universidades El *U.S. News & World Report* proporciona periódicamente su clasificación de universidades en Estados Unidos, y en un año reciente las clasificaciones para Princeton, Harvard y Yale fueron 1, 2 y 3, respectivamente.

22. Encuesta de salida Para la elección presidencial de 2016, ABC News realizó una encuesta de salida en la que se pidió a los votantes identificar el partido político (demócrata, republicano, etcétera) por el que votaron.

23. Colores de M&Ms Colores de M&Ms (rojo, naranja, amarillo, café, azul, verde) listados en el Conjunto de datos 27 "Pesos de M&M" del apéndice B.

24. Tiempos de servicio en la comida rápida En un estudio de los tiempos de servicio en la comida rápida, un investigador registra los intervalos de tiempo que inician cuando los clientes hacen su pedido y terminan cuando lo reciben.

25. Béisbol El estadístico de béisbol Bill James registra los años en los que un equipo de la Liga Nacional gana la Serie Mundial de béisbol.

26. Calificaciones de una película El autor de este libro calificó la película *Guerra de las galaxias: El despertar de la fuerza* con 5 estrellas en una escala de 5 estrellas.

27. Plomo en la sangre Los niveles de plomo en la sangre bajo, medio y alto, se usan para describir a los sujetos en el Conjunto de datos 7 "CI y Plomo" del apéndice B.

28. Temperaturas corporales Las temperaturas corporales (en grados Fahrenheit) listadas en el Conjunto de datos 3 "Temperaturas corporales" del apéndice B.

En los ejercicios 29 a 32, identifique el nivel de medición de los datos como nominal, ordinal, de intervalo o de razón. Asimismo, explique qué es incorrecto en el cálculo dado.

29. Súper Tazón El primer Súper Tazón al que asistió el autor de este libro fue el XLVIII. En la primera jugada del partido, la defensa de Seattle anotó con un balón suelto. Los jugadores defensivos llevaban las camisetas 31, 28, 41, 56, 25, 54, 69, 50, 91, 72 y 29. La media (el promedio) de esos números es 49.6.

30. Números de Seguridad Social Como parte de un proyecto en una clase de estadística, los estudiantes reportan los últimos cuatro dígitos de sus números de Seguro Social, y la media (el promedio) de esos dígitos es 4.7.

31. Temperaturas Mientras este ejercicio se escribe, hay 80°F en la casa del autor y 40°F en Auckland, Nueva Zelanda; por lo que hay el doble de calor en la casa del autor que en Auckland, Nueva Zelanda.

32. Clasificaciones de universidades Mientras esto se escribe, *U.S. News & World Report* clasificaba a las universidades del país, incluyendo los siguientes resultados: Princeton (1), Harvard (2), Yale (3) y Columbia (4). La diferencia entre Princeton y Harvard es igual a la diferencia entre Yale y Columbia.

1-2 Más allá de lo básico

33. Contabilizable En cada uno de los siguientes casos, categorice la naturaleza de los datos con una de las siguientes descripciones: (1) discreta porque el número de valores posibles es finito; (2) discreta porque el número de valores posibles es infinito pero contabilizable; (3) continua porque el número de valores posibles es infinito y no contabilizable.

a. Longitudes exactas de los pies de los miembros de la banda los Monkees.

b. Tamaños de zapato de los miembros de la banda de los Monkees (por ejemplo 9, 9½, etcétera).

c. El número de discos vendidos por la banda de los Monkees.

d. El número de monos que se sientan frente a un teclado antes de que uno de ellos toque aleatoriamente la letra de la canción “Daydream Believer”.

1-3 Recopilación de datos muestrales

Concepto clave Cuando se utiliza la estadística en un estudio, la planificación es muy importante y resulta esencial utilizar un método apropiado para recopilar los datos muestrales. Esta sección incluye comentarios sobre varios métodos y procedimientos de muestreo. De particular importancia es el método consistente en utilizar una *muestra aleatoria simple*. En el resto del presente libro se hará uso frecuente de este método de muestreo.

Mientras lea esta sección, recuerde lo siguiente:

Si los datos muestrales no se recopilan de manera apropiada, pueden ser tan inútiles que ni la aplicación de innumerables trucos estadísticos pueda rescatarlos.

PARTE 1 Fundamentos del diseño de experimentos y recopilación de datos muestrales

La regla de oro La aleatorización en los grupos de tratamiento con placebos suele denominarse la “regla de oro”, debido a su eficacia. (Un placebo, por ejemplo una píldora de azúcar, no tiene efecto medicinal). En el siguiente ejemplo se describe cómo se utilizó la regla de oro en el experimento médico más grande jamás realizado.

EJEMPLO 1 Experimento de la vacuna Salk

En 1954, se diseñó un experimento para probar la eficacia de la vacuna Salk en la prevención de la poliomielitis, que había matado o paralizado a miles de niños. Por selección aleatoria se asignaron 401,914 niños a dos grupos: (1) 200,745 niños recibieron un tratamiento consistente en inyecciones de la vacuna Salk; (2) a 201,229 niños se les inyectó un placebo que no contenía ningún medicamento. Los niños fueron asignados al grupo con tratamiento o con placebo mediante un proceso de selección aleatoria, equivalente a lanzar una moneda. Entre los niños que recibieron la vacuna Salk, 33 desarrollaron posteriormente polio paralizante, y entre los niños a los que se administró un placebo, 115 la desarrollaron.

El ejemplo 1 describe un *experimento* porque a los sujetos se les dio un tratamiento, pero las consideraciones éticas, de costo, tiempo, etcétera, a veces prohíben el uso de un experimento. Nunca querríamos llevar a cabo un experimento de conducción y mensajes de texto en el que solicitamos a los sujetos enviar textos mientras conducen; algunos de ellos podrían morir. Sería mucho mejor observar los resultados de los accidentes ocurridos para entender los efectos de conducir mientras se envían textos. Vea las siguientes definiciones.

DEFINICIONES

En un **experimento** aplicamos algún *tratamiento* y después procedemos a observar sus efectos sobre los individuos. (Los individuos en los experimentos se llaman **unidades experimentales**, y con frecuencia se denominan **sujetos** cuando son personas).

En un **estudio observacional**, medimos y registramos características específicas, pero no intentamos *modificar* los individuos bajo estudio.

Los experimentos suelen ser mejores que los estudios observacionales porque cuando son bien planificados reducen la posibilidad de obtener resultados afectados por alguna variable que no forma parte del estudio. Una *variable de confusión* es aquella que afecta a las variables involucradas en el estudio, pero no está incluida en éste.

EJEMPLO 2 Helado y ahogamientos

Estudio observacional: Observar los datos históricos para concluir que el helado causa ahogamientos (con base en datos que muestran que el aumento en las ventas de helado está asociado con el incremento en los ahogamientos). El error es no considerar la variable de la temperatura que se encuentra oculta y el no ver que a medida que la temperatura aumenta, las ventas de helados se incrementan y los ahogamientos aumentan porque hay más gente que se mete al agua a refrescarse.

Experimento: Realizar un *experimento* donde a un grupo se le administra helado y a otro no. Veríamos que la tasa de víctimas de ahogamiento es casi la misma en ambos grupos, por lo que el consumo de helado no tiene efecto sobre los ahogamientos. Aquí, el experimento es claramente mejor que el estudio observacional.

Diseño de experimentos

El buen diseño de los experimentos incluye la *réplica*, el *estudio a ciegas* y la *aleatorización*.

- **Réplica** se refiere a la repetición de un experimento en más de un individuo. Un buen uso de la repetición requiere tamaños de muestra suficientemente grandes para que sea posible ver los efectos de los tratamientos. En el experimento Salk del ejemplo 1, se usaron tamaños de muestra suficientemente grandes, por lo que los investigadores pudieron ver que la vacuna Salk era efectiva.
- El **estudio a ciegas** se utiliza cuando el sujeto no sabe si está recibiendo un tratamiento o un placebo. Es una forma de evitar el **efecto placebo**, que ocurre cuando un sujeto no tratado informa una mejora en los síntomas. El experimento Salk del ejemplo 1 fue doblemente a ciegas, lo que significa que existieron dos niveles: (1) los niños que recibían la inyección no sabían si estaban recibiendo la vacuna Salk o un placebo, y (2) los médicos que dieron las inyecciones y evaluaron los resultados tampoco lo sabían. Se utilizaron claves para que los investigadores pudieran evaluar objetivamente la eficacia de la vacuna Salk.
- La **aleatorización** se utiliza cuando los individuos son asignados a diferentes grupos a través de un proceso de selección aleatoria, como en el experimento de la vacuna Salk del ejemplo 1. La lógica detrás de la aleatorización es utilizar el azar como una forma de crear dos grupos similares. La siguiente definición se refiere a una manera común y efectiva de recolectar datos muestrales utilizando la aleatorización.

Ensayos clínicos contra estudios observacionales

En un artículo del *New York Times* acerca de la terapia hormonal para mujeres, la reportera Denise Grady escribió sobre un informe de tratamientos probados en ensayos clínicos que involucraron sujetos aleatoriamente asignados a un grupo de tratamiento o a otro grupo que no recibió tratamiento. Tales ensayos clínicos aleatorios suelen conocerse como la “regla de oro” para la investigación médica. En contraste, los estudios observacionales pueden involucrar a pacientes que deciden someterse a algún tratamiento. Los sujetos que deciden recibir tratamientos suelen estar más saludables que los demás, por lo que el grupo de tratamiento puede parecer más exitoso simplemente porque involucra pacientes más sanos, no necesariamente porque el tratamiento sea efectivo. Los investigadores criticaron los estudios observacionales acerca de la terapia hormonal para mujeres argumentando que los resultados podrían hacer ver que el tratamiento es más eficaz de lo que realmente es.

DEFINICIÓN

Una **muestra aleatoria simple** de *n* sujetos se selecciona de modo que cada *muestra* posible *del mismo tamaño n* tiene la misma posibilidad de ser elegida. (Con frecuencia, una muestra aleatoria simple se denomina muestra aleatoria, pero en sentido estricto una *muestra aleatoria* tiene el requisito más débil de que todos los miembros de la población tengan la misma posibilidad de ser seleccionados. Esta distinción no es tan importante en el presente texto. Vea el ejercicio “Muestra aleatoria simple contra muestra aleatoria”).

A lo largo de este libro se utilizarán diversos procedimientos estadísticos, y a menudo existirá el requisito de recolectar una *muestra aleatoria simple*, tal como se acaba de definir.

A diferencia del muestreo descuidado o casual, el muestreo aleatorio generalmente requiere una planificación y ejecución muy cuidadosa. Wayne Barber, del Chemeketa Community College, está en lo cierto cuando dice a sus estudiantes que “la aleatoriedad necesita ayuda”.

Otros métodos de muestreo Además del muestreo aleatorio simple, a continuación se presentan algunos otros métodos de muestreo comúnmente utilizados en las encuestas. La figura 1-3 ilustra estos diferentes métodos de muestreo.

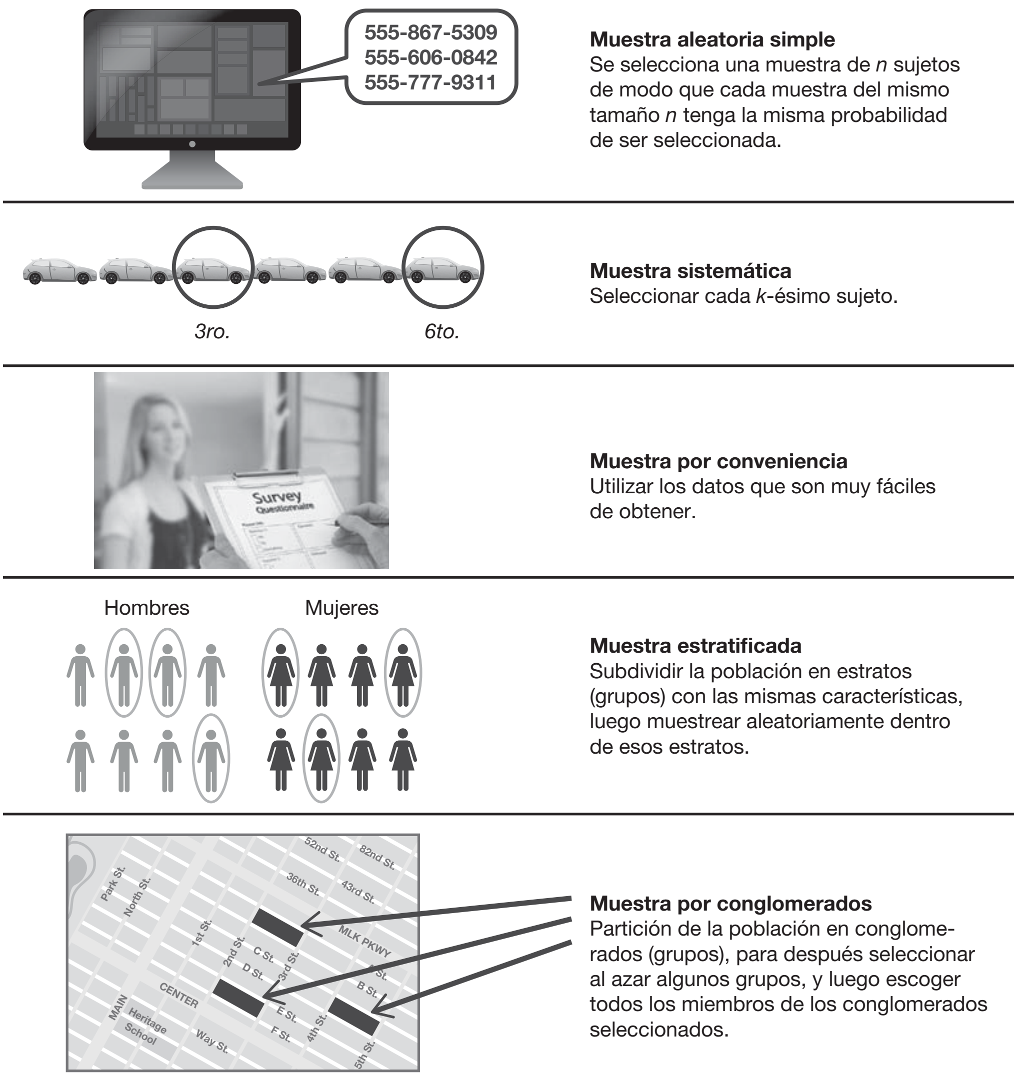


FIGURA 1-3 Métodos comunes de muestreo

Los efectos Hawthorne y del experimentador

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

DEFINICIONES

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

Con el **muestreo por comodidad, de conveniencia o sin norma**, simplemente utilizamos los datos que son muy fáciles de obtener.

En el **muestreo estratificado**, subdividimos la población en al menos dos subgrupos diferentes (o estratos) de modo que los sujetos dentro del mismo subgrupo compartan las mismas características (como el género). A continuación se extrae una muestra de cada subgrupo (o estrato).

En el **muestreo por conglomerados**, primero dividimos el área de la población en secciones (o conglomerados). Después seleccionamos aleatoriamente algunos de esos grupos y elegimos todos los miembros de los grupos seleccionados.

Muestreo en etapas múltiples Con frecuencia, los encuestadores profesionales y los investigadores gubernamentales recopilan datos utilizando una combinación de los métodos de muestreo anteriores. En un diseño muestral en etapas múltiples, los encuestadores seleccionan una muestra en diferentes etapas, y cada una de ellas puede utilizar distintos métodos de muestreo, como en el ejemplo siguiente.

EJEMPLO 3 Diseño muestral en etapas múltiples

Las estadísticas de desempleo del gobierno de Estados Unidos se basan en encuestas a domicilio. No es práctico examinar personalmente cada hogar en una muestra aleatoria simple, ya que estarían dispersos por todo el país. En cambio, la Oficina de Censos de Estados Unidos y la Oficina de Estadísticas Laborales colaboran para llevar a cabo un estudio llamado Encuesta de la Población Actual. Una encuesta reciente incorpora un diseño muestral en etapas múltiples, siguiendo en general los pasos descritos a continuación:

1. La totalidad de Estados Unidos se divide en 2,007 regiones llamadas *unidades primarias de muestreo* (PSU por sus siglas en inglés). Las unidades de muestreo primarias son áreas metropolitanas, condados grandes o combinaciones de condados más pequeños. Las 2,007 unidades primarias de muestreo se agrupan en 824 estratos.
2. En cada uno de los 824 estratos se selecciona una de las unidades de muestreo primarias de manera que la probabilidad de selección sea proporcional al tamaño de la población en cada unidad de muestreo primario.
3. En cada una de las 824 unidades de muestreo primario seleccionadas, los datos del censo se usan para identificar un *distrito de enumeración* del censo, cada uno con aproximadamente 300 hogares. Los distritos de enumeración se seleccionan al azar.
4. En cada uno de los distritos de enumeración seleccionados, los grupos de aproximadamente cuatro direcciones (contiguas siempre que sea posible) se seleccionan al azar.
5. Una persona responsable en cada uno de los 60,000 hogares seleccionados es entrevistada sobre la situación laboral de cada miembro del hogar de 16 años de edad o más.

Este diseño muestral en etapas múltiples incluye una combinación de muestreo aleatorio, estratificado y de conglomerados en diferentes etapas. El resultado final es un diseño muestral muy complicado, pero es mucho más práctico, menos costoso y más rápido que usar un diseño más sencillo, como una muestra aleatoria simple.

PARTE 2 Más allá de lo básico sobre el diseño de experimentos y la recolección de datos muestrales

En la parte 2 de esta sección se analizan diferentes tipos de estudios observacionales y distintas maneras de diseñar experimentos.

Estudios observacionales. Las siguientes definiciones identifican la terminología estándar utilizada en revistas profesionales para diferentes tipos de estudios observacionales. Estas definiciones se ilustran en la figura 1-4.

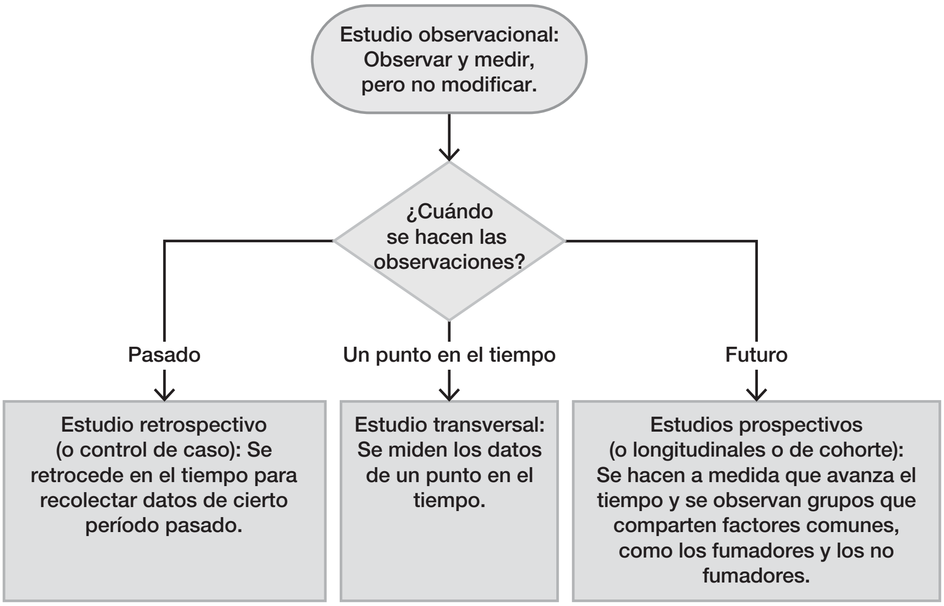


FIGURA 1-4 Tipos de estudios observacionales

DEFINICIONES

En un **estudio transversal**, los datos se observan, se miden y se recolectan en un momento dado, no durante un período determinado.

En un **estudio retrospectivo** (o de **control de caso**), se recolectan datos correspondienes a un periodo del pasado (a través del análisis de registros, entrevistas, etcétera).

En un estudio **prospectivo** (o **longitudinal** o **de cohorte**), los datos se recolectan en el futuro a partir de grupos que comparten factores comunes (estos grupos se denominan *cohortes*).

Experimentos En un experimento se presenta **confusión** cuando podemos ver algún efecto, pero no podemos identificar el factor específico que lo causó, como en el estudio observacional del helado y los ahogamientos del ejemplo 2. Vea también el mal diseño experimental ilustrado en la figura 1-5(a), donde puede ocurrir confusión cuando el grupo de tratamiento de mujeres muestra fuertes resultados positivos. Debido a que el grupo de tratamiento consta de mujeres y el grupo placebo está formado por hombres, la confusión ha ocurrido porque no podemos determinar si el tratamiento o el género de los sujetos causaron los resultados positivos. El experimento de la vacuna Salk en el ejemplo 1 ilustra un método para controlar el efecto de la variable de tratamiento: Utilizar un *diseño experimental completamente aleatorio*,

¿Las mujeres ganan menos que los hombres?

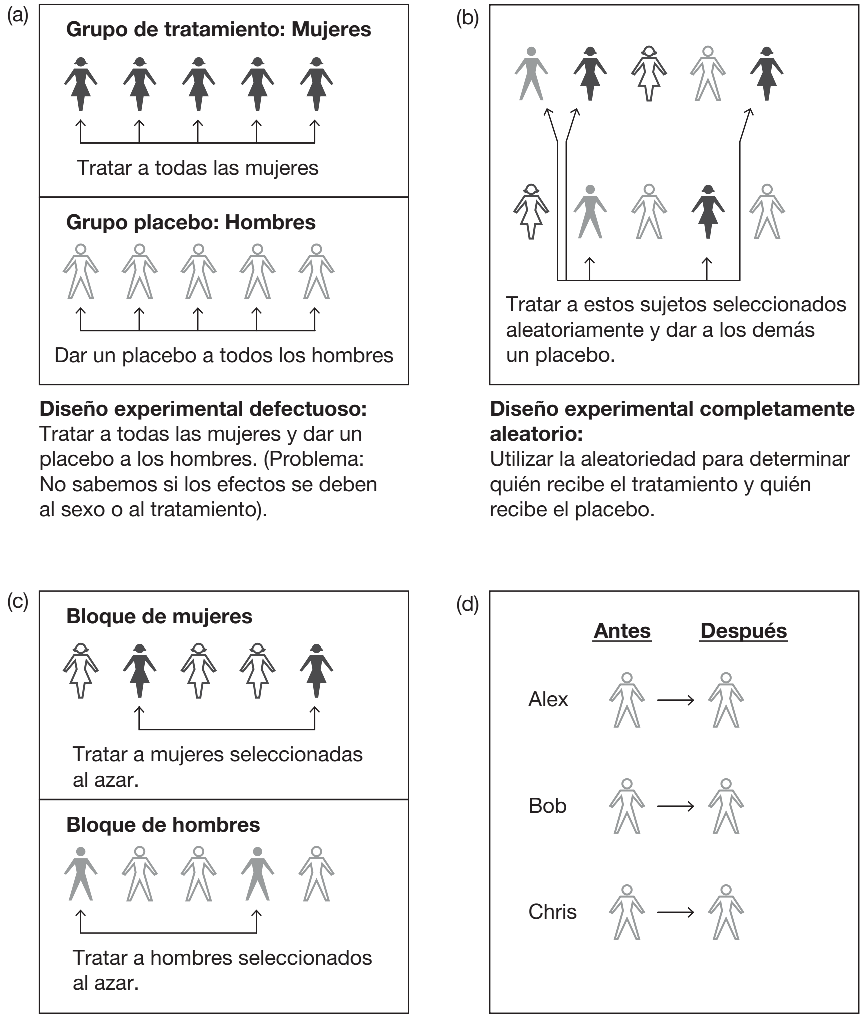
La evidencia de la Oficina de Censos y la Oficina de Estadísticas Laborales indican que las mujeres ganan alrededor de 77% de lo que ganan los hombres. Jillian Berman informó en el *Huffington Post* que la compañía PayScale utilizó los datos de los salarios de sus millones de usuarios del sitio web para concluir que los hombres y las mujeres ganan aproximadamente lo mismo cuando comienzan sus carreras, pero los hombres tienden a ganar más a medida que avanzan. Afirmó que, según el estudio, “las mujeres que trabajan en diversos puestos no gerenciales ganan alrededor de 98% de lo que ganan los hombres en promedio”. Berman señala que esta conclusión se basa en los datos de PayScale, que los usuarios del sitio web reportan en encuestas en línea, y no en datos obtenidos de las oficinas gubernamentales. El estudio PayScale explica factores como la educación y las responsabilidades laborales. Este estudio parece confirmar que las mujeres ocupan desproporcionadamente menos puestos de trabajo de alto nivel y desproporcionadamente más puestos de trabajo de bajo nivel, por lo que resulta claro que existe una brecha de género.

Valor de una vida estadística

El *valor de una vida estadística* (VSL, por sus siglas en inglés) es una medida rutinariamente calculada y utilizada para tomar decisiones en campos tales como la medicina, los seguros, la salud ambiental y la seguridad en el transporte. Al momento de escribir este libro, el valor de una vida estadística era de 6.9 millones de dólares.

Muchas personas se oponen al concepto de poner un valor a una vida humana, pero la palabra estadística en "valor de una vida estadística" se utiliza para asegurar que no se equipare con el verdadero valor de una vida humana. Algunas personas sostienen legítimamente que una vida humana es invaluable, pero otros afirman que hay condiciones en las que es imposible o impracticable salvar todas las vidas, por lo que de alguna manera debe asignarse un valor a una vida humana para poder tomar decisiones racionales y sanas. No lejos de la casa del autor de este texto, se modificó una avenida a un costo de alrededor de 3 millones de dólares con el fin de mejorar la seguridad en un lugar donde anteriormente habían fallecido automovilistas en accidentes de tránsito. En el análisis costo-beneficio que condujo a esta mejora en la seguridad, seguramente se consideró el valor de una vida estadística.

utilizando el azar para asignar sujetos al grupo de tratamiento y al grupo placebo. Un diseño experimental completamente aleatorio es uno de los siguientes métodos que se utilizan para controlar los efectos de las variables.



Diseño experimental defectuoso: Tratar a todas las mujeres y dar un placebo a los hombres. (Problema: No sabemos si los efectos se deben al sexo o al tratamiento).

Diseño experimental completamente aleatorio: Utilizar la aleatoriedad para determinar quién recibe el tratamiento y quién recibe el placebo.

Diseño de bloques aleatorizados:
1. Formar un bloque de mujeres y un bloque de hombres.
2. Dentro de cada bloque, seleccionar aleatoriamente los sujetos a tratar.

Diseño de pares relacionados: Obtener mediciones de los mismos sujetos antes y después de algún tratamiento.

FIGURA 1-5 Diseños de experimentos

Diseño experimental completamente aleatorio: Se asignan sujetos a diferentes grupos de tratamiento a través de un proceso de selección aleatoria, como se ilustra en la figura 1-5(b).

Diseño de bloques aleatorizados: Vea la figura 1-5(c). Un bloque es un grupo de sujetos que son similares, pero los bloques difieren de maneras que pueden afectar el resultado del experimento. Se utiliza el procedimiento ilustrado en la figura 1-5(c):

1. Formar bloques (o grupos) de sujetos con características similares.
2. Asignar aleatoriamente tratamientos a los sujetos dentro de cada bloque.

Por ejemplo, al diseñar un experimento para probar la efectividad de los tratamientos con aspirina contra enfermedades del corazón, podríamos formar un bloque de hombres y un bloque de mujeres, porque se sabe que los corazones de los hombres y las mujeres pueden comportarse de manera diferente. Al controlar el género, este diseño de bloques aleatorizados elimina el género como una posible fuente de confusión.

Un diseño de bloques aleatorizados utiliza la misma idea básica que el muestreo estratificado, pero el primero se utiliza en el diseño de experimentos, mientras que el muestreo estratificado se usa en encuestas.

Diseño de pares relacionados: Se comparan dos grupos de tratamiento (por ejemplo tratamiento y placebo) utilizando sujetos relacionados en pares que de alguna manera tengan características similares, como en los siguientes casos.

- Antes/Después: Los pares relacionados pueden consistir en mediciones de sujetos antes y después de algún tratamiento, como se ilustra en la figura 1-5(d). Cada sujeto produce una medición "antes" y una medición "después", y cada par de mediciones antes/después es un par relacionado.
- Gemelos: Una prueba de la pasta dentífrica Crest utiliza parejas de gemelos, donde un gemelo utiliza Crest y el otro utiliza otra pasta de dientes.

Diseño rigurosamente controlado: Se asignan cuidadosamente los sujetos a diferentes grupos de tratamiento, de manera que los que reciben cada tratamiento sean similares de un modo importante para el experimento. Esto puede ser extremadamente difícil de implementar, y con frecuencia no se tiene la seguridad de haber tomado en cuenta todos los factores relevantes.

Errores de muestreo

En estadística, es posible utilizar un buen método de muestreo y hacer todo correctamente y, no obstante, obtener resultados erróneos. No importa cuán bien se planee y ejecute el proceso de recolección de muestras, es probable que haya algún error en los resultados. Aquí se describen los diferentes tipos de errores de muestreo.

DEFINICIONES

Un **error de muestreo** (o **error de muestreo aleatorio**) ocurre cuando la muestra ha sido seleccionada con un método aleatorio, pero hay una discrepancia entre el resultado de la muestra y el resultado de la población real; tal error es el resultado de las fluctuaciones probabilísticas de la muestra.

Un **error no muestral** es el resultado de un error humano, incluyendo factores tales como registros incorrectos de datos, errores computacionales, preguntas con redacción sesgada, datos falsos proporcionados por los encuestados, conclusiones sesgadas o aplicación de métodos estadísticos que no son apropiados para las circunstancias.

Un **error de muestreo no aleatorio** es el resultado de utilizar un método de muestreo que no es aleatorio, como una muestra por comodidad o por conveniencia, o una muestra de respuesta voluntaria.

El diseño experimental requiere mucho más pensamiento y cuidado de lo que es posible describir en esta sección relativamente breve. Tomar un curso completo sobre el diseño de experimentos es un buen comienzo para aprender mucho más sobre este importante tema.

Dificultades de las encuestas

Las encuestas constituyen un negocio enorme y creciente en Estados Unidos, pero sus resultados pueden verse comprometidos por muchos factores. Cada vez más personas se niegan a responder; actualmente, el promedio de la tasa de respuesta es de aproximadamente 22%, en comparación con 36% alrededor del año 2000. Un número creciente de personas son más difíciles de encontrar porque utilizan teléfonos celulares (no directorios); alrededor de 15% de los adultos ahora tienen teléfonos celulares y no teléfonos fijos, y tienden a ser más jóvenes que el promedio. Hay problemas obvios asociados con las encuestas que preguntan a los sujetos sobre el uso de drogas, el robo o el comportamiento sexual, y se presenta un sesgo de deseabilidad social cuando los encuestados no son honestos porque no quieren ser vistos negativamente por la persona que realiza la entrevista.

1-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Tratamiento contra el dolor de espalda En un estudio diseñado para probar la eficacia del paracetamol (o acetaminofén) como tratamiento para el dolor de espalda baja, se asignaron aleatoriamente 1643 pacientes a uno de tres grupos: (1) los 547 sujetos en el grupo placebo recibieron píldoras sin medicamentos; (2) 550 sujetos estaban en un grupo que recibía píldoras con paracetamol tomadas a intervalos regulares; (3) 546 sujetos estaban en un grupo que recibía píldoras con paracetamol para que las tomaran cada vez que fuera necesario para aliviar el dolor. (Vea "Eficacia del paracetamol en el tratamiento del dolor agudo de espalda baja", *"Eficacy of Paracetamol for Acute Low-Back Pain"*). ¿Es este estudio un experimento o un estudio observacional? Explique.

2. Estudio a ciegas ¿Qué significa cuando decimos que el estudio citado en el ejercicio 1 fue "doblemente a ciegas"?

3. Réplica ¿De qué manera específica se aplicó la repetición en el estudio citado en el ejercicio 1?

4. Método de muestreo Los pacientes incluidos en el estudio citado en el ejercicio 1 eran personas "que buscaban atención para el dolor lumbar directamente o en respuesta a un anuncio comunitario". ¿Qué tipo de muestreo describe mejor la forma en que se eligieron los 1643 sujetos: muestra aleatoria simple, muestra sistemática, muestra de conveniencia, muestra estratificada, muestra por conglomerados? ¿El método de muestreo parece afectar adversamente la calidad de los resultados?

Los ejercicios 5 a 8 se refieren al estudio de una asociación entre qué oreja se utiliza para las llamadas con teléfono celular y si el sujeto es zurdo o diestro. El estudio se publicó en "Dominancia hemisférica y uso del celular", **"Hemispheric Dominance and Cell Phone Use"** ***de Seidman et al., JAMA Otolaryngology - Head & Neck Surgery, vol. 139, núm. 5. El estudio comenzó con una encuesta enviada por correo electrónico a 5000 personas pertenecientes a un grupo de otología en línea y se recibió la respuesta a 717 encuestas (*****otología** ***se refiere al oído y la audición).***

5. Método de muestreo ¿Qué tipo de muestreo describe mejor la forma en que se eligieron los 717 sujetos: muestra aleatoria simple, muestra sistemática, muestra de conveniencia, muestra estratificada, muestra por conglomerados? ¿El método de muestreo parece afectar adversamente la calidad de los resultados?

6. Experimento o estudio observacional ¿Es el estudio un experimento o un estudio observacional? Explique.

7. Tasa de respuesta ¿Qué porcentaje de las 5000 encuestas fueron respondidas? ¿Parece que esa tasa de respuesta es baja? En general, ¿cuál es el problema con una tasa de respuesta muy baja?

8. Método de muestreo Suponga que la población está formada por todos los estudiantes que se encuentran actualmente en su clase de estadística. Describa cómo obtener una muestra de seis estudiantes para que el resultado sea una muestra de cada uno de los siguientes tipos.

a. Muestra aleatoria simple

b. Muestra sistemática

c. Muestra estratificada

d. Muestra por conglomerados

En los ejercicios 9 a 20, identifique el tipo de muestreo utilizado: aleatorio, sistemático, de conveniencia, estratificado o por conglomerados.

9. Densidad de cormoranes Las densidades de población de las aves llamadas cormoranes fueron estudiadas usando el "método de transecto lineal" con observaciones aéreas realizadas mediante vuelos a lo largo de la costa del Lago Hurón y la recolección de datos muestrales a intervalos de 20 km (con base en datos del *Journal of Great Lakes Research*).

10. Sexualidad de las mujeres La sexualidad de las mujeres fue analizada en el libro de Shere Hite *Women and Love: A Cultural Revolution*. Sus conclusiones se basaron en datos muestrales consistentes en 4,500 encuestas enviadas por correo como respuesta a 100,000 cuestionarios enviados a mujeres.

11. Encuesta OVNI En una encuesta de Kelton Research, 1114 estadounidenses mayores de 18 años recibieron llamadas telefónicas después de que sus números fueron generados aleatoriamente por una computadora, y 36% de los encuestados dijeron que creían en la existencia de ovnis.

12. Encuesta en clase El autor de este libro encuestó una muestra de la población de su clase de estadística mediante la identificación de los grupos de hombres y mujeres, para después seleccionar al azar a cinco estudiantes de cada uno de los dos grupos.

13. Conducción Un alumno del autor llevó a cabo una encuesta sobre los hábitos de conducción mediante la selección aleatoria de tres clases diferentes y la aplicación de la encuesta a todos los estudiantes que salían de esas clases.

14. Estudio de acupuntura En un estudio de tratamientos para el dolor de espalda, 641 sujetos fueron asignados al azar a cuatro grupos de tratamiento con acupuntura individualizada, acupuntura estándar, acupuntura simulada y el tratamiento habitual (con base en los datos de un ensayo aleatorizado que compara la acupuntura, la acupuntura simulada y el tratamiento habitual para el dolor crónico de espalda baja, de Cherkin *et al*., *Archives of Internal Medicine* vol. 169, núm. 9).

15. Diccionario El autor de este texto recopiló datos muestrales seleccionando al azar cinco libros de cada una de las siguientes categorías: ciencia, ficción e historia. Después identificó el número de páginas en los libros.

16. Tasas de deforestación Se utilizan satélites para recolectar datos muestrales con el fin de estimar las tasas de deforestación. La Evaluación de Recursos Forestales de la Organización de las Naciones Unidas (ONU) para la Agricultura y la Alimentación utiliza un método para seleccionar una muestra de un cuadrado de 10 km de ancho en cada intersección de 1° de latitud y longitud.

17. Prueba de Lipitor En un ensayo clínico del medicamento para el colesterol Lipitor (atorvastatina), los sujetos fueron divididos en grupos que recibieron placebo o dosis de Lipitor de 10 mg, 20 mg, 40 mg u 80 mg. Los sujetos fueron asignados aleatoriamente a los diferentes grupos de tratamiento (con base en datos de Pfizer Inc.).

18. Encuestas de salida Durante las últimas elecciones presidenciales, CNN llevó a cabo una encuesta de salida en la que se seleccionaron mesas de votación aleatoriamente y se encuestó a todos los votantes a la salida de éstas.

19. Encuesta de *Literary Digest* En 1936, la revista *Literary Digest* envió cuestionarios por correo a 10 millones de personas y obtuvo 2,266,566 respuestas. Las respuestas indicaban que Alf Landon ganaría las elecciones presidenciales. No lo hizo.

20. Resistencia de una autopista El Departamento de Transporte del Estado de Nueva York evaluó la calidad de una autopista que cruza todo el estado mediante pruebas a muestras recolectadas a intervalos regulares de 1 milla.

Pensamiento crítico: ¿Qué es erróneo? ***En los ejercicios 21 a 28, determine si el estudio es un experimento o un estudio observacional, y después identifique un problema importante del estudio.***

21. Noticias en línea En una encuesta realizada por *USA Today*, 1465 usuarios de Internet decidieron responder a la siguiente pregunta publicada en la edición electrónica del periódico: "¿Las noticias en línea son tan satisfactorias como las noticias impresas y televisivas?". Cincuenta y dos por ciento de los encuestados respondieron "sí".

22. Estudio de salud de los médicos El Estudio de la Salud de los Médicos incluyó a 22,071 médicos varones. Con base en selecciones aleatorias, 11,037 de ellos fueron tratados con aspirina y los restantes 11,034 recibieron placebos. El estudio terminó pronto porque se hizo evidente que la aspirina redujo el riesgo de infartos al miocardio en una cantidad sustancial.

23. Beber y conducir Un investigador de un consorcio de compañías de seguros planea probar los efectos de la bebida en la capacidad de conducción al seleccionar aleatoriamente a 1000 conductores y luego asignarlos al azar a dos grupos: Un grupo de 500 conducirá en Nueva York después de no consumir alcohol y el segundo grupo conducirá en la misma ciudad después de consumir tres tragos del whisky bourbon Jim Beam.

24. Presión sanguínea Una investigadora médica realizó un ensayo para encontrar la diferencia en la presión sanguínea sistólica entre los estudiantes de ambos sexos que tenían 12 años de edad. Seleccionó aleatoriamente a cuatro hombres y cuatro mujeres para su estudio.

25. Agresividad de conductores En la prueba de un tratamiento diseñado para reducir la agresividad de los conductores en Estados Unidos, el plan original era utilizar una muestra de 500 conductores seleccionados al azar en todo el país. Los directores del programa saben que obtendrán una muestra sesgada si limitan su estudio a los conductores en la ciudad de Nueva York, por lo que planean compensar ese sesgo utilizando una muestra más grande de 3000 conductores en dicha ciudad.

26. Programa de pérdida de peso Atkins Un investigador independiente probó la eficacia del programa de pérdida de peso Atkins seleccionando al azar a 1000 sujetos que utilizan ese programa. Se pidió a cada uno de los sujetos que reportara su peso antes de la dieta y después de ésta.

27. Investigación sobre delincuencia Un sociólogo ha creado una breve encuesta que se aplicará a 2000 adultos seleccionados al azar entre la población de Estados Unidos. Las dos primeras preguntas de la encuesta son: (1) ¿Alguna vez ha sido víctima de un delito grave? (2) ¿Alguna vez ha sido condenado por un delito grave?

28. Medicamentos Investigación y Fabricación Farmacéutica de Estados Unidos quiere información sobre el consumo de varios medicamentos. Un investigador independiente realiza una encuesta enviando 10,000 cuestionarios a adultos seleccionados al azar en Estados Unidos, y recibe 152 respuestas.

1-3 Más allá de lo básico

En los ejercicios 29 a 32, indique si el estudio observacional utilizado es transversal, retrospectivo o prospectivo.

29. Estudio de la salud de las enfermeras II La fase II del estudio de salud de las enfermeras inició en 1989 con 116,000 enfermeras registradas. El estudio está en curso.

30. Estudio de la salud del corazón Se seleccionaron muestras de sujetos con y sin enfermedades cardíacas, luego los investigadores retrocedieron en el tiempo para determinar si tomaban aspirina de forma regular.

31. Estudio sobre la marihuana Los investigadores de los Institutos Nacionales de Salud quieren determinar las tasas actuales de consumo de marihuana entre los adultos que viven en los estados donde se ha legalizado su uso. Realizan una encuesta a 500 adultos en esos estados.

32. Estudio Framingham sobre el corazón El Estudio Framingham sobre el corazón inició en 1948 y está en curso. Se centra en las enfermedades cardiacas.

En los ejercicios 33 a 36, identifique cuál de los siguientes diseños es el más adecuado para el experimento dado: diseño completamente aleatorio, diseño aleatorio de bloques o diseño de pares relacionados.

33. Lunesta Lunesta es un medicamento diseñado para tratar el insomnio. En un ensayo clínico de Lunesta, las cantidades de sueño cada noche se miden antes y después de que los sujetos han sido tratados con el fármaco.

34. Lipitor Se está planeando un ensayo clínico de los tratamientos con Lipitor para determinar si sus efectos sobre la presión arterial diastólica son diferentes para hombres y mujeres.

35. Vacuna contra el Nilo Occidental Actualmente, no existe una vacuna aprobada para la prevención de la infección por el virus del Nilo Occidental. Se está planeando un ensayo clínico de una posible vacuna que incluya sujetos tratados con la vacuna, así como a sujetos a los que se les administrará un placebo.

36. Vacuna contra el VIH La Red de Ensayos de VIH está llevando a cabo un estudio para probar la eficacia de dos diferentes vacunas experimentales contra el VIH. Los sujetos constarán de 80 pares de gemelos. Para cada par de gemelos, uno de los sujetos se tratará con la vacuna de ADN y el otro gemelo se tratará con la vacuna de vector adenoviral.

37. Muestra aleatoria simple contra muestra aleatoria Consulte la definición de muestra aleatoria simple en la página 27 y su definición adjunta de muestra aleatoria incluida entre paréntesis. Determine si cada uno de los siguientes casos es una muestra aleatoria simple y una muestra aleatoria.

a. Una clase de estadística con 36 estudiantes se dispone de manera que hay 6 filas con 6 estudiantes en cada fila y las filas están numeradas de 1 a 6. Se lanza un dado y una muestra consiste en todos los estudiantes en la fila correspondiente al resultado del dado.

b. Para la misma clase descrita en el inciso (a), los 36 nombres de los estudiantes se escriben en 36 tarjetas individuales. Las tarjetas se barajan y se extraen los seis nombres de la parte superior.

d. Para la misma clase descrita en el inciso (a), se seleccionan los seis estudiantes más jóvenes.

Examen rápido del capítulo

1. Hospitales En un estudio de los nacimientos en el Estado de Nueva York, se recolectaron datos de cuatro hospitales codificados de la siguiente manera: (1) Centro Médico de Albany, (1438) Centro Hospitalario de Bellevue, (66) Hospital General Olean, (413) Hospital Strong Memorial. ¿Tiene sentido calcular el promedio (la media) de los números 1, 1438, 66 y 413?

2. Hospitales ¿Cuál de las siguientes opciones describe mejor el nivel de medición de los números 1, 1438, 66 y 413 del ejercicio 1: nominal, ordinal, de intervalo, de razón?

3. Pesos al nacer En el mismo estudio citado en el ejercicio 1, el peso al nacer de los recién nacidos se da en gramos. ¿Son estos pesos datos discretos o datos continuos?

4. Pesos al nacer ¿Los pesos al nacer descritos en el ejercicio 3 son datos cuantitativos o datos categóricos?

5. Pesos al nacer ¿Cuál de las siguientes opciones describe mejor el nivel de medición de los pesos al nacer descritos en el ejercicio 3: nominal, ordinal, de intervalo, de razón?

6. Estadístico/parámetro En una encuesta AARP aplicada a 1019 adultos seleccionados al azar, se pidió a los encuestados que identificaran el número de tarjetas de crédito que poseían, y 26% dijeron que no tenían tarjetas de crédito. ¿Es el valor de 26% un dato estadístico o un parámetro?

7. Encuesta AARP Con referencia a la encuesta descrita en el ejercicio 6, debido a que los 1019 sujetos estuvieron de acuerdo en responder, ¿constituyen una muestra de respuesta voluntaria?

8. Estudio observacional o experimento ¿Los datos descritos en el ejercicio 6 son el resultado de un estudio observacional o de un experimento?

9. Estudio de la salud de los médicos En el Estudio de la Salud de los Médicos, algunos de los sujetos fueron tratados con aspirina mientras que otros recibieron un placebo. Para los sujetos de este experimento, ¿qué es el estudio a ciegas?

10. Muestreo En un estudio estadístico, ¿cuál de los siguientes tipos de muestras suele ser mejor: muestra de conveniencia, muestra de respuesta voluntaria, muestra aleatoria simple, muestra sesgada?

Ejercicios de repaso

1. ¿Qué es erróneo? En una encuesta de la Asociación Americana de Optometría se seleccionaron 1009 adultos al azar y se les pidió identificar lo que más les preocupa perder. 51% de los encuestados eligió la "vista". ¿Qué es erróneo aquí?

2. Pago en la primera cita *USA Today* publicó esta pregunta en la versión electrónica de su periódico: "¿Deberían los hombres pagar la cuenta en la primera cita?" De los 1148 sujetos que decidieron responder, 857 dijeron que "sí".

a. ¿Qué hay de erróneo en esta encuesta?

b. ¿Es el valor de 85% un estadístico o un parámetro?

c. ¿La encuesta es un experimento o un estudio observacional?

3. Conocimiento del diseño muestral En el estudio "Efectos cardiovasculares de la triiodotironina intravenosa en pacientes sometidos a cirugía de injerto de bypass de la arteria coronaria" ["*Cardiovascular Effects of Intravenous Triiodothyronine in Pactients Undergoing Coronary Artery Bypass Graft Surgery", Journal of the American Medical Association* (JAMA), vol. 275, núm. 9], los autores explican que los pacientes fueron asignados a uno de tres grupos: (1) un grupo tratado con triiodotironina, (2) un grupo tratado con bolo salino normal y dopamina, y (3) un grupo placebo que recibió solución salina normal. Los autores resumen el diseño de la muestra como "aleatorizado y doblemente a ciegas". Describa el significado de "aleatorizado" y "doblemente a ciegas" en el contexto de este estudio.

4. Divorcios y margarina Un estudio mostró que existe una correlación muy alta entre la tasa de divorcios en Maine y el consumo *per cápita* de margarina en Estados Unidos. ¿Es posible concluir que cualquiera de esas dos variables es causante de la otra?

5. Muestra aleatoria simple ¿Cuál de las siguientes opciones son muestras aleatorias simples?

a. Mientras las píldoras de Lipitor se fabrican, un plan de control de la calidad consiste en seleccionar cada píldora número 500 y probarla para confirmar que contiene 80 mg de atorvastatina.

b. Para probar una diferencia de género en la forma en que los hombres y las mujeres realizan compras en línea, Gallup encuesta a 500 hombres y 500 mujeres seleccionadas aleatoriamente.

c. Se obtiene una lista de los 10,877 adultos en el condado de Trinity, California: la lista está numerada de 1 a 10,877 y luego se usa una computadora para generar aleatoriamente 250 números diferentes entre 1 y 10,877. La muestra consta de los adultos correspondientes a los números seleccionados.

6. Ley de Defensa del Matrimonio Las dos preguntas siguientes son esencialmente iguales. ¿Es posible que la diferencia en la redacción pueda afectar la manera en que las personas responden?

- ¿Está usted a favor de la "Ley de defensa del matrimonio"?
- ¿Está usted a favor de una ley para los niveles federal y estatal, en la que sólo los matrimonios heterosexuales deben ser reconocidos?

7. Universidades en Estados Unidos En la actualidad, hay 1612 universidades en Estados Unidos, y el número de estudiantes a tiempo completo es de 13,203,477.

a. ¿El número de estudiantes de tiempo completo en las diferentes universidades es discreto o continuo?

b. ¿Cuál es el nivel de medición para el número de estudiantes de tiempo completo en las universidades? (nominal, ordinal, de intervalo, de razón).

c. ¿Qué hay de erróneo en encuestar a los estudiantes universitarios enviando cuestionarios a 10,000 de ellos seleccionados aleatoriamente?

d. Si se seleccionan aleatoriamente 50 estudiantes universitarios de tiempo completo en cada uno de los 50 estados, ¿qué tipo de muestra se obtiene? (aleatoria, sistemática, de conveniencia, estratificada, por conglomerados).

e. Si se seleccionan aleatoriamente cuatro universidades y se examinan todos sus estudiantes de tiempo completo, ¿qué tipo de muestra se obtiene? (aleatoria, sistemática, de conveniencia, estratificada, por conglomerados).

8. Porcentajes

a. Las etiquetas de las barras energéticas de proteína U-Turn incluyen la afirmación de que estas barras contienen "125% menos grasa que las principales marcas de dulces de chocolate" (con base en la revista *Consumer Reports*) ¿Qué es erróneo en esa afirmación?

b. En una encuesta de Pew Research Center sobre conducción, 58% de los 1182 encuestados dijeron que les gusta conducir. ¿Cuál es el número real de encuestados que dijeron que les gusta conducir?

c. En una encuesta de Pew Research Center sobre conducción, 331 de los 1182 encuestados dijeron que conducir es algo que hacen por necesidad. ¿Qué porcentaje de encuestados dijo que conducir es algo que hacen por necesidad?

9. Tipos de datos En cada uno de los siguientes casos, identifique el nivel de medición de los datos muestrales (nominal, ordinal, de intervalo, de razón) y el tipo de muestreo utilizado para obtener los datos (aleatorio, sistemático, de conveniencia, estratificado, por conglomerados).

a. En el Centro Médico de Albany, se selecciona cada décimo recién nacido y se mide su temperatura corporal (grados Fahrenheit).

b. En cada uno de los 50 estados, se seleccionan 50 electores al azar y se identifican sus afiliaciones a los partidos políticos.

e. Un encuestador detiene a cada persona que pasa por la puerta de su oficina y le pide que califique la última película que vio (en una escala de 1 estrella a 4 estrellas).

10. Significancia estadística y significancia práctica El Technogene Research Group ha desarrollado un procedimiento diseñado para aumentar la probabilidad de que un bebé nazca siendo niña. En un ensayo clínico de su procedimiento, nacieron 236 niñas en 450 parejas diferentes. Si el método no tiene ningún efecto, hay aproximadamente un 15% de probabilidad de ocurrencia de tales resultados extremos. ¿Parece que el procedimiento tiene significancia estadística? ¿Parece que el procedimiento tiene significancia práctica?

Ejercicios de repaso acumulativo

Del capítulo 2 al capítulo 14, los ejercicios de repaso acumulativo incluyen temas de los capítulos anteriores. Para este capítulo, se presentan algunos ejercicios de calentamiento con la calculadora, que incluyen expresiones similares a las encontradas a lo largo del libro. Utilice su calculadora para encontrar los valores indicados.

1. Pesos al nacer A continuación se listan los pesos (en gramos) de recién nacidos en el Hospital del Centro Médico Albany. ¿Qué valor se obtiene al sumar esos pesos y al dividir el total por el número de pesos? (Este resultado, llamado la *media*, se analiza en el capítulo 3). ¿Qué puede destacarse de estos valores y qué nos dice sobre cómo se midieron los pesos?

3600	1700	4000	3900	3100	3800	2200	3000

2. Seis hijos Jule Cole es una de las fundadoras de Mabel's Labels y es madre de seis hijos. La probabilidad de que seis bebés seleccionados al azar sean todas niñas se encuentra al calcular 0.5^6. Encuentre ese valor.

3. Persona más alta Robert Wadlow (1918-1940) se considera la persona más alta que ha vivido jamás. La expresión siguiente convierte su altura de 272 cm a una puntuación estandarizada. Encuentre este valor y redondee el resultado a dos decimales. Estas puntuaciones estandarizadas se consideran significativamente altas si son mayores de 2 o 3. ¿Es el resultado insignificantemente alto?

$$\frac{272 - 176}{6}$$

4. Temperatura corporal La expresión dada se usa para determinar la probabilidad de que la temperatura promedio (media) del cuerpo humano sea diferente del valor de 98.6°F usado comúnmente. Encuentre el valor dado y redondee el resultado a dos decimales.

$$\frac{98.2 - 98.6}{\frac{0.62}{\sqrt{106}}}$$

5. Determinación del tamaño de muestra La expresión dada se utiliza para determinar el tamaño de muestra necesario para estimar la proporción de estudiantes universitarios que tienen el conocimiento necesario para tomar un curso de estadística. Encuentre el valor y redondee el resultado al número entero más cercano.

$$\frac{1.96^2 \cdot 0.25}{0.03^2}$$

6. Desviación estándar Una manera de obtener una aproximación gruesa del valor de una desviación estándar de datos muestrales es encontrar el rango y luego dividirlo por 4. El rango es la diferencia entre el valor muestral más alto y el más bajo. Si se usa este enfoque, ¿qué valor se obtiene de los datos muestrales listados en el ejercicio 1 "Pesos al nacer"?

7. Desviación estándar La desviación estándar es un concepto extremadamente importante que se introduce en el capítulo 3. Con base en los datos muestrales del ejercicio 1, "Pesos al nacer", la siguiente expresión muestra una parte del cálculo de la desviación estándar. Evalúe esta expresión. (Por fortuna, las calculadoras y el software están diseñados para ejecutar automáticamente tales expresiones, por lo que nuestro trabajo futuro con desviaciones estándar no estará cargado con cálculos complicados).

$$\frac{(3600 - 3162.5)^2}{7}$$

8. Desviación estándar La expresión dada se usa para calcular la desviación estándar de tres temperaturas corporales seleccionadas aleatoriamente. Realice el cálculo y redondee el resultado a dos decimales.

$$\sqrt{\frac{(98.4 - 98.6)^2 + (98.6 - 98.6)^2 + (98.8 - 98.6)^2}{3 - 1}}$$

Notación científica. ***En los ejercicios 9 a 12, las expresiones dadas están diseñadas para producir resultados que se expresan en forma de notación científica. Por ejemplo, el resultado mostrado por la calculadora de 1.23E5 puede expresarse como 123,000, y el resultado de 1.23E-4 puede expresarse como 0.000123. Realice la operación indicada y exprese el resultado como un número ordinario, es decir sin la notación científica.***

9. 0.4^8 **10.** 9^{11} **11.** 6^{14} **12.** 03^{12}

Proyecto de tecnología

1. Datos faltantes El enfoque de este proyecto es descargar un conjunto de datos y manipularlo para evitar los datos faltantes.

a. Primero, descargue el Conjunto de datos 3, "Temperaturas corporales", en el apéndice B que se encuentra en www.pearsonenespañol.com/triola. Elija el formato de descarga que coincida con su tecnología.

b. Algunos procedimientos estadísticos, como los relacionados con la correlación y la regresión (que se estudian en capítulos posteriores), requieren datos que consisten en pares de valores relacionados, y tales procedimientos ignoran los pares en los que falte al menos uno de los valores de datos de un par relacionado. Suponga que queremos llevar a cabo análisis de correlación y regresión en las dos últimas columnas del conjunto de datos 3: temperaturas corporales medidas a las 8 AM el día 2 y de nuevo a las 12 AM el día 2. Para estas dos últimas columnas, identifique las filas con al menos un valor faltante. Tenga en cuenta que en algunas tecnologías, como las calculadoras TI-83/84 Plus, los datos faltantes deben estar representados por una constante como −9 o 999.

c. Aquí hay dos estrategias diferentes para reconfigurar el conjunto de datos y trabajar con los datos faltantes en las dos últimas columnas (suponiendo que se requieran parejas relacionadas de datos sin valores faltantes):

i. Eliminación manual Resaltar las filas con al menos un valor faltante en las dos últimas columnas y luego eliminar esas filas. Esto puede ser tedioso si hay muchas filas con datos faltantes y esas filas están intercaladas en vez de ser filas adyacentes.

ii. Ordenar La mayoría de las tecnologías tienen una función de ordenar que permite reorganizar todas las filas utilizando una columna en particular como base para la clasificación (las calculadoras Tl-83/84 Plus no tienen este tipo de entidad de ordenación). El resultado es que todas las filas permanecen iguales pero están en un orden diferente. Primero, utilice la función Ordenar de la tecnología para reorganizar todas las filas utilizando la columna "8 AM día 2" como base para la ordenación (de modo que todos los valores faltantes en la columna "8 AM día 2" estén al principio); después, resalte y elimine todas las filas con valores faltantes en la columna "8 AM día 2". A continuación, utilice la función Ordenar de la tecnología para reorganizar todas las filas utilizando la columna "12 AM día 2" como base para la ordenación (de modo que todos los valores faltantes en la columna "12 AM día 2" estén al principio); luego resalte y elimine todas las filas con valores faltantes en la columna "12 AM día 2". Las filas restantes incluirán parejas de temperatura corporal, y tales filas serán adecuadas para análisis como la correlación y la regresión. Imprima el conjunto de datos reconfigurados que resulta.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿Los directores de orquesta masculinos realmente viven más tiempo?

Varios informes de los medios de comunicación hicieron la interesante observación de que los directores de orquesta masculinos viven más tiempo que otros hombres. John Amaral escribió en *Awaken* que los directores de orquesta “viven más tiempo que casi cualquier otro grupo de personas por tres a siete años”. Robert Levine escribió en Polyphonic.org que viven más “porque están de pie mientras trabajan”. Algunos proporcionaron otras explicaciones para este fenómeno, refiriéndose a menudo a la actividad cardiovascular. Pero, ¿los directores de orquesta masculinos realmente viven más tiempo que otros grupos de hombres? Las respuestas posibles pueden investigarse en Internet. Consideremos también lo siguiente.

Análisis

1. Considere la afirmación de que “los directores de orquesta masculinos viven más tiempo”. Identifique el grupo específico que supuestamente vive menos que el de los directores. ¿Este otro grupo está formado por hombres elegidos aleatoriamente de la población general?

2. Es razonable suponer que los hombres no se convierten en directores de orquesta hasta que han alcanzado por lo menos la edad de 40 años. Cuando se comparan los períodos de vida de los directores masculinos, ¿debemos compararlos con otros varones de la población general, o debemos compararlos con otros varones que vivieron hasta por lo menos 40 años de edad? Explique.

3. Sin ninguna discapacidad, los hombres califican para el seguro médico Medicare si tienen 65 años o más y cumplen con algunos otros requisitos. Si comparamos los intervalos de vida de los varones en Medicare con los intervalos de vida de varones seleccionados al azar de la población general, ¿por qué encontraríamos que los varones de Medicare tienen períodos de vida más largos?

4. Explique en detalle cómo diseñar un estudio para recolectar datos y determinar si es engañoso afirmar que los directores de orquesta masculinos viven más tiempo. ¿Debería el estudio ser un experimento o un estudio observacional?

Actividades en equipo

1. Actividad en clase En grupos de tres o cuatro alumnos, diseñe un experimento para determinar si las tasas de pulso de los estudiantes universitarios son las mismas mientras los estudiantes están de pie y sentados. Realice el experimento y recopile los datos. Guarde los datos para que puedan analizarse con los métodos presentados en los siguientes capítulos.

2. Actividad en clase En grupos de tres o cuatro alumnos, elabore una breve encuesta que incluya sólo unas pocas preguntas que se puedan formular rápidamente. Considere algunas preguntas objetivas junto con algunas que estén sesgadas, como la primera pregunta que se muestra a continuación.

- ¿Debe su universidad obligar a todos los estudiantes a pagar una cuota de $100 por actividades?
- ¿Debe su universidad financiar actividades cobrando una cuota de $100?

Realice la encuesta y trate de detectar el efecto que la redacción sesgada tiene en las respuestas.

3. Actividad en clase Identifique los problemas con un correo de la revista *Consumer Reports* que incluía un cuestionario anual sobre automóviles y otros productos de consumo. También se incluía una solicitud de una contribución voluntaria de dinero y una votación para la junta directiva. Las respuestas debían enviarse a vuelta de correo en sobres que requerían sellos postales.

4. Actividad fuera de clase Encuentre un reporte de una encuesta que haya usado una muestra de respuesta voluntaria. Describa por qué es muy posible que los resultados no reflejen con exactitud la población.

5. Actividad fuera de clase Encuentre una revista científica con un artículo que incluya un análisis estadístico de un experimento. Describa y comente el diseño del experimento. Identifique un tema particular abordado por el estudio, y determine si los resultados fueron estadísticamente significativos. Determine si esos mismos resultados tienen significancia práctica.

2

EXPLORACIÓN DE DATOS CON TABLAS Y GRÁFICAS

PROBLEMA DEL CAPÍTULO

Restaurantes de comida rápida: ¿Cuál es el más rápido?

Una atractiva característica de los restaurantes de comida rápida es ¡que son rápidos! Para seguir siendo competitivos, los restaurantes de comida rápida deben no sólo proporcionar una buena experiencia culinaria, sino también deben hacerlo tan rápido como sus competidores. El conjunto de datos 25 "Comida rápida" del apéndice B muestra los tiempos de servicio en el auto (medidos en segundos), que se obtuvieron de muestras de clientes en diferentes restaurantes. En la tabla 2-1 se listan las 50 mediciones del tiempo de servicio de la primera columna del conjunto de datos 25. Quien pueda obtener conclusiones significativas simplemente con mirar esos datos es una persona excepcionalmente rara. En este capítulo presentamos métodos que se enfocan en organizar y resumir los datos y utilizar gráficas que permitan entender las características importantes de los datos, especialmente su distribución. Estos métodos nos ayudarán a comparar los restaurantes.

TABLA 2-1 Tiempos de servicio en el auto (segundos) para los almuerzos en McDonald's

107	139	197	209	281	254	163	150	127	308	206	187	169	83	127	133	140
143	130	144	91	113	153	255	252	200	117	167	148	184	123	153	155	154
100	117	101	138	186	196	146	90	144	119	135	151	197	171	190	169	

OBJETIVOS DEL CAPÍTULO

Este capítulo y el siguiente se centran en las características más importantes de los datos, entre las que se incluyen:

Características de los datos

1. **Centro:** Un valor representativo que indica dónde se encuentra el centro del conjunto de datos.
2. **Variación:** Una medida de qué tanto varían los valores de los datos.
3. **Distribución:** La naturaleza o forma de la dispersión de los datos en el rango de valores (por ejemplo, en forma de campana).
4. **Datos atípicos:** Valores muestrales que están muy alejados de la gran mayoría de los demás valores de la muestra.
5. **Tiempo:** Cualquier cambio en las características de los datos a través del tiempo.

Este capítulo proporciona herramientas que permiten conocer los datos al organizarlos, resumirlos y representarlos de manera que sea posible observar sus características más importantes. Los objetivos del capítulo son:

2-1 Distribuciones de frecuencias para organizar y resumir datos

- Desarrollar la capacidad de resumir datos en el formato de una distribución de frecuencias y una distribución de frecuencias relativas.
- Para una distribución de frecuencias, identificar los valores de la anchura de clase, la marca de clase, los límites de clase y las fronteras de clase.

2-2 Histogramas

- Desarrollar la capacidad de representar la distribución de datos en el formato de un histograma o un histograma de frecuencias relativas.
- Examinar un histograma e identificar las distribuciones comunes, incluyendo una distribución uniforme y una distribución normal.

2-3 Gráficas que informan y gráficas que engañan

- Desarrollar la capacidad de graficar datos utilizando un diagrama de puntos, una gráfica de tallo y hojas, una gráfica de series de tiempo, una gráfica de Pareto, un gráfico circular y un gráfico de polígono de frecuencias.
- Determinar cuándo una gráfica es engañosa debido al uso de un eje sin cero o un pictograma que utiliza un objeto de área o volumen para datos unidimensionales.

2-4 Diagramas de dispersión, correlación y regresión

- Desarrollar la capacidad de trazar un diagrama de dispersión de datos pareados.
- Analizar un diagrama de dispersión para determinar si parece haber una correlación entre dos variables.

2-1 Distribuciones de frecuencias para organizar y resumir datos

TABLA 2-2 Tiempos de servicio en el auto para los almuerzos en McDonald's

Tiempo (segundos)	Frecuencia
75-124	11
125-174	24
175-224	10
225-274	3
275-324	2

Concepto clave Cuando se trabaja con grandes conjuntos de datos, una *distribución de frecuencias* (o *tabla de frecuencias*) suele ser útil para la organización y el resumen de los datos. Una distribución de frecuencias nos ayuda a comprender la naturaleza de la *distribución* de un conjunto de datos.

DEFINICIÓN

Una **distribución de frecuencias** (o **tabla de frecuencias**) indica cómo un conjunto de datos se divide en varias categorías (o clases) al listar todas las categorías junto con el número de valores de los datos (frecuencias) que hay en cada una.

Utilicemos los tiempos de servicio de McDonald's para el almuerzo (en segundos) listados en la tabla 2-1. Por otra parte, la tabla 2-2 es una distribución de frecuencias que resume los tiempos de servicio. La **frecuencia** para una clase particular es el número de valores originales que caen en esa clase. Por ejemplo, la primera clase de la tabla 2-2 tiene una frecuencia de 11, por lo que 11 de los tiempos de servicio están entre 75 y 124 segundos, inclusive.

Los siguientes términos estándar se usan con frecuencia para la elaboración de distribuciones y gráficas de frecuencias.

DEFINICIONES

Los **límites inferiores de clase** son las cifras más pequeñas que pueden pertenecer a cada una de las clases. (La tabla 2-2 tiene límites inferiores de clase de 75, 125, 175, 225 y 275).

Los **límites superiores de clase** son las cifras más grandes que pueden pertenecer a cada una de las clases. (La tabla 2-2 tiene límites superiores de clase de 124, 174, 224, 274 y 324).

Las **fronteras de clase** son las cifras que se utilizan para separar las clases, pero sin los espacios creados por los límites de clase. La figura 2-1 muestra los espacios creados por los límites de clase de la tabla 2-2. Se observa que los valores 124.5, 174.5, 224.5 y 274.5 están en los puntos medios de esos espacios. Si se sigue el patrón de las fronteras de clase, puede verse que la frontera de clase inferior es 74.5 y la frontera de clase superior es 324.5. La lista completa de las fronteras de clase es 74.5, 124.5, 174.5, 224.5, 274.5 y 324.5.

Las **marcas de clase** son los valores en el punto medio de las clases. La tabla 2-2 tiene las marcas de clase 99.5, 149.5, 199.5, 249.5 y 299.5. Cada marca de clase se puede encontrar al sumar el límite inferior de clase más el límite superior de clase y dividir la suma por 2.

Anchura de clase es la diferencia entre dos límites inferiores de clase consecutivos (o dos fronteras inferiores de clase consecutivas) en una distribución de frecuencias. La tabla 2-2 usa una anchura de clase de 50. (Los dos primeros límites inferiores de clase son 75 y 125, y su diferencia es 50).

PRECAUCIÓN Encontrar la anchura de clase correcta puede ser complicado. Para la anchura de clase, no cometa el error más común de usar la diferencia entre un límite inferior de clase y un límite superior de clase. Observe en la tabla 2-2 que la anchura de clase es 50, no 49.

PRECAUCIÓN Con respecto a las fronteras de clase, recuerde que dividen la diferencia entre el final de una clase y el comienzo de la siguiente, como se muestra en la figura 2-1.

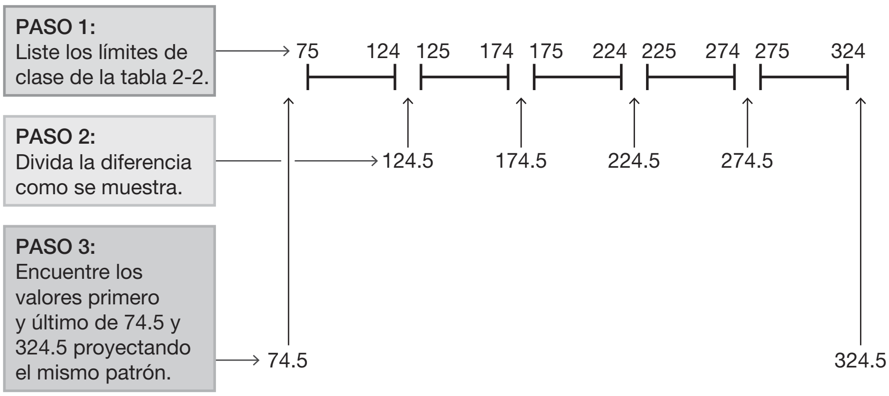


FIGURA 2-1 Determinación de las fronteras de clase a partir de los límites de clase de la tabla 2-2

Procedimiento para elaborar una distribución de frecuencias

Las distribuciones de frecuencias se elaboran para (1) resumir grandes conjuntos de datos, (2) observar la distribución e identificar los valores atípicos, y (3) tener una base para producir gráficas (como los *histogramas* que se introducen en la sección 2-2). Las distribuciones de frecuencias pueden generarse mediante software, pero a continuación se indican los pasos para elaborarlas manualmente:

1. Seleccione el número de clases, normalmente entre 5 y 20. El número de clases puede verse afectado por la conveniencia de utilizar números redondeados. (De acuerdo con la "Regla de Sturges", el número ideal de clases para una distribución de frecuencias puede aproximarse por $1 + (\log n)/(\log 2)$ donde n es el número de valores de los datos. En este libro no seguimos esa regla).
2. Calcule la anchura de clase.

$$\text{Anchura de clase} \approx \frac{(\text{valor máximo de datos}) - (\text{valor mínimo de datos})}{\text{número de clases}}$$

Redondee este resultado para obtener un número conveniente. (Por lo general, es mejor redondear *hacia arriba*). El uso de un número específico de clases no es demasiado importante, y generalmente se recomienda cambiar el número de clases de manera que se tengan valores convenientes para los límites de clase.

3. Elija el valor para el primer límite inferior de clase utilizando el valor mínimo o un valor conveniente por debajo del mínimo.
4. A partir del primer límite inferior de clase y de la anchura de clase, liste los demás límites inferiores de clase. (Esto se hace sumando la anchura de clase al primer límite inferior de clase para obtener el segundo límite inferior de clase. Después se suma la anchura de clase al segundo límite inferior de clase para obtener el tercer límite inferior de clase, y así sucesivamente).

Ni teléfonos ni bañeras

Muchos análisis estadísticos deben considerar las características cambiantes de las poblaciones a través del tiempo. A continuación se presentan algunas observaciones de la vida en Estados Unidos hace 100 años:

- 8% de los hogares tenía un teléfono.
- 14% de los hogares tenía una bañera.
- La esperanza de vida media era de 47 años.
- El salario medio por hora era de 22 centavos.
- Había aproximadamente 230 asesinatos en todo Estados Unidos.

Aunque estas observaciones de hace 100 años indican un marcado contraste con Estados Unidos de hoy, los análisis estadísticos deben considerar siempre características cambiantes de la población que podrían tener efectos más sutiles.

5. Liste los límites inferiores de clase en una columna vertical y después determine e introduzca los límites superiores de clase.
6. Tome cada valor de datos individual y coloque una marca de registro en la clase apropiada. Agregue las marcas de registro para encontrar la frecuencia total de cada clase.

Al elaborar una distribución de frecuencias, asegúrese de que las clases no se superpongan. Cada uno de los valores originales debe pertenecer exactamente a una clase. Incluya todas las clases, incluso aquellas con una frecuencia de cero. Trate de usar la misma anchura para todas las clases, aunque en ocasiones es imposible evitar intervalos abiertos, como "65 años o más".

Autores identificados

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

EJEMPLO 1 Tiempos de servicio para los almuerzos en McDonald's

A partir de los tiempos de servicio para los almuerzos en McDonald's de la tabla 2-1, siga el procedimiento anterior para elaborar la distribución de frecuencias mostrada en la tabla 2-2. Utilice cinco clases.

SOLUCIÓN

Paso 1: Seleccione 5 como el número de clases deseadas.

Paso 2: Calcule la anchura de clase como se muestra a continuación. Observe que redondeamos de 45 a 50, que es un número más cómodo de usar.

$$\text{Anchura de clase} \approx \frac{(\text{valor máximo de datos}) - (\text{valor mínimo de datos})}{\text{número de clases}}$$

$$= \frac{308 - 83}{5} = 45 \approx 50 \text{ (redondeado a un número más conveniente)}$$

Paso 3: El valor mínimo de los datos es 83, que no es un punto de inicio muy conveniente, así que busque un valor más conveniente por debajo de 83, por ejemplo 75, como el primer límite inferior de clase. (Podríamos haber usado 80 o 50 en su lugar).

Paso 4: Sume la anchura de clase de 50 al valor inicial de 75 para obtener el segundo límite inferior de clase de 125. Continúe sumando la anchura de clase de 50 hasta tener cinco límites inferiores de clase. Por lo tanto, los límites inferiores de clase son 75, 125, 175, 225 y 275.

75–
125–
175–
225–
275–

Paso 5: Liste los límites inferiores de clase en forma vertical, como se muestra en el margen. A partir de esta lista, identifique los límites superiores de clase correspondientes como 124, 174, 224, 214 y 324.

Paso 6: Introduzca una marca de registro para cada valor de datos en la clase apropiada. A continuación, sume las marcas de registro para encontrar las frecuencias mostradas en la tabla 2-2.

SU TURNO Resuelva el ejercicio 11 "Old Faithful".

Datos categóricos Hasta ahora hemos analizado distribuciones de frecuencias utilizando sólo conjuntos de datos cuantitativos, pero también es posible utilizar las distribuciones de frecuencias para resumir datos categóricos (o cualitativos, o de atributo), como lo ilustra el ejemplo 2.

EJEMPLO 2 Visitas a las salas de emergencia por lesiones en deportes y actividades recreativas

En la tabla 2-3 se presentan los datos de las siete principales fuentes de lesiones que resultaron en una visita a la sala de urgencias de un hospital en un año reciente (con base en datos de los Centros para el Control de Enfermedades). Los nombres de las actividades son datos categóricos al nivel nominal de medición, pero podemos crear la distribución de frecuencias como se muestra. Resulta sorprendente ver que el ciclismo está en la parte superior de esta lista, pero esto no significa que el ciclismo sea la más peligrosa de las actividades; hay muchas más personas que montan bicicleta que las que juegan fútbol americano, manejan un vehículo todo terreno o hacen cualquiera de las otras actividades listadas.

TABLA 2-3 Visitas anuales a la sala de urgencias por lesiones en deportes y recreación

Actividad	Frecuencia
Ciclismo	26,212
Fútbol americano	25,376
Juegos infantiles	16,706
Baloncesto	13,987
Fútbol	10,436
Béisbol	9,634
Vehículo todo terreno	6,337

En cifras

14: El número de formas de las narices humanas, a partir de un estudio de Abrahim Tamir que fue publicado en el *Journal of Craniofacial Surgery*.

Distribución de frecuencias relativas

Una variación de la distribución de frecuencias básica es una **distribución de frecuencias relativas** o **distribución de frecuencias porcentuales**, en la que cada frecuencia de clase se sustituye por una frecuencia relativa (o proporción) o porcentaje. En este texto usamos el término "distribución de frecuencias relativas", tanto si empleamos frecuencias relativas como porcentajes. Las frecuencias relativas y porcentajes se calculan de la siguiente manera.

$$\text{Frecuencias relativas para una clase} = \frac{\text{frecuencia para una clase}}{\text{suma de todas las frecuencias}}$$

$$\text{Porcentaje para una clase} = \frac{\text{frecuencia para una clase}}{\text{suma de todas las frecuencias}} \times 100\%$$

La tabla 2-4 es un ejemplo de una distribución de frecuencias relativas. Es una variación de la tabla 2-2 en la que cada frecuencia de clase se sustituye por el valor porcentual correspondiente. Debido a que hay 50 valores de datos, se divide cada frecuencia de clase por 50, y después se multiplica por 100%. La primera clase de la tabla 2-2 tiene una frecuencia de 11, así que se divide 11 por 50 para obtener 0.22 y luego se multiplica por 100% para obtener 22%. La suma de los porcentajes debe ser 100%, con una discrepancia mínima permitida para los errores de redondeo, por lo que una suma como 99% o 101% es aceptable. La suma de los porcentajes de la tabla 2-4 es del 100%.

La suma de los porcentajes en una distribución de frecuencias relativas debe ser muy cercana a 100% (con un pequeño margen para errores de redondeo).

TABLA 2-4 Distribución de frecuencias relativas de los tiempos de servicio para el almuerzo en McDonald's

Tiempo (segundos)	Frecuencias relativas
75-124	22%
125-174	48%
175-224	20%
225-274	6%
275-324	4%

Distribución de frecuencias acumuladas

Otra variación de una distribución de frecuencias es una **distribución de frecuencias acumuladas** en la que la frecuencia para cada clase es la suma de las frecuencias para la misma y todas las anteriores. La tabla 2-5 es la distribución de frecuencias acumuladas de la tabla 2-2. A partir de las frecuencias originales de 11, 24, 10, 3, 2, se suma 11 + 24 para obtener la segunda frecuencia acumulada de 35, luego se suma 11 + 24 + 10 para obtener la tercera, y así sucesivamente. Observe en la tabla 2-5 que además del uso de frecuencias acumuladas, los límites de clase son reemplazados por expresiones "menor que", las cuales describen los nuevos intervalos de valores.

TABLA 2-5 Distribución de frecuencias acumuladas de los tiempos de servicio para el almuerzo en McDonald's

Tiempo (segundos)	Frecuencia acumulada
Menor que 125	11
Menor que 175	35
Menor que 225	45
Menor que 275	48
Menor que 325	50

Pensamiento crítico: uso de distribuciones de frecuencias para entender los datos

Al comienzo de esta sección observamos que una distribución de frecuencias puede ayudarnos a entender la *distribución* de un conjunto de datos, que es la naturaleza o forma de la dispersión de los datos sobre el rango de valores (por ejemplo, en forma de campana). Con frecuencia, en estadística es importante determinar si los datos tienen una *distribución nor-*

Gráficas de crecimiento actualizadas

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

mal. (Las distribuciones normales se analizan de manera amplia en el capítulo 6.) Los datos que tienen una distribución aproximadamente normal se caracterizan por una distribución de frecuencias con las siguientes características.

Distribución normal

1. Las frecuencias comienzan bajas, luego aumentan a una o dos frecuencias altas y luego disminuyen a una frecuencia baja.
2. La distribución es aproximadamente simétrica: las frecuencias que preceden a la frecuencia máxima deben ser aproximadamente una imagen especular de aquellas que siguen de la frecuencia máxima.

La tabla 2-6 satisface estas dos condiciones. Las frecuencias comienzan en un nivel bajo, aumentan hasta el máximo de 30 y luego disminuyen a una frecuencia baja. Además, las frecuencias de 2 y 8 que preceden al máximo son una imagen especular de las frecuencias 8 y 2 que siguen al máximo. Por lo general, los conjuntos de datos reales no son tan perfectos como en la tabla 2-6, y es necesario usar el juicio para determinar si la distribución está "suficientemente cerca" de satisfacer esas dos condiciones. (Más adelante se presentan otros procedimientos objetivos).

TABLA 2-6 Distribución de frecuencias que muestra una distribución normal

Tiempo	Frecuencia	Distribución normal
75-124	2	⟵ Las frecuencias comienzan bajas, …
125-174	8	
175-224	30	⟵ Aumentan hasta este máximo, …
225-274	8	
275-324	2	⟵ Disminuyen hasta volverse bajas de nuevo.

Análisis de los últimos dígitos En el ejemplo 3 se ilustra este principio:

En ocasiones, las frecuencias de los últimos dígitos revelan cómo se recolectaron o midieron los datos.

EJEMPLO 3 Exploración de datos: ¿cómo se midió el pulso?

Al examinar los pulsos medidos a 2219 adultos incluidos en la Encuesta Nacional de Salud y Exámenes, se identifican los últimos dígitos de los pulsos registrados y la distribución de frecuencias para esos últimos dígitos es la presentada en la tabla 2-7. Aquí hay que observar algo: todos los últimos dígitos son números *pares*. Si los pulsos se contaron durante 1 minuto completo, seguramente habría un gran número de ellos que terminen con un dígito *impar*. ¿Entonces qué pasó?

Una explicación razonable es que a pesar de que los pulsos son el número de latidos en 1 minuto, probablemente fueron contados durante 30 segundos y el número de latidos se duplicó. (Las tasas de pulso originales no son todos múltiplos de 4, por lo que podemos descartar un procedimiento de contar durante 15 segundos y luego multiplicar por 4). El análisis de estos últimos dígitos nos revela el método utilizado para obtener los datos.

En muchas encuestas es posible determinar si a los sujetos encuestados se les pidió *reportar* algunos valores, como sus estaturas o pesos, porque desproporcionadamente muchos valores terminan en 0 o 5. Esto es una pista sólida de que el encuestado está redondeando en lugar de ser medido físicamente. ¡Cosas fascinantes!

SU TURNO Resuelva el ejercicio 17 "Análisis de los últimos dígitos".

TABLA 2-7 Últimos dígitos de los pulsos de la Encuesta Nacional de Salud y Exámenes

Último dígito de los pulsos	Frecuencia
0	455
1	0
2	461
3	0
4	479
5	0
6	425
7	0
8	399
9	0

Brechas En el ejemplo 4 se ilustra este principio:

La presencia de brechas puede sugerir que los datos son de dos o más poblaciones diferentes.

Lo contrario a este principio no es verdadero, porque los datos de diferentes poblaciones no necesariamente producen espacios.

EJEMPLO 4 Exploración de datos: ¿qué nos indica un espacio?

La tabla 2-8 es una distribución de frecuencias de los pesos (en gramos) de monedas seleccionadas al azar. El examen de las frecuencias revela *una gran brecha* entre las monedas de centavo más ligeras y las más pesadas. Esto sugiere que tenemos dos poblaciones distintas: los centavos hechos antes de 1983 son 95% de cobre y 5% de zinc, mientras que los centavos hechos después de 1983 son 2.5% de cobre y 97.5% de zinc, lo que explica la gran diferencia entre los centavos más ligeros y los más pesados representados en la tabla 2-8.

SU TURNO Resuelva el ejercicio 18 "Análisis de los últimos dígitos" y determine si hay una brecha. Si es así, dé una explicación razonable.

TABLA 2-8 Monedas seleccionadas al azar

Peso (en gramos) de los centavos	Frecuencia
2.40-2.49	18
2.50-2.59	19
2.60-2.69	0
2.70-2.79	0
2.80-2.89	0
2.90-2.99	2
3.00-3.09	25
3.10-3.19	8

Comparaciones En el ejemplo 5 se ilustra este principio:

La combinación de dos o más distribuciones de frecuencias relativas en una tabla facilita en gran medida las comparaciones entre los datos.

EJEMPLO 5 Comparación de McDonald's y Dunkin' Donuts

La tabla 2-9 muestra las distribuciones de frecuencias relativas para los tiempos de servicio de los almuerzos servidos en el auto (en segundos) para McDonald's y Dunkin' Donuts. Debido a las grandes diferencias en sus menús, podríamos esperar que los tiempos de servicio fueran muy diferentes. Al comparar las frecuencias relativas en la tabla 2-9, se observa que hay diferencias importantes. Los tiempos de servicio de Dunkin' Donuts parecen ser inferiores a los de McDonald's. Esto no es demasiado sorprendente, dado que probablemente muchas de las órdenes en Dunkin' Donuts consisten en un café y una dona.

TABLA 2-9 Tiempos de servicio para el almuerzo en McDonald's y Dunkin' Donuts

Tiempo (en segundos)	McDonald's	Dunkin' Donuts
25-74		22%
75-124	22%	44%
125-174	48%	28%
175-224	20%	6%
225-274	6%	
275-324	4%	

SU TURNO Resuelva el ejercicio 19 "Ganadores del Oscar".

CENTRO DE TECNOLOGÍA

Distribuciones de frecuencia

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

A menudo, las distribuciones de frecuencias son fáciles de obtener después de generar un histograma, como se describe en la sección 2-2. Con Statdisk, por ejemplo, es posible generar un histograma con un punto de inicio y una anchura de clase deseados y luego marcar "Bar Labels" para ver la frecuencia de cada clase. Si no se utilizan histogramas, la opción "Sort" (para ordenar los datos) permite ver los valores máximo y mínimo que se utilizan para calcular la anchura de clase. Una vez que se establecen los límites de clase, es fácil encontrar la frecuencia para cada clase usando datos ordenados. Cada paquete de software estadístico incluye una función de ordenación.

2-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

Tabla para el ejercicio 1

Tiempos de servicio para las cenas en McDonald's

Tiempo (s)	Frecuencia
60-119	7
120-179	22
180-239	14
240-299	2
300-359	5

Tabla para el ejercicio 4

Estatura (cm)	Frecuencia relativa
130-144	23%
145-159	25%
160-174	22%
175-189	27%
190-204	28%

1. Tiempos de servicio para las cenas en McDonald's Consulte la tabla anexa que resume los tiempos de servicio (en segundos) de las cenas en McDonald's. ¿Cuántas personas se incluyen en el resumen? ¿Es posible identificar los valores exactos de todos los tiempos de servicio originales?

2. Tiempos de servicio para las cenas en McDonald's Consulte la distribución de frecuencias anexa. ¿Qué problema se crearía al utilizar clases de 60 a 120, 120 a 180, ..., 300 a 360?

3. Distribución de frecuencias relativas Use los porcentajes para elaborar la distribución de frecuencias relativas correspondiente a la distribución de frecuencias anexa de los tiempos de servicio para las cenas en McDonald's.

4. ¿Qué hay de erróneo? Se sabe que las estaturas de los varones adultos tienen una distribución normal, como se describe en esta sección. Un investigador afirma haber seleccionado aleatoriamente a varones adultos y haber medido sus estaturas con la distribución de frecuencia relativa resultante que se muestra aquí. Identifique dos fallas importantes con base en estos resultados.

En los ejercicios 5 a 8, identifique la anchura de clase, los puntos medios de clase y los límites de clase para la distribución de frecuencias dada. También identifique el número de individuos incluidos en el resumen. Las distribuciones de frecuencias se basan en datos reales del apéndice B.

5.

Edad (en años) de ganadoras del Oscar como mejor actriz	Frecuencia
20-29	29
30-39	34
40-49	14
50-59	3
60-69	5
70-79	1
80-89	1

6.

Edad (en años) de ganadores del Oscar como mejor actor	Frecuencia
20-29	1
30-39	28
40-49	36
50-59	15
60-69	6
70-79	1

7.

Conteo de plaquetas sanguíneas en hombres	Frecuencia
0-99	1
100-199	51
200-299	90
300-399	10
400-499	0
500-599	0
600-699	1

8.

Conteo de plaquetas sanguíneas en mujeres	Frecuencia
100-199	25
200-299	92
300-399	28
400-499	0
500-599	2

Distribuciones normales. ***En los ejercicios 9 y 10, use una interpretación flexible de los criterios para determinar si una distribución de frecuencias se aproxima a una distribución normal. Dé una explicación breve.***

9. Mejores actrices Considere la distribución de frecuencias del ejercicio 5.

10. Mejores actores Considere la distribución de frecuencias del ejercicio 6.

Elaboración de distribuciones de frecuencias. ***En los ejercicios 11 a 18, utilice los datos indicados para elaborar la distribución de frecuencias. (Los datos de los ejercicios 13 a 16 se pueden descargar de* www.pearsonenespañol.com/triola*).***

11. Old Faithful A continuación se listan los tiempos de duración (en segundos) de las erupciones del géiser Old Faithful en el Parque Nacional de Yellowstone. Utilice estos tiempos para elaborar una distribución de frecuencias. Use una anchura de clase de 25 segundos e inicie en un límite inferior de clase de 125 segundos.

125 203 205 221 225 229 233 233 235 236 236 237 238 238 239 240 240
240 240 241 241 242 242 242 243 243 244 245 245 245 245 246 246 248
248 248 249 249 250 251 252 253 253 255 255 256 257 258 262 264

12. Tornados A continuación se listan las intensidades en la escala F de los tornados recientes en Estados Unidos. Elabore una distribución de frecuencias. ¿Las intensidades parecen tener una distribución normal?

0 4 0 0 1 1 1 0 0 0 1 2 0 1 1 0 1 0 1 1 1 1 0
0 1 0 0 1 0 0 1 1 1 3 0 0 0 2 0 3 0 0 0 0 0

13. Tiempos de servicio para almuerzos en Burger King Consulte el conjunto de datos 25 "Comida rápida" y utilice los tiempos de servicio en auto para los almuerzos en Burger King. Inicie en un límite inferior de clase de 70 segundos y utilice una anchura de clase de 40 segundos.

14. Tiempos de servicio para cenas en Burger King Consulte el conjunto de datos 25 "Comida rápida" y utilice los tiempos de servicio en auto para las cenas en Burger King. Inicie en un límite inferior de clase de 30 segundos y utilice una anchura de clase de 40 segundos.

15. Tiempos de servicio para almuerzos en Wendy's Consulte el conjunto de datos 25 "Comida rápida" y utilice los tiempos de servicio en auto para los almuerzos en Wendy's. Inicie en un límite inferior de clase de 70 segundos y utilice una anchura de clase de 80 segundos. ¿Parece que la distribución es una distribución normal?

16. Tiempos de servicio para cenas en Wendy's Consulte el conjunto de datos 25 "Comida rápida" y use los tiempos de servicio para las cenas en Wendy's. Inicie en un límite inferior de clase de 30 segundos y utilice una anchura de clase de 40 segundos. Use una interpretación flexible de una distribución normal y diga si esta distribución parece ser una distribución normal.

17. Análisis de los últimos dígitos El autor obtuvo las estaturas de sus estudiantes de estadística como parte de un experimento realizado en clase. A continuación se listan los últimos dígitos de esas estaturas. Elabore una distribución de frecuencias con 10 clases. Con base en la distribución, ¿parece que las estaturas fueron reportadas o realmente medidas? ¿Qué puede saber usted de la exactitud de los resultados?

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

18. Análisis de los últimos dígitos Los pesos de los encuestados se registraron como parte de la Entrevista del Sondeo de la Salud en California. A continuación se listan los últimos dígitos de los pesos de 50 encuestados seleccionados al azar. Elabore una distribución de frecuencias con 10 clases. Con base en la distribución, ¿parece que los pesos son reportados o realmente medidos? ¿Qué puede saber usted de la exactitud de los resultados?

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

Frecuencias relativas para comparaciones. ***En los ejercicios 19 y 20, elabore las distribuciones de frecuencias relativas y responda las preguntas.***

19. Ganadores del Oscar Elabore una tabla (similar a la tabla 2-9 de la página 47) que incluya frecuencias relativas basadas en las distribuciones de frecuencia de los ejercicios 5 y 6, y después compare las edades de las actrices y los actores ganadoras del Oscar. ¿Hay diferencias notables?

20. Conteos de plaquetas sanguíneas Elabore una tabla (similar a la tabla 2-9 de la página 47) que incluya frecuencias relativas basadas en las distribuciones de frecuencia de los ejercicios 7 y 8, y después compárelas. ¿Hay diferencias notables?

Distribuciones de frecuencia acumuladas. ***En los ejercicios 21 y 22, elabore la distribución de frecuencias acumuladas correspondiente a la distribución de frecuencias del ejercicio indicado.***

21. Ejercicio 5 (Edad de la ganadora del Oscar a la mejor actriz).

22. Ejercicio 6 (Edad del ganador al Oscar al mejor actor).

Datos categóricos. ***En los ejercicios 23 y 24, utilice los datos categóricos dados para elaborar la distribución de frecuencias relativas.***

23. Ensayo clínico Cuando se administró XELJANZ (tofacitinib) como parte de un ensayo clínico para un tratamiento de artritis reumatoide, 1336 sujetos recibieron dosis de 5 mg del fármaco, y a continuación se muestra el número de reacciones adversas: 57 tuvieron dolores de cabeza, 2l tuvieron hipertensión, 60 infecciones del tracto respiratorio, 51 nasofaringitis y 53 diarrea. ¿Alguna de estas reacciones adversas parece ser mucho más común que las otras? (*Sugerencia:* Encuentre las frecuencias relativas usando solamente las reacciones adversas, no el número total de sujetos tratados).

24. Nacimientos Los nacimientos naturales seleccionados aleatoriamente de cuatro hospitales en el Estado de Nueva York ocurrieron en los días de la semana (en orden de lunes a domingo) con las siguientes frecuencias: 52, 66, 72, 57, 57, 43, 53. ¿Parece que los nacimientos ocurren en los días de la semana con la misma frecuencia?

Conjuntos grandes de datos. ***Los ejercicios 25 a 28 implican grandes conjuntos de datos, por lo que debe usarse la tecnología. Las listas completas de los datos no se incluyen en el apéndice B, pero pueden descargarse del sitio web* www.pearsonenespañol.com/triola*. Utilice los datos indicados y elabore la distribución de frecuencias.***

25. Presión arterial sistólica Utilice la presión arterial sistólica de 300 sujetos incluida en el conjunto de datos 1 "Datos corporales". Use una anchura de clase de 20 mm Hg e inicie en un límite inferior de clase de 80 mm Hg. ¿Parece que la distribución de frecuencias es una distribución normal?

26. Presión arterial diastólica Utilice la presión arterial diastólica de 300 sujetos incluida en el conjunto de datos 1 "Datos corporales". Use una anchura de clase de 15 mm Hg e inicie con un límite inferior de clase de 40 mm Hg. ¿Parece que la distribución de frecuencias es una distribución normal?

27. Magnitudes de terremoto Utilice las magnitudes de 600 terremotos incluidas en el conjunto de datos 21 "Terremotos". Use una anchura de clase de 0.5 e inicie con un límite inferior de clase de 1.00. ¿Parece que la distribución de frecuencias es una distribución normal?

28. Profundidades de terremoto Utilice las profundidades (en km) de 600 terremotos incluidas en el conjunto de datos 21 "Terremotos". Use una anchura de clase de 10.0 km e inicie con un límite inferior de clase de 0.0 km. ¿Parece que la distribución de frecuencias es una distribución normal?

2-1 Más allá de lo básico

29. Interpretación de efectos de los valores atípicos Consulte, en el conjunto de datos 30 "Latas de aluminio" del apéndice B, las cargas axiales de latas de aluminio de 0.0111 pulgadas de grosor. Una carga axial es la fuerza a la cual la parte superior de una lata se colapsa. La carga de 504 lb es un valor atípico porque está muy alejado de todos los demás valores. Elabore una distribución de frecuencias que incluya el valor de 504 lb, y después otra distribución de frecuencias sin el valor de 504 libras. En ambos casos, inicie la primera clase en 200 libras y utilice una anchura de clase de 20 lb. Indique una generalización sobre el efecto de un valor atípico en una distribución de frecuencias.

2-2 Histogramas

PARTE 1 Conceptos básicos de los histogramas

Concepto clave Una distribución de frecuencias es una herramienta útil para resumir datos e investigar su distribución; una herramienta incluso mejor es un *histograma*, una gráfica más fácil de interpretar que una tabla de números.

> **DEFINICIÓN**
>
> Un **histograma** es una gráfica que consiste en barras adyacentes de igual anchura dibujadas (a menos que haya espacios en los datos). La escala horizontal representa clases de valores cuantitativos, y la escala vertical representa sus frecuencias. Las alturas de las barras corresponden a los valores de frecuencia.

Usos importantes de un histograma

- Despliega visualmente la forma de la *distribución* de los datos.
- Muestra la ubicación del *centro* de los datos.
- Muestra la *dispersión* de los datos.
- Identifica los valores *atípicos*.

Un histograma es en esencia una gráfica de una distribución de frecuencias. Por ejemplo, la figura 2-2 muestra el histograma generado por Minitab correspondiente a la distribución de frecuencias dada en la tabla 2-2 de la página 42.

Las frecuencias de clase deben usarse para la escala vertical, la cual se debe etiquetar como en la figura 2-2. No hay un acuerdo universal sobre el procedimiento para seleccionar los valores que se usarán para ubicar las barras a lo largo de la escala horizontal, pero es común utilizar los límites de clase (como se muestra en la figura 2-2) o los puntos medios de clase o los límites de clase o algún otro valor. Con frecuencia, resulta más sencillo usar los puntos medios de clase para la escala horizontal. Por lo general, los histogramas se pueden generar usando software.

Histograma de frecuencias relativas

Un **histograma de frecuencias relativas** tiene la misma forma y escala horizontal que un histograma, pero la escala vertical utiliza frecuencias relativas (como porcentajes o proporciones) en vez de frecuencias reales. La figura 2-3 es el histograma de frecuencias relativas correspondiente a la figura 2-2.

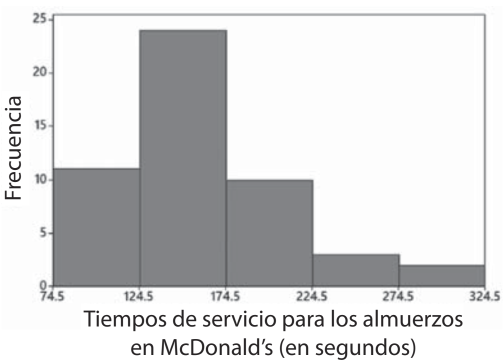


FIGURA 2-2 **Histograma de frecuencias relativas del tiempo de servicio para los almuerzos en auto en McDonald's (en segundos)**

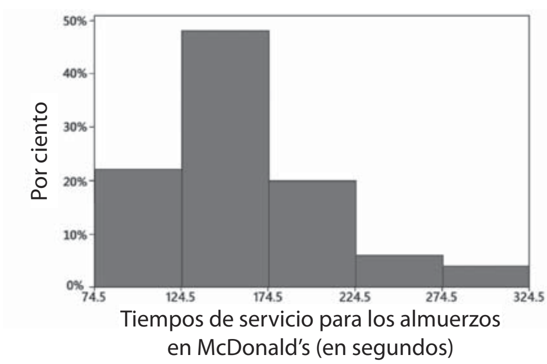


FIGURA 2-3 **Histograma del tiempo de servicio para los almuerzos en auto en McDonald's (en segundos)**

Pensamiento crítico: Interpretación de histogramas

Aunque la creación de histogramas es lo más divertido que los seres humanos pueden hacer, el objetivo final es *entender* las características de los datos. Se pueden explorar datos analizando el histograma para ver qué se puede aprender de ellos con “CVDOT”: el centro de los datos, la variación (que se analizará de manera extensa en la sección 3-2), la forma de la distribución, si hay valores atípicos (*outliers*, en inglés)(muy alejados de los demás valores) y el tiempo (si hay algún cambio en las características de los datos a lo largo del tiempo). Al examinar la figura 2-2, se observa que el histograma está centrado alrededor de 160 o 170 segundos, los valores varían desde alrededor de 75 segundos hasta 325, y la distribución es *muy aproximada* a una forma de campana. No hay valores atípicos y cualquier cambio en el tiempo es irrelevante para estos datos.

Formas comunes de distribución

Los histogramas que se muestran en la figura 2-4 representan cuatro formas comunes de distribución.

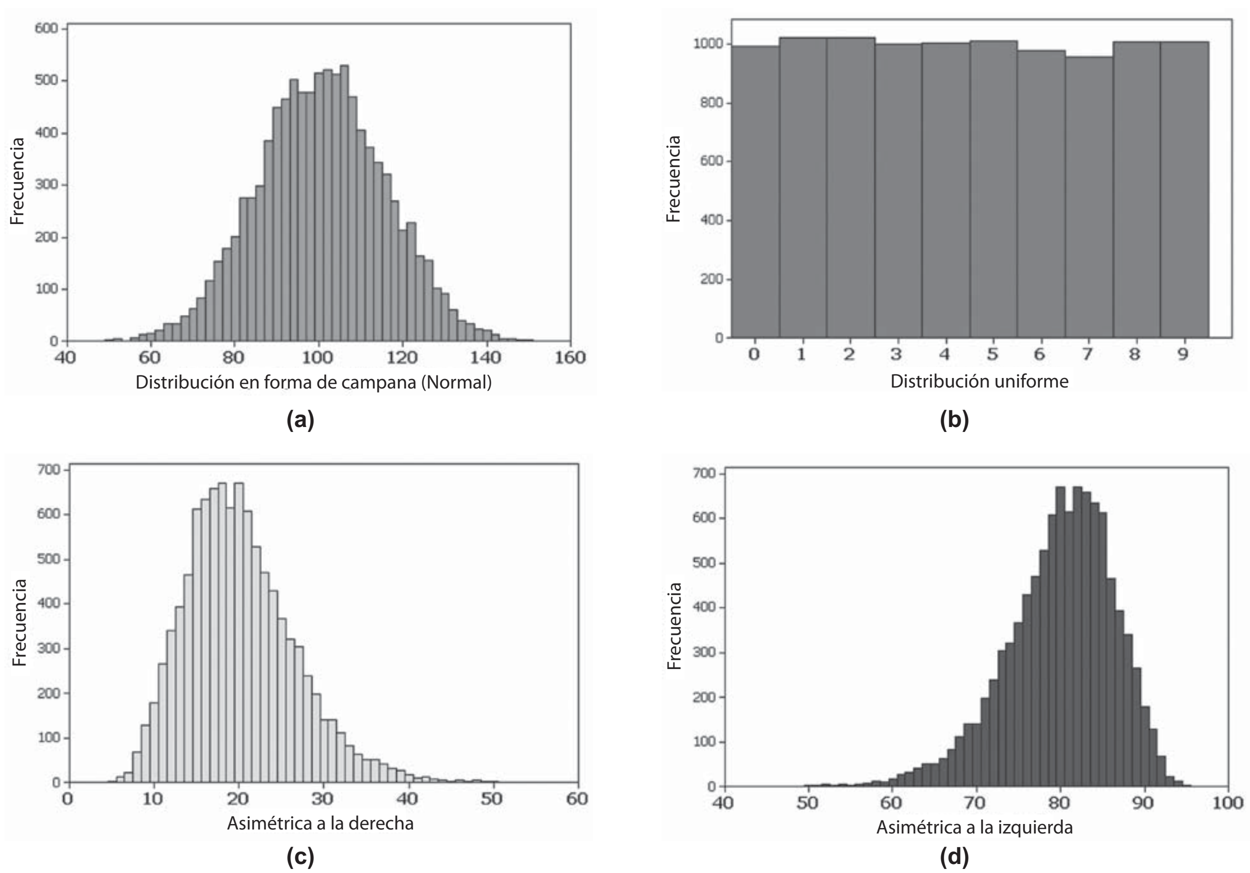


FIGURA 2-4 Distribuciones comunes

Distribución normal

Cuando se grafica como un histograma, una distribución normal tiene una forma de “campana” similar a la superpuesta en la figura 2-5. Muchos métodos estadísticos requieren que los datos muestrales provengan de una población con distribución aproximadamente normal, y a menudo se puede usar un histograma para juzgar si este requisito se cumple (aunque existen métodos más avanzados y menos subjetivos para determinar si la distribución es una distribución normal). Las gráficas cuantilares normales son muy útiles para evaluar la normalidad: vea la parte 2 de esta sección.

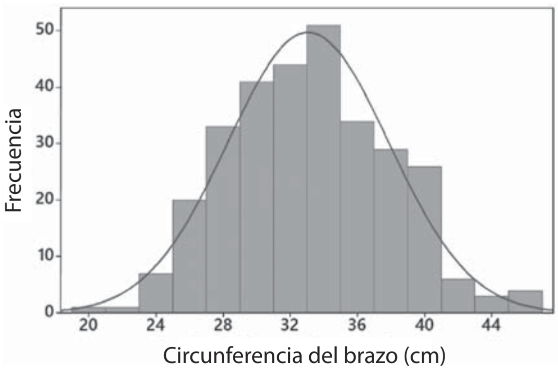


FIGURA 2-5 Distribución en forma de campana de las circunferencias del brazo
Debido a que este histograma tiene una forma aproximada de campana, se dice que los datos tienen una *distribución normal*. (En el capítulo 6 se dará una definición más rigurosa).

En cifras

2.5 trillones de bytes: Cantidad de datos que generamos el año pasado. (Un trillón es un 1 seguido por 18 ceros).

Distribución uniforme

Los diferentes valores posibles se producen con aproximadamente la misma frecuencia, por lo que las alturas de las barras en el histograma son aproximadamente uniformes, como en la figura 2-4(b) que muestra los resultados de los dígitos de las loterías estatales.

Asimetría

Una distribución de datos es asimétrica si se extiende más hacia un lado que hacia el otro. Los datos **asimétricos a la derecha** (también llamados *positivamente asimétricos*) tienen una cola derecha más larga, como en la figura 2-4(c). Los ingresos anuales de los estadounidenses adultos son positivamente asimétricos. Los datos **asimétricos a la izquierda** (también llamados *negativamente asimétricos*) tienen una cola izquierda más larga, como en la figura 2-4(d). Los datos de la duración de la vida en los seres humanos son asimétricos a la izquierda. (Aquí hay un truco de mnemotecnia para recordar la asimetría: una distribución asimétrica a la derecha se parece a los dedos del pie derecho, y una asimétrica a la izquierda se parece a los dedos del pie izquierdo). Las distribuciones asimétricas a la derecha son más comunes que las asimétricas a la izquierda porque a menudo es más fácil obtener valores excepcionalmente grandes que valores muy pequeños. Con los ingresos anuales, por ejemplo, es imposible obtener valores por debajo de cero, pero hay algunas personas que ganan millones o miles de millones de dólares al año. Por lo tanto, los ingresos anuales tienden a ser asimétricos a la derecha.

Para recordar la asimetría:

Asimetría a la izquierda:	**Se asemeja a los dedos del pie izquierdo**
Asimetría a la derecha:	**Se asemeja a los dedos del pie derecho**

PARTE 2 Evaluación de la normalidad con gráficas cuantilares normales

Algunos métodos realmente importantes que se presentan en capítulos subsecuentes tienen el requisito de que los datos muestrales provengan de una población con distribución normal. Los histogramas pueden ser útiles para determinar si se satisface tal requisito de normalidad, pero no son muy útiles con conjuntos de datos pequeños. En la sección 6-5 se analizan métodos para *evaluar la normalidad*, es decir, para determinar si los datos muestrales provienen de una población normalmente distribuida. Dicha sección incluye un procedimiento para trazar *gráficas cuantilares normales*, que son fáciles de generar usando tecnologías como Statdisk, Minitab, XLSTAT, StatCrunch o una calculadora TI-83/84 Plus. La interpretación de una gráfica cuantilar normales se basa en los siguientes criterios:

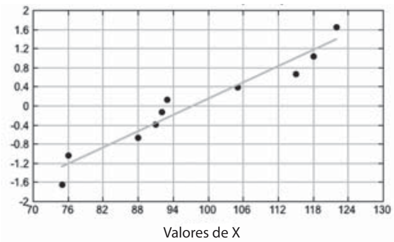
Criterios para evaluar la normalidad con una gráfica cuantilar normal

Distribución normal: La distribución de una población es *normal* si el patrón de los puntos en la gráfica cuantilar normal está razonablemente cerca de una línea recta y los puntos no muestran un patrón sistemático diferente a un patrón lineal.

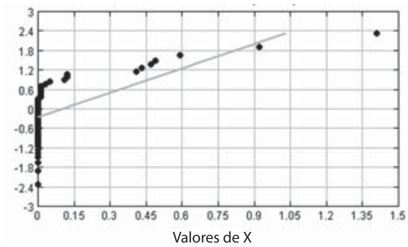
Distribución no normal: La distribución de la población *no* es normal si la gráfica cuantilar normal cumple una o dos de las siguientes condiciones:

- Los puntos no se encuentran razonablemente cerca de un patrón de línea recta.
- Los puntos muestran algún *patrón sistemático* diferente a un patrón de línea recta.

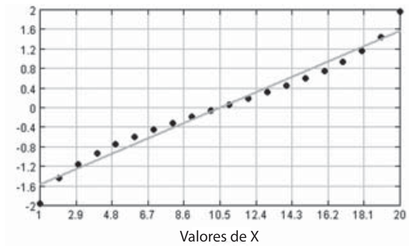
Los siguientes son ejemplos de gráficas cuantilares normales. Los procedimientos para crear tales gráficas se describen en la sección 6-5.



Distribución normal: Los puntos están razonablemente cerca de un patrón lineal, y no hay otro patrón sistemático diferente.



Distribución no normal: Los puntos no se encuentran razonablemente cerca de una línea recta.



Distribución no normal: Los puntos muestran un patrón sistemático que no es un patrón de línea recta.

CENTRO DE TECNOLOGÍA

Histogramas

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Data** del menú superior.
2. Seleccione **Histogram** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Plot**.
5. Marque **Bar Labels** en *Plot Options* para ver la frecuencia de cada clase.
6. Marque **User Defined** en *Plot Options* para utilizar su propia anchura de clase y punto de inicio.

Sugerencia: Este procedimiento también es una manera fácil de identificar frecuencias en una distribución de frecuencias.

Minitab

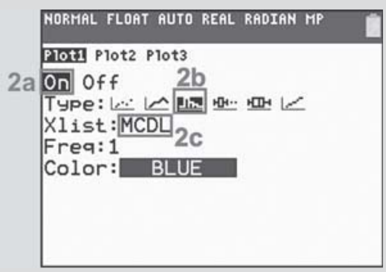
1. Haga clic en **Graph** del menú superior.
2. Seleccione **Histogram** en el menú desplegable.
3. Seleccione el histograma **Simple** y haga clic en **OK**.
4. Haga clic en la columna de datos deseada, luego haga clic en **Select** y después en **OK**.
5. Cambie la anchura de clase predeterminada y el punto de inicio según sea necesario haciendo clic con el botón derecho en el eje horizontal y seleccionando **Edit X Scale**.
   – Seleccione la pestaña **Scale** para introducir la ubicación de las marcas de escala.
   – Seleccione la pestaña **Binning** para introducir los puntos medios de clase.

StatCrunch

1. Haga clic en **Graph** del menú superior.
2. Seleccione **Histogram** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Para personalizar el histograma, introduzca el punto de inicio deseado y la anchura de clase en **Bins**.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Abra el menú **STAT PLOTS** pulsando 2ND, Y=.
2. Presione ENTER para acceder a la pantalla de configuración *Plot 1* como se muestra:
   a. Seleccione **ON** y presione ENTER.
   b. Seleccione la opción de la gráfica de barras y presione **ENTER**.
   c. Introduzca el nombre de la lista que contiene los datos.
3. Presione ZOOM, luego 9 (ZoomStat) para generar el histograma predeterminado.
4. Presione TRACE y use ◁ ▷ para ver las fronteras de clase y las frecuencias de cada clase.
5. Presione WINDOW para personalizar la anchura y las fronteras de clase. Presione GRAPH para ver el histograma.



CENTRO DE TECNOLOGÍA *continuación*

Histogramas

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Excel

Es extremadamente difícil generar histogramas en Excel; se debe utilizar el complemento XLSTAT:

1. Seleccione la pestaña **XLSTAT** en la cinta de opciones.
2. Haga clic en el botón **Visualizing Data**.
3. Seleccione **Histograms** en el menú desplegable.
4. Introduzca el rango de celdas que contiene los datos deseados. Haga clic en **Sample labels** si la primera celda contiene un nombre de datos.
5. Haga clic en **OK** para generar un histograma predeterminado.

Sugerencia: Para personalizar, introduzca las fronteras de clase deseadas en una columna, seleccione la pestaña **Options**, haga clic en **User Defined** e ingrese el rango de celdas que contiene las fronteras en el cuadro.

2-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Estaturas Las estaturas de los varones adultos se distribuyen normalmente. Si se selecciona aleatoriamente una gran muestra de estaturas de varones adultos y las estaturas se ilustran en un histograma, ¿cuál será la forma de ese histograma?

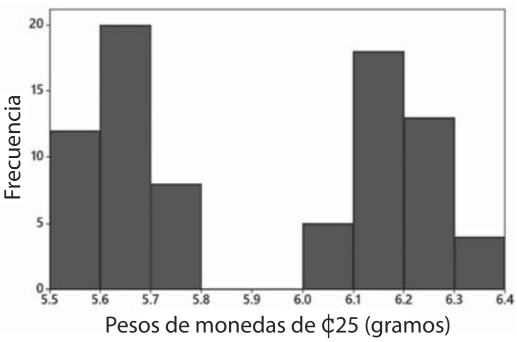
2. Más estaturas La población de las estaturas de los varones adultos se distribuye normalmente. Si obtenemos una muestra de respuesta voluntaria de 5000 de esas estaturas, ¿un histograma de los datos muestrales tendrá forma de campana?

3. Conteo de plaquetas sanguíneas A continuación se listan conteos de plaquetas sanguíneas (1000 células/μl) seleccionados aleatoriamente en adultos de Estados Unidos. ¿Por qué no tiene sentido elaborar un histograma para este conjunto de datos?

191	286	263	193	193	215	162	646	250	386

4. Conteo de plaquetas sanguíneas Si se recolecta una muestra de conteos de plaquetas sanguíneas mucho mayor que la muestra del ejercicio 3, y si dicha muestra incluye un único valor atípico, ¿cómo aparecerá ese valor atípico en un histograma?

Interpretación de un histograma. ***En los ejercicios 5 a 8, responda las preguntas con referencia al siguiente histograma generado en Minitab, que representa los pesos (en gramos) de monedas de veinticinco centavos listados en el conjunto de datos 29 "Pesos de monedas" del apéndice B (los gramos son en realidad unidades de masa y los valores mostrados en la escala horizontal están redondeados).***



5. Tamaño de muestra ¿Cuál es el número aproximado de monedas de veinticinco centavos representado en las tres barras que se encuentran más a la izquierda?

6. Anchura de clase y límites de clase Indique los valores aproximados de la anchura de clase y los límites de clase inferior y superior de la clase representada en la barra más alejada hacia la izquierda.

7. Histograma de frecuencias relativas ¿Cómo cambiaría la forma del histograma si la escala vertical utilizara frecuencias relativas expresadas en porcentajes en vez de los conteos de frecuencias reales que se muestran aquí?

8. Espacio ¿Cuál es una explicación razonable para el espacio entre las monedas con pesos de 5.5 gramos a 5.8 gramos y el grupo de monedas con pesos de 6.0 gramos y 6.4 gramos? (*Sugerencia:* Considere las columnas de las monedas de 25 centavos en el conjunto de datos 29 "Pesos de monedas" del apéndice B).

Elaboración de histogramas. ***En los ejercicios 9 a 16, trace los histogramas y responda las preguntas dadas.***

9. Old Faithful Utilice la distribución de frecuencias del ejercicio 11 en la sección 2-1 de la página 49 para elaborar un histograma. ¿La gráfica parece ser de una población con distribución normal?

10. Tornados Utilice la distribución de frecuencias del ejercicio 12 en la sección 2-1 de la página 49 para elaborar un histograma. ¿El histograma parece ser asimétrico? Si es así, identifique el tipo de asimetría.

11. Tiempos de servicio para el almuerzo en Burger King Utilice la distribución de frecuencias del ejercicio 13 en la sección 2-1 de la página 49 para elaborar un histograma. ¿El histograma parece ser asimétrico? Si es así, identifique el tipo de asimetría.

12. Tiempos de servicio para la cena en Burger King Utilice la distribución de frecuencias del ejercicio 1.1 en la sección 2-1 de la página 49 para elaborar un histograma. Utilice una interpretación estricta de los criterios para una distribución normal, ¿parece que el histograma representa los datos de una población con distribución normal?

13. Tiempos de servicio para el almuerzo en Wendy's Utilice la distribución de frecuencias del ejercicio 15 en la sección 2-1 de la página 49 para elaborar un histograma. ¿El histograma parece ser asimétrico? Si es así, identifique el tipo de asimetría.

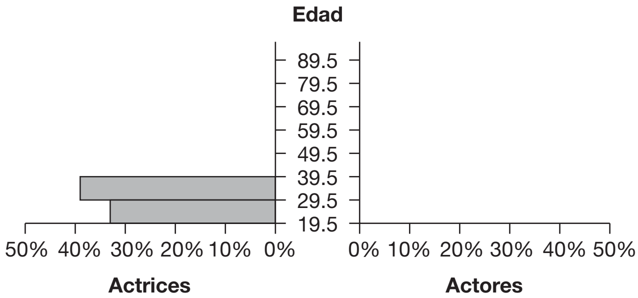
14. Tiempos de servicio para la cena en Wendy's Utilice la distribución de frecuencias del ejercicio 16 en la sección 2-1 de la página 49 para elaborar un histograma. Utilice una interpretación estricta de los criterios para una distribución normal, ¿parece que el histograma representa datos de una población con distribución normal?

15. Análisis de los últimos dígitos Utilice la distribución de frecuencias del ejercicio 17 en la sección 2-1 de la página 49 para elaborar un histograma. ¿Qué puede deducirse de la distribución de los dígitos? Específicamente, ¿las estaturas parecen ser reportadas o realmente medidas?

16. Análisis de los últimos dígitos Utilice la distribución de frecuencias del ejercicio 18 en la sección 2-1 de la página 49 para elaborar un histograma. ¿Qué puede deducirse de la distribución de los dígitos? Específicamente, ¿las estaturas parecen ser reportadas o realmente medidas?

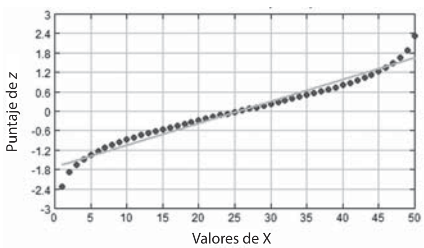
2-2 Más allá de lo básico

17. Histogramas de frecuencias relativas espalda con espalda Cuando se usan histogramas para comparar dos conjuntos de datos, en ocasiones es difícil hacer comparaciones con ellos. Un histograma de frecuencias relativas espalda con espalda tiene un formato que facilita en gran medida la comparación. En lugar de frecuencias, se deben usar frecuencias relativas (porcentajes o proporciones) para que las comparaciones no sean difíciles cuando hay diferentes tamaños de muestra. Utilice las distribuciones de frecuencias relativas de las edades de las actrices y los actores ganadores del Oscar en el ejercicio 19 de la sección 2-1 en la página 49 y complete los histogramas de frecuencias relativas espalda con espalda que se muestran a continuación. Después, utilice el resultado para comparar los dos conjuntos de datos.

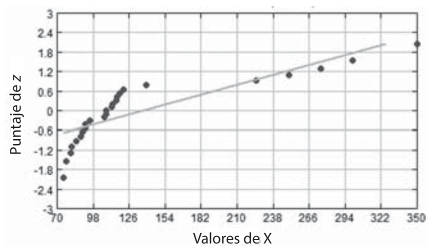


18. Interpretación de las gráficas cuantilares normales ¿Cuáles de las siguientes gráficas cuantilares normales parecen representar datos de una población con distribución normal? Explique.

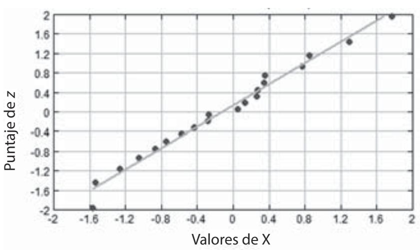
a.



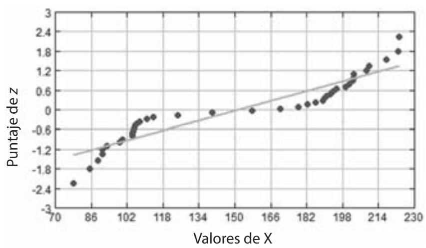
b.



c.



d.



2-3 Gráficas que informan y gráficas que engañan

Concepto clave En la sección 2-2 se introdujo el histograma, y esta sección presenta otras gráficas comunes que fomentan la comprensión de los datos. También se estudian algunas gráficas inexactas porque crean impresiones sobre los datos que, de alguna manera, son engañosas o erróneas.

La era de las gráficas encantadoras y primitivas dibujadas a mano ha pasado a la historia, y ahora la tecnología proporciona herramientas poderosas para generar una gran variedad de gráficas. ¡Aquí vamos!

Gráficas que informan

Gráficas de puntos

Una **gráfica de puntos** consiste en un gráfico de datos *cuantitativos* en el que cada valor de datos se representa como un punto sobre una escala horizontal de valores. Los puntos que representan valores iguales se apilan.

Características de una gráfica de puntos

- Muestra la forma de la distribución de los datos.
- Por lo general, es posible recrear la lista original de datos.

El poder de una gráfica

Con ventas anuales cercanas a los $13 mil millones y con alrededor de 50 millones de usuarios, el fármaco Lipitor (nombre genérico, atorvastatina) de Pfizer se ha convertido en el medicamento de prescripción más redituable y más utilizado de la historia. Al inicio de su desarrollo, Lipitor se comparó con otros fármacos (Zocor [simvastatina], Mevacor [lovastatina], Lescol [fluvastatina] y Pravachol [pravastatina]) en un proceso que implicó ensayos controlados. El resumen del informe incluyó una gráfica que mostraba una curva del Lipitor con un incremento más pronunciado que las curvas de los otros medicamentos, lo cual demostraba visualmente que Lipitor era más eficaz para reducir el colesterol que los otros fármacos. Pat Kelly, que en ese entonces era ejecutiva de marketing de nivel superior en Pfizer, declaró: “Nunca olvidaré cuando vi esa gráfica [...] En ese momento dije ‘¡Ah!, ahora sé de qué se trata’. ¡Podemos comunicar esto!”. La Administración de Alimentos y Medicamentos (FDA) de Estados Unidos aprobó el Lipitor y permitió a Pfizer incluir la gráfica con cada prescripción. El personal de ventas de la empresa también distribuyó la gráfica entre los médicos.

EJEMPLO 1 Gráfica de puntos de los pulsos en varones

La figura 2-6 muestra una gráfica de puntos de los pulsos (pulsaciones por minuto) de los varones del conjunto de datos 1 “Datos corporales” del apéndice B. Los dos puntos apilados encima de la posición 50 indican que dos de los pulsos son de 50. (En esta gráfica de puntos, la escala horizontal permite sólo números pares, pero los pulsos originales son todos pares).

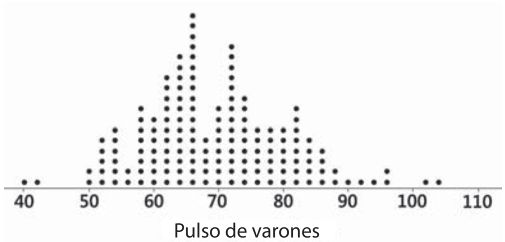


FIGURA 2-6 Gráfica de puntos de los pulsos de varones

SU TURNO Resuelva el ejercicio 5 “Pulsos”.

Diagramas de tallo y hojas

Un **diagrama de tallo y hojas** representa datos *cuantitativos* separando cada valor en dos partes: el tallo (por ejemplo el dígito más a la izquierda) y las hojas (como el dígito más a la derecha). A menudo, los mejores diagramas tallo y hojas se obtienen al redondear primero los valores de los datos originales. Además, los diagramas de tallo y hojas se pueden *expandir* para incluir más filas, o bien *condensar* para incluir menos, como en el ejercicio 21 “Diagramas de tallo y hojas expandidos”.

Características de un diagrama de tallo y hojas

- Muestra la forma de la distribución de los datos.
- Conserva los valores de los datos originales.
- Los datos muestrales aparecen ordenados.

EJEMPLO 2 Diagrama de tallo y hojas de los pulsos de varones

El siguiente diagrama muestra los pulsos de los varones en el conjunto de datos 1 “Datos corporales” del apéndice B. El pulso más bajo de 40 se separa en el tallo 4 y la hoja 0. Los tallos y las hojas se ordenan de manera creciente, no el orden establecido por la lista original. Si el tallo se coloca en forma horizontal, es posible ver la distribución de los pulsos de la misma manera que se vería en un histograma o una gráfica de puntos.

```
4 | 02  ←  Los pulsos son 40 y 42
5 | 0022222444444668888888
6 | 000000222222222224444444444444466666666666666666688888
7 | 00000000222222222222224444444444666666888888
8 | 000000222222222444446666688
9 | 02466  ←  Los pulsos son 90, 92, 94, 96, 96
10| 24
```

SU TURNO Resuelva el ejercicio 7 “Pulsos”.

Gráfica de series de tiempo

Una **gráfica de series de tiempo** es una gráfica de *datos dados en series* de tiempo, los cuales son datos cuantitativos recopilados en diferentes momentos, por ejemplo cada mes o cada año.

Característica de una gráfica de series de tiempo

- Revela información sobre tendencias a través del tiempo.

EJEMPLO 3 Gráfica de series de tiempo de bajas de agentes de la ley

La gráfica de series de tiempo que se muestra en la figura 2-7 presenta el número anual de muertes de agentes de la ley en Estados Unidos. Vea que ocurrió un pico en 2001, el año de los ataques terroristas del 11 de septiembre. Con excepción de los datos de 2001, parece haber una ligera tendencia a la baja.

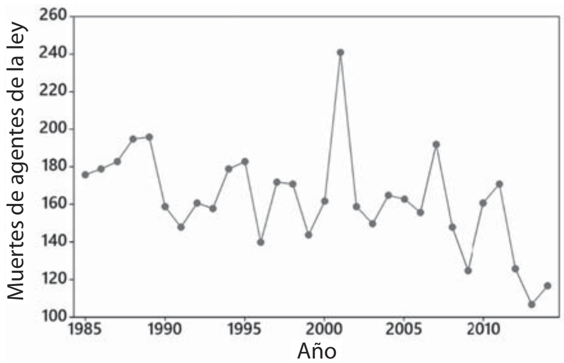


FIGURA 2-7 Gráfica de series de tiempo de las bajas de agentes de la ley

SU TURNO Resuelva el ejercicio 9 "Brecha de género en el salario".

Florence Nightingale

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

Gráficas de barras

Una **gráfica de barras** utiliza barras de igual anchura para mostrar las frecuencias de las categorías de datos *categóricos* (o cualitativos). Las barras pueden o no estar separadas por pequeños espacios.

Característica de una gráfica de barras

- Muestra la distribución relativa de datos categóricos para que sea más fácil comparar las diferentes categorías.

Gráficos de Pareto

Una **gráfica de Pareto** es una gráfica de barras para datos categóricos, con la estipulación añadida de que *las barras se ordenan de manera descendente* de acuerdo con las frecuencias; por ello las barras disminuyen de altura de izquierda a derecha.

Características de una gráfica de Pareto

- Muestra la distribución relativa de datos categóricos para que sea más fácil comparar las diferentes categorías.
- Orienta la atención hacia las categorías más importantes.

EJEMPLO 4 Gráfica de Pareto para los robos de vehículos acuáticos

La figura 2-8 muestra los tipos de vehículos acuáticos más robados en un año reciente. Se puede ver que los robos de motos acuáticas son el problema más grave.

continúa

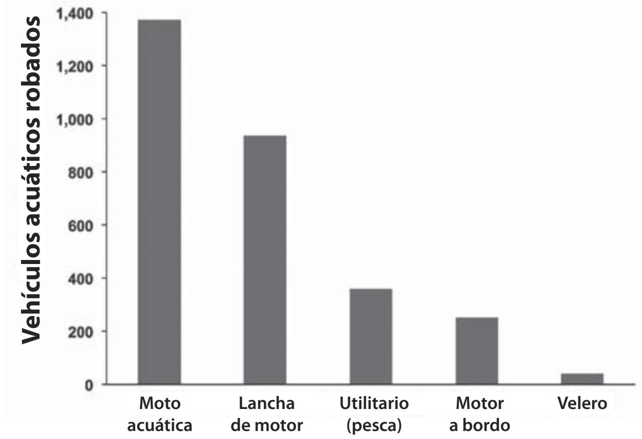


FIGURA 2-8 **Gráfica de Pareto de vehículos acuáticos robados**

SU TURNO Resuelva el ejercicio 11 "Rectificaciones en revistas".

Gráficas circulares

Una **gráfica circular** es una gráfica muy común que representa datos categóricos como rebanadas de un círculo; el tamaño de cada rebanada es proporcional al conteo de frecuencias para la categoría. Aunque las gráficas circulares son muy comunes, no son tan efectivas como las gráficas de Pareto.

Característica de una gráfica circular

- Muestra la distribución de datos categóricos en un formato de uso común.

EJEMPLO 5 Gráfica circular de vehículos acuáticos robados

La figura 2-9 es una gráfica circular con los mismos datos de robo del ejemplo 4. El trazado de una gráfica circular implica cortar el círculo en las proporciones apropiadas que representen frecuencias relativas. Por ejemplo, la categoría de motos acuáticas es el 46% del total, por lo que el segmento que representa las motos acuáticas debe ser 46% del total (con un ángulo central de $0.46 \times 360° = 166°$).

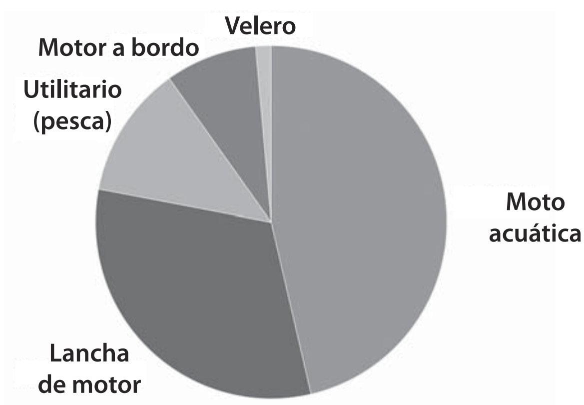


FIGURA 2-9 **Gráfica circular de vehículos acuáticos robados**

SU TURNO Resuelva el ejercicio 13 "Rectificaciones en revistas".

La gráfica de Pareto en la figura 2-8 y la gráfica circular de la figura 2-9 muestran los mismos datos de diferentes formas, pero la primera muestra de mejor manera los tamaños relativos de los diferentes componentes. El experto en gráficos Edwin Tufte hace la siguiente sugerencia:

> **Nunca utilice gráficas circulares porque desperdician tinta en componentes que no son datos y carecen de una escala apropiada.**

Polígono de frecuencias

Un **polígono de frecuencias** utiliza segmentos de línea conectados a puntos situados directamente encima de los valores de los puntos medios de clase. Un polígono de frecuencias es muy similar a un histograma, pero el polígono de frecuencias utiliza segmentos de línea en vez de barras.

Una variación del polígono de frecuencias básico es el **polígono de frecuencias relativas**, que utiliza frecuencias relativas (proporciones o porcentajes) en la escala vertical. Una ventaja de los polígonos de frecuencias relativas es que dos o más de ellos se pueden combinar en una sola gráfica para facilitar la comparación, como en la figura 2-11 (página 62).

EJEMPLO 6 Polígono de frecuencias de los tiempos de servicio para el almuerzo en McDonald's

Vea en la figura 2-10 el polígono de frecuencias correspondiente a los tiempos de servicio para el almuerzo en McDonald's, que se resumen en la distribución de frecuencias de la tabla 2-2 en la página 42. Las alturas de los puntos corresponden a las frecuencias de clase, los segmentos de línea se extienden hacia la derecha, y la gráfica comienza y termina sobre el eje horizontal.

FIGURA 2-10 Polígono de frecuencias de los tiempos de servicio para el almuerzo en McDonald's

SU TURNO Resuelva el ejercicio 15 "Old Faithful".

EJEMPLO 7 Polígono de frecuencias relativas: Tiempos de servicio para el almuerzo en McDonald's

La figura 2-11 muestra los polígonos de frecuencias relativas de los tiempos de servicio para el almuerzo en auto en McDonald's y Dunkin' Donuts. Se observa que los tiempos de servicio de Dunkin' Donuts son generalmente más bajos (más alejados hacia la izquierda en la gráfica) que los de McDonald's. Esto era de esperarse, dada la distinta naturaleza de sus menús.

continúa

En cifras

2,295,882,327: Cantidad de usuarios de Internet en el mundo.

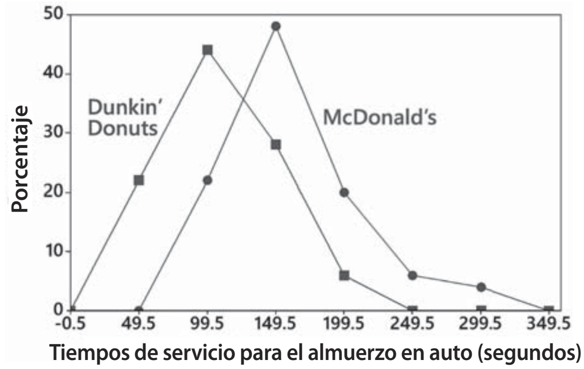


FIGURA 2-11 **Polígonos de frecuencias relativas para McDonald's y Dunkin' Donuts**

Gráficas que engañan

Por lo general, las gráficas engañosas se usan para mentirle a la gente, y realmente no deseamos que los estudiantes de estadística estén entre las personas susceptibles a tales engaños. Las gráficas se deben trazar de una manera justa y objetiva. Se debe dejar que los lectores hagan sus propios juicios, en vez de manipularlos mediante gráficas engañosas. A continuación presentamos dos de las maneras en que las gráficas suelen usarse para representar los datos de manera engañosa.

Eje vertical sin cero

Una gráfica engañosa común implica el uso de una escala vertical que comienza en algún valor mayor que cero para exagerar las diferencias entre los grupos.

EJE SIN CERO: Siempre examine cuidadosamente una gráfica para ver si el eje vertical empieza en algún punto distinto de cero, de modo que las diferencias sean exageradas.

EJEMPLO 8 Eje sin cero

Las figuras 2-12(a) y 2-12(b) se basan en los mismos datos de un ensayo clínico con OxyContin (oxycodone), un medicamento usado para tratar el dolor de moderado a severo. Los resultados de ese ensayo clínico incluyeron el porcentaje de sujetos que experimentaron náuseas en un grupo de tratamiento con OxyContin y otro grupo al que se administró un placebo.

Si se utiliza una escala vertical que comienza en 10% en lugar de 0%, como en la figura 2-12(a), se exagera groseramente la diferencia entre los dos grupos. La figura 2-12(a) hace parecer que aquellos que usan OxyContin experimentan náuseas a una tasa que es aproximadamente 12 veces mayor que la tasa de los sujetos que toman un placebo, pero la figura 2-12(b) muestra que la relación verdadera es de aproximadamente 2:1, no 12:1. Tal vez alguien quiere desalentar el uso recreativo de OxyContin influyendo en las personas para que piensen que el problema con las náuseas es mucho mayor de lo que realmente es. El objetivo puede ser honesto, pero el uso de una gráfica engañosa no es la forma correcta de lograr ese objetivo.

SU TURNO Resuelva el ejercicio 17 "Vehículos auto-conducidos"

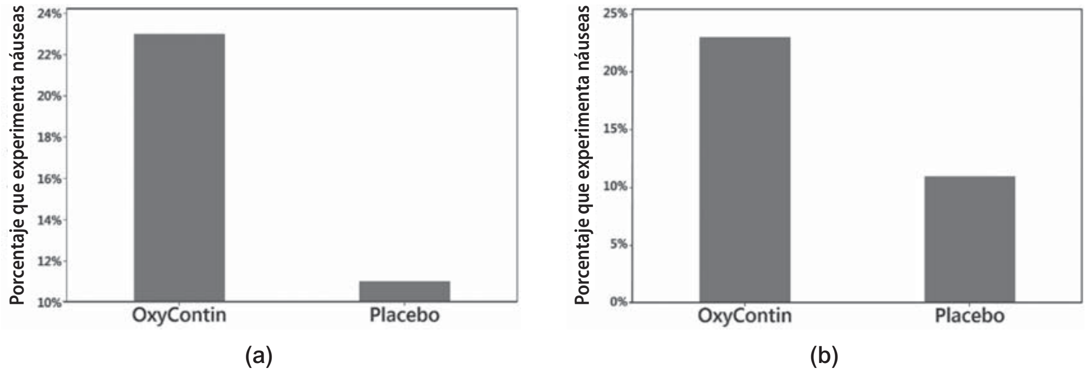


FIGURA 2-12 Náuseas en un ensayo clínico

Pictogramas

El trazado de *pictogramas* suele ser engañoso. Los datos que son de naturaleza unidimensional (como las cantidades presupuestarias) se representan a menudo con objetos bidimensionales (como billetes de dólar) u objetos tridimensionales (como pilas de monedas, casas o barriles). Mediante el uso de pictogramas, los dibujantes pueden crear falsas impresiones que distorsionan groseramente las diferencias mediante el uso de sencillos principios de geometría básica: (1) al duplicar cada lado de un cuadrado, su área no sólo se duplica; aumenta por un factor de *cuatro*; (2) cuando se duplica cada lado de un cubo, su volumen no sólo se duplica, sino que aumenta en un factor de *ocho*.

PICTOGRAMAS: Al examinar los datos representados con un pictograma, determine si el gráfico es engañoso porque los objetos de área o volumen se usan para representar cantidades que son realmente unidimensionales. (Los histogramas y las gráficas de barras representan datos unidimensionales con barras bidimensionales, pero usan barras con el mismo ancho para que el gráfico no sea engañoso.)

EJEMPLO 9 Pictograma de fumadores

Observe la figura 2-13 y note que el cigarrillo más grande es aproximadamente dos veces más largo, dos veces más alto y dos veces más grueso que el cigarrillo más pequeño, por lo que el volumen del cigarrillo mayor es aproximadamente *ocho veces* el volumen del menor. (Los datos provienen de los Centros para el Control y la Prevención de Enfermedades). El cigarrillo mayor *parece* ser ocho veces más grande que el cigarrillo menor, pero los porcentajes reales muestran que la tasa del 37% de fumadores en 1970 es aproximadamente *dos veces* el 18% de 2013.

1970: 37% de los estadounidenses fumaban.

2013: 18% de los estadounidenses fumaban.

FIGURA 2-13 Tabaquismo en adultos estadounidenses

SU TURNO Resuelva el ejercicio 19 “Costo de dar a luz”.

Pensamientos conclusivos

Además de las gráficas que se han analizado en esta sección, hay muchas otras gráficas útiles, algunas de las cuales aún no han sido creadas. El mundo necesita desesperadamente más gente que pueda crear gráficas originales que nos aclaren la naturaleza de los datos. En *The Visual Display of Quantitative Information*, Edward Tufte ofrece los siguientes principios:

- Para conjuntos de datos pequeños de 20 valores o menos, utilice una tabla en vez de una gráfica.
- Una gráfica de datos debe hacer que el lector se concentre en la verdadera naturaleza de los datos, no en otros elementos, como características de diseño llamativas pero distractoras.
- No distorsione los datos; elabore la gráfica para revelar la verdadera naturaleza de los datos.
- La mayor parte de la tinta en una gráfica debe utilizarse para los datos, no para otros elementos de diseño.

CENTRO DE TECNOLOGÍA

Capacidades gráficas

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

En lugar de enumerar instrucciones para cada tipo de gráfica, las siguientes listas identifican las gráficas que se pueden generar mediante el uso de las diferentes tecnologías.

Statdisk

- Histogramas
- Gráficas circulares
- Diagramas de dispersión

Minitab

- Histogramas
- Gráficas de puntos
- Diagramas de tallo y hojas
- Gráficas de series de tiempo
- Gráficas de barras
- Gráficas de Pareto
- Gráficas circulares
- Polígonos de frecuencias
- Diagramas de dispersión

StatCrunch

- Histogramas
- Gráficas de puntos
- Diagramas de tallo y hojas
- Gráficas de barras
- Gráficas circulares
- Diagramas de dispersión

Calculadora TI-83/84 Plus

- Histogramas
- Gráficas de series de tiempo
- Polígonos de frecuencias
- Diagramas de dispersión

Excel

- Histogramas
- Gráficas de series de tiempo
- Gráficas de barras
- Gráficas de Pareto
- Gráficas circulares
- Diagramas de dispersión

2-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Temperaturas corporales A continuación se listan temperaturas corporales (°F) de adultos sanos. ¿Por qué una gráfica de estos datos no sería muy efectiva para ayudarnos a entender los datos?

98.6	98.6	98.0	98.0	99.0	98.4	98.4	98.4	98.4	98.6

2. Datos de respuesta voluntaria Si tenemos una gran muestra de respuesta voluntaria consistente en los pesos de sujetos que optaron por responder a una encuesta publicada en Internet, ¿puede una gráfica ayudar a superar la deficiencia de que la muestra sea de respuesta voluntaria?

3. Ética Hay datos que muestran que fumar es perjudicial para la salud. Dado que se podría ayudar a las personas y salvar vidas mediante la reducción del tabaquismo, ¿es ético graficar los datos de una manera que sea engañosa al exagerar los riesgos para la salud del tabaquismo?

4. CVDOT La sección 2-1 introdujo características importantes de los datos resumidas por el acrónimo CVDOT. ¿Qué características representan esas letras y cuál gráfica proporciona el mejor discernimiento de la última de esas características?

Gráficas de puntos. ***En los ejercicios 5 y 6, elabore la gráfica de puntos.***

5. Pulsos A continuación se listan los pulsos (pulsaciones por minuto) de mujeres seleccionadas del conjunto de datos 1 "Datos corporales" en el apéndice B. Todas estos pulsos son números pares. ¿Hay algún pulso que parezca ser un valor atípico? ¿Cuál es ese valor?

80 94 58 66 56 82 78 86 88 56 36 66 84 76 78 64 66 78 60 64

6. Presión arterial diastólica A continuación se listan las mediciones de presión arterial diastólica (mm Hg) de mujeres seleccionadas del conjunto de datos 1 "Datos corporales" en el apéndice B. Todos los valores son números pares. ¿Existen valores atípicos? Si es así, identifique sus valores.

62	70	72	88	70	66	68	70	82	74	90	62
70	76	90	86	60	78	82	78	84	76	60	64

Diagramas de tallo y hojas. ***En los ejercicios 7 y 8, elabore el diagrama de tallo y hojas.***

7. Pulsos Considere los datos listados en el ejercicio 5. ¿Cómo se ordenan los datos en el diagrama de tallo y hojas?

8. Presión arterial diastólica Considere los datos listados en el ejercicio 6. Identifique los dos valores que están más cerca del centro si los datos se ordenan de menor a mayor. (Estos valores se usan a menudo para encontrar la *mediana*, que se define en la sección 3-1).

Gráficas de series de tiempo. ***En los ejercicios 9 y 10, elabore la gráfica de series de tiempo.***

9. Brecha de género en el salario A continuación se listan las medianas de las ganancias de las mujeres como porcentaje de las ganancias medianas de los hombres en los últimos años a partir de 1990. ¿Hay una tendencia? ¿Cómo parece afectar a las mujeres?

71.6	69.9	70.8	71.5	72.0	71.4	73.8	74.2	73.2	72.3	73.7
76.3	76.6	75.5	76.6	77.0	76.9	77.8	77.1	77.0	77.4	77.0

10. Jonrones A continuación se listan las cantidades de jonrones conectados en las Ligas Mayores de Béisbol, cada año, a partir de 1990 (listados en orden por filas). ¿Hay una tendencia?

3317	3383	3038	4030	3306	4081	4962	4640	5064	5528	5693	5458
5059	5207	5451	5017	5386	4957	4878	5042	4613	4552	4934	4661

Gráficas de Pareto. ***En los ejercicios 11 y 12, elabore la gráfica de Pareto.***

11. Rectificaciones en revistas En un estudio de las rectificaciones en revistas biomédicas, 436 fueron por error, 201 por plagio, 888 por fraude, 291 por publicaciones duplicadas y 287 por otras causas (con base en los datos de "Recuento de mala praxis para la mayoría de rectificaciones de revistas científicas" (*Misconducts Accounts for the Majority of Retracted Scientific Publications*", de Fang, Steen, Casadevall, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, núm. 3). Entre tales rectificaciones, ¿la mala conducta (fraude, duplicación, plagio) parece ser un factor importante?

12. Obtención de un empleo En una encuesta, a los sujetos que buscaban un empleo se les preguntó a quién debían enviar una nota de agradecimiento después de tener una entrevista de trabajo. Los resultados fueron los siguientes: 210 dijeron que sólo a la persona con la que pasaron la mayor parte del tiempo, 396 dijeron que a todos los que conocieron, 40 dijeron que sólo a la persona de más alto nivel, 15 dijeron que a la persona con la que tuvieron la mejor conversación y 10 dijeron que no envían notas de agradecimiento (basado en datos de TheLadders.com). Comente los resultados.

Gráficas circulares. ***En los ejercicios 13 y 14, elabore la gráfica circular.***

13. Rectificaciones en revistas Utilice los datos del ejercicio 11 “Rectificaciones en revistas”.

14. Obtención de un empleo Utilice los datos del ejercicio 12 “Obtención de un empleo”.

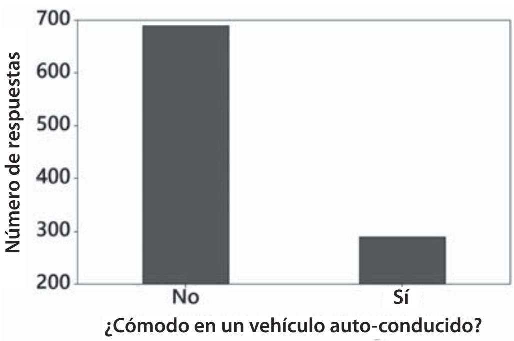
Polígono de frecuencias. ***En los ejercicios 15 y 16, elabore los polígonos de frecuencias.***

15. Old Faithful Utilice la distribución de frecuencias del ejercicio 11 en la sección 2-1 de la página 49 para trazar un polígono de frecuencias. ¿La gráfica sugiere que la distribución es asimétrica? ¿De qué manera lo es?

16. Tornados Utilice la distribución de frecuencias del ejercicio 12 en la sección 2-1 de la página 49 para trazar un polígono de frecuencias. ¿La gráfica sugiere que la distribución es asimétrica? ¿De qué manera lo es?

Gráficas engañosas. ***En los ejercicios 17 a 20, identifique por qué la gráfica es engañosa.***

17. Vehículos auto-conducidos En una encuesta a adultos, se preguntó a los sujetos si se sentían cómodos en un vehículo auto-conducido. La gráfica adjunta muestra los resultados (con base en datos de TE Connectivity).



18. Tarifa del metro En 1986, la tarifa del metro de la Ciudad de Nueva York era de $1, y el costo actual es de $2.50, así que el precio de 1986 se ha multiplicado por 2.5. En la gráfica adjunta, el billete grande es 2.5 veces más alto y 2.5 veces más ancho que el billete pequeño.

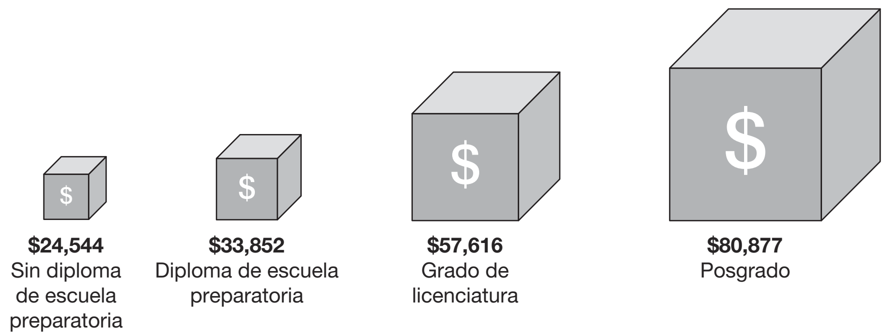
Tarifa del metro en 1986

Tarifa actual del metro

19. Costo de dar a luz De acuerdo con la Agencia para la Investigación de la Salud y el Proyecto de Costo y Utilización de la Salud de Calidad, el costo típico de un parto por cesárea es de $4500 y el de un parto vaginal es de $2600. Observe la siguiente ilustración.

Costo del parto por cesárea: $4500 **Costo del parto vaginal: $2600**

20. Ingresos y grados académicos La gráfica adjunta presenta trabajadores con varios grados académicos, junto con sus niveles de ingreso.



2-3 Más allá de lo básico

21. Diagramas de tallo y hojas expandidas Un diagrama de tallo y hojas se puede *condensar* combinando las filas adyacentes. Podríamos usar un tallo de "6-7" en vez de tallos separados de 6 y 7. Cada fila en el diagrama de tallo y hojas condensado debe incluir un asterisco para separar los dígitos asociados con los diferentes valores del tallo. Un diagrama de tallo y hojas se puede *expandir* subdividiendo las filas en aquellas con hojas que tienen dígitos de 0 a 4 y las que tienen hojas con dígitos de 5 a 9. Si se consideran las temperaturas corporales a partir de las 12 AM del día 2 listadas en el conjunto de datos 3 "Temperaturas corporales" en el apéndice B, se observa que las tres primeras filas de un diagrama de tallo y hojas expandido tienen tallos de 96 (para hojas entre 5 y 9 inclusive), 97 (para hojas entre 0 y 4 inclusive) y 97 (para hojas entre 5 y 9 inclusive). Elabore el diagrama completo de tallo y hojas expandido para las temperaturas corporales a partir de las 12 AM del día 2, que se listan en el conjunto de datos 3 "Temperaturas corporales" en el apéndice B.

2-4 Diagramas de dispersión, correlación y regresión

Concepto clave En esta sección se presenta el análisis de datos muestrales *pareados*. En la parte 1 de esta sección se analiza la *correlación* y el papel de una gráfica llamada *gráfica de dispersión*. En la parte 2 se proporciona una introducción al uso del *coeficiente de correlación lineal*. En la parte 3 se hace un breve análisis de la *regresión lineal*, que implica la ecuación y la gráfica de la recta que mejor se ajusta a los datos muestrales emparejados.

Todos los principios estudiados en esta sección se analizan con mayor detalle en el capítulo 10, pero esta sección sirve como una introducción rápida a algunos conceptos importantes de la correlación y la regresión. La presente sección no incluye detalles para ejecutar cálculos manuales, los cuales se realizan con poca frecuencia. Las instrucciones para utilizar la tecnología y obtener los resultados deseados se incluyen en el capítulo 10.

PARTE 1 Diagrama de dispersión y correlación

Nuestro objetivo en esta sección es explorar si existe una *correlación*, o asociación, entre dos variables. Comenzamos con definiciones básicas.

DEFINICIONES

Existe una **correlación** entre dos variables cuando los valores de una variable están de alguna manera asociados con los valores de la otra.

Existe una **correlación lineal** entre dos variables cuando hay una correlación y los puntos graficados de los datos pareados dan como resultado un patrón que puede aproximarse mediante una línea recta.

Un **diagrama de dispersión** (o **gráfica de dispersión**) es un diagrama de datos cuantitativos pareados (x, y) con un eje x horizontal y un eje vertical y; el eje horizontal se utiliza para la primera variable (x) y el eje vertical se usa para la segunda variable (y).

PRECAUCIÓN La presencia de una correlación entre dos variables no es evidencia de que una de las variables *cause* la otra. Podríamos encontrar una correlación entre el consumo de cerveza y el peso, pero no podemos concluir a partir de la evidencia estadística que beber cerveza tiene un efecto directo sobre el peso.

¡La correlación no implica causalidad!

Un diagrama de dispersión puede ser muy útil para determinar si existe una correlación (o relación) entre dos variables. (Este tema se analiza de manera extensa durante el estudio de la correlación en la sección 10-1).

EJEMPLO 1 Correlación: Circunferencia de la cintura y del brazo

El conjunto de datos 1 “Datos corporales” en el apéndice B incluye circunferencias de la cintura (cm) y del brazo (cm) de sujetos adultos seleccionados al azar. La figura 2-14 es un diagrama de dispersión de las mediciones de cintura y brazo pareadas. Los puntos muestran un patrón de valores crecientes de izquierda a derecha. Este patrón sugiere que existe una correlación o relación entre las circunferencias de la cintura y las del brazo.

SU TURNO Resuelva el ejercicio 7 “Peso del automóvil y consumo de combustible”.

EJEMPLO 2 Sin correlación: Peso y pulso

El conjunto de datos 1 “Datos corporales” en el apéndice B incluye pesos (kg) y pulsaciones (latidos por minuto) de sujetos adultos seleccionados al azar. La figura 2-15 es un diagrama de dispersión de las mediciones de peso y pulso emparejadas. Los puntos en la figura 2-15 no muestran ningún patrón obvio, y esta carencia de un patrón sugiere que no hay correlación o relación entre los pesos y los pulsos.

SU TURNO Resuelva el ejercicio 8 “Estaturas de padres e hijos”.

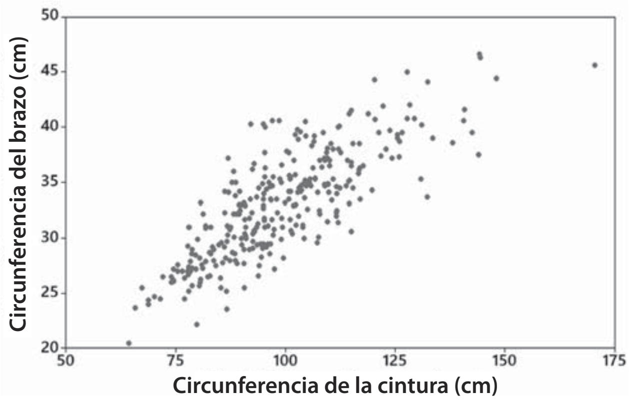


FIGURA 2-14 Circunferencias de cintura y brazo

Correlación: El patrón distintivo de los puntos graficados sugiere que hay una correlación entre las circunferencias de la cintura y las del brazo.

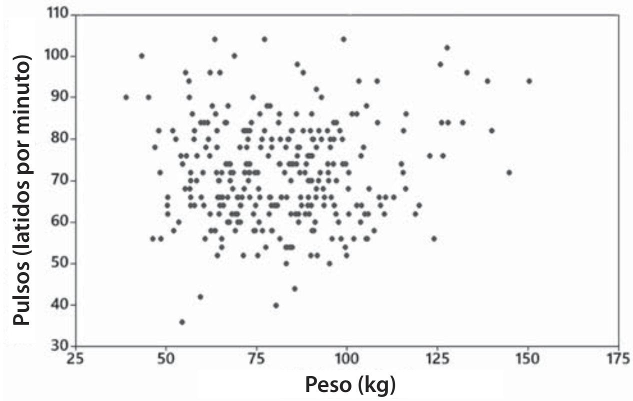


FIGURA 2-15 Pesos y pulsos

Sin correlación: Los puntos trazados no muestran un patrón distintivo, por lo que parece que no hay correlación entre los pesos y los pulsos.

EJEMPLO 3 Conglomerados y un espacio

Considere el diagrama de dispersión de la figura 2-16. Representa los datos apareados que consisten en el peso (gramos) y el año de fabricación de cada uno de 72 centavos. Este diagrama de dispersión muestra dos grupos muy distintos separados por un espacio, lo que puede explicarse por la inclusión de dos poblaciones diferentes: Las monedas de antes de 1983 son 97% de cobre y 3% de zinc, pero las monedas posteriores a 1983 son 2.5% de cobre y 97.5% de zinc. Si se ignora la característica de los conglomerados, podríamos pensar incorrectamente que hay una relación entre el peso de un centavo y el año en que se hizo. Si examinamos los dos grupos por separado, vemos que *no* parece haber una relación entre los pesos de los centavos y los años en que se produjeron.

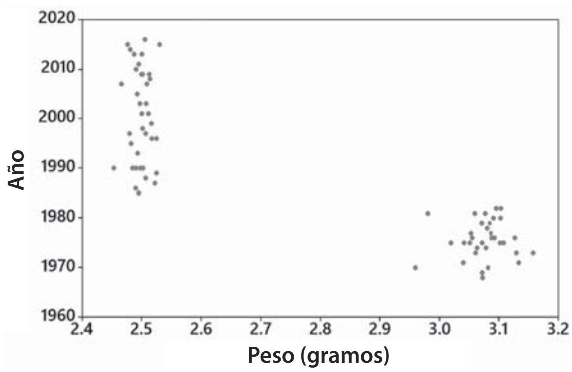


FIGURA 2-16 Pesos de centavos y años de producción

Los tres ejemplos anteriores implican tomar decisiones sobre una correlación basada en juicios subjetivos de diagramas de dispersión, pero la parte 2 presenta el coeficiente de *correlación lineal* como una medida que puede ayudarnos a tomar tales decisiones de manera más objetiva. Mediante el uso de datos pareados, es posible calcular el valor del *coeficiente de correlación lineal r*.

PARTE 2 Coeficiente de correlación lineal *r*

DEFINICIÓN

El **coeficiente de correlación lineal** se expresa con *r*, y mide la fuerza de la asociación lineal entre dos variables.

El valor de un coeficiente de correlación lineal *r* se puede calcular manualmente aplicando la fórmula 10-1 o la fórmula 10-2 que se encuentran en la sección 10-l de la página 473, pero en la práctica, *r* se encuentra casi siempre utilizando software estadístico o una calculadora adecuada.

Utilización de *r* para determinar la correlación

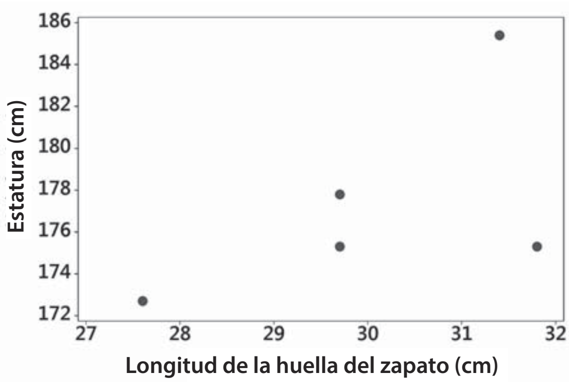
El valor calculado del coeficiente de correlación lineal está siempre entre -1 y 1. Si *r* es cercano a -1 o cercano a 1, parece haber una correlación, pero si *r* es cercano a 0, no parece haber una correlación lineal. Para los datos representados en el diagrama de dispersión de la figura 2-14, $r = 0.802$ (algo cercano a 1), y los datos en el diagrama de dispersión de la figura 2-15 resultan en $r = 0.082$ (muy cercano a 0). Estas descripciones de "cercano a" -1 o 1 o 0 son vagas, pero hay otros criterios objetivos. Por ahora se utilizará una tabla de valores especiales (tabla 2-11) para decidir si existe una correlación lineal. Vea el siguiente ejemplo que ilustra la interpretación del coeficiente de correlación lineal *r*.

EJEMPLO 4 ¿Correlación entre las longitudes de las huellas del zapato y las estaturas?

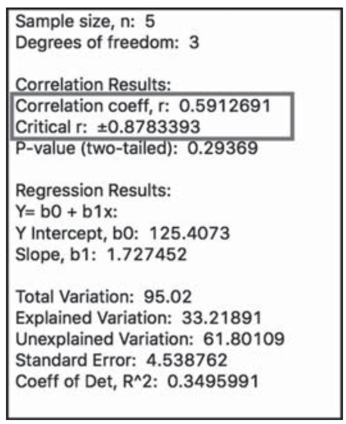
Considere los datos de la tabla 2-10 (con datos del conjunto de datos 2 "Pie y estatura" del apéndice B). A partir del diagrama de dispersión adjunto de datos pareados en la tabla 2-10, no está muy claro si existe una correlación lineal. La pantalla de resultados de Statdisk muestra que el coeficiente de correlación lineal tiene el valor de $r = 0.591$ (redondeado).

TABLA 2-10 Longitudes de la huella del zapato y estatura de hombres

Longitud de la huella del zapato (cm)	29.7	29.7	31.4	31.8	27.6
Estatura (cm)	175.3	177.8	185.4	175.3	172.7



Statdisk


Sample size, n: 5
Degrees of freedom: 3

Correlation Results:
Correlation coeff, r: 0.5912691
Critical r: ±0.8783393
P-value (two-tailed): 0.29369

Regression Results:
Y= b0 + b1x:
Y Intercept, b0: 125.4073
Slope, b1: 1.727452

Total Variation: 95.02
Explained Variation: 33.21891
Unexplained Variation: 61.80109
Standard Error: 4.538762
Coeff of Det, R^2: 0.3495991

SU TURNO Resuelva el ejercicio 9 "Coeficiente de correlación lineal".

En el ejemplo 4, se sabe por la pantalla de Statdisk que al utilizar los cinco pares de datos de la tabla 2-10, el coeficiente de correlación lineal se calcula como $r = 0.591$. Utilice los siguientes criterios para interpretar dichos valores.

Uso de la tabla 2-11 para interpretar r: Considere los valores críticos de la tabla 2-11 como positivos y negativos y dibuje una gráfica similar a la figura 2-17. Utilice los valores de la tabla para determinar si un valor de un coeficiente de correlación lineal r es "cercano a" 0 o "cercano a" −1 o "cercano a" 1 aplicando los siguientes criterios:

Correlación Si el coeficiente de correlación lineal calculado r está en la región de la cola izquierda o derecha más allá del valor de la tabla para esa cola, concluya que hay evidencia suficiente para apoyar la afirmación de una correlación lineal.

Sin correlación Si el coeficiente de correlación lineal calculado r está entre los dos valores críticos, concluya que no hay evidencia suficiente para apoyar la afirmación de una correlación lineal.

TABLA 2-11 Valores críticos del coeficiente de correlación lineal *r*

Número de pares de datos *n*	Valor crítico de *r*
4	0.950
5	**0.878**
6	0.811
7	0.754
8	0.707
9	0.666
10	0.632
11	0.602
12	0.576

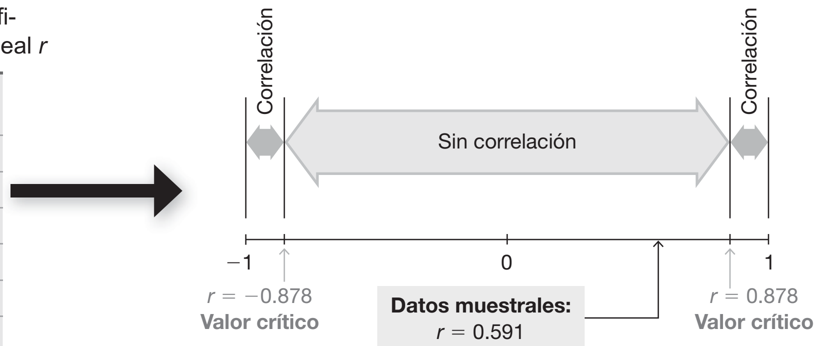


FIGURA 2-17 Valores críticos de la Tabla 2-11 y el valor calculado de *r*

La figura 2-17 muestra que el coeficiente de correlación lineal de $r = 0.591$ calculado a partir de los datos de muestrales pareados es un valor que se encuentra entre los valores críticos de $r = -0.878$ y $r = 0.878$ (se encuentran en la tabla 2-11). La figura 2-17 muestra que podemos considerar que el valor de $r = 0.591$ es cercano a 0 en lugar de estar cerca de −1 o cerca de 1. Por lo tanto, no tenemos evidencia suficiente para concluir que hay una correlación lineal entre las longitudes de la huella del zapato y las estaturas de los hombres.

Valores de *P* para determinar la correlación lineal

En el ejemplo 4, se utilizó el valor calculado del coeficiente de correlación lineal $r = 0.591$ y se comparó con los valores críticos de r de ±0.878 encontrados en la tabla 2-11. (Vea la figura 2-17). En el mundo real de las aplicaciones estadísticas, el uso de estas tablas es casi obsoleto. La sección 10-1 describe un enfoque más común que se basa en "valores de *P*" en lugar de tablas. La pantalla de Statdisk que acompaña al ejemplo 4 muestra que el valor de *P* es 0.29369, o 0.294, una vez que se redondea. Los valores de *P* se introducen en el capítulo 8, pero a continuación se da una definición preliminar adecuada para el contexto de esta sección:

DEFINICIÓN

Si no hay correlación lineal entre dos variables, el ***valor de P*** es la probabilidad de obtener datos muestrales pareados con un coeficiente de correlación lineal r que sea al menos tan extremo como el obtenido a partir de los datos muestrales emparejados.

Con base en el ejemplo 4 y la pantalla de resultados de Statdisk que muestran un valor P de 0.294, sabemos que hay una probabilidad de 0.294 (o una probabilidad de 29.4%) de obtener un coeficiente de correlación lineal $r = 0.591$ o más extremo, suponiendo que no hay correlación lineal entre la longitud de la huella del zapato y la estatura. (Los valores de r que son "más extremos" que 0.591 son los mayores que 0.591 y los menores que -0.591).

Interpretación de un valor *P*. El valor P de 0.294 del ejemplo 4 es alto. Muestra que existe una alta probabilidad de obtener un coeficiente de correlación lineal de $r = 0.591$ (o más extremo) por casualidad cuando no hay correlación lineal entre las dos variables. Debido a que la probabilidad de obtener $r = 0.591$ o un valor más extremo es tan alta (29.4% de probabilidad), concluimos que no hay suficiente evidencia para afirmar que existe una correlación lineal entre las longitudes de la huella del zapato y las estaturas de los hombres.

Sólo un valor de *P pequeño*, como 0.05 o menor (o una probabilidad de 5% o menos), sugiere que *no* es probable que los resultados de la muestra ocurran por casualidad cuando no hay correlación lineal, por lo que un valor de *P* pequeño apoya la conclusión de que existe una correlación lineal entre las dos variables.

EJEMPLO 5 ¿Correlación entre las longitudes de la huella del zapato y las estaturas?

En el ejemplo 4 se usaron sólo cinco pares de datos del conjunto 2 "Pie y estatura" en el apéndice B. Si se utilizan las longitudes de la huella del zapato y las estaturas de los 40 elementos listados en el conjunto de datos 2 del apéndice B, se obtiene el diagrama de dispersión mostrado en la figura 2-18 y los resultados de Minitab mostrados en la pantalla adjunta. El diagrama de dispersión muestra un patrón distintivo en lugar de tener puntos dispersos por todas partes. Además, se observa que el valor del coeficiente de correlación lineal es $r = 0.813$, y el valor de P es 0.000 redondeado a tres decimales. Debido a que el valor P de 0.000 es *pequeño*, se tiene evidencia suficiente para concluir que existe una correlación lineal entre las longitudes de las huellas del zapato y las estaturas.

En el ejemplo 4, con sólo cinco pares de datos, no teníamos suficiente evidencia para concluir que existe tal correlación lineal, pero con 40 pares de datos, aquí existe evidencia suficiente para concluir la existencia de correlación lineal.

Minitab

```
Pearson correlation of Shoe Print Length and Height = 0.813
P-Value = 0.000
```

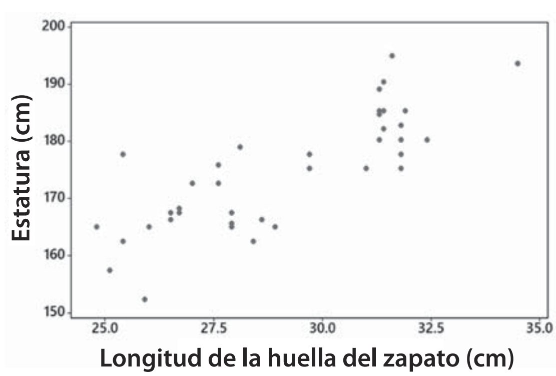


FIGURA 2-18 Diagrama de dispersión de 40 pares de datos

Muertes de policías en persecuciones automovilísticas

USA Today investigó el informe anual sobre el número de policías muertos durante persecuciones automovilísticas. Se encontró que la Oficina Federal de Investigaciones (FBI) contabilizó 24 muertes en los últimos 35 años, pero otros registros muestran que hubo 371 muertes durante ese período. El reportero de *USA Today* Thomas Frank escribió que "el bajo conteo es uno de los ejemplos más extremos de la incapacidad del gobierno federal para rastrear con precisión las muertes violentas y ha llevado al FBI a minimizar el peligro de que la policía persiga a los automovilistas". Aparentemente, el FBI categorizaba estas muertes como accidentes automovilísticos en lugar de designarlas como muertes de policías ocurridas durante una persecución automovilística.

PARTE 3 Regresión

Cuando concluimos que parece haber una correlación lineal entre dos variables (como en el ejemplo 5), podemos encontrar la ecuación de la recta que mejor se ajusta a los datos muestrales, y esa ecuación se puede utilizar para predecir el valor de una variable cuando se da un valor específico de la otra variable. Con base en los resultados del ejemplo 5, podemos predecir la estatura de alguien dada la longitud de la huella de su zapato (que puede haber sido encontrada en una escena del crimen).

En vez de usar el formato de la ecuación lineal $y = mx + b$ que aprendimos en los cursos previos de matemáticas, usamos el formato siguiente.

DEFINICIÓN

Dada una colección de datos muestrales pareados, la **línea de regresión** (o *línea de mejor ajuste* o *línea de mínimos cuadrados*) es la recta que "mejor" se ajusta a la dispersión de los datos. (El criterio específico para la "mejor" línea recta de ajuste es la propiedad de "mínimos cuadrados" descrita en la sección 10-2).

La **ecuación de regresión**

$$\hat{y} = b_0 + b_1x$$

describe algebraicamente la línea de regresión.

La sección 10-2 proporciona una buena razón para usar el formato $\hat{y} = b_0 + b_1x$ en vez del formato $y = mx + b$. La sección 10-2 también proporciona fórmulas que podrían utilizarse para identificar los valores de la intersección en y, b_0, y de la pendiente, b_1, pero los valores se encuentran normalmente mediante el uso de un software estadístico o una calculadora adecuada.

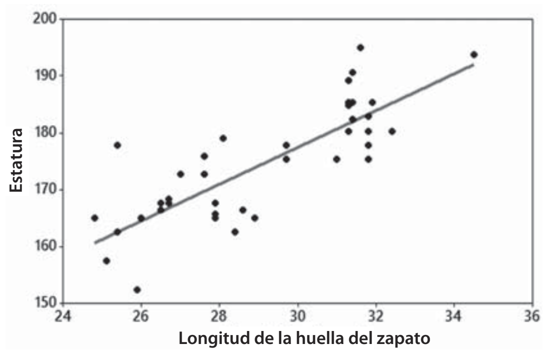
EJEMPLO 6 Línea de regresión

El ejemplo 5 incluyó un diagrama de dispersión de los 40 pares de longitudes de huella del zapato y estaturas del conjunto de datos 2 "Pie y estatura" del apéndice B. La figura 2-19 de la página siguiente muestra ese mismo diagrama de dispersión, pero incluye la gráfica de la línea de regresión. También se muestra la pantalla de Statdisk para los 40 pares de datos.

A partir de la pantalla de Statdisk, se observa que la forma general de la ecuación de regresión tiene un intersección y de $b_0 = 80.9$ (redondeada) y una pendiente $b_1 = 3.22$ (redondeada), por lo que la ecuación de la línea de regresión mostrada en la figura 2-19 es $\hat{y} = 80.9 + 3.22x$. Podría ser útil expresar esta ecuación más claramente usando los nombres de las variables:

$$\text{Estatura} = 80.9 + 3.22\ (\text{Longitud de la huella del zapato})$$

Observe que la *ecuación* muestra la intersección y de 80.9 que no aparece en la escala vertical de la *gráfica*. La escala vertical más a la izquierda de la figura 2-19 no es el eje real de y que pasa por 0 en el eje x. Si la gráfica se extendiera a la izquierda, la línea de regresión intersecaría al eje y real a la altura de $y = 80.9$ cm.



Statdisk

Correlation Results:
Correlation coeff, r: 0.812948
Critical r: ±0.3120061
P-value (two-tailed): 0.000

Regression Results:
Y= b0 + b1x:
Y Intercept, b0: 80.93041
Slope, b1: 3.218561

FIGURA 2-19 Línea de regresión

2-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Correlación lineal En esta sección usamos r para expresar el valor del coeficiente de correlación lineal. ¿Por qué nos referimos a este coeficiente de correlación como *lineal*?

2. Causalidad Un estudio ha demostrado que existe una correlación entre el peso corporal y la presión arterial. Los pesos corporales más altos están asociados con los niveles más altos de presión arterial. ¿Podemos concluir que el aumento de peso es una causa del aumento de la presión arterial?

3. Diagrama de dispersión ¿Qué es un diagrama de dispersión y cómo nos ayuda?

4. Estimación de *r* En cada uno de los siguientes casos, estime el valor del coeficiente de correlación lineal r para los datos pareados que se dan y que corresponden a 50 adultos seleccionados al azar.

a. Sus estaturas se miden en pulgadas (x) y esas mismas estaturas se registran en centímetros (y).

b. Se miden sus puntuaciones de IQ (x) y sus estaturas en centímetros (y).

c. Se miden sus pulsos (x) y sus puntuaciones de IQ (y).

d. Se miden sus estaturas en centímetros (x) y se listan de nuevo las mismas estaturas, pero con un signo negativo precediendo cada dato (y).

Diagrama de dispersión ***En los Ejercicios 5 a 8, use los datos muestrales para trazar un diagrama de dispersión. Utilice la primera variable para el eje x. Con base en el diagrama de dispersión, ¿qué conclusión obtiene sobre una correlación lineal?***

5. Volumen cerebral e IQ La tabla lista los volúmenes del cerebro (cm^3) y las puntuaciones de IQ para cinco varones (del conjunto de datos 8 "IQ y tamaño del cerebro" del apéndice B).

Volumen del cerebro (cm^3)	1173	1067	1347	1029	1204
IQ	101	93	94	97	113

6. Medidas de un oso La tabla lista los tamaños de tórax (distancia alrededor del pecho en pulgadas) y pesos (libras) que se midieron en osos anestesiados (conjunto de datos 9 "Mediciones de un oso" en el apéndice B).

Pecho (pulg.)	26	45	54	49	35	41	41
Peso (lb)	80	344	416	348	166	220	262

7. Peso del auto y consumo de combustible La tabla lista los pesos (en libras) y los consumos de combustible en carretera (mpg) para los autos Hyundai Elantra, Nissan Altima, VW Passat, Buick Lucerne, Mercury Grand Marquis, Honda Civic y Honda Accord.

Peso (lb)	2895	3215	3465	4095	4180	2740	3270
Carretera (mpg)	33	31	29	25	24	36	30

8. Estaturas de padres e hijos La tabla lista las alturas de padres y de sus primeros hijos (obtenido de Francis Galton).

Altura del padre (pulg.)	73.0	75.5	75.0	75.0	75.0	74.0	74.0	73.0	73.0	78.5
Altura del primer hijo (pulg.)	74.0	73.5	71.0	70.5	72.0	76.5	74.0	71.0	72.0	73.2

Coeficiente de correlación lineal. ***En los ejercicios 9 a 12, se proporciona el coeficiente de correlación lineal r. Utilice la tabla 2-11 en la página 71 para encontrar los valores críticos de r. Con base en una comparación del coeficiente de correlación lineal r y los valores críticos, ¿qué concluye usted acerca de una correlación lineal?***

9. Considerando los datos del ejercicio 5 "Volumen cerebral e IQ", el coeficiente de correlación lineal es $r = 0.127$.

10. Considerando los datos del ejercicio 6 "Mediciones de un oso", el coeficiente de correlación lineal es $r = 0.980$.

11. Considerando los datos del ejercicio 7 "Peso del auto y consumo de combustible", el coeficiente de correlación lineal es $r = -0.987$.

12. Considerando los datos del ejercicio 8 "Estaturas de padres e hijos", el coeficiente de correlación lineal es $r = -0.017$.

2-4 Más allá de lo básico

Valores de *P*. ***En los ejercicios 13 a 16, escriba un enunciado que interprete el valor P e incluya una conclusión acerca de la correlación lineal.***

13. Considerando los datos del ejercicio 5 "Volumen cerebral e IQ" el valor de P es 0.839.

14. Considerando los datos del ejercicio 6 "Mediciones de un oso" el valor de P es 0.000.

15. Considerando los datos del ejercicio 7 "Peso del auto y consumo de combustible", el valor de P es 0.000.

16. Usando los datos del ejercicio 8 "Estaturas de padres e hijos", el valor de P es 0.963.

Examen rápido del capítulo

1. Galletas Revise la distribución de frecuencias adjunta que resume el número de chispas de chocolate en cada galleta de una muestra de Chips Ahoy regulares (del conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B). ¿Cuál es la anchura de clase? ¿Es posible identificar los valores de los datos originales?

Chispas de chocolate	Frecuencia
18-20	6
21-23	11
24-26	18
27-29	4
30-32	1

2. Galletas Con base en la misma distribución de frecuencias del ejercicio 1, identifique las fronteras y los límites de la primera clase.

3. Galletas Con base en la misma distribución de frecuencias del ejercicio 1, ¿cuántas galletas se incluyen?

4. Galletas Se crea un diagrama de tallo y hojas para las mismas galletas resumidas en el ejercicio 1, y la primera fila de es diagrama es 1|99. Identifique los valores representados por esa fila del diagrama de tallo y hojas.

5. Computadoras Como gerente de control de calidad en Texas Instruments, usted encuentra que los defectos en las calculadoras tienen varias causas, incluyendo maquinaria desgastada, errores humanos, suministros incorrectos y maltratos durante el embalaje. ¿Cuál de las siguientes gráficas sería la mejor para describir las causas de los defectos: histograma, diagrama de dispersión, gráfica de Pareto, gráfica de puntos, gráfica circular?

6. Distribución de la riqueza En los últimos años, ha habido mucha discusión sobre la distribución de la riqueza entre los adultos de Estados Unidos. Si usted planea realizar una investigación original obteniendo de alguna manera el monto de la riqueza de 3000 adultos seleccionados al azar, ¿qué gráfica sería la mejor para ilustrar la *distribución* de la riqueza?

7. Ensayo de salud En una investigación de la relación entre las presiones arteriales sistólica y diastólica de las mujeres adultas, ¿cuál de las siguientes gráficas es más útil: histograma, gráfica circular, diagrama de dispersión, diagrama de tallo y hojas, gráfica de puntos?

8. Lotería En el juego de lotería Play 4 de Florida, cada día se seleccionan aleatoriamente cuatro dígitos entre 0 y 9 inclusive. Normalmente esperamos que cada uno de los 10 dígitos ocurra alrededor de 1/10 de las veces, y un análisis de los resultados del año pasado muestra que así sucedió. Debido a que los resultados son lo que normalmente esperamos, ¿es correcto decir que la distribución de los dígitos seleccionados es una distribución normal?

9. Cinturones de seguridad La compañía Cinturones de Seguridad Beams fabrica ... bueno, usted ya sabe. Cuando se examina una muestra de cinturones de seguridad en relación con su punto de ruptura (medida en kilogramos), se exploran los datos muestrales. Identifique la característica importante de los datos que falta en la siguiente lista: centro, distribución, valores atípicos, características cambiantes en el tiempo.

10. Cinturones de seguridad Se trazará un histograma a partir de los puntos de ruptura medidos (en libras) de los cinturones de seguridad para automóvil probados. Identifique dos características clave de un histograma de esos valores que sugieren que los datos tienen una *distribución normal*.

Ejercicios de repaso

1. Distribución de frecuencias de temperaturas corporales Elabore una distribución de frecuencias de las 20 temperaturas corporales (°F) que se listan a continuación. (Estos datos pertenecen al conjunto de datos 3 "Temperaturas corporales" del apéndice B.) Utilice una anchura de clase de 0.5 °F y un valor inicial de 97.0 °F.

97.1	97.2	97.5	97.6	97.6	97.8	98.0	98.0	98.2	98.2
98.2	98.3	98.4	98.6	98.6	98.7	98.7	98.9	99.1	99.4

2. Histograma de temperaturas corporales Elabore el histograma que corresponde a la distribución de frecuencias del ejercicio 1. Utilice los puntos medios de clase para la escala horizontal. ¿El histograma sugiere que los datos provienen de una población con distribución normal? ¿Por qué sí o por qué no?

3. Gráfica de puntos de temperaturas corporales Elabore una gráfica de puntos de las temperaturas corporales listadas en el ejercicio 1. ¿Cuál es la mejor manera de ilustrar la distribución de los datos: el histograma del ejercicio 2 o la gráfica de puntos?

4. Diagrama de tallo y hojas de temperaturas corporales Elabore un diagrama de tallo y hojas de las temperaturas corporales listadas en el ejercicio 1. ¿Existen valores atípicos?

5. Temperaturas corporales A continuación se listan las temperaturas de nueve varones medidos a las 8 AM y de nuevo a las 12 AM (del conjunto de datos 3 "Temperaturas corporales" en el apéndice B). Elabore un diagrama de dispersión y, con base en él, diga si parece haber una relación entre las temperaturas de las 8 AM y de las 12 AM.

8 AM	98.0	97.0	98.6	97.4	97.4	98.2	98.2	96.6	97.4
12 AM	98.0	97.6	98.8	98.0	98.8	98.8	97.6	98.6	98.6

6. Medio ambiente

a. Después de recopilar las temperaturas medias (promedio) globales en cada uno de los últimos 100 años, queremos trazar la gráfica más apropiada para estos datos. ¿Qué gráfica sería la mejor?

b. Después de recopilar las temperaturas medias (promedio) globales y la cantidad de emisiones de monóxido de carbono en los últimos 100 años, queremos trazar una gráfica para investigar la asociación entre esas dos variables. ¿Qué gráfica es la mejor?

c. Una investigación de las fuentes de monóxido de carbono incluye vehículos de motor, hornos, fuegos, centrales eléctricas que queman carbón y humo de tabaco. Si queremos trazar una gráfica que ilustre la importancia relativa de estas fuentes, ¿qué gráfica es la mejor?

7. Como que es hora de hacer este ejercicio En una encuesta Marista aplicada a adultos, las siguientes son las palabras o frases que los sujetos encuentran más molestas en una conversación (junto con sus frecuencias de respuesta): como que o como algo (127); nada más digo (81); ya sabes (104); como sea (219); obviamente (35). Elabore una gráfica circular. Identifique una desventaja de una gráfica circular.

8. Como sea Utilice los mismos datos del ejercicio 7 para trazar una gráfica de Pareto. ¿Qué gráfica ilustra de mejor manera los datos: la gráfica de Pareto o la gráfica circular?

Ejercicios de repaso acumulado

En los ejercicios 1 a 6, utilice los siguientes datos, que son los tiempos totales de juego en casa (horas) para todos los equipos de Grandes Ligas en un año reciente (con base en los datos de Baseball Prospectus).

236	237	238	239	241	241	242	245	245	245	246	247	247	248	248
249	250	250	250	251	252	252	253	253	258	258	258	260	262	264

1. Distribución de frecuencias Elabore una distribución de frecuencias. Utilice una anchura de clase de 5 horas y un tiempo de inicio de 235 horas.

2. Distribución de frecuencias Para la distribución de frecuencias del ejercicio 1, busque lo siguiente.

a. Límites de la primera clase

b. Fronteras de la primera clase

c. Punto medio de la primera clase

3. Histograma Elabore el histograma correspondiente a la distribución de frecuencias del ejercicio 1. Para los valores en el eje horizontal, utilice los puntos medios de clase. ¿Cuál de las siguientes expresiones describe de mejor manera la distribución: uniforme, normal, asimétrica a la izquierda, asimétrica a la derecha?

4. Gráfica engañosa Suponga que desea crear el histograma para el ejercicio 3 de manera que exagere las diferencias entre los tiempos. Describa cómo se puede modificar el histograma del ejercicio 3 para lograr esa exageración.

5. Diagrama de tallo y hojas Utilice los tiempos totales de juego para crear un diagrama de tallo y hojas. ¿Qué revela este diagrama sobre la distribución de los datos?

6. Tipo de datos

a. Los tiempos de juego listados se redondean al número entero más cercano. Antes de redondear, ¿los tiempos exactos de juego son datos discretos o datos continuos?

b. ¿Para los tiempos listados, los datos son categóricos o cuantitativos?

c. Identifique el nivel de medición de los tiempos listados: nominal, ordinal, de intervalo o de razón.

continúa

d. ¿Cuál de las siguientes opciones describe mejor los datos muestrales: muestra de respuesta voluntaria, muestra aleatoria, muestra de conveniencia, muestra simple?

e. Los tiempos de juego totales listados son de un año reciente, y se dispone de los datos de todos los años hasta 1950. Dado que los tiempos listados son parte de una colección más grande de tiempos, ¿los datos constituyen una muestra o una población?

Proyecto de tecnología

En este capítulo se estableció que los días de las gráficas encantadoras y primitivas dibujadas a mano habían quedado muy atrás, y la tecnología ahora proporciona herramientas poderosas para generar una gran variedad de gráficas. Por lo tanto, este proyecto sirve como una buena preparación para las presentaciones profesionales que inevitablemente se harán en el futuro.

Los conjuntos de datos completos del apéndice B se pueden descargar de www.pearsonenespañol.com/triola. Statdisk ya incluye los conjuntos de datos. Pueden abrirse mediante paquetes de software estadístico, como Statdisk. Minitab, Excel, SPSS y JMP. Utilice uno de estos paquetes para abrir el conjunto de datos 4 "Nacimientos". Utilice los métodos de este capítulo para explorar y comparar el peso al nacer de las niñas y de los niños. (Debido a que las unidades son gramos, en realidad son medidas de masa y no de peso).

- Obtenga una copia impresa de los dos histogramas.
- Describa las naturalezas de las dos distribuciones (uniforme, normal, asimétrica a la izquierda, asimétrica a la derecha), e identifique los posibles valores atípicos.
- Escriba una breve descripción de sus resultados.

Sugerencia: Los géneros están codificados como 1 para masculino y 0 para femenino, así que ordene todas las filas utilizando la columna de género como base para la clasificación. Después se pueden separar las filas de los niños y de las niñas, usando las funciones de copiar/pegar y cortar del software.

DE LOS DATOS A LA DECISIÓN

Tiempos de servicio en auto de los restaurantes de comida rápida: ¿Quién es el mejor?

El conjunto de datos 25 "Comida rápida" en el apéndice B incluye los tiempos de servicio para el almuerzo y la cena en los restaurantes McDonald's, Burger King, Wendy's y Dunkin' Donuts. Varios ejemplos y ejercicios en este capítulo usan algunos de esos tiempos de servicio. Para este proyecto, utilice todos los tiempos del conjunto.

Pensamiento crítico

Utilice los métodos de este capítulo para responder las siguientes preguntas.

1. ¿Cuál de los cuatro restaurantes parece tener el tiempo de servicio en auto más rápido para el almuerzo?

2. ¿Cuál de los cuatro restaurantes parece tener el tiempo de servicio en auto más rápido para la cena?

3. ¿Los tiempos de servicio en auto para el almuerzo parecen ser diferentes de los tiempos de servicio en auto para la cena? Explique.

4. De acuerdo con las opciones de menú disponibles en los diferentes restaurantes, ¿alguno de ellos tiene una ventaja inherente en relación con los tiempos de servicio? Explique.

5. Considerando las diferencias en las opciones de menú, ¿hay un restaurante que parezca ser más eficiente que los demás? Explique.

Actividades en equipo

1. Actividad fuera de la clase El problema del capítulo usa los tiempos de servicio medidos en el conjunto de datos 25 "Comida rápida" del apéndice B. Vaya a uno o más restaurantes de comida rápida y recopile sus propios tiempos de servicio. Compare los resultados con los encontrados en el conjunto de datos 25 del apéndice B.

2. Actividad en clase Utilice un paquete de galletas con chispas de chocolate; cada estudiante debe recibir dos o tres galletas. Proceda a contar el número de chispas de chocolate en cada galleta. No todas las chispas son visibles, por lo que deben utilizarse "pruebas destructivas" a través de un proceso que implica el consumo. Registre el número de chispas de chocolate para cada galleta y combine todos los resultados. Elabore una distribución de frecuencias, un histograma, una gráfica de puntos y un diagrama de tallo y hojas con los resultados. Dado que las galletas se hicieron a través de un proceso de producción en masa, podríamos esperar que el número de chispas de chocolate por galleta no varíe mucho. ¿Indican eso los resultados? Explique.

3. Actividad en clase En la clase, cada estudiante debe registrar dos frecuencias de pulso contando el número de sus latidos cardíacos en 1 minuto. El primer pulso se debe medir mientras el estudiante está sentado; el segundo, mientras está de pie. Utilizando las pulsaciones contadas mientras están sentados, elabore una distribución de frecuencias y un histograma para los pulsos de los varones, y luego otra distribución de frecuencias e histograma para los pulsos de las mujeres. Utilizando las pulsaciones medidas mientras están de pie, elabore una distribución de frecuencias y un histograma para los pulsos de los varones, y luego otra distribución de frecuencias e histograma para los pulsos de las mujeres. Compare los resultados. ¿Los hombres y las mujeres parecen tener pulsos diferentes? ¿Las pulsaciones contadas al estar sentados parecen ser diferentes de los pulsos mientras están de pie? Utilice una gráfica apropiada para determinar si existe una relación entre el pulso sentado y el pulso de pie.

4. Actividad fuera de la clase Busque periódicos y revistas para encontrar un ejemplo de una gráfica que sea engañosa. Describa por qué lo es. Vuelva a dibujar la gráfica para que represente la información correctamente.

5. Actividad fuera de la clase Encuentre la gráfica de Charles Joseph Minard que describe la marcha de Napoleón a Moscú y viceversa, y explique por qué Edward Tufte dice que "puede ser la mejor gráfica jamás dibujada". (Vea *The Visual Display of Quantitative Information* de Edward Tufte, Graphics Press).

6. Actividad fuera de la clase En *The Visual Display of Quantitative Information* de Edward Tufte (Graphics Press), encuentre la gráfica que apareció en *American Education*, y explique por qué Tufte dice que "esta puede ser la peor gráfica jamás impresa". Elabore una gráfica que resulte efectiva para representar los mismos datos.

3

DESCRIPCIÓN, EXPLORACIÓN Y COMPARACIÓN DE DATOS

PROBLEMA DEL CAPÍTULO

¿Qué compañía tiene la mejor velocidad (de transferencia) de datos para teléfonos inteligentes en los aeropuertos?

El conjunto de datos 32, “Velocidades de datos en aeropuertos” del apéndice B, lista las velocidades (de transferencia) de datos medidos por RootMetrics en 50 aeropuertos de Estados Unidos considerando las cuatro principales compañías proveedoras del servicio en ese país (Verizon, Sprint, AT&T y T-Mobile). Todas las velocidades se dan en unidades de megabits (o 1 millón de bits) por segundo, expresados como Mbps. Debido a que las velocidades de datos originales listadas en el conjunto de datos 32 incluyen números decimales como 38.5 Mbps, las gráficas de puntos con los datos no modificados serían un poco complicadas y no tan útiles, pero si redondeamos todos los conjuntos de datos originales, obtendremos la gráfica de puntos mostrada en la figura 3-1. (Una observación de la escala horizontal en la figura 3-1 revela que las velocidades de datos originales han sido redondeadas al entero par más cercano por el software utilizado para crear las gráficas de puntos). Si se utilizan las mismas cuatro escalas horizontales y se apilan las cuatro gráficas de puntos, las comparaciones resultan mucho más sencillas.

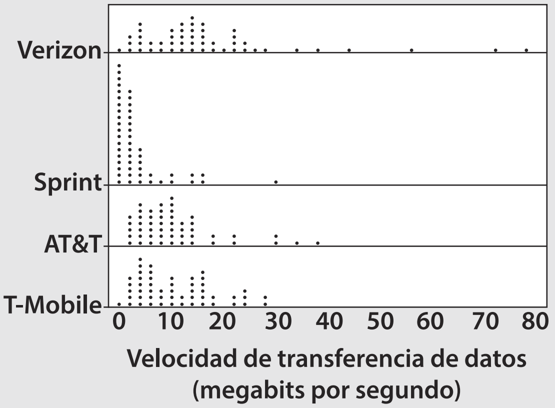


FIGURA 3-1 **Gráfica de puntos de las velocidades de datos para Smartphone**

El análisis de la figura 3-1 sugiere que Verizon tiene el mejor desempeño global, con velocidades de datos que tienden a ser más altas que las de las otras tres compañías. Pero en vez de confiar únicamente en interpretaciones subjetivas de una gráfica como la de la figura 3-1, este capítulo introduce medidas que son esenciales para cualquier estudio de estadística: la media, la mediana, la desviación estándar y la varianza, que se encuentran entre los datos estadísticos más importantes en el estudio de esta materia. Éstos se usarán para describir, explorar y comparar las velocidades de datos medidos de Verizon, Sprint, AT&T y T-Mobile como se listan en el conjunto de datos 32.

OBJETIVOS DEL CAPÍTULO

Pensamiento crítico e interpretación: más allá de las fórmulas y la aritmética

En este curso moderno de estadística no es tan importante memorizar fórmulas o efectuar cálculos manuales de aritmética. Es posible obtener resultados con una calculadora o software para que podamos concentrarnos en el sentido práctico de los resultados a través del pensamiento crítico. Aunque el presente capítulo incluye pasos detallados para procedimientos importantes, no siempre es necesario dominarlos. No obstante, a menudo es útil llevar a cabo unos cuantos cálculos manuales antes de usar la computadora, con el fin de incentivar la comprensión.

Los métodos y herramientas presentados en este capítulo se llaman con frecuencia métodos de **estadística descriptiva**, puesto que resumen o *describen* las características relevantes de los datos. En los capítulos subsecuentes usamos la **estadística inferencial** para hacer *inferencias* , o generalizaciones, sobre las poblaciones. Los siguientes son los objetivos del capítulo:

3-1 Medidas de tendencia central

- Desarrollar la capacidad de medir el centro de los datos mediante la determinación de la media, la mediana, la moda y la mitad del rango.
- Determinar si un valor atípico tiene un efecto sustancial sobre la media y la mediana.

3-2 Medidas de variación

- Desarrollar la capacidad de medir la variación en un conjunto de datos muestrales mediante la determinación de los valores del rango, la varianza y la desviación estándar.
- Desarrollar la capacidad de interpretar los valores de la desviación estándar aplicando la *regla práctica del rango* para determinar si un valor particular es *significativamente bajo* o *significativamente alto*.

3-3 Medidas de posición relativa y gráficas de caja

- Desarrollar la capacidad de calcular una puntuación z y utilizar el resultado para determinar si un valor dado x es *significativamente bajo* o *significativamente alto*.
- Identificar valores de los *percentiles* y *cuartiles* de un conjunto de datos.
- Desarrollar la capacidad de construir una gráfica de caja a partir de un conjunto de datos.

3-1 Medidas de tendencia central

Concepto clave El enfoque de esta sección está en obtener un valor que mida el *centro* de un conjunto de datos. En particular, se presentan medidas de tendencia central, incluyendo la *media* y la *mediana*. El objetivo aquí no es sólo encontrar el valor de cada medida de tendencia central, sino también interpretar esos valores. La parte 1 de la presente sección incluye conceptos básicos que deben ser entendidos antes de considerar la parte 2.

PARTE 1 Conceptos básicos de las medidas de tendencia central

En la parte 1 de esta sección introducimos la media, la mediana, la moda y la mitad del rango como diferentes medidas de tendencia central. Tales medidas se utilizan ampliamente para proporcionar valores representativos que "resumen" los conjuntos de datos.

En cifras

$3.19: Monto medio dejado por el ratón de los dientes, con base en una encuesta realizada por Visa. El 10% de los niños con menos suerte no obtiene nada.

DEFINICIÓN

Una **medida de tendencia central** es un valor en medio o en el centro de un conjunto de datos.

Existen diferentes métodos para medir el centro, por lo que tenemos distintas definiciones para ellos. Comenzamos con la media.

Media

Por lo general, la media (o media aritmética) es la más importante de las mediciones numéricas usadas para describir datos, y es lo que la mayoría de las personas llama *promedio*.

DEFINICIÓN

La **media** (o **media aritmética**) de un conjunto de datos es la medida de tendencia central que se encuentra al sumar todos los valores de los datos y dividir el total por el número de datos.

Propiedades importantes de la media

- Las medias muestrales de una misma población tienden a variar menos que otras medidas de tendencia central.
- La media de un conjunto de datos utiliza todos los valores de los datos.
- Una desventaja de la media es que un solo valor extremo (atípico) puede cambiar el valor de la media en forma sustancial. (Con base en la siguiente definición, puede decirse que la media no es *resistente*).

DEFINICIÓN

Un dato estadístico es **resistente** si la presencia de valores extremos (atípicos) no ocasiona que éste cambie mucho.

Cálculo y notación de la media

La definición de la media puede expresarse como en la fórmula 3-1, donde la letra griega Σ (sigma mayúscula) indica que los valores de los datos deben sumarse, por lo que Σx representa la suma de todos los valores de los datos. El símbolo n expresa el **tamaño de la muestra**, que es el número de valores de datos.

FÓRMULA 3-1

$$\text{Media} = \frac{\Sigma x \leftarrow \text{suma de todos los valores de datos}}{n \quad \leftarrow \text{número de valores de datos}}$$

Si los datos son una *muestra* de una población, la media se expresa con $\bar{x}$ (que se pronuncia "x barra"); si los datos son la población entera, la media se expresa mediante μ (letra griega mu minúscula).

NOTACIÓN *Sugerencia:* Los estadísticos muestrales se representan usualmente mediante caracteres latinos, como $\bar{x}$, y los parámetros de población por medio de letras griegas, como μ.

Σ	expresa la *suma* de un conjunto de valores de datos.
x	es la *variable* que generalmente se usa para representar los valores de datos individuales.
n	representa el número de valores de datos en una *muestra*.
N	representa el número de valores de datos en 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 en una *población*.

EJEMPLO 1 Media

El conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B incluye medidas de las velocidades de datos para teléfonos inteligentes en cuatro compañías. Encuentre la media de los cinco primeros datos de velocidad para Verizon: 38.5, 55.6, 22.4, 14.1 y 23.1 (todos en megabits por segundo, o Mbps).

SOLUCIÓN

La media se calcula mediante la fórmula 3-1. Primero sume los valores de datos, luego divida por el número de valores:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{38.5 + 55.6 + 22.4 + 14.1 + 23.1}{5} = \frac{153.7}{5}$$
$$= 30.74 \text{ Mbps}$$

La media de las cinco primeras velocidades de datos para Verizon es de 30.74 Mbps.

SU TURNO Encuentre la media en el ejercicio 5 "Números de jugadores de fútbol americano".

PRECAUCIÓN Nunca utilice el término *promedio* cuando se refiera a una medida de tendencia central. Esa palabra se utiliza a menudo para la media, pero en ocasiones se usa para otras medidas del centro. Los estadísticos no emplean el término *promedio* y en el resto de este libro no se usará para hacer referencia a una medida de tendencia central específica. El término *promedio* tampoco es utilizado por la comunidad estadística o las revistas profesionales.

Paradoja del tamaño de la clase

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

Lo que la mediana no es

El biólogo de Harvard, Stephen Jay Gould escribió: “La mediana no es el mensaje”. Así describe lo que aprendió cuando se enteró que tenía mesotelioma abdominal, una forma de cáncer. Fue a la biblioteca para aprender más y se sorprendió al descubrir que el mesotelioma era incurable (la mediana de supervivencia es de sólo ocho meses después de ser descubierto). Gould escribió lo siguiente: “Sospecho que la mayor parte de las personas, sin formación en estadística, leería esta afirmación como ‘probablemente estaré muerto en ocho meses’; conclusión que debe evitarse, ya que no es así, y más considerando que la actitud (en la lucha contra el cáncer) importa tanto”. Gould procedió a interpretar cuidadosamente el valor de la mediana. Sabía que su oportunidad de vivir más tiempo que la mediana era buena porque era joven, su cáncer fue diagnosticado de manera temprana y tendría el mejor tratamiento médico. También razonó que algunas personas podrían vivir mucho más de ocho meses, y no vio ninguna razón por la que no pudiera estar en ese grupo. Armado con esta interpretación reflexiva de la mediana y una fuerte actitud positiva, Gould vivió 20 años después de su diagnóstico. Murió de otro cáncer no relacionado con el mesotelioma.

Mediana

La mediana se puede considerar de manera general como un “valor medio” en el sentido de que aproximadamente la mitad de los valores en un conjunto de datos son menores y la mitad son mayores que la mediana. La siguiente definición es más precisa.

DEFINICIÓN

La **mediana** de un conjunto de datos es la medida de tendencia central que indica el *valor intermedio*, cuando los datos originales se presentan en orden de magnitud creciente (o decreciente).

Propiedades importantes de la mediana

- La mediana no cambia por mucho cuando se incluyen sólo unos pocos valores extremos, por lo que la mediana es una medida de tendencia central *resistente*.
- La mediana no utiliza directamente todos los valores de datos. (Por ejemplo, si el valor mayor se cambia por uno mucho más grande, la mediana no cambia).

Cálculo y notación de la mediana

La mediana de una muestra se denomina a veces $\tilde{x}$ (que se pronuncia “x tilde”) o M o Med; no existe una notación generalmente aceptada y no hay un símbolo especial para la mediana de una población. Para encontrar la mediana, primero *ordene* los valores (póngalos en orden) y luego siga uno de los siguientes dos procedimientos:

1. Si el número de valores de datos es *impar*, la mediana es el número ubicado en el intermedio exacto de la lista ordenada.
2. Si el número de valores de datos es *par*, la mediana se obtiene calculando la media de los dos números intermedios de la lista ordenada.

EJEMPLO 2 Mediana con un número *impar* de valores de datos

Encuentre la mediana de las cinco primeras velocidades de datos para Verizon: 38.5, 55.6, 22.4, 14.1 y 23.1 (todas en megabits por segundo, o Mbps).

SOLUCIÓN

Primero ordene en forma ascendente los valores de datos, como se muestra a continuación:

14.1 22.4 23.1 38.5 55.6

Como hay 5 valores de datos, tal cantidad es un número impar (5), por lo que la mediana es el número intermedio exacto de la lista ordenada, que es 23.1 Mbps. Por lo tanto, la mediana es 23.10 Mbps. Observe que la *mediana* de 23.10 Mbps es diferente de la *media* de 30.74 Mbps que se encontró en el ejemplo 1. Note también que el resultado de 23.10 Mbps sigue la regla de redondeo que se proporciona más adelante en esta sección.

SU TURNO Encuentre la mediana en el ejercicio 5 “Números de jugadores de fútbol americano”.

EJEMPLO 3 Mediana con un número *par* de valores de datos

Repita el ejemplo 2 después de incluir la sexta velocidad de datos de 24.5 Mbps. Es decir, encuentre la mediana de las siguientes velocidades de datos: 38.5, 55.6, 22.4, 14.1, 23.1, 24.5 (todo en Mbps).

SOLUCIÓN

Primero ordene los valores de manera ascendente:

14.1 22.4 23.1 24.5 38.5 55.6

Debido a que el número de valores de datos es par (6), la mediana se obtiene calculando la media de los dos números intermedios, que son 23.1 y 24.5.

$$\text{Mediana} = \frac{23.1 + 24.5}{2} = \frac{47.6}{2} = 23.80 \text{ Mbps}$$

La mediana es 23.80 Mbps.

SU TURNO Encuentre la mediana en el ejercicio 7 "Valor neto de celebridades".

Moda

La moda no se utiliza mucho con datos cuantitativos, pero es la única medida de tendencia central que puede usarse con datos cualitativos (que consisten solamente en nombres, etiquetas o categorías).

En cifras

Mohammed: El nombre más común en el mundo.

DEFINICIÓN

La **moda** de un conjunto de datos es el (los) valor(es) que ocurre(n) con mayor frecuencia.

Propiedades importantes de la moda

- La moda se puede encontrar con datos cualitativos.
- Un conjunto de datos puede tener una moda, o múltiples modas, o no tener ninguna.

Determinación de la moda: Un conjunto de datos puede tener una moda, más de una moda, o ninguna moda.

- Cuando dos valores de datos ocurren con la misma mayor frecuencia, cada uno es una moda y se dice que el conjunto de datos es **bimodal**.
- Cuando más de dos valores de datos ocurren con la misma mayor frecuencia, cada uno es una moda y se dice que el conjunto de datos es **multimodal**.
- Cuando ningún valor de datos se repite, se dice que **no hay moda**.
- Cuando usted pide helado con su pastel, se dice que está "a la moda".

EJEMPLO 4 Moda

Encuentre la moda de las siguientes velocidades de datos para Sprint (en Mbps):

0.2 0.3 0.3 0.3 0.6 0.6 1.2

SOLUCIÓN

La moda es de 0.3 Mbps, porque es la velocidad de datos que ocurre con más frecuencia (tres veces).

SU TURNO Encuentre la moda en el ejercicio 7 "Valor neto de celebridades".

En el ejemplo 4, la moda es un valor único. A continuación se presentan otras circunstancias posibles:

Dos modas: Las velocidades de datos (Mbps) de 0.3, 0.3, 0.6, 4.0 y 4.0 tienen dos modas: 0.3 Mbps y 4.0 Mbps.

Sin moda: Las velocidades de datos (Mbps) de 0.3, 1.1, 2.4, 4.0 y 5.0 no tienen moda porque ningún valor se repite.

Mitad del rango

Otra medida de tendencia central es la mitad del rango.

DEFINICIÓN

La **mitad del rango** de un conjunto de datos es la medida de tendencia central que consiste en el valor que está a la mitad entre los valores máximo y mínimo del conjunto de datos original. Se encuentra al sumar el valor máximo y el valor mínimo de los datos y después dividir esa suma entre 2, como se muestra en la siguiente fórmula:

$$\text{Mitad del rango} = \frac{\text{valor máximo de datos} + \text{valor mínimo de datos}}{2}$$

Propiedades importantes de la mitad del rango

- Debido a que la mitad del rango utiliza sólo los valores máximo y mínimo, es muy sensible a esos extremos y, por lo tanto, *no es resistente*.
- En la práctica, la mitad del rango se utiliza con poca frecuencia, pero tiene tres características redentoras:

1. Es muy fácil de calcular.
2. Ayuda a reforzar la muy importante idea de que hay varias maneras de definir el centro de un conjunto de datos.
3. En ocasiones, su valor se utiliza incorrectamente para la mediana, así que la confusión puede reducirse al definir claramente la mitad del rango junto con la mediana.

EJEMPLO 5 **Mitad del rango**

Encuentre la mitad del rango de las siguientes velocidades de datos para Verizon del ejemplo 1: 38.5, 55.6, 22.4, 14.1 y 23.1 (todas en Mbps)

SOLUCIÓN

La mitad del rango se encuentra de la siguiente manera:

$$\text{Mitad del rango} = \frac{\text{valor máximo de datos} + \text{valor mínimo de datos}}{2}$$

$$= \frac{55.6 + 14.1}{2} = 34.85 \text{ Mbps}$$

La mitad del rango es 34.85 Mbps.

SU TURNO Encuentre la mitad del rango en el ejercicio 5 “Números de jugadores de fútbol americano”.

Redondeo de medidas de tendencia central

A menudo, cuando se calculan medidas de tendencia central, es necesario redondear el resultado. Se utiliza la siguiente regla.

Reglas de redondeo para medidas de tendencia central

- **Para la media, la mediana y la mitad del rango, incluya un decimal más que los presentes en el conjunto original de valores.**
- **Para la moda, deje el valor sin redondear** (porque los valores de la moda son iguales que algunos de los valores de los datos originales).

Cuando aplique las reglas de redondeo, redondee sólo la respuesta final, *no los valores intermedios producidos durante los cálculos*. Por ejemplo, la media de 2, 3 y 5 es 3.333333..., que se redondea a 3.3; esto es, con un decimal más que los valores originales de 2, 3 y 5. Otro ejemplo es: la media de 80.4 y 80.6 es 80.50 (un decimal más que los usados para los valores originales). Debido a que la moda es uno o más de los valores originales, no la redondeamos; simplemente usamos los mismos valores originales que resultaron ser modas.

Pensamiento crítico

Siempre es posible calcular las medidas de tendencia central a partir de una muestra de números, pero es necesario considerar si hacerlo tiene sentido. En la sección 1-2 se estableció que no tiene sentido hacer cálculos numéricos con datos al nivel nominal de medición, ya que estos datos constan sólo de nombres, etiquetas o categorías, por lo que los datos estadísticos como la media y la mediana carecen de significado. También se debe pensar en el método de muestreo utilizado para recopilar los datos. Si el método de muestreo no es sólido, los datos estadísticos obtenidos pueden ser muy engañosos.

EJEMPLO 6 Pensamiento crítico y medidas de tendencia central

Considere cada una de las siguientes situaciones ilustrativas en las que la media y la mediana *no* son estadísticos significativos.

a. Códigos postales del Gateway Arch en San Luis, la Casa Blanca, la División de la Fuerza Aérea del Pentágono, el Edificio Empire State y la Estatua de la Libertad: 63102, 20500, 20330, 10118, 10004. (Los códigos postales no miden ni cuentan nada. Los números son sólo etiquetas para ubicaciones geográficas).

b. Clasificaciones de universidades nacionales seleccionadas: Harvard, Yale, Duke, Dartmouth y Brown (por *US News & World Report*): 2, 3, 7, 10, 14. (Las clasificaciones reflejan una ordenación, pero no miden ni cuentan nada).

c. Números en las camisetas de la defensiva titular de los Halcones marinos de Seattle cuando ganaron el Súper Tazón XLVIII: 31, 28, 41, 56, 25, 54, 69, 50, 91, 72, 29. (Los números en las camisetas de fútbol americano no miden ni cuentan nada, son sólo sustitutos de los nombres).

d. Primeros cinco ingresos de directores de empresa (en millones de dólares): 131.2, 66.7, 64.4, 53.3, 51.5. (Estas listas de "los primeros 5" o "los primeros 10" incluyen datos que no son en absoluto representativos de toda la población).

e. Las 50 edades medias calculadas a partir de las medias en cada uno de los 50 estados. (Si se calcula la media de esos 50 valores, el resultado no es la edad media de las personas en todo el territorio de Estados Unidos y se debe tener en cuenta los tamaños de la población de los 50 estados, como se describe en la introducción de la *media ponderada* en la parte 2 de esta sección).

SU TURNO Para el ejercicio 5 "Números de jugadores de fútbol americano", determine por qué la media y la mediana no son significativas.

Con la idea de describir, explorar y comparar datos, proporcionamos la tabla 3-1 en la página 88, que resume las diferentes medidas de tendencia central para las velocidades de datos para teléfonos inteligentes que se mencionan en el problema del capítulo. Los datos se listan en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B. La figura 3-1 de la página 81 sugiere que Verizon tiene las mayores velocidades y la comparación de medias y medianas en la tabla 3-1 también lo indica. En los siguientes capítulos se describen otras herramientas que pueden utilizarse para una comparación efectiva.

Un error de redondeo cambia un récord mundial

Los errores de redondeo a menudo pueden tener resultados desastrosos. Justin Gatlin estaba eufórico cuando estableció el récord mundial como la persona en correr los 100 metros en el menor tiempo (9.76 segundos). Sin embargo, su tiempo récord duró sólo cinco días, cuando se corrigió a 9.77 segundos y empató el récord mundial en lugar de romperlo. Su tiempo real fue de 9.766 segundos, y debería haberse redondeado a 9.77 segundos, pero la persona que tomó el tiempo no sabía que tenía que presionar un botón para obtener el redondeo. El representante de Gatlin dijo que el atleta estaba muy perturbado y que el incidente era una "gran vergüenza para la IAAF (International Association of Athletics Federations) y para nuestro deporte".

TABLA 3-1 Comparación de las velocidades de datos para teléfonos inteligentes (Mbps) en los aeropuertos

	Verizon	Sprint	AT&T	T-Mobile
Media	17.60	3.71	10.70	10.99
Mediana	13.90	1.60	8.65	9.70
Moda	4.5, 11.1	0.3	2.7	3.2, 4.4, 5.1, 13.3, 15.0, 16.7, 27.3
Mitad del rango	39.30	15.30	19.80	14.00

PARTE 2 Más allá de lo básico en las medidas de tendencia central

Cálculo de la media a partir de una distribución de frecuencias

La fórmula 3-2 es el mismo cálculo para la media que se presentó en la parte 1, pero incorpora el siguiente método: cuando trabajamos con datos resumidos en una distribución de frecuencias, hacemos posibles los cálculos asumiendo que todos los valores muestrales en cada clase son iguales al punto medio de dicha clase. La fórmula 3-2 no es realmente un concepto nuevo; es simplemente una variación de la fórmula 3-1 (media).

FÓRMULA 3-2 MEDIA A PARTIR DE UNA DISTRIBUCION DE FRECUENCIAS

Primero multiplique cada frecuencia y el punto medio de la clase; luego sume los productos.

$$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f} \quad \text{(El resultado es una aproximación)}$$

Suma de frecuencias (igual a n)

El ejemplo 7 ilustra el procedimiento para encontrar la media a partir de una distribución de frecuencias.

EJEMPLO 7 Cálculo de la media a partir de una distribución de frecuencias

Las dos primeras columnas de la tabla 3-2 mostradas aquí son las mismas de la distribución de frecuencias de la tabla 2-2 del capítulo 2. Utilice la distribución de frecuencias en las dos primeras columnas de la tabla 3-2 para encontrar la media.

TABLA 3-2 Tiempos de servicio para el almuerzo en McDonald's

Tiempo (segundos)	Frecuencia f	Punto medio de la clase x	$f \cdot x$
75-124	11	99.5	1094.5
125-174	24	149.5	3588.0
175-224	10	199.5	1995.0
225-274	3	249.5	748.5
275-324	2	299.5	599.0
Totales:	$\Sigma f = 50$		$\Sigma(f \cdot x) = 8025.0$

SOLUCIÓN

Recuerde que al trabajar con datos resumidos en una distribución de frecuencias, se hacen posibles los cálculos asumiendo que todos los valores muestrales en cada clase son iguales al punto medio de dicha clase. Por ejemplo, considere el intervalo de la primera clase de 75 a 124 con una frecuencia de 11. Asumimos que cada uno de los 11 tiempos de servicio es de 99.5 segundos (el punto medio de la clase). Con el tiempo de servicio de 99.5 segundos repetido 11 veces, tenemos un total de $99.5 \cdot 11 = 1094.5$, como se muestra en la última columna de la tabla 3-2. Después sumamos los resultados para encontrar la sumatoria de todos los valores muestrales.

La fila inferior de la tabla 3-2 muestra los dos componentes que se requieren para calcular la media (como en la fórmula 3-2) $\Sigma f = 50$ y $\Sigma(f \cdot x) = 8025.0$. Calculamos la media usando la fórmula 3-2 como sigue:

$$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f} = \frac{8025.0}{50} = 160.5 \text{ segundos}$$

El resultado de $\bar{x} = 160.5$ segundos es una *aproximación* porque se basa en el uso de los valores de los puntos medios de las clases en lugar de la lista original de tiempos de servicio. La media de 160.2 segundos encontrada mediante el uso de todos los tiempos de servicio originales es un resultado más preciso.

SU TURNO Resuelva el ejercicio 29 “Edades de las mejores actrices”.

Cálculo de una media ponderada

Cuando a los diferentes valores de datos x se les asignan pesos w distintos, podemos calcular la media ponderada, que está dada por la fórmula 3-3.

FÓRMULA 3-3

$$\text{Media ponderada: } \bar{x} = \frac{\Sigma(w \cdot x)}{\Sigma w}$$

La fórmula 3-3 implica primero multiplicar cada peso w por el valor correspondiente x, luego sumar los productos, y finalmente dividir el total por la suma de los pesos, Σw.

EJEMPLO 8 Cálculo del promedio de calificaciones

En su primer semestre de la universidad, una alumna del autor tomó cinco cursos. Sus calificaciones finales, junto con el número de créditos para cada curso, fueron A (3 créditos), A (4 créditos), B (3 créditos), C (3 créditos) y F (1 crédito). El sistema de clasificación asigna puntos de calidad a las calificaciones con letras de la siguiente manera: A = 4; B = 3; C = 2; D = 1; F = 0. Calcule su promedio de calificaciones.

SOLUCIÓN

Use los números de créditos como pesos: $w = 3, 4, 3, 3, 1$. Reemplace las calificaciones con letras de A, A, B, C y F con los puntos de calidad correspondientes: $x = 4, 4, 3, 2, 0$. Ahora usamos la fórmula 3-3 como se muestra a continuación. El resultado es un promedio de 3.07 en el primer semestre. (Al utilizar la regla de redondeo precedente, el resultado debe redondearse a 3.1, pero es común que los promedios de calificaciones se redondeen a dos decimales).

$$\begin{aligned}\bar{x} &= \frac{\Sigma(w \cdot x)}{\Sigma w} \\ &= \frac{(3 \times 4) + (4 \times 4) + (3 \times 3) + (3 \times 2) + (1 \times 0)}{3 + 4 + 3 + 3 + 1} \\ &= \frac{43}{14} = 3.07\end{aligned}$$

SU TURNO Resuelva el ejercicio 33 “Media ponderada”.

CENTRO DE TECNOLOGÍA

Ejemplos de pantallas de estadística descriptiva

Acceda a los complementos de software, videos y conjuntos de datos en www.pearsonenespañol.com/triola

Las siguientes pantallas se basan en las velocidades de datos de Verizon dentro del conjunto de datos 32 “Velocidades de datos en aeropuertos”.

Statdisk

Explore Data - Column 2

Sample Size, n: 50
Mean: 17.598
Median: 13.9
Midrange: 39.3
RMS: 23.6877
Variance, s^2: 256.5484
Standard Deviation, s: 16.01713
Mean Absolute Deviation: 10.66528
Range: 77
Coefficient of Variance: 91.02%

Minimum: 0.8
1st Quartile: 7.9
2nd Quartile: 13.9
3rd Quartile: 21.5
Maximum: 77.8

Sum: 879.9
Sum of Squares: 28055.35

95% CI for the Mean:
13.046 < mean <22.15

95% CI for the Standard Deviation:
13.3796 < SD < 19.9595

95% CI for the Variance:
179.0149 < VAR < 398.3804

TI-83/4 Plus

NORMAL FLOAT AUTO REAL RADIAN MP

1-Var Stats
$\bar{x}$=17.598
Σx=879.9
Σx²=28055.35
Sx=16.01712719
σx=15.85614695
n=50
minX=.8
↓Q₁=7.9

NORMAL FLOAT AUTO REAL RADIAN MP

1-Var Stats
↑Sx=16.01712719
σx=15.85614695
n=50
minX=.8
Q₁=7.9
Med=13.9
Q₃=21.5
maxX=77.8

Minitab

Descriptive Statistics: VERIZON

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
VERIZON	50	0	17.60	2.27	16.02	0.80	7.85	13.90	21.68	77.80

StatCrunch

Summary statistics:

Column	n	Mean	Variance	Std. dev.	Std. err.	Median	Range	Min	Max	Q1	Q3
VERIZON	50	17.598	256.54836	16.017127	2.2651638	13.9	77	0.8	77.8	7.9	21.5

Herramienta de análisis de datos de Excel

VERIZON	
Mean	17.598
Standard Error	2.26516385
Median	13.9
Mode	11.1
Standard Deviation	16.01712719
Sample Variance	256.5483633
Kurtosis	5.485898877
Skewness	2.193633453
Range	77
Minimum	0.8
Maximum	77.8
Sum	879.9
Count	50

Suplemento XLSTAT Excel

Statistic	VERIZON
Nbr. of observations	50
Minimum	0.800
Maximum	77.800
Freq. of minimum	1
Freq. of maximum	1
Range	77.000
1st Quartile	8.400
Median	13.900
3rd Quartile	21.400
Sum	879.900
Mean	17.598
Variance (n)	251.417
Variance (n-1)	256.548
Standard deviation (n)	15.856
Standard deviation (n-1)	16.017

CENTRO DE TECNOLOGÍA

Estadística descriptiva

Acceda a los complementos tecnológicos, videos y conjuntos de datos de Elementary Statistic, 13E en www.pearsonenespañol.com/triola

Statdisk

1. Haga clic en **Data** en el menú superior.
2. Seleccione **Explore Data-Descriptive Statistics** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Evaluate** para ver los estadísticos descriptivos.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y después elija **Display Descriptive Statistics**.
3. Haga doble clic en la columna de datos deseada para que aparezca en la ventana *Variables*.
4. Haga clic en **OK** para ver los estadísticos descriptivos.

SUGERENCIA: Haga clic en el botón **Statistics** arriba de **OK** para seleccionar los estadísticos individuales que desea mostrar.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Summary Stats** en el menú desplegable y después elija **Columns**.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Compute!** para ver los estadísticos descriptivos.

SUGERENCIA: Personalice los estadísticos descriptivos al seleccionar los elementos bajo **Statistics**.

Calculadora T1-83/84 Plus

1. Presione STAT, luego seleccione **CALC** en el menú superior.
2. Seleccione **1-Var Stats** y pulse ENTER.
3. Introduzca el nombre de la lista que incluye los datos deseados (por ejemplo, L1).
4. Seleccione **Calculate** y pulse ENTER para ver los estadísticos descriptivos.

SUGERENCIA: Presione para ver los estadísticos adicionales que no se desplegaron en la pantalla inicial.

Excel

Complemento XLSTAT

1. Haga clic en la ficha **XLSTAT** de la barra de opciones y después haga clic en **Describing Data**.
2. Seleccione **Descriptive Statistics** en el menú desplegable.
3. Marque la casilla **Quantitative Data** e introduzca el intervalo de datos deseado. Si la primera fila de datos contiene una etiqueta, también marque la casilla **Sample labels**.
4. Haga clic en **OK** para ver los estadísticos descriptivos.

Complemento Excel Data Analysis

1. Haga clic en la ficha **Data** de la barra de opciones y después seleccione **Data Analysis** en el menú superior.
2. Seleccione **Descriptive Statistics** bajo *Analysis Tools*.
3. Introduzca el rango de datos deseado para el **Inpute Range**. Si la primera fila de datos contiene una etiqueta, también marque la casilla **Labels in First Row**.
4. Marque la etiqueta **Summary Statistics** y haga clic en **OK** para ver los estadísticos descriptivos.

3-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Promedio El extinto sitio web IncomeTaxlist.com indicaba que el ingreso anual “promedio” en Florida era de $35,031. ¿Cuál es el papel del término *promedio* en estadística? ¿Debería usarse otro término en lugar de *promedio*?

2. ¿Qué hay de erróneo? *USA Today* publicó una lista consistente en el impuesto estatal por cada galón de gasolina. Si se suman las 50 cantidades de impuestos estatales y luego se divide la suma por 50, se obtiene 27.3 centavos. ¿Es el valor de 27.3 centavos el importe medio del impuesto estatal sobre ventas pagado por todos los conductores estadounidenses? ¿Por qué sí o por qué no?

3. Medidas de tendencia central ¿En qué sentido la media, la mediana, la moda y la mitad del rango son medidas del “centro”?

4. Medidas resistentes Las siguientes son cuatro de las velocidades de datos Verizon (Mbps) de la figura 3-1: 13.5, 10.2, 21.1, 15.1. Encuentre la media y la mediana de estos cuatro valores. A continuación, determine la media y la mediana después de incluir un quinto valor de 142, que es un valor atípico. (Una de las velocidades de datos de Verizon es 14.2 Mbps, pero 142 se utiliza aquí como un error resultante de una entrada con un punto decimal faltante). Compare los dos conjuntos de resultados. ¿Qué tanto se vio afectada la media por la inclusión del valor atípico? ¿Cuánto fue afectada la mediana por la inclusión del valor atípico?

Pensamiento crítico. ***En los ejercicios 5 a 20, tenga cuidado con las pequeñas trampas. Cada uno de estos ejercicios implica alguna característica capciosa. Encuentre la (a) media (b) mediana, (c) moda, (d) mitad del rango, y luego responda la pregunta dada.***

5. Números de jugadores de fútbol americano A continuación se listan los números de camiseta de 11 jugadores seleccionados al azar del equipo de los Halcones marinos de Seattle, cuando ganaron el Súper Tazón XLVIII. ¿Qué nos dicen los resultados?

89 91 55 7 20 99 25 81 19 82 60

6. Pesos de jugadores de fútbol americano A continuación se listan los pesos en libras de 11 jugadores seleccionados al azar del equipo de los Halcones marinos de Seattle, cuando ganaron el Súper Tazón XLVIII (los mismos jugadores del ejercicio anterior). ¿Es probable que los resultados sean representativos de todos los jugadores de la NFL?

189 254 235 225 190 305 195 202 190 252 305

7. Valor neto de celebridades A continuación se listan los mayores valores netos (en millones de dólares) de las celebridades. Las celebridades son Tom Cruise, Will Smith, Robert De Niro, Drew Carey, George Clooney, John Travolta, Samuel L. Jackson, Larry King, Demi Moore y Bruce Willis. ¿Qué nos dicen los resultados sobre la población de todas las celebridades? Sobre la base de la naturaleza de las cantidades, ¿qué se puede inferir acerca de su precisión?

250 200 185 165 160 160 150 150 150 150

8. Lo que pasa en Las Vegas... A continuación se listan los precios en dólares por una noche en diferentes hoteles ubicados en Las Vegas Boulevard (the "Strip"). Si usted decide hospedarse en uno de estos hoteles, ¿qué estadístico es más relevante, además de las medidas de tendencia central? Aparte del precio, identifique otro factor importante que afectaría su elección.

212 77 121 104 153 264 195 244

9. Huracanes A continuación se listan las cantidades de huracanes que se produjeron en el Atlántico cada año. Los datos se dan en orden anual, a partir del año 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de tendencia central?

8 9 8 7 9 15 5 6 8 4 12 7 8 2

10. Chícharos en una vaina Los biólogos han realizado experimentos para determinar si una falta de bióxido de carbono en el suelo afecta a los fenotipos de los chícharos. A continuación se listan los códigos del fenotipo, donde 1 = amarillo liso, 2 = verde liso, 3 = amarillo corrugado y 4 = verde corrugado. ¿Se pueden obtener las medidas de tendencia central para estos valores? ¿Los resultados tienen sentido?

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

11. Precios de televisores A continuación se listan los precios de venta (en dólares) de televisores de 60 pulgadas o más, los cuales fueron calificados como "las mejores compras" por la revista *Consumer Reports*. ¿Los datos estadísticos resultantes son representativos de la población de todos los televisores de 60 pulgadas o más? Si usted decide comprar uno de estos televisores, ¿qué dato estadístico es más relevante, además de las medidas de tendencia central?

1800 1500 1200 1500 1400 1600 1500 950 1600 1150 1500 1750

12. Radiación del teléfono celular A continuación se listan las tasas de absorción de radiación medidas (en W/kg) correspondientes a los siguientes teléfonos celulares: iPhone 5S, BlackBerry Z30, Sanyo Vero, Optimus V, Droid Razr, Nokia N97, Samsung Vibrant, Sony Z750a, Kyocera Kona, LG G2 y Virgin Mobile Supreme. Los datos provienen de la Comisión Federal de Comunicaciones (FCC, por sus siglas en inglés). Los medios de comunicación a menudo informan sobre los peligros de la radiación de los teléfonos celulares como una causa de cáncer. La FCC tiene un estándar de que la tasa de absorción de un teléfono celular debe ser de 1.6 W/kg o menos. Si usted está planeando comprar un teléfono celular, ¿es alguna de las medidas de tendencia central el estadístico más importante? ¿Hay otro dato estadístico que sea más relevante? Si es así, ¿cuál?

1.18 1.41 1.49 1.04 1.45 0.74 0.89 1.42 1.45 0.51 1.38

13. Cafeína en bebidas A continuación se indican las cantidades medidas de cafeína (mg por 12 onzas de bebida) obtenidas en una lata de cada una de 20 marcas (7-UP, A & W Root Beer, Cherry Coke, ...). ¿Son los estadísticos representativos de la población de todas las latas de las mismas 20 marcas consumidas por los estadounidenses?

0 0 34 34 34 45 41 51 55 36 47 41 0 0 53 54 38 0 41 47

14. Muertes de bomberos A continuación se listan las cantidades de heroicos bomberos que perdieron sus vidas en Estados Unidos anualmente mientras combatían incendios forestales. Las cantidades están ordenadas por año, comenzando en 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de tendencia central?

20 18 23 30 20 12 24 9 25 15 8 11 15 34

15. Longitud de los pies A continuación se listan las longitudes en pulgadas de los pies de mujeres del ejército seleccionadas al azar, medidos en el sondeo antropométrico de 1988 (ANSUR, abreviatura en inglés). Estos datos estadísticos, ¿son representativos de la población actual de todas las mujeres del ejército?

10.4 9.3 9.1 9.3 10.0 9.4 8.6 9.8 9.9 9.1 9.1

16. Universidades más caras A continuación se listan los costos anuales (en dólares) de matrícula y colegiaturas en las 10 universidades más caras de Estados Unidos para un año reciente (con base en datos de *US News & World Report*). Las universidades listadas en orden son Columbia, Vassar, Trinidad, George Washington, Carnegie Mellon, Wesleyan, Tulane, Bucknell, Oberlin y Union. ¿Qué nos dice esta "lista de las 10 primeras" sobre esos costos para la población de todas las universidades estadounidenses?

49,138 47,890 47,510 47,343 46,962 46,944 46,930 46,902 46,870 46,785

17. Anillo de diamante A continuación se listan las cantidades en dólares que cuestan los paquetes para propuestas matrimoniales en los diferentes estadios de las Ligas Mayores de Béisbol. Cinco de los equipos no permiten propuestas. ¿Existen valores atípicos?

39	50	50	50	55	55	75	85	100	115	175	175	200
209	250	250	350	400	450	500	500	500	500	1500	2500	

18. Ventas de álbumes de discos LP de vinilo A continuación se listan las ventas anuales de discos de vinilo en Estados Unidos (millones de unidades). Las cantidades de álbumes vendidos se presentan en orden cronológico, y la última entrada representa el año más reciente. ¿Las medidas de tendencia central nos dan alguna información sobre una tendencia cambiante a lo largo del tiempo?

0.3	0.6	0.8	1.1	1.1	1.4	1.4	1.5	1.2	1.3	1.4	1.2	0.9	0.9
1	1.9	2.5	2.8	3.9	4.6	6.1							

19. Fumadores de California En la encuesta Entrevista sobre Salud en California, se entrevistó a adultos seleccionados al azar. Una de las preguntas fue acerca de cuántos cigarrillos fumaban al día, y los resultados se listan a continuación para 50 encuestados seleccionados aleatoriamente. ¿Qué tan bien reflejan los resultados los hábitos del tabaquismo de los adultos en California?

9	10	10	20	40	50	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0										

20. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a las mujeres que evaluaran el atractivo de sus acompañantes masculinos; a continuación se lista una muestra de los resultados (1 = no atractivo, 10 = extremadamente atractivo). ¿Se pueden utilizar los resultados para describir el atractivo de la población de varones adultos?

5 8 3 8 6 10 3 7 9 8 5 5 6 8 8 7 3 5 5 6 8 7 8 8 8 7

En los ejercicios 21 a 24, encuentre la media y la mediana para cada una de las dos muestras, luego compare los dos conjuntos de resultados.

21. Presión arterial Una muestra de mediciones de presión arterial se toma del conjunto de datos 1 "Datos corporales" del apéndice B, y los valores (mm Hg) se relacionan de manera que 10 sujetos tienen medidas sistólicas y diastólicas. (La presión sistólica es una medida de la fuerza de la sangre empujada a través de las arterias, mientras que la presión diastólica es una medida de la presión arterial mientras el corazón está en reposo entre latidos). ¿Son las medidas de tendencia central los mejores estadísticos que pueden obtenerse de estos datos?

Sistólica:	118	128	158	96	156	122	116	136	126	120
Diastólica:	80	76	74	52	90	88	58	64	72	82

22. Robo de parquímetros A continuación se listan los montos (en millones de dólares) recaudados en parquímetros por Brinks y otras empresas en la ciudad de Nueva York durante períodos similares. Se utilizó un conjunto de datos más amplio para condenar a cinco empleados de Brinks por hurto mayor. Los datos fueron proporcionados por el abogado de la ciudad de Nueva York, y se pueden encontrar en el sitio web de *Data and Story Library* (DASL). ¿Los datos limitados que aparecen aquí muestran evidencia de robo por los empleados de Brinks?

El contratista de recaudación fue Brinks	1.3	1.5	1.3	1.5	1.4	1.7	1.8	1.7	1.7	1.6
El contratista de recaudación no fue Brinks	2.2	1.9	1.5	1.6	1.5	1.7	1.9	1.6	1.6	1.8

23. Pulsos A continuación se listan los pulsos (latidos por minuto) de muestras de hombres y mujeres adultas (del conjunto de datos 1 "Datos corporales" en el apéndice B). ¿Parece haber una diferencia?

Hombres:	86	72	64	72	72	54	66	56	80	72	64	64	96	58	66
Mujeres:	64	84	82	70	74	86	90	88	90	90	94	68	90	82	80

24. Filas en el banco Los tiempos de espera (en segundos) de los clientes en el Banco de Ahorro de Madison se registran con dos configuraciones: línea de clientes única y líneas de clientes individuales. Examine cuidadosamente los datos para determinar si hay una diferencia entre los dos conjuntos de datos que no sea evidente a partir de una comparación de las medidas de tendencia central. Si es así, ¿cuál es?

Línea única	390	396	402	408	426	438	444	462	462	462
Líneas individuales	252	324	348	372	402	462	462	510	558	600

Grandes conjuntos de datos del apéndice B. ***En los ejercicios 25 a 28, considere el conjunto de datos indicado en el apéndice B. Use software o una calculadora para encontrar las medias y las medianas.***

25. Tornados Utilice las medidas en la escala F de tornados listadas en el conjunto de datos 22 "Tornados" del apéndice B. Entre los 500 tornados, ¿a cuántos les faltan mediciones de la escala F? (*Precaución:* En algunas tecnologías, los datos faltantes se representan mediante una constante como –9 o 9999).

26. Terremotos Utilice las magnitudes (en la escala de Richter) de los 600 terremotos listados en el conjunto de datos 21 "Terremotos" del apéndice B. En 1989, el área de la Bahía de San Francisco fue golpeada por un terremoto de 7.0 en la escala de Richter. Ese terremoto ocurrió durante el período de calentamiento para el tercer juego de la Serie Mundial de béisbol. ¿Es la magnitud del terremoto de la Serie Mundial un valor atípico cuando se considera en el contexto de los datos muestrales dados en el conjunto de datos 21? Explique.

27. Temperaturas corporales Considere el conjunto de datos 3 "Temperaturas corporales" en el apéndice B y use las temperaturas corporales para las 12:00 AM del día 2. ¿Los resultados respaldan o contradicen la creencia común de que la temperatura corporal media es de 98.6 °F?

28. Nacimientos Use los pesos al nacer (en gramos) de los 400 bebés listados en el conjunto de datos 4 "Nacimientos" en el apéndice B. Examine la lista de pesos al nacer y haga una observación sobre esos números. ¿Cómo afecta esta observación la forma en que los resultados deben redondearse?

En los ejercicios 29 a 32, encuentre la media de los datos resumidos en la distribución de frecuencias. Además compare las medias calculadas con las medias reales obtenidas utilizando la lista original de valores de datos, que son las siguientes: (Ejercicio 29) 36.2 años; (Ejercicio 30) 44.1 años; (Ejercicio 31) 224.3; (Ejercicio 32) 255.1.

29.

Edad (en años) de la mejor actriz al ganar el Oscar	Frecuencia
20-29	29
30-39	34
40-49	14
50-59	3
60-69	5
70-79	1
80-89	1

30.

Edad (en años) del mejor actor al ganar el Oscar	Frecuencia
20-29	1
30-39	28
40-49	36
50-59	15
60-69	6
70-79	1

31.

Conteo de plaquetas en la sangre de hombres (1000 células/μl)	Frecuencia
0-99	1
100-199	51
200-299	90
300-399	10
400-499	0
500-599	0
600-699	1

32.

Conteo de plaquetas en la sangre de mujeres (1000 células/μl)	Frecuencia
100-199	25
200-299	92
300-399	28
400-499	0
500-599	2

33. Media ponderada Un alumno del autor obtuvo calificaciones de A, C, B, A y D. Estos cursos tenían las siguientes cantidades correspondientes de horas de crédito: 3, 3, 3, 4 y 1. El sistema de calificación asigna puntos a las calificaciones con letras como sigue: A = 4; B = 3; C = 2; D = 1; F = 0. Calcule el promedio de calificaciones y redondee el resultado con dos decimales. Si la lista del decano requiere un promedio de 3.00 o mayor, ¿este estudiante entrará a la lista del decano?

34. Media ponderada Una alumna del autor obtuvo calificaciones de 63, 91, 88, 84 y 79 en sus cinco exámenes regulares de estadística. Su calificación en el examen final fue 86 y en su proyecto de clase obtuvo 90. Su calificación combinada en las tareas fue 70. Los cinco exámenes regulares representan 60% de la calificación final, el examen final 10%, el proyecto 15% y las tareas 15%. ¿Cuál es su calificación media ponderada? ¿Qué calificación con letra obtuvo (A, B, C, D o F)? Suponga que una media de 90 o más es una A, una media de 80 a 89 es una B, y así sucesivamente.

3-1 Más allá de lo básico

35. Grados de Libertad Cinco frecuencias de pulso aleatoriamente seleccionadas del conjunto de datos 1 "Datos corporales" en el apéndice B tienen una media de 78.0 latidos por minuto. Cuatro de los pulsos son 82, 78, 56 y 84.

a. Encuentra el valor faltante.

b. Necesitamos elaborar una lista de *n* valores que tengan una media específica conocida. Somos libres de seleccionar los valores que deseemos para algunos de los *n* valores. ¿Cuántos de los *n* valores pueden asignarse libremente antes de que los valores restantes queden determinados? (El resultado se conoce como el *número de grados de libertad*).

36. Datos censurados El conjunto de datos 15 "Presidentes" del apéndice B muestra el número de años que vivieron los presidentes estadounidenses después de su primera toma de poder. Hasta la fecha, cinco de los presidentes siguen vivos y después de su primera asunción han vivido 37, 25, 21, 13 y 5 años. Podemos usar los valores de 37+, 25+, 21+, 13+ y 5+, donde los signos positivos indican que el valor real es igual o mayor que el valor actual. (Se dice que estos valores están *censurados* en el momento actual en que se compiló a lista). Si se usan los valores en el conjunto de datos 15 y se ignoran los presidentes que todavía están vivos, ¿cuál es la media? Si utiliza los valores dados en el conjunto de datos 15 junto con los valores adicionales de 37+, 25+, 21+, 13+ y 5+, ¿qué sabemos acerca de la media? ¿Los dos resultados difieren mucho?

37. Media recortada Debido a que la media es muy sensible a los valores extremos, se dice que no es una medida de tendencia central resistente. Al eliminar algunos valores bajos y altos, la **media recortada** se hace más resistente. Para encontrar la media recortada del 10% para un conjunto de datos, primero ordene los datos, luego elimine el 10% de los valores inferiores y el 10% de los valores superiores, luego calcule la media de los valores restantes. Utilice las cargas axiales (en libras) de las latas de aluminio que se listan a continuación (del conjunto de datos 30 “Latas de aluminio” en el apéndice B) para latas de 0.0111 pulgadas de espesor. Una carga axial es la fuerza a la que la tapa de una lata colapsa. Identifique cualquier valor atípico, luego compare la mediana, la media, la media recortada del 10% y la media recortada del 20%.

247	260	268	273	276	279	281	283	284	285	286	288
289	291	293	295	296	299	310	504				

38. Media armónica La **media armónica** se utiliza a menudo como una medida de tendencia central para conjuntos de datos que consisten en tasas de cambio, como velocidades. Se encuentra al dividir el número de valores n por la suma de los recíprocos de todos los valores, expresados como

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

(Ningún valor puede ser cero). El autor condujo 1163 millas para ir a una conferencia en Orlando, Florida. En el viaje de ida, el autor se detuvo durante la noche, y la velocidad media de principio a fin fue de 38 millas por hora. Durante el viaje de regreso, se detuvo sólo por comida y combustible, y la velocidad media desde el principio hasta el final fue de 56 millas por hora. Encuentre la media armónica de 38 mi/h y 56 mi/h para encontrar la verdadera velocidad “media” del viaje de ida y vuelta.

39. Media geométrica La **media geométrica** se utiliza a menudo en los negocios y la economía para encontrar tasas de cambio promedio, tasas de crecimiento promedio o razones medias. Para encontrar la media geométrica de n valores (donde todos son positivos), primero multiplique los valores, luego encuentre la raíz n-ésima del producto. Por un período de 6 años, el dinero depositado en certificados anuales de depósito tenía tasas de interés anual de 5.154%, 2.730%, 0.488%, 0.319%, 0.313% y 0.268%. Identifique el porcentaje de crecimiento único que es igual que las cinco tasas de crecimiento consecutivas, calculando la media geométrica de 1.05154, 1.02730, 1.00488, 1.00319, 1.00313 y 1.00268.

40. Media cuadrática La **media cuadrática** (o la **raíz cuadrada media**, o **R.C.M**.) se utiliza en aplicaciones físicas, como sistemas de distribución de energía. La media cuadrática de un conjunto de valores se obtiene al elevar al cuadrado cada valor, sumar esos cuadrados, dividir la suma por el número de valores n, y luego obtener la raíz cuadrada del resultado, como se indica a continuación:

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

Encuentre la R.C.M. de los siguientes voltajes medidos en corrientes domésticas: 0, 60, 110, –110, –60, 0. ¿Cómo se compara el resultado con la media?

41. Mediana Cuando los datos se resumen en una distribución de frecuencias, la mediana se puede encontrar identificando primero la clase mediana, que es la clase que contiene la mediana. Asumimos entonces que los valores de esa clase están uniformemente distribuidos e interpolamos. Si n expresa la suma de todas las frecuencias de clase, y m expresa la suma de las frecuencias de clase que preceden a la clase mediana, la mediana se puede estimar como se muestra a continuación.

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

Utilice este procedimiento para encontrar la mediana de la distribución de frecuencias dada en la tabla 3-2 de la página 88. ¿Cuánto se aleja este resultado de la mediana encontrada en la lista original de tiempos de servicio en McDonald’s del conjunto de datos 25 “Comida rápida” en el apéndice B?

3-2 Medidas de variación

Concepto clave La variación es el tema más importante en estadística, por lo que ésta es la sección más importante del libro. La presente sección introduce tres medidas importantes de variación: el *rango*, la *desviación estándar* y la *varianza*. Estos estadísticos son números, pero nuestro enfoque no es sólo calcular esos números, sino desarrollar la capacidad de *interpretarlos* y *comprenderlos*. Esta sección no es un estudio de aritmética; se trata de entender e interpretar medidas de variación, en especial la desviación estándar.

SUGERENCIA DE ESTUDIO La parte 1 de esta sección introduce conceptos básicos de la variación, y la parte 2 presenta conceptos adicionales relacionados con la desviación estándar. Ambas partes incluyen fórmulas para el cálculo, pero no dedican mucho tiempo a la enseñanza de fórmulas o la realización de cálculos aritméticos. En cambio, se centran en *entender* e *interpretar* los valores de la desviación estándar.

PARTE 1 Conceptos básicos de variación

Para visualizar la propiedad de variación, vea la figura 3-2, que ilustra los tiempos de espera (segundos) de los clientes de un banco en dos condiciones diferentes: (1) Todos los clientes ingresan a una sola fila que alimenta a diferentes cajeros; (2) todos los clientes eligen unirse a la línea en uno de *varios* cajeros. Verifique esta observación importante: los tiempos de espera con la línea única (gráfica superior) tienen menos *variación* que los tiempos de espera con múltiples líneas (gráfica inferior). *Ambos conjuntos de tiempos de espera tienen la misma media de 100.0 segundos, la misma mediana de 100.0 segundos y la misma moda de 100 segundos*. Esas medidas de tendencia central no "ven" la diferencia en la variación.

Para mantener nuestras reglas de redondeo tan consistentes y tan simples como sea posible, redondearemos las medidas de variación usando la siguiente regla:

REGLA DE REDONDEO PARA MEDIDAS DE VARIACION. Al redondear el valor de una medida de variación, conserve un decimal más que los presentes en el conjunto original de datos.

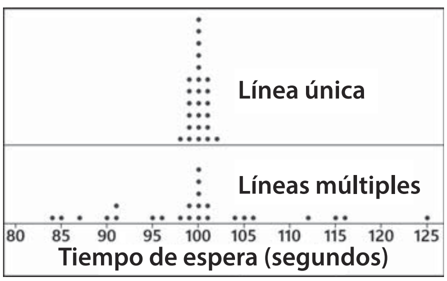


FIGURA 3-2 **Gráficas de puntos para los tiempos de espera (segundos) con una línea individual y líneas múltiples**

Mejora de la calidad Esta ilustración de una línea única y líneas múltiples es maravillosa porque los bancos realmente cambiaron de múltiples líneas a una sola línea no porque los hiciera más eficientes, no porque los tiempos de espera de los clientes se redujeran, sino porque los clientes son más felices con tiempos de espera con *menos variación*. El cambio no afectó las medidas de tendencia central, pero los bancos instituyeron el cambio para reducir la variación. Un objetivo importante de los negocios y la industria es el siguiente: ***Mejorar la calidad al reducir la variación***.

Rango

Comenzaremos con el rango porque es rápido y fácil de calcular, pero no es tan importante como otras medidas de variación.

> **DEFINICIÓN**
>
> El **rango** de un conjunto de valores de datos es la diferencia entre el valor máximo de datos y el valor mínimo de datos.
>
> **Rango = (valor máximo de datos) − (valor mínimo de datos)**

Propiedad importante del rango

- El rango utiliza sólo los valores máximo y mínimo de los datos, por lo que es muy sensible a los valores extremos. El rango no es *resistente*.
- Debido a que el rango utiliza sólo los valores máximo y mínimo, no toma en cuenta todos los valores y, por lo tanto, no refleja realmente la variación entre todos los valores de los datos.

EJEMPLO 1 Rango

Encuentre el rango de estas velocidades de datos (en Mbps) para Verizon: 38.5, 55.6, 22.4, 14.1, 23.1. (Estas son las primeras cinco velocidades de datos de Verizon listadas en el conjunto de datos 32 “Velocidad de datos en aeropuertos” del apéndice B).

SOLUCIÓN

El rango se encuentra restando el valor más bajo del valor más grande, así que se obtiene

$$\text{Rango} = (\text{valor máximo}) - (\text{valor mínimo}) = 55.6 - 14.1 = 41.50 \text{ Mbps}$$

El rango de 41.50 Mbps se muestra con un decimal más que los presentes en los valores de datos originales.

SU TURNO Encuentre el rango en el ejercicio 5 “Números de jugadores de fútbol americano”.

Desviación estándar de una muestra

La *desviación estándar* es la medida de variación más comúnmente utilizada en estadística.

> **DEFINICIÓN**
>
> La **desviación estándar** de un conjunto de valores muestrales, expresada por s, es una medida de cuánto se desvían los valores de datos de la media. Se calcula utilizando la fórmula 3-4 o 3-5. La fórmula 3-5 es solamente una versión diferente de la fórmula 3-4; ambas son algebraicamente iguales.

La desviación estándar encontrada a partir de datos muestrales es un dato estadístico expresado por s, y la desviación estándar determinada a partir de los datos de la población es un parámetro que se expresa con σ. La fórmula de σ es ligeramente diferente porque en la división se usa el tamaño de población N en lugar de $n - 1$. La desviación estándar σ de la población se estudiará más adelante.

Notación

s = desviación estándar *muestral*

σ = desviación estándar *poblacional*

¿Tienes un segundo?

La unidad de tiempo de un segundo ahora se define como “la duración de 9,192,631,770 periodos de la radiación correspondiente a la transición entre dos niveles hiperfinos del estado estable del átomo de cesio-133”. Esta es la primera vez que un segundo se define mediante el comportamiento de los átomos, y no con base en el movimiento de la Tierra; esto da como resultado una exactitud de ±1 segundo en 10,000,000 de años, la unidad de medida más precisa en uso hasta ahora. Debido a su gran exactitud, la definición de un segundo se está utilizando para definir otras cantidades, como el metro. Antes el metro se definía como 1/10,000,000 de la distancia, a lo largo de la superficie de la Tierra, entre el Polo Norte y el Ecuador (pasando por París). En la actualidad, el metro se define como la longitud de la distancia que recorre la luz en el vacío durante un intervalo de tiempo de 1/299,792,458 segundos.

En lo que respecta a los aparatos para medir el tiempo, se ha visto que la desviación estándar tradicional es inaceptable debido a la característica de una tendencia que consiste en una media que cambia con el paso del tiempo. Por ello, se usan otras medidas especiales de la variación, como la varianza Allan, la varianza total y TheoH.

Sin relación con la estadística, pero no menos interesante es el hecho de que los anuncios de relojes suelen mostrar un reloj con una hora cercana a las 10:10. Ese tiempo permite que la marca sea visible, y crea una imagen subliminal de una cara feliz. La hora de las 10:10 ha sido el estándar de la industria desde la década de 1940.

FÓRMULA 3-4

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} \quad \text{desviación estándar muestral}$$

FÓRMULA 3-5

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}} \quad \text{fórmula modificada para la desviación estándar muestral (utilizada por calculadoras y software)}$$

Más adelante explicaremos el razonamiento detrás de estas fórmulas, pero por ahora recomendamos el uso de la fórmula 3-4 para un ejemplo o dos, y después aprender a encontrar los valores de desviación estándar usando una calculadora o software.

Propiedades importantes de la desviación estándar

- La desviación estándar es una medida de cuánto se desvían los valores de datos de la *media*.
- El valor de la desviación estándar s nunca es negativo. Es cero sólo cuando todos los valores de datos son exactamente iguales.
- Los mayores valores de s indican mayores cantidades de variación.
- La desviación estándar s puede aumentar dramáticamente con uno o más valores atípicos.
- Las unidades de la desviación estándar s (como minutos, pies, libras) son las mismas que las unidades de los valores de datos originales.
- La desviación estándar muestral s es un **estimador sesgado** de la desviación estándar σ de la población, lo que significa que los valores de la desviación estándar muestral s no se centran en torno al valor de σ. (Esto se explica en la parte 2).

El ejemplo 2 ilustra un cálculo usando la fórmula 3-4 porque esta expresión señala de mejor manera que la desviación estándar se basa en las desviaciones de los valores muestrales con respecto a la media.

Más acciones, menos riesgo

En su libro *Investments*, los autores Zvi Bodie, Alex Kane y Alan Marcus afirman que “la desviación estándar promedio de los rendimientos de carteras compuestas por un solo tipo de acciones fue de 0.554. El riesgo promedio disminuye rápidamente cuando aumenta el número de acciones incluidas en la cartera”. También señalan que, con 32 acciones, la desviación estándar es de 0.325, lo que indica mucho menos variación y riesgo. Los autores destacan que con sólo unas cuantas acciones, una cartera tiene alto grado de riesgo “específico de una empresa”, lo que significa que el riesgo puede atribuirse a la escasa cantidad de acciones implicadas. Con más de 30 acciones hay muy poco riesgo específico asociado con una sola empresa; en tal situación, casi todo el riesgo es “riesgo de mercado”, atribuible al mercado global de acciones. Además, señalan que estos principios son “sólo una aplicación de la bien conocida ley de promedios”.

EJEMPLO 2 Cálculo de la desviación estándar con la fórmula 3-4

Utilice la fórmula 3-4 para encontrar la desviación estándar de estas velocidades de datos (en Mbps) para Verizon: 38.5, 55.6, 22.4, 14.1, 23.1.

SOLUCIÓN

La columna izquierda de la tabla 3-3 resume el procedimiento general para encontrar la desviación estándar utilizando fórmula 3-4, y la columna derecha ilustra ese procedimiento para los valores muestrales 38.5, 55.6, 22.4, 14.1, 23.1. El resultado que se muestra en la tabla 3-3 es de 16.45 Mbps, que se redondea a un decimal más que los presentes en la lista original de valores muestrales. Además, las unidades para la desviación estándar son las mismas que las unidades de los datos originales. Debido a que los datos originales tienen unidades de Mbps, la desviación estándar es de 16.45 Mbps.

SU TURNO Encuentre la desviación estándar en el ejercicio 5 “Números de jugadores de fútbol americano”

¿Dónde están los bateadores de 0.400?

El último beisbolista que bateó más de 0.400 fue Ted Williams, quien promedió 0.406 en 1941. Hubo promedios por arriba de 0.400 en 1876, 1879, 1887, 1894, 1895, 1896, 1897, 1899, 1901, 1911, 1920, 1922, 1924, 1925 y 1930, pero ninguno desde 1941. ¿Será que ya no existen grandes bateadores? El fallecido Stephen Jay Gould, de la Universidad de Harvard, señaló que el promedio de bateo medio se mantuvo 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 argumentaba que las estrellas de hoy son tan buenas como las del pasado, pero que los mejores lanzadores actuales mantienen los promedios de bateo por debajo de 0.400.

TABLA 3-3

Procedimiento general para determinar la desviación estándar con la fórmula 3-4	Ejemplo específico usando los siguientes valores muestrales: 38.5, 55.6. 22.4. 14.1, 23.1
Paso 1: Calcule la media $\overline{x}$.	La suma de 38.5, 55.6, 22.4, 14.1, 23.1 es 153.7; por lo tanto: $\overline{x} = \frac{\Sigma x}{n} = \frac{38.5 + 55.6 + 22.4 + 14.1 + 23.1}{5} = \frac{153.7}{5} = 30.74$
Paso 2: Reste la media de cada valor muestral individual. [El resultado es una lista de desviaciones de la forma $(x - \overline{x})$.]	Reste la media de 30.74 de cada valor muestral para obtener estas desviaciones respecto a la media: 7.76, 24.86, −8.34, −16.64, −7.64
Paso 3: Eleve al cuadrado cada una de las desviaciones obtenidas en el paso 2. [Esto produce números de la forma $(x - \overline{x})^2$.]	Los cuadrados de las desviaciones del paso 2 son: 60.2176, 618.0196, 69.5556, 276.8896, 58.3696.
Paso 4: Sume todos los cuadrados obtenidos en el paso 3. El resultado es $\Sigma(x - \overline{x})^2$.	La suma de los cuadrados del paso 3 es 1083.0520.
Paso 5: Divida el total del paso 4 por el número $n - 1$, que es 1 menos que el número total de valores muestrales presentes.	Con $n = 5$ valores de datos, $n - 1 = 4$, por lo que dividimos 1083.0520 entre 4 para obtener: $\frac{1083.0520}{4} = 270.7630.$
Paso 6: Encuentre la raíz cuadrada del resultado del paso 5. El resultado es la desviación estándar, expresada por s.	La desviación estándar es $\sqrt{270.7630} = 16.4548777$. Al redondear el resultado, obtenemos $s = 16.45$ Mbps.

EJEMPLO 3 Cálculo de la desviación estándar con la fórmula 3-5

Utilice la fórmula 3-5 para encontrar la desviación estándar de las velocidades de datos (Mbps) para Verizon en el ejemplo 1 (38.5, 55.6, 22.4, 14.1, 23.1).

SOLUCIÓN

Los componentes necesarios en la fórmula 3-5 son:

$n = 5$ (porque hay 5 valores muestrales)

$\Sigma x = 153.7$ (obtenida al sumar los valores originales de la muestra)

$\Sigma x^2 = 5807.79$ (obtenida al sumar los cuadrados de los valores muestrales, $38.5^2 + 55.6^2 + 22.4^2 + 14.1^2 + 23.1^2 = 5807.79$)

Con la fórmula 3-5, obtenemos

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}} = \sqrt{\frac{5(5807.79) - (153.7)^2}{5(5 - 1)}} = \sqrt{\frac{5415.26}{20}} = 16.45 \text{ Mbps}$$

El resultado $s = 16.45$ Mbps es el mismo que el del ejemplo 2.

SU TURNO Encuentre la desviación estándar en el ejercicio 5 "Números de jugadores de fútbol americano".

Regla práctica del rango para entender la desviación estándar

La ***regla práctica del rango*** es una herramienta básica pero simple para entender e interpretar la desviación estándar. Se basa en el principio de que, para muchos conjuntos de datos, la gran mayoría (alrededor de 95%) de los valores muestrales se encuentran dentro de 2 desviaciones estándar de la media. Podríamos mejorar la precisión de esta regla teniendo en cuenta factores como el tamaño de la muestra y la distribución, pero aquí sacrificamos la precisión en busca de simplicidad. El concepto de *significancia* que se presenta enseguida será mejorado en capítulos posteriores, especialmente aquellos que incluyen el tema de las pruebas de hipótesis, que también se llaman pruebas de significancia. La siguiente regla práctica del rango se basa en la media poblacional μ y en la desviación estándar poblacional σ, pero para muestras grandes y representativas, podríamos usar $\bar{x}$ y s en su lugar.

Regla práctica del rango para identificar valores significativos

Los valores ***significativamente bajos*** son $\mu - 2\sigma$ o inferiores.

Los valores **significativamente altos** son $\mu + 2\sigma$ o superiores.

Los **valores no significativos** están entre $(\mu - 2\sigma)$ y $(\mu + 2\sigma)$.

Vea la figura 3-3, que ilustra los criterios anteriores.

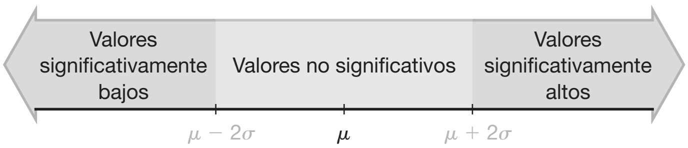


FIGURA 3-3 Regla práctica del rango para identificar los valores significativos

Regla práctica del rango para estimar un valor de la desviación estándar *s*

Para estimar en forma aproximada la desviación estándar de una colección de datos muestrales conocidos, utilice

$$s \approx \frac{\text{rango}}{4}$$

EJEMPLO 4 Regla práctica del rango para interpretar *s*

Si se consideran las 50 velocidades de datos de Verizon listadas en el conjunto de datos 32 “Velocidades de datos en aeropuertos” del apéndice B, la media es 17.60 Mbps y la desviación estándar es 16.02 Mbps. Utilice la regla práctica del rango para encontrar los límites que separan los valores significativamente bajos o significativamente altos; después determine si la velocidad de datos de 77.8 Mbps es significativamente alta.

SOLUCIÓN

Con una media de 17.60 y una desviación estándar de 16.02, utilizamos la regla práctica del rango para encontrar los límites que separan los valores significativamente bajos o significativamente altos, de la siguiente manera:

Los ***valores significativamente bajos*** son $(17.60 - 2 \times 16.02)$ o inferiores.

Por lo tanto, los valores significativamente bajos son -14.44 Mbps o menores.

Los ***valores significativamente altos*** son $(17.60 + 2 \times 16.02)$ o superiores.

Por lo tanto, los valores significativamente altos son 49.64 o mayores.

Los ***valores no significativos*** están entre -14.44 Mbps y 49.64 Mbps.

INTERPRETACIÓN

Con base en estos resultados, esperamos que las velocidades de datos típicas en aeropuertos para Verizon estén entre –14.44 Mbps y 49.64 Mbps. Debido a que el valor dado de 77.8 Mbps queda por encima de 49.64 Mbps, podemos considerarlo significativamente alto.

SU TURNO Resuelva el ejercicio 33 “Pulsos de mujeres”.

Variación en las caras

Los investigadores han comentado que "si todos luciéramos más o menos igual, habría un caos total". Se han estudiado las mediciones del cuerpo humano y se ha encontrado que los rasgos faciales varían más que los demás rasgos corporales y la mayor variación ocurre dentro del triángulo formado por los ojos y la boca. Los investigadores afirman que nuestra variación facial desempeña un papel importante en la evolución humana (vea "Morphological and Population Genomic Evidence That Human Faces Have Evolved to Signal Individual Identity", de Sheehan y Nachman, *Nature Communications*, vol. 5, núm. 4800).

EJEMPLO 5 **Regla práctica del rango para estimar *s***

Utilice la regla práctica del rango para estimar la desviación estándar de la muestra de 50 velocidades de datos para Verizon listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B. Estos 50 valores tienen un mínimo de 0.8 Mbps y un máximo de 77.8 Mbps.

SOLUCIÓN

La regla práctica del rango indica que podemos estimar la desviación estándar encontrando el rango y dividiéndolo por 4. Con un mínimo de 0.8 y un máximo de 77.8, la desviación estándar s puede estimarse como sigue:

$$s \approx \frac{\text{rango}}{4} = \frac{77.8 - 0.8}{4} = 19.25 \text{ Mbps}$$

INTERPRETACIÓN

El valor real de la desviación estándar es $s = 16.02$ Mbps, por lo que la estimación de 19.25 Mbps está cerca del resultado exacto. Debido a que esta estimación se basa sólo en los valores mínimo y máximo, es generalmente una estimación aproximada que podría estar alejada a una distancia considerable.

SU TURNO Resuelva el ejercicio 29 "Estimación de la desviación estándar".

Desviación estándar de una población

La definición de desviación estándar y las fórmulas 3-4 y 3-5 se aplican a la desviación estándar de datos *muestrales*. Para calcular la desviación estándar σ (sigma minúscula) de una *población*, se usa una fórmula ligeramente diferente: en lugar de dividir por $n - 1$, se divide por el tamaño N de la población, como se muestra:

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

Debido a que generalmente tratamos con datos muestrales, usaremos la fórmula 3-4, en la cual dividimos por $n - 1$. Muchas calculadoras dan tanto la desviación estándar muestral como la poblacional, pero usan varias notaciones.

PRECAUCIÓN Cuando utilice una calculadora para encontrar la desviación estándar, identifique la notación utilizada por su calculadora para obtener la desviación estándar *muestral*, no la desviación estándar *poblacional*.

Varianza de una muestra y una población

Hasta ahora, hemos utilizado el término *variación* como una descripción general de la cantidad en que los valores varían entre sí. (Los términos *dispersión* y *esparcimiento* se usan ocasionalmente en vez de *variación*). El término *varianza* tiene un significado específico.

DEFINICIÓN

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

- Varianza muestral: s^2 = cuadrado de la desviación estándar s.
- Varianza poblacional: σ^2 = cuadrado de la desviación estándar poblacional σ.

Notación A continuación se presenta un resumen de la notación para la desviación estándar y la varianza:

s = desviación estándar *muestral*

s^2 = varianza *muestral*

σ = desviación estándar *poblacional*

σ^2 = varianza *poblacional*

Nota: Los artículos de revistas e informes profesionales a menudo utilizan SD para la desviación estándar y VAR para la varianza.

Propiedades importantes de la varianza

- Las unidades de la varianza son los *cuadrados* de las unidades de los valores de datos originales. (Si los valores de datos originales están en pies, la varianza tendrá unidades de pies^2; si están en segundos, la varianza tendrá unidades de seg^2).
- El valor de la varianza puede aumentar dramáticamente con la inclusión de valores atípicos. (La varianza no es *resistente*).
- El valor de la varianza nunca es negativo. Es cero sólo cuando todos los valores de datos son el mismo número.
- La varianza muestral s^2 es un **estimador no sesgado** de la varianza poblacional σ^2, como se describe en la parte 2 de esta sección.

La varianza es un estadístico utilizado en algunos métodos estadísticos, pero para nuestros propósitos actuales, la varianza tiene la seria desventaja de usar unidades que son *diferentes de las unidades del conjunto original de datos*. Esto dificulta entender la varianza en cuanto a su relación con el conjunto de datos original. Debido a esta propiedad, es mejor centrarse primero en la desviación estándar al intentar desarrollar una comprensión de la variación.

PARTE 2 Más allá de lo básico de la variación

En la parte 2 nos centramos en dar sentido a la desviación estándar, de modo que no sea un número misterioso desprovisto de significado práctico. Comenzamos abordando preguntas comunes que se relacionan con la desviación estándar.

¿Por qué la desviación estándar está definida por la fórmula 3-4?

Al medir la variación en un conjunto de datos muestrales, tiene sentido comenzar con las cantidades individuales en las que los valores se desvían de la media. Para un valor de datos particular x, la cantidad de desviación es $x - \bar{x}$. Tiene sentido combinar de alguna manera esas desviaciones en un número que pueda servir como una medida de la variación. La suma de las desviaciones no es buena, porque dicha suma siempre será cero. Para obtener un dato estadístico que mida la variación, es necesario evitar la cancelación de números negativos y positivos. Un método es agregar valores absolutos, como en $\Sigma|x - \bar{x}|$. Si encontramos la media de esa suma, obtenemos la **desviación media absoluta** (o DMA), que es la distancia media de los datos con respecto a la media:

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

¿Por qué no utilizar la desviación absoluta media en lugar de la desviación estándar?

El cálculo de la desviación absoluta media usa valores absolutos, por lo que emplea una operación que no es “algebraica”. La utilización de valores absolutos sería simple, pero crearía dificultades algebraicas en los métodos de estadística inferencial que se estudian en capítulos posteriores. La desviación estándar tiene la ventaja de utilizar sólo operaciones algebraicas.

Debido a que se basa en la raíz cuadrada de una suma de cuadrados, la desviación estándar es muy similar a las fórmulas de distancia encontradas en álgebra. Hay muchos casos en que un procedimiento estadístico se basa en una suma de cuadrados similar. En consecuencia, en vez de utilizar valores absolutos, se elevan al *cuadrado* todas las desviaciones $(x - \bar{x})$ para que sean no negativas, y esos cuadrados se utilizan para calcular la desviación estándar.

¿Por qué dividir por $n - 1$? Después de encontrar todos los valores individuales de $(x - \bar{x})^2$ los combinamos encontrando su suma. Luego dividimos por $n - 1$ porque sólo hay $n - 1$ valores que se pueden asignar sin restricción. Con una media dada, podemos usar cualquier número para los primeros $n - 1$ valores, pero el último valor estará determinado automáticamente. Con la división por $n - 1$, las varianzas muestrales s^2 tienden a centrarse alrededor del valor de la varianza poblacional σ^2; con la división por n, las varianzas muestrales s^2 tienden a *subestimar* el valor de la varianza poblacional σ^2.

¿Cómo encontramos el sentido de un valor de desviación estándar? La parte 1 de esta sección incluyó la regla práctica del rango para interpretar un valor conocido de una desviación estándar o estimar el valor de una desviación estándar. (Vea los ejemplos 4 y 5). Otros dos métodos para interpretar la desviación estándar son la regla empírica y el teorema de Chebyshev.

Regla empírica (o 68-95-99.7) para datos con una distribución en forma de campana

Un concepto útil para interpretar el valor de una desviación estándar es la **regla empírica**. Esta regla establece que *para los conjuntos de datos que tienen una distribución aproximadamente en forma de campana*, se aplican las siguientes propiedades. (Vea la figura 3-4).

- Aproximadamente 68% de todos los valores caen dentro de 1 desviación estándar de la media.
- Aproximadamente 95% de todos los valores caen dentro de 2 desviaciones estándar de la media.
- Aproximadamente 99.7% de todos los valores caen dentro de 3 desviaciones estándar de la media.

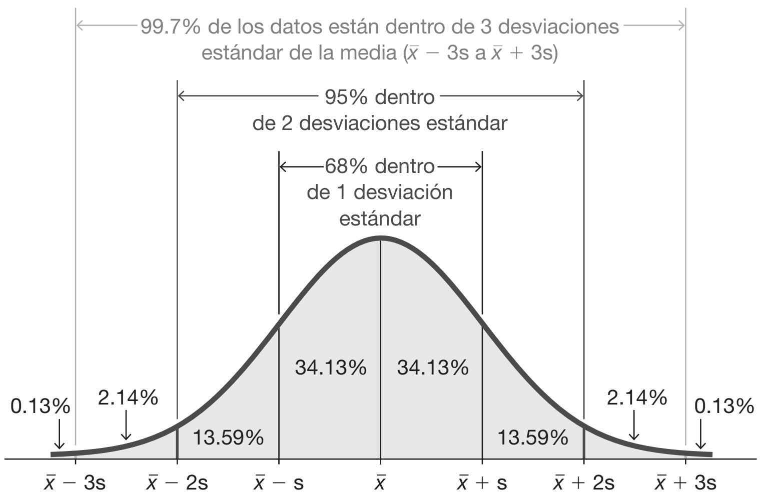


FIGURA 3-4 La regla empírica

EJEMPLO 6 La regla empírica

Las puntuaciones de IQ tienen una distribución en forma de campana con una media de 100 y una desviación estándar de 15. ¿Qué porcentaje de las puntuaciones de IQ está entre 70 y 130?

SOLUCIÓN

La clave para resolver este problema es reconocer que las puntuaciones de 70 y 130 están, cada una, a exactamente dos desviaciones estándar de la media de 100, como se muestra a continuación:

$$2 \text{ desviaciones estándar} = 2s = 2(15) = 30$$

Por lo tanto, 2 desviaciones estándar desde la media es

$$100 - 30 = 70$$
$$\text{o} \quad 100 + 30 = 130$$

La regla empírica nos dice que alrededor de 95% de todos los valores están dentro de 2 desviaciones estándar de la media, por lo que alrededor de 95% de todos los puntajes de IQ están entre 70 y 130.

SU TURNO Resuelva el ejercicio 41 "La regla empírica".

Otro concepto útil para entender o interpretar el valor de una desviación estándar es el **teorema de Chebyshev**. La regla empírica se aplica sólo a conjuntos de datos con distribuciones en forma de campana, pero el teorema de Chebyshev se aplica a *cualquier* conjunto de datos. Desafortunadamente, los resultados del teorema de Chebyshev son sólo aproximados. Debido a que los resultados son límites inferiores ("al menos"), el teorema de Chebyshev tiene una utilidad limitada.

Teorema de Chebyshev

La proporción de cualquier conjunto de datos comprendidos 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 $K = 3$, obtenemos las siguientes afirmaciones:

- Al menos 3/4 (o 75%) de todos los valores se encuentran dentro de 2 desviaciones estándar de la media.
- Al menos 8/9 (u 89%) de todos los valores se encuentran dentro de 3 desviaciones estándar de la media.

EJEMPLO 7 Teorema de Chebyshev

Las puntuaciones de IQ tienen una media de 100 y una desviación estándar de 15. ¿Qué podemos concluir a partir del teorema de Chebyshev?

SOLUCIÓN

Si se aplica el teorema de Chebyshev con una media de 100 y una desviación estándar de 15, podemos llegar a las siguientes conclusiones:

- Al menos 3/4 (o 75%) de las puntuaciones de IQ están dentro de 2 desviaciones estándar de la media (entre 70 y 130).
- Al menos 8/9 (u 89%) de todos los puntajes de IQ están dentro de 3 desviaciones estándar de la media (entre 55 y 145).

SU TURNO Resuelva el ejercicio 43 "Teorema de Chebyshev".

Comparación de la variación en diferentes muestras o poblaciones

Una buena práctica consiste en comparar dos desviaciones estándar muestrales sólo cuando las medias de la muestra son aproximadamente iguales. Al comparar la variación en muestras o poblaciones con medias muy diferentes, es mejor utilizar el *coeficiente de variación*. Utilice también el coeficiente de variación para comparar la variación de dos muestras o poblaciones con diferentes escalas o unidades de valores, como la comparación de la variación de las *alturas* de los hombres y los *pesos* de los hombres. (Vea el ejemplo 8).

DEFINICIÓN

El **coeficiente de variación** (o ***CV***) para un conjunto de datos muestrales o poblacionales no negativos, expresado como porcentaje, describe la desviación estándar en relación con la media, y está dado por:

Muestra	**Población**
$CV = \frac{s}{\bar{x}} \cdot 100$	$CV = \frac{\sigma}{\mu} \cdot 100$

REGLA DEL REDONDEO PARA EL COEFICIENTE DE VARIACION Redondee el coeficiente de variación a un decimal (como en 25.3%).

EJEMPLO 8 Velocidades de datos de Verizon y magnitudes de terremoto

Compare la variación de las 50 velocidades de datos de Verizon listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B y las magnitudes de 600 terremotos en el conjunto de datos 21 "Terremotos" del apéndice B. Para las velocidades de datos de Verizon $\bar{x} = 17.60$ Mbps y $s = 16.02$ Mbps; para las magnitudes de terremotos, $\bar{x} = 2.572$ y $s = 0.651$. Observe que queremos comparar la variación entre *velocidades de datos* con la variación entre *magnitudes de terremotos*.

SOLUCIÓN

Podemos comparar las desviaciones estándar si se utilizan las mismas escalas y unidades, y si las dos medias son aproximadamente iguales; pero aquí tenemos diferentes escalas y diferentes unidades de medida, por lo que usaremos los coeficientes de variación:

$$\text{Velocidades de datos de Verizon: } CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{16.02 \text{ Mbps}}{17.60 \text{ Mbps}} \cdot 100\% = 91.0\%$$

$$\text{Magnitudes de terremotos: } CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{0.651}{2.572} \cdot 100\% = 25.3\%$$

Ahora podemos ver que las velocidades de datos de Verizon (con $CV = 91.0\%$) varían considerablemente más que las magnitudes de los terremotos (con $CV = 25.3\%$).

Estimadores sesgados y no sesgados

La desviación estándar muestral s es un **estimador sesgado** de la desviación estándar poblacional σ, lo que significa que los valores de la desviación estándar muestral s *no* tienden a centrarse alrededor del valor de la desviación estándar poblacional σ. Aunque los valores individuales de s podrían ser iguales o superiores a σ, los valores de s generalmente tienden a *subestimar* el valor de σ. Por ejemplo, considere una prueba de IQ diseñada para que la desviación estándar de la población sea 15. Si repite el proceso de selección aleatoria de 100 sujetos, dándoles pruebas de IQ y calculando la desviación estándar de la muestra s en cada caso, las desviaciones estándar que obtendrá tenderán a ser inferiores a 15, que es la desviación estándar de la población. No hay corrección que nos permita fijar el sesgo para todas las distribuciones de datos. Existe una corrección que nos permite fijar el sesgo para las poblaciones normalmente distribuidas, pero rara vez se utiliza porque es demasiado compleja y hace modificaciones relativamente menores.

La varianza muestral s^2 es un **estimador no sesgado** de la varianza poblacional σ^2, lo que significa que los valores de s^2 tienden a centrarse alrededor del valor de σ^2 en lugar de tender sistemáticamente a sobreestimarlo o subestimarlo. Considere una prueba de IQ diseñada para que la varianza de la población sea 225. Si repite el proceso de selección aleatoria de 100 sujetos, dándoles pruebas de IQ y calculando la varianza muestral s^2 en cada caso, las varianzas muestrales que obtendrá tenderán a centrarse alrededor 225, que es la varianza de la población.

Los estimadores sesgados y los estimadores no sesgados se analizarán con mayor detalle en la sección 6-3.

CENTRO DE TECNOLOGÍA

Medidas de variación

Acceda a los complementos de software, videos y conjuntos de datos en www.pearsonenespañol.com/triola

Statdisk, Minitab, StatCrunch, Excel y la Calculadora TI-83/84 Plus pueden utilizarse para realizar los cálculos importantes de esta sección. Use los mismos procedimientos de estadística descriptiva dados al final de la sección 3-1 en la página 91.

3-2 Habilidades básicas y conceptos

Conocimiento estadístico y pensamiento crítico

1. Regla práctica del rango para estimar s Los 20 volúmenes cerebrales (en cm^3) del conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B varían de un mínimo de 963 cm^3 a un máximo de 1439 cm^3. Utilice la regla práctica del rango para estimar la desviación estándar s y compare el resultado con la desviación estándar exacta de 124.9 cm^3.

2. Regla práctica del rango para interpretar s Los 20 volúmenes cerebrales (en cm^3) del conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B tienen una media de 1126.0 cm^3 y una desviación estándar de 124.9 cm^3. Utilice la regla práctica del rango para identificar los límites que separan los valores significativamente bajos y significativamente altos. Para tales datos, ¿un volumen cerebral de 1440 cm^3 sería significativamente alto?

3. Varianza Los 20 sujetos utilizados en el conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B tienen pesos con una desviación estándar de 20.0414 kg ¿Cuál es la varianza de sus pesos? Asegúrese de incluir las unidades apropiadas con el resultado.

4. Símbolos Identifique los símbolos utilizados para cada uno de los siguientes estadísticos: (a) desviación estándar muestral; (b) desviación estándar poblacional; (c) varianza muestral; (d) varianza poblacional.

En los ejercicios 5 a 20, encuentre el* rango*, la* varianza *y la* desviación estándar *para los datos muestrales dados. Incluya las unidades apropiadas (como "minutos") en sus resultados. (Los mismos datos se usaron en la sección 3-1, donde encontramos medidas de tendencia central, aquí determinaremos medidas de variación). Luego responda las preguntas dadas.

5. Números de jugadores de fútbol americano A continuación se listan los números de camiseta de 11 jugadores seleccionados aleatoriamente del equipo de los Halcones marinos de Seattle cuando ganaron el Súper Tazón XLVIII. ¿Qué nos dicen los resultados?

89 91 55 7 20 99 25 81 19 82 60

6. Pesos de jugadores de fútbol americano A continuación se listan los pesos en libras de 11 jugadores seleccionados al azar del equipo de los Halcones marinos de Seattle, cuando ganaron el Súper Tazón XLVIII (los mismos jugadores del ejercicio anterior). ¿Es probable que las medidas de variación sean típicas de todos los jugadores de la NFL?

189 254 235 225 190 305 195 202 190 252 305

7. Valor neto de celebridades A continuación se listan los mayores valores netos (en millones de dólares) de las celebridades. Las celebridades son Tom Cruise, Will Smith, Robert De Niro, Drew Carey, George Clooney, John Travolta, Samuel L. Jackson, Larry King, Demi Moore y Bruce Willis. ¿Son las medidas de variación típicas para todas las celebridades?

250 200 185 165 160 160 150 150 150 150

8. Lo que pasa en Las Vegas... A continuación se listan los precios en dólares por una noche en diferentes hoteles ubicados en Las Vegas Boulevard (the “Strip”). ¿Qué tan útiles son las medidas de variación para alguien que busca una habitación?

212 77 121 104 153 264 195 244

9. Huracanes A continuación se lista la cantidad de huracanes que se produjeron en el Atlántico cada año. Los datos se dan en orden anual, a partir del año 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de variación?

8 9 8 7 9 15 5 6 8 4 12 7 8 2

10. Chícharos en una vaina Los biólogos han realizado experimentos para determinar si una falta de bióxido de carbono en el suelo afecta a los fenotipos de los chícharos. A continuación se listan los códigos del fenotipo, donde 1 = amarillo liso, 2 = verde liso, 3 = amarillo corrugado y 4 = verde corrugado. ¿Se pueden obtener las medidas de variación para estos valores? ¿Los resultados tienen sentido?

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

11. Precios de televisores A continuación se listan los precios de venta (en dólares) de televisores de 60 pulgadas o más, los cuales fueron calificados como “las mejores compras” por la revista *Consumer Reports*. ¿Es probable que las medidas de variación sean típicas para todos los televisores de 60 pulgadas o más?

1800 1500 1200 1500 1400 1600 1500 950 1600 1150 1500 1750

12. Radiación del teléfono celular A continuación se listan las tasas de absorción de radiación medidas (en W/kg) correspondientes a los siguientes teléfonos celulares: iPhone 5S, BlackBerry Z30, Sanyo Vero, Optimus V, Droid Razr, Nokia N97, Samsung Vibrant, Sony Z750a, Kyocera Kona, LG G2 y Virgin Mobile Supreme. Los datos provienen de la Comisión Federal de Comunicaciones. Si a un ejemplar de cada modelo se le mide la absorción de radiación y los resultados se utilizan para encontrar las medidas de variación, ¿son los resultados típicos de la población de teléfonos celulares que están en uso?

1.18 1.41 1.49 1.04 1.45 0.74 0.89 1.42 1.45 0.51 1.38

13. Cafeína en bebidas A continuación se indican las cantidades medidas de cafeína (mg por 12 onzas de bebida) obtenidas en una lata de cada una de 20 marcas (7-UP, A & W Root Beer, Cherry Coke, ...). ¿Son estos datos estadísticos representativos de la población de todas las latas de las mismas 20 marcas consumidas por los estadounidenses?

0 0 34 34 34 45 41 51 55 36 47 41 0 0 53 54 38 0 41 47

14. Muertes de bomberos A continuación se listan las cantidades de heroicos bomberos que perdieron sus vidas en Estados Unidos cada año, mientras combatían incendios forestales. Las cantidades están ordenadas anualmente, comenzando con el año 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de variación?

20 18 23 30 20 12 24 9 25 15 8 11 15 34

15. Longitud de los pies A continuación se listan las longitudes en pulgadas de los pies de mujeres del ejército seleccionadas al azar, medidos en el sondeo antropométrico de 1988 (ANSUR, abreviatura en inglés). ¿Son los estadísticos representativos de la población actual de todas las mujeres del ejército?

10.4 9.3 9.1 9.3 10.0 9.4 8.6 9.8 9.9 9.1 9.1

16. Universidades más caras A continuación se listan los costos anuales (en dólares) de matrícula y colegiaturas en las 10 universidades más caras de Estados Unidos para un año reciente (con base en datos de *US News & World Report*). Las universidades listadas en orden son Columbia, Vassar, Trinidad, George Washington, Carnegie Mellon, Wesleyana, Tulane, Bucknell, Oberlin y Union. ¿Qué nos dice esta "lista de las 10 primeras" sobre la variación entre los costos para la población de todas las universidades estadounidenses?

49,138 47,890 47,510 47,343 46,962 46,944 46,930 46,902 46,870 46,785

17. Anillo de diamante A continuación se listan las cantidades en dólares que cuestan los paquetes para propuestas matrimoniales en los diferentes estadios de las Ligas Mayores de Béisbol. Cinco de los equipos no permiten propuestas. ¿Existen valores atípicos, y es probable que tengan mucho efecto sobre las medidas de variación?

39	50	50	50	55	55	75	85	100	115	175	175	200
209	250	250	350	400	450	500	500	500	500	1500	2500	

18. Ventas de álbumes de discos LP de vinilo A continuación se listan las ventas anuales de discos de vinilo en Estados Unidos (millones de unidades). Las cantidades de álbumes vendidos se presentan en orden cronológico, y la última entrada representa el año más reciente. ¿Las medidas de variación nos dan alguna información sobre una tendencia cambiante a lo largo del tiempo?

0.3 0.6 0.8 1.1 1.1 1.4 1.4 1.5 1.2 1.3 1.4 1.2 0.9 0.9 1.0 1.9 2.5 2.8 3.9 4.6 6.1

19. Fumadores de California En la Encuesta Entrevista sobre Salud en California, se entrevistó a adultos seleccionados al azar. Una de las preguntas fue acerca de cuántos cigarrillos fuman al día, y los resultados se listan a continuación para 50 encuestados seleccionados aleatoriamente. ¿Qué tan bien reflejan los resultados los hábitos del tabaquismo de los adultos en California?

9	10	10	20	40	50	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0										

20. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a las mujeres que evaluaran el atractivo de sus acompañantes masculinos, y a continuación se lista una muestra de los resultados (1 = no atractivo, 10 = extremadamente atractivo). ¿Se pueden utilizar los resultados para describir la variación entre los atractivos de la población de varones adultos?

5 8 3 8 6 10 3 7 9 8 5 5 6 8 8 7 3 5 5 6 8 7 8 8 8 7

En los ejercicios 21 a 24, determine el* coeficiente de variación *para cada una de las dos muestras; luego compare la variación. (Los mismos datos se usaron en la sección 3-1).

21. Presión arterial Una muestra de mediciones de presión arterial se toma del conjunto de datos 1 "Datos corporales" del apéndice B, y los valores (mm Hg) se relacionan de manera que 10 sujetos tienen medidas sistólicas y diastólicas.

Sistólica:	118	128	158	96	156	122	116	136	126	120
Diastólica:	80	76	74	52	90	88	58	64	72	82

22. Robo de parquímetros A continuación se listan los montos (en millones de dólares) recaudados en parquímetros por Brinks y otras empresas en la ciudad de Nueva York durante períodos de tiempo similares. Se utilizó un conjunto de datos más amplio para condenar a cinco empleados de Brinks por hurto mayor. Los datos fueron proporcionados por el abogado de la Ciudad de Nueva York, y se pueden encontrar en el sitio web de DASL. ¿Las dos muestras parecen tener diferentes cantidades de variación?

El contratista de recaudación fue Brinks	1.3	1.5	1.3	1.5	1.4	1.7	1.8	1.7	1.7	1.6
El contratista de recaudación no fue Brinks	2.2	1.9	1.5	1.6	1.5	1.7	1.9	1.6	1.6	1.8

23. Pulsos A continuación se listan los pulsos (latidos por minuto) de muestras de hombres y mujeres adultas (del conjunto de datos 1 "Datos corporales" en el apéndice B). ¿Parece haber una diferencia?

Hombres:	86	72	64	72	72	54	66	56	80	72	64	64	96	58	66
Mujeres:	64	84	82	70	74	86	90	88	90	90	94	68	90	82	80

24. Filas en el banco Los tiempos de espera (en segundos) de los clientes en el Banco de Ahorro de Madison se registran con dos configuraciones: línea de clientes única; líneas de clientes individuales.

Línea única	390	396	402	408	426	438	444	462	462	462
Líneas individuales	252	324	348	372	402	462	462	510	558	600

Grandes conjuntos de datos del apéndice B. ***En los ejercicios 25 a 28, considere el conjunto de datos indicado en el apéndice B. Use un software o una calculadora para encontrar el*** **rango,** ***la*** **varianza** ***y la*** **desviación estándar.** ***Exprese las respuestas usando unidades apropiadas, como "minutos".***

25. Tornados Utilice las medidas en la escala F de tornados listadas en el conjunto de datos 22 "Tornados" del apéndice B. Tenga cuidado de considerar los datos faltantes.

26. Terremotos Utilice las magnitudes (en la escala de Richter) de los 600 terremotos listados en el conjunto de datos 21 "Terremotos" del apéndice B. En 1989, el área de la Bahía de San Francisco fue golpeada por un terremoto que midió 7.0 en la escala de Richter. Si añadimos ese valor de 7.0 a los listados en el conjunto de datos, ¿cambian mucho las medidas de variación?

27. Temperaturas corporales Considere el conjunto de datos 3 "Temperaturas corporales" en el apéndice B y use las temperaturas corporales para las 12:00 AM del día 2.

28. Nacimientos Use los pesos al nacer (en gramos) de los 400 bebés listados en el conjunto de datos 4 "Nacimientos" en el apéndice B. Examine la lista de pesos al nacer y haga una observación sobre esos números. ¿Cómo afecta esta observación la forma en que los resultados deben redondearse?

Estimación de la desviación estándar con la regla práctica del rango. ***En los ejercicios 29 a 32, consulte los datos del ejercicio indicado. Después de encontrar el*** **rango** ***de los datos, utilice la regla práctica del rango para calcular el valor de la*** **desviación estándar.** ***Compare el resultado con la desviación estándar calculada con todos los datos.***

29. Ejercicio 25 "Tornados"

30. Ejercicio 26 "Terremotos"

31. Ejercicio 27 "Temperaturas corporales"

32. Ejercicio 28 "Nacimientos"

Identificación de valores significativos con la regla práctica del rango. ***En los ejercicios 33 a 36, use la regla práctica del rango para identificar los límites que separan los*** **valores significativamente bajos** ***y*** **significativamente altos.**

33. Pulso de mujeres Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, las mujeres tienen pulsos con una media de 74.0 latidos por minuto y una desviación estándar de 12.5 latidos por minuto. ¿Un pulso de 44 latidos por minuto es significativamente bajo o significativamente alto? (Todas estas pulsaciones se miden en reposo).

34. Pulso de hombres Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, los hombres tienen pulsos con una media de 69.6 latidos por minuto y una desviación estándar de 11.3 latidos por minuto. ¿Un pulso de 50 latidos por minuto es significativamente bajo o significativamente alto? (Todas estas pulsaciones se miden en reposo). Explique.

35. Longitud de los pies Con base en el conjunto de datos 2 "Pies y alturas" en el apéndice B, los varones adultos tienen longitudes de pie con una media de 27.32 cm y una desviación estándar de 1.29 cm. ¿Una longitud de los pies de un adulto de 30 cm es significativamente baja o significativamente alta? Explique.

36. Temperaturas corporales Con base en el conjunto de datos 3 "Temperaturas corporales" en el apéndice B, las temperaturas corporales de los adultos tienen una media de 98.20 °F y una desviación estándar de 0.62 °F. (Se usan los datos de las 12 AM en el día 2). ¿Una temperatura corporal de un adulto de 100 °F es significativamente baja o significativamente alta?

Determinación de la desviación estándar a partir de una distribución de frecuencias. ***En los ejercicios 37 a 40, considere la distribución de frecuencias en el ejercicio dado y encuentre la*** **desviación estándar** ***usando la fórmula siguiente, donde x representa el punto medio de la clase, f es la frecuencia de la clase y n el número total de valores muestrales. Además, compare las desviaciones estándar calculadas con estas desviaciones estándar obtenidas usando la fórmula 3-4 con la lista original de valores de datos: (ejercicio 37) 11.5 años; (ejercicio 38) 8.9 años; (ejercicio 39) 59.5; (Ejercicio 10) 65.4.***

Desviación estándar para la distribución de frecuencias

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

37.

Edad (en años) de la mejor actriz al ganar el Oscar	Frecuencia
20-29	29
30-39	34
40-49	14
50-59	3
60-69	5
70-79	1
80-89	1

38.

Edad (en años) del mejor actor al ganar el Oscar	Frecuencia
20-29	1
30-39	28
40-49	36
50-59	15
60-69	6
70-79	1

39.

Conteo de plaquetas en la sangre de hombres (1000 células/μl)	Frecuencia
0-99	1
100-199	51
200-299	90
300-399	10
400-499	0
500-599	0
600-699	1

40.

Conteo de plaquetas en la sangre de mujeres (1000 células/μl)	Frecuencia
100-199	25
200-299	92
300-399	28
400-499	0
500-599	2

41. La regla empírica Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, los conteos de plaquetas en la sangre de mujeres tienen una distribución en forma de campana con una media de 255.1 y una desviación estándar de 65.4. (Todas las unidades se dan en 1000 células/μL). Use la regla empírica para determinar cuál es el porcentaje aproximado de mujeres con conteo de plaquetas

a. dentro de 2 desviaciones estándar de la media, o entre 124.3 y 385.9.

b. entre 189.7 y 320.5.

42. La regla empírica Con base en el conjunto de datos 3 "Temperaturas corporales" del apéndice B, las temperaturas corporales de adultos sanos tienen una distribución en forma de campana con una media de 98.20 °F y una desviación estándar de 0.62 °F. Use la regla empírica para determinar cuál es el porcentaje aproximado de adultos sanos con temperatura corporal

a. dentro de 1 desviación estándar de la media, o entre 97.58 °F y 98.82 °F.

b. entre 96.34 °F y 100.06 °F.

43. Teorema de Chebyshev Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, los conteos de plaquetas en la sangre de mujeres tienen una distribución en forma de campana con una media de 255.1 y una desviación estándar de 65.4. Utilice el teorema de Chebyshev para determinar qué sabemos sobre el porcentaje de mujeres con conteo de plaquetas dentro de 3 desviaciones estándar de la media. ¿Cuáles son los conteos de plaquetas mínimo y máximo que están dentro de 3 desviaciones estándar de la media?

44. Teorema de Chebyshev Con base en el conjunto de datos 3 "Temperaturas corporales" del apéndice B, las temperaturas corporales de adultos sanos tienen una distribución en forma de campana con una media de 98.20 °F y una desviación estándar de 0.62 °F (si se consideran los datos de las 12 AM en el día 2). Utilice el teorema de Chebyshev para determinar qué sabemos sobre el porcentaje de adultos sanos con temperaturas corporales que están dentro de 2 desviaciones estándar de la media. ¿Cuáles son las temperaturas corporales mínima y máxima que están dentro de 2 desviaciones estándar de la media?

3-2 Más allá de lo básico

45. ¿Por qué dividir por $n - 1$? Considere que una población consiste en los valores 9 cigarrillos, 10 cigarrillos y 20 cigarrillos fumados en un día (con base en los datos de la encuesta Entrevista de Salud en California). Suponga que se seleccionan al azar muestras de dos valores con reemplazo de esta población. (Es decir, se reemplaza un valor seleccionado antes de realizar la segunda selección).

a. Encuentre la varianza σ^2 de la población {9 cigarrillos, 10 cigarrillos, 20 cigarrillos}.

b. Después de listar las nueve posibles muestras posibles de dos valores seleccionados con reemplazo, encuentre la varianza muestral s^2 (que incluye la división por $n - 1$) para cada una de ellas; luego encuentre la media de las nueve varianzas muestrales s^2.

c. Para cada una de las nueve diferentes muestras posibles de dos valores seleccionados con reemplazo, encuentre la varianza tratando cada muestra como si fuera una población (usando la fórmula para la varianza poblacional, que incluye la división por n); después determine la media de esas nueve varianzas poblacionales.

d. ¿Qué enfoque resulta en valores que son mejores estimaciones de σ^2: el del inciso (b) o el del inciso (c)? ¿Por qué? Al calcular las varianzas muestrales, ¿debería usarse la división por n o por $n - 1$?

e. Los incisos anteriores muestran que s^2 es un estimador no sesgado de σ^2. ¿Es s un estimador no sesgado de σ? Explique.

46. Desviación media absoluta Utilice la misma población de {9 cigarrillos, 10 cigarrillos, 20 cigarrillos} del ejercicio 45. Muestre que cuando se seleccionan al azar muestras de tamaño 2 con reemplazo, las muestras tienen desviaciones medias absolutas que no se centran alrededor del valor de la desviación media absoluta de la población. ¿Qué indica esto acerca de una desviación media absoluta muestral como estimador de la desviación media absoluta de una población?

3-3 Medidas de posición relativa y gráficas de caja

Concepto clave Esta sección presenta las medidas de posición relativa, que son números que indican la ubicación de los valores de datos en relación con los demás valores dentro del mismo conjunto de datos. El concepto más importante en esta sección es la *puntuación z*, que se utilizará a menudo en los siguientes capítulos. También estudiaremos los cuartiles y percentiles, que son estadísticos comunes, así como un nuevo tipo de gráfica estadística llamada gráfica de caja.

PARTE 1 Fundamentos de las puntuaciones z, percentiles, cuartiles y gráficas de caja

Puntuaciones z

Una puntuación z se encuentra al convertir un valor a una escala estandarizada, como se establece en la siguiente definición. Esta definición establece que una puntuación z corresponde al número de desviaciones estándar que separan a un dato de la media. Utilizaremos ampliamente las puntuaciones z en el capítulo 6 y en capítulos posteriores.

DEFINICIÓN

Una **puntuación *z*** (o **puntuación estándar** o **valor estandarizado**) es el número de desviaciones estándar que un valor dado x se encuentra por arriba o por debajo de la media. Se calcula utilizando las siguientes expresiones:

Muestra		**Población**
$z = \dfrac{x - \bar{x}}{s}$	o	$z = \dfrac{x - \mu}{\sigma}$

REGLA DE REDONDEO PARA LAS PUNTUACIONES *z* Redondee z a dos posiciones decimales (como en 2.31)

Esta regla de redondeo está motivada por el formato de tablas estándar en las que los puntajes se expresan con dos decimales, como en la tabla A-2 del apéndice A. El ejemplo 1 ilustra cómo se pueden usar las puntuaciones z para comparar valores, incluso si proceden de diferentes poblaciones.

Propiedades importantes de las puntuaciones *z*

1. Una puntuación z es el número de desviaciones estándar que un valor dado x está por arriba o por debajo de la media.
2. Las puntuaciones z se expresan como números sin unidades de medida.
3. Un valor de datos es *significativamente bajo*, si su puntuación z es menor o igual a -2 o el valor es significativamente alto si su puntuación z es mayor o igual a $+2$.
4. Si un valor de datos individual es menor que la media, su puntuación correspondiente es un número negativo.

EJEMPLO 1 Comparación del peso de un bebé y la temperatura corporal de un adulto

¿Cuál de los siguientes dos valores de datos es más extremo en relación con el conjunto de datos del que procede?

- El peso de 4000 g de un bebé recién nacido (entre 400 pesos con media muestral $\bar{x} =$ 3152.0 g y desviación estándar muestral $s = 693.4$ g)
- La temperatura de 99 °F de un adulto (entre 106 adultos con media muestral $\bar{x} = 98.20$ °F y desviación estándar muestral $s = 0.62$ °F)

SOLUCIÓN

El peso de 4000 g y la temperatura corporal de 99 °F pueden estandarizarse convirtiendo cada uno de ellos a puntuaciones z como se muestra a continuación.

4000 g de peso al nacer:

$$z = \frac{x - \bar{x}}{s} = \frac{4000\text{ g} - 3152.0\text{ g}}{693.4\text{ g}} = 1.22$$

99 °F de temperatura corporal:

$$z = \frac{x - \bar{x}}{s} = \frac{99\,°\text{F} - 98.20\,°\text{F}}{0.62\,°\text{F}} = 1.29$$

INTERPRETACIÓN

Las puntuaciones z muestran que el peso al nacer de 4000 g está 1.22 desviaciones estándar por arriba de la media, y la temperatura corporal de 99 °F está 1.29 desviaciones estándar por arriba de la media. Debido a que la temperatura del cuerpo está más alejada de la media, es el valor más extremo. Una temperatura corporal de 99 °F es ligeramente más extrema que un peso al nacer de 4000 g.

SU TURNO Resuelva el ejercicio 13 "Los hombres más altos y más bajos".

Uso de puntuaciones z para identificar valores significativos En la sección 3-2 usamos la regla práctica del rango para concluir que un valor es significativamente bajo o significativamente alto si está por lo menos 2 desviaciones estándar alejado de la media. Se deduce que los valores significativamente bajos tienen puntuaciones z menores o iguales a -2 y que los valores significativamente altos tienen puntuaciones z mayores o iguales a $+2$, como se ilustra en la figura 3-5. Si se usa este criterio con los dos valores individuales usados en el ejemplo 1, se observa que ninguno de los dos valores es significativo porque ambas puntuaciones z están entre -2 y $+2$.

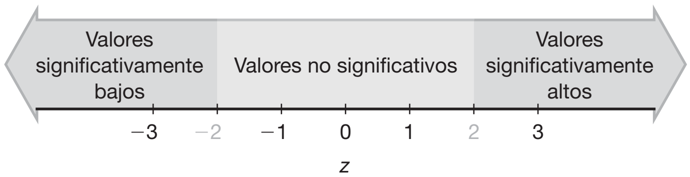


FIGURA 3-5 Interpretación de puntuaciones z

Los valores significativos son aquellos con puntuaciones $z \leq -2.00$ o ≥ 2.00.

EJEMPLO 2 ¿Un conteo de plaquetas de 75 es significativamente bajo?

El conteo de plaquetas más bajo en el conjunto de datos 1 "Datos corporales" en el apéndice B es 75. (El conteo de plaquetas se mide en 1000 células/μl). ¿Es ese valor significativamente bajo? Con base en los conteos de plaquetas del conjunto de datos 1 del apéndice B, suponga que los conteos de plaquetas tienen una media de $\bar{x} = 239.4$ y una desviación estándar de $s = 64.2$.

SOLUCIÓN

El conteo de plaquetas de 75 se convierte en una puntuación z como se muestra a continuación:

$$z = \frac{x - \bar{x}}{s} = \frac{75 - 239.4}{64.2} = -2.56$$

INTERPRETACIÓN

El conteo de plaquetas de 75 se convierte en la puntuación z de -2.56. Consulte la figura 3-5 para ver que $z = -2.56$ es menor que -2, por lo que el conteo de plaquetas de 75 es significativamente bajo. (El conteo bajo de plaquetas se llama trombocitopenia, no por falta de un término mejor).

SU TURNO Resuelva el ejercicio 9 "ACT".

Una puntuación z es una medida de posición, en el sentido de que describe la ubicación de un valor (en términos de desviaciones estándar) con relación a la media. Los percentiles y cuartiles son otras medidas de posición útiles para comparar valores dentro del mismo conjunto de datos o entre diferentes conjuntos de datos.

Percentiles

Los percentiles son un tipo de *cuantiles* —o *fractiles*— que dividen los datos en grupos con aproximadamente el mismo número de valores en cada grupo.

DEFINICIÓN

Los **percentiles** son medidas de ubicación, expresadas como P_1, P_2, . . . , P_{99}, que dividen un conjunto de datos en 100 grupos con aproximadamente el 1% de los valores en cada grupo.

El percentil 50, denominado P_{50}, tiene aproximadamente el 50% de los valores de datos por debajo de él y aproximadamente el 50% de los valores de datos por encima de él, por lo que el percentil 50 es igual que la mediana. No hay un acuerdo universal sobre el procedimiento que debe usarse para calcular los percentiles, pero describiremos procedimientos relativamente simples para (1) encontrar el percentil de un valor de datos y (2) convertir un percentil en su valor de datos correspondiente. Comenzamos con el primer procedimiento.

Determinación del percentil de un valor de datos

El proceso de encontrar el percentil que corresponde a un valor de datos particular x está dado por lo siguiente (redondee el resultado al número entero más próximo):

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

EJEMPLO 3 Determinación de un percentil

En la tabla 3-4 se listan las mismas velocidades de datos para teléfonos celulares de Verizon que aparecen en el conjunto de datos 32 “Velocidades de datos en aeropuertos”, pero en la tabla 3-4 las velocidades de datos se colocan en orden creciente. Encuentre el percentil para la velocidad de datos de 11.8 Mbps.

TABLA 3-4 Velocidades de datos en aeropuertos para Verizon *ordenadas* (Mbps)

0.8	1.4	1.8	1.9	3.2	3.6	4.5	4.5	4.6	6.2
6.5	7.7	7.9	9.9	10.2	10.3	10.9	11.1	11.1	11.6
11.8	12.0	13.1	13.5	13.7	14.1	14.2	14.7	15.0	15.1
15.5	15.8	16.0	17.5	18.2	20.2	21.1	21.5	22.2	22.4
23.1	24.5	25.7	28.5	34.6	38.5	43.0	55.6	71.3	77.8

SOLUCIÓN

En la lista ordenada de velocidades de datos en aeropuertos de la tabla 3-4, se observa que hay 20 velocidades de datos menores que 11.8 Mbps, así que

$$\text{Percentil de } 11.8 = \frac{20}{50} \cdot 100 = 40$$

INTERPRETACIÓN

Una velocidad de datos de 11.8 Mbps está en el percentil 40. Esto puede interpretarse de manera flexible: una velocidad de datos de 11.8 Mbps separa el 40% de los valores más bajos del 60% de los valores más altos. Se tiene $P_{40} = 11.8$ Mbps.

SU TURNO Resuelva el ejercicio 17 “Percentiles”.

El ejemplo 3 muestra cómo convertir de un valor muestral dado al percentil correspondiente. Existen varios métodos diferentes para el procedimiento inverso de convertir un percentil dado al valor correspondiente en el conjunto de datos. El procedimiento que utilizaremos se resume en la figura 3-6, que utiliza la siguiente notación.

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. Este es el mismo método básico que se utiliza en la creación del Índice de Precios al Consumidor (IPC), el cual se basa en un promedio ponderado de bienes y servicios adquiridos por consumidores comunes. Mientras que las puntuaciones estándar y los percentiles nos permiten comparar valores diferentes ignorando cualquier elemento del tiempo, los números índice, como el ICR y el IPC, nos permiten comparar el valor de alguna variable con su valor en un periodo tomado como base. El valor de un número índice es el valor actual dividido entre el valor base, multiplicado por 100.

Notación

n número total de valores en el conjunto de datos

k percentil que se utiliza (ejemplo: para el percentil 25, $k = 25$).

L localizador que da la *posición* de un valor (ejemplo: para el duodécimo valor en la lista ordenada, $L = 12$).

P_k k-ésimo percentil (ejemplo: P_{25} es el percentil 25).

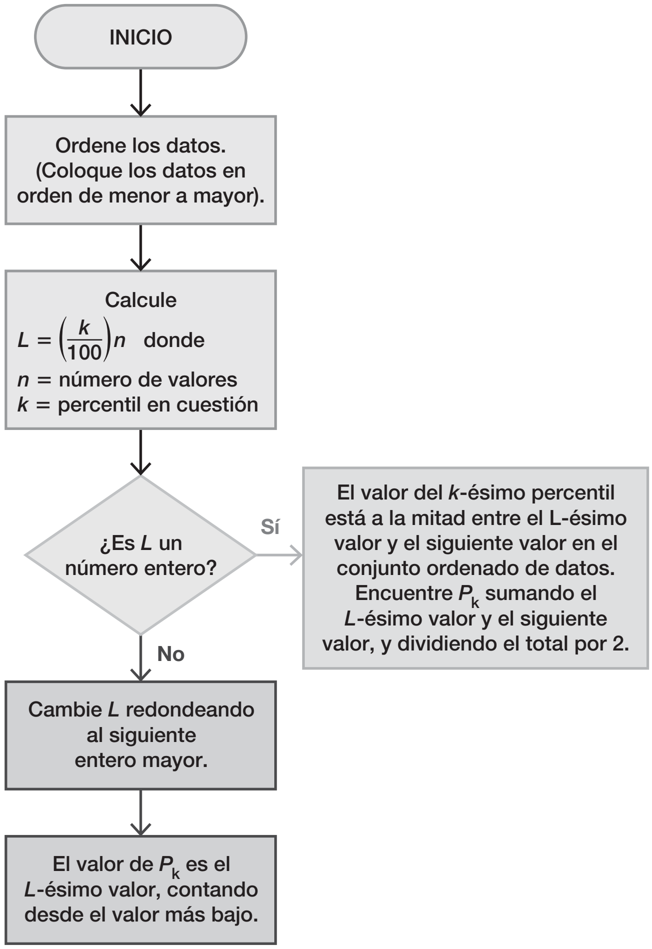


FIGURA 3-6 Conversión del *k*-ésimo percentil al valor de datos correspondiente

EJEMPLO 4 Conversión de un percentil a un valor de datos

Consulte las velocidades de datos ordenadas en la tabla 3-4 y use el procedimiento de la figura 3-6 para encontrar el valor del vigésimo quinto percentil, P_{25}.

SOLUCIÓN

A partir de la figura 3-6, se observa que los datos de la muestra ya están ordenados, por lo que podemos proceder a encontrar el valor del localizador L. En este cálculo usamos $k = 25$ porque estamos tratando de encontrar el valor del percentil 25. Utilizamos $n = 50$ porque hay 50 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{25}{100} \cdot 50 = 12.5$$

Puesto que $L = 12.5$ no es un número entero, pasamos al siguiente cuadro inferior de la figura 3-6, donde cambiamos L al redondear de 12.5 al siguiente entero mayor: 13. (En este libro redondeamos de la manera habitual, pero éste es uno de los dos casos en los que redondeamos hacia arriba). En el cuadro inferior vemos que el valor de P_{25} es el valor número 13, contando desde el valor más bajo. En la tabla 3-4, el valor 13 es 7.9. Es decir, $P_{25} = 7.9$ Mbps. A grandes rasgos, alrededor de 25% de las velocidades de datos son menores que 7.9 Mbps y 75% de ellas son mayores que 7.9 Mbps.

SU TURNO Resuelva el ejercicio 23 "Cuartil".

EJEMPLO 5 Conversión de un percentil a un valor de datos

Consulte las velocidades de datos ordenadas de la tabla 3-4. Utilice la figura 3-6 para encontrar el percentil 40, expresada por P_{40}.

SOLUCIÓN

Con referencia a la figura 3-6, vemos que los datos muestrales ya están ordenados, por lo que podemos proceder a calcular el valor del localizador L. En este cálculo, usamos $k = 40$ porque estamos tratando de encontrar el valor del percentil 40, y utilizamos $n = 50$ porque hay 50 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{40}{100} \cdot 50 = 20$$

Puesto que $L = 20$ es un número entero, procedemos a la caja de la figura 3-6 ubicada a la derecha. Ahora vemos que el valor del percentil 40 está a la mitad entre el L-ésimo valor (vigésimo) y el siguiente valor en el conjunto original de datos. Es decir, el valor del percentil 40 está a la mitad entre el valor 20 y el valor 21. El valor 20 en la tabla 3-4 es 11.6 y el valor 21 es 11.8, por lo que el valor a la mitad entre ellos es 11.7 Mbps. Concluimos que el percentil 40 es $P_{40} = 11.7$ Mbps.

SU TURNO Resuelva el ejercicio 21 "Percentil".

Cuartiles

De la misma manera que hay 99 percentiles que dividen los datos en 100 grupos, hay tres cuartiles que dividen los datos en cuatro grupos.

DEFINICIÓN

Los **cuartiles** son medidas de ubicación, denominadas Q_1, Q_2 y Q_3, que dividen un conjunto de datos en cuatro grupos con aproximadamente el 25% de los valores en cada uno.

A continuación se describen los cuartiles de manera más precisa que en la definición anterior:

Q_1 (Primer cuartil): mismo valor que P_{25}. Separa el 25% de los valores inferiores ordenados del 75% superior. (Para ser más precisos, al menos 25% de los valores ordenados son menores o iguales a Q_1, y al menos 75% de los valores son mayores o iguales a Q_1).

Q_2 (Segundo cuartil): igual a P_{50} e igual a la mediana. Separa el 50% de los valores inferiores ordenados del 50% superior.

Q_3 (Tercer cuartil): igual a P_{75}. Separa el 75% de los valores inferiores ordenados del 25% superior. (Para ser más precisos, al menos 75% de los valores ordenados son menores o iguales a Q_3 y al menos 25% de los valores son mayores o iguales a Q_3).

Niveles Nielsen para estudiantes universitarios

Los niveles Nielsen representan una de las medidas más importantes del teleauditorio y afectan a miles de millones de dólares en publicidad televisiva. En el pasado, los hábitos televisivos de los estudiantes universitarios eran ignorados, con el resultado de que un gran segmento de la audiencia joven quedara excluido. Nielsen Media Research está ahora incluyendo a los estudiantes universitarios que no viven en casa.

Algunos programas de televisión tienen gran atractivo para los televidentes en el grupo de edad de 18 a 24 años, y los niveles de audiencia de estos programas han aumentado sustancialmente con la inclusión de los estudiantes universitarios. Para los varones, la difusión del *fútbol americano del domingo por la noche* a través de la NBC tuvo un aumento del 20% después de incluir a los estudiantes universitarios. Los niveles más altos de audiencia se traducen en mayores ganancias por mayores cobros a los patrocinadores comerciales. Estos niveles también dan reconocimiento a los estudiantes universitarios, lo cual afecta la programación que reciben.

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

La determinación de los valores de los cuartiles se puede lograr con el mismo procedimiento utilizado para encontrar percentiles. Simplemente utilice las relaciones mostradas al margen. En el ejemplo 4 se encontró que $P_{25} = 7.9$ Mbps, por lo que se deduce que $Q_1 = 7.9$ Mbps.

PRECAUCIÓN Así como no hay un acuerdo universal sobre un procedimiento para encontrar percentiles, tampoco existe un solo procedimiento para calcular cuartiles, y las diferentes tecnologías a menudo dan resultados diferentes. Si utiliza una calculadora o un software que incluyan cuartiles, puede obtener resultados que difieran un poco de las respuestas obtenidas utilizando los procedimientos descritos aquí.

En secciones anteriores de este capítulo describimos varios estadísticos, incluyendo la media, la mediana, la moda, el rango y la desviación estándar. Algunos otros estadísticos se definen utilizando cuartiles y percentiles, como los siguientes:

Rango intercuartil (o RIQ) $= Q_3 - Q_1$

Rango semi-intercuartil $= \dfrac{Q_3 - Q_1}{2}$

Cuartil medio $= \dfrac{Q_3 + Q_1}{2}$

Rango percentil 10-90 $= P_{90} - P_{10}$

Resumen de 5 números y gráfica de caja

Los valores del mínimo, el máximo y los tres cuartiles (Q_1, Q_2, Q_3) se utilizan para el resumen de 5 números y la construcción de gráficas de caja.

DEFINICIÓN

Para un conjunto de datos, el **resumen de 5 números** consta de los siguientes cinco valores:

1. Mínimo
2. Primer cuartil, Q_1
3. Segundo cuartil, Q_2 (igual a la mediana)
4. Tercer cuartil, Q_3
5. Máximo

EJEMPLO 6 Determinación de un resumen de 5 números

Utilice las velocidades de datos en aeropuertos para Verizon de la tabla 3-4 para encontrar el resumen de 5 números.

SOLUCIÓN

Debido a que las velocidades de datos en aeropuertos para Verizon de la tabla 3-4 están ordenadas, es fácil ver que el mínimo es 0.8 Mbps y el máximo es 77.8 Mbps. El valor del primer cuartil es $Q_1 = 7.9$ Mbps (del ejemplo 4). La mediana es igual a Q_2, y es 13.9 Mbps. Además, podemos encontrar que $Q_3 = 21.5$ Mbps usando el mismo procedimiento para encontrar P_{75} (tal como lo indica la figura 3-6). Por lo tanto, el resumen de 5 números es 0.8, 7.9, 13.9, 21.5 y 77.8 (todos en unidades de Mbps).

SU TURNO Encuentre el resumen de 5 números en el ejercicio 29 "Citas rápidas".

Los valores del resumen de 5 números se utilizan para la construcción de una gráfica de caja, definido como sigue.

DEFINICIÓN

Una **gráfica de caja** (o **diagrama de caja y bigotes**) es una gráfica de un conjunto de datos que consiste en una línea que se extiende desde el valor mínimo hasta el valor máximo, y una caja con líneas dibujadas en el primer cuartil Q_1, la mediana y el tercer cuartil Q_3. (Vea la figura 3-7).

Procedimiento para elaborar una gráfica de caja

1. Encuentre el resumen de 5 números (valor mínimo, Q_1, Q_2, Q_3, valor máximo).
2. Construya un segmento de línea que se extienda desde el valor mínimo hasta el valor máximo de los datos.
3. Construya una caja (rectángulo) que se extienda de Q_1 a Q_3, y dibuje una línea en la caja sobre el valor de Q_2 (la mediana).

PRECAUCIÓN Debido a que no hay un acuerdo universal sobre los procedimientos para encontrar cuartiles, y como las gráficas de caja se basan en cuartiles, los diversos paquetes y aplicaciones de software pueden producir gráficas distintas.

EJEMPLO 7 Elaboración de una gráfica de caja

Utilice las velocidades de datos en aeropuertos para Verizon listadas en la tabla 3-4 para construir una gráfica de caja.

SOLUCIÓN

La gráfica de caja usa el resumen de 5 números encontrado en el ejemplo 6: 0.8, 7.9, 13.9, 21.5 y 77.8 (todos en unidades de Mbps). La figura 3-7 es la gráfica de caja que representa las velocidades de datos en aeropuertos para Verizon listadas en la tabla 3-4.

FIGURA 3-7 Gráfica de caja de las velocidades de datos en aeropuertos para Verizon (Mbps)

SU TURNO Construya la gráfica de caja en el ejercicio 29 “Citas rápidas”.

Asimetría A menudo, una gráfica de caja se puede usar para identificar asimetrías. Recuerde que en la sección 2-2 establecimos que una distribución de datos es **asimétrica** si se extiende más hacia un lado que hacia el otro. En un histograma de datos asimétrico a la derecha (también denominado *positivamente asimétrico*), hay una cola derecha más larga que indica que relativamente pocos valores de datos son altos; la mayoría de los valores se encuentran a la izquierda. La gráfica de caja de la figura 3-7 muestra que los datos son asimétricos a la derecha y la mayoría de los valores se encuentran a la izquierda.

Debido a que la forma de una gráfica de caja está determinada por los valores del resumen de cinco números, no es una gráfica de la distribución de los datos, y no muestra información tan detallada como un histograma o un diagrama de tallo y hojas. Sin embargo, las gráficas de caja son útiles para comparar dos o más conjuntos de datos. Cuando utilice dos o más gráficas de caja para comparar diferentes conjuntos de datos, grafíquelos en la misma escala para facilitar las comparaciones. Los métodos que se estudian más adelante en este libro permiten analizar las comparaciones de los conjuntos de datos más formalmente que las conclusiones subjetivas basadas en una gráfica. Siempre es aconsejable construir gráficas adecuadas, como histogramas, gráficas de puntos y gráficas de caja, pero no debemos confiar únicamente en juicios subjetivos basados en gráficas.

EJEMPLO 8 Comparación de las velocidades de datos de Verizon, Sprint, AT&T y T-Mobile

El problema del capítulo se refiere a las velocidades de datos para teléfonos inteligentes en 50 aeropuertos, y las velocidades se miden para los operadores Verizon, Sprint, AT&T y T-Mobile. Utilice la misma escala para construir las cuatro gráficas de caja correspondientes; luego compare los resultados.

SOLUCIÓN

Los diagramas de caja generados por Statdisk que se muestran en la figura 3-8 sugieren que las velocidades de datos de Verizon son generalmente más rápidas.

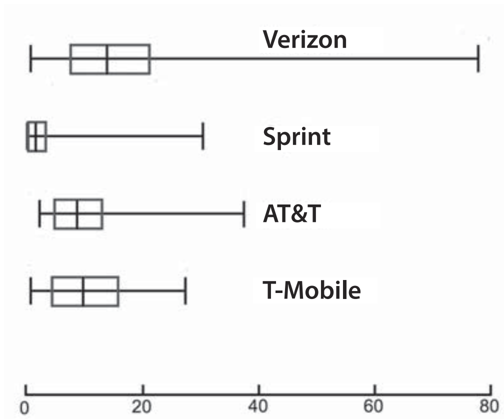


FIGURA 3-8 Gráficas de caja de las velocidades de datos en aeropuertos para diferentes operadores

SU TURNO Resuelva el ejercicio 33 “Pulsos”.

Valores atípicos

Cuando se analizan datos, es importante identificar y considerar los valores atípicos porque pueden afectar significativamente los valores de algunos datos estadísticos importantes (como la media y la desviación estándar) y también pueden tener un gran impacto en los métodos importantes que se estudian más adelante en este libro. En el capítulo 2 describimos los valores atípicos como valores muestrales que están muy alejados de la gran mayoría de los demás valores de un conjunto de datos, pero esa descripción es vaga y no proporciona criterios objetivos específicos. La parte 2 de esta sección incluye una descripción de las *gráficas de caja modificadas* junto con una definición más precisa de los valores atípicos utilizados en el contexto de la creación de este tipo de gráficas.

PRECAUCIÓN Cuando analice datos, siempre identifique los valores atípicos y considere sus efectos, los cuales pueden ser sustanciales.

PARTE 2 Valores atípicos y gráficas de caja modificadas

Hemos observado que la descripción de los valores atípicos es algo imprecisa, pero con la intención de construir *gráficas de caja modificadas*, podemos considerar valores atípicos como valores de datos que cumplen con criterios específicos basados en los cuartiles y el rango intercuartil. (A menudo se denomina RIQ $= Q_3 - Q_1$).

Identificación de valores atípicos para las gráficas de caja modificadas

1. Encuentre los cuartiles Q_1, Q_2 y Q_3.
2. Determine el rango intercuartil (RIQ), donde RIQ $= Q_3 - Q_1$.
3. Evalúe $1.5 \times$ RIQ.
4. **En una gráfica de caja modificada, un valor de datos es *atípico* si está**
   por arriba de Q_3, en una cantidad superior a $1.5 \times$ RIQ
   o por debajo de Q_1, en una cantidad superior a $1.5 \times$ RIQ

Gráficas de caja modificadas

Las gráficas descritas anteriormente se denominan **gráficas de caja esqueléticas** (o **regulares**), pero algunos paquetes de software estadístico proporcionan gráficas modificadas, que representan valores extremos como puntos especiales. Una **gráfica de caja modificada** es una gráfica de caja regular construida con las siguientes modificaciones: (1) Un símbolo especial (un asterisco o un punto) se utiliza para identificar los valores atípicos según se definieron anteriormente, y (2) la línea horizontal sólida sólo se extiende hasta el valor de datos mínimo que no es atípico y hasta el valor de datos máximo que tampoco es atípico. (*Nota: los ejercicios que implican gráficas de caja modificadas se encuentran solamente en los ejercicios "Más allá de lo básico"*).

EJEMPLO 9 Construcción de una gráfica de caja modificada

Utilice las velocidades de datos en aeropuertos para Verizon en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B para construir una gráfica de caja modificada.

SOLUCIÓN

Comencemos con los cuatro pasos descritos anteriormente para identificar los valores atípicos en una gráfica de caja modificada.

1. A partir de las velocidades de datos para Verizon, los tres cuartiles son $Q_1 = 7.9$, $Q_2 = 13.9$ (mediana) y $Q_3 = 21.5$. (Todos los valores se dan en Mbps y estos cuartiles se encontraron en el ejemplo 6).
2. El rango intercuartil es RIQ $= Q_3 - Q_1 = 21.5 - 7.9 = 13.6$.
3. $1.5 \times$ RIQ $= 1.5 \times 13.6 = 20.4$.
4. Cualesquiera valores atípicos están arriba de $Q_3 = 21.5$ por más de 20.4, o debajo de $Q_1 = 7.9$ por más de 20.4. Esto significa que cualquier valor atípico es mayor que 41.9 o menor que -12.5 (lo que es imposible; entonces, aquí no hay valores atípicos en el extremo inferior).

Ahora es posible examinar las velocidades de datos en aeropuertos para Verizon originales a fin de identificar las velocidades superiores a 41.9, y se encuentra: 43.0, 55.6, 71.3 y 77.8, que son los únicos valores atípicos.

Ahora podemos construir la gráfica de caja modificada que se muestra en la figura 3-9 de la página siguiente. Ahí, los cuatro valores atípicos se identifican como puntos especiales, los tres cuartiles se muestran como en una gráfica regular y la línea horizontal se extiende desde el valor de datos más bajo que no es atípico (0.8) hasta el valor de datos más alto que tampoco es atípico (38.5).

continúa

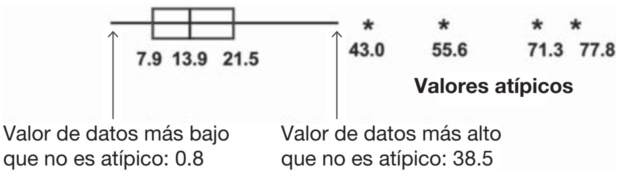


FIGURA 3-9 **Gráfica de caja modificada de las velocidades de datos en aeropuertos para Verizon (Mbps)**

SU TURNO Resuelva el ejercicio 37 “Valores atípicos y gráficas de caja modificadas”.

PRECAUCIÓN Debido a que no existe un acuerdo universal sobre los procedimientos para encontrar cuartiles, y como las gráficas de caja modificadas se basan en cuartiles, las diferentes tecnologías pueden producir gráficas de caja modificadas distintas.

CENTRO DE TECNOLOGÍA

Gráficas de caja, resumen de 5 números, valores atípicos

Acceda a los complementos técnicos, videos y conjuntos de datos en www.pearsonenespañol.com/triola

Statdisk

Resumen de 5 números
Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

1. Haga clic en **Data** del menú superior.
2. Seleccione **Boxplot** en el menú desplegable.
3. Seleccione las columnas de datos deseadas.
4. Haga clic en **Boxplot** o **Modified Boxplot**.

Valores atípicos
Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la siguiente manera:

1. Haga clic en **Data** del menú superior.
2. Seleccione **Sort Data** en el menú desplegable.
3. Haga clic en **Sort** después de elegir las opciones deseadas en el menú de ordenación.
4. Examine los valores mínimo y máximo para determinar si están muy alejados de los demás valores.

Minitab

Resumen de 5 números
Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

1. Haga clic en **Graph** del menú superior
2. Seleccione **Boxplot** en el menú desplegable
3. Seleccione la opción **Simple** para una o varias gráficas de caja, luego haga clic en **OK**.
4. Haga doble clic en la(s) columna(s) de datos deseada(s) para que aparezca en la ventana de *Graph variables* y luego haga clic en **OK**.

Valores atípicos
Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la siguiente manera:

1. Haga clic en **Data** del menú superior.
2. Seleccione **Sort** en el menú desplegable.
3. Haga doble clic en la columna de datos deseada para que aparezca en la ventana de *Sort Column(s)*.
4. Haga clic en el cuadro *By column* y seleccione la misma columna de datos. Haga clic en **OK**.
5. Examine los valores mínimo y máximo para determinar si están muy alejados de otros valores.

CENTRO DE TECNOLOGÍA *continuación*

Gráficas de caja, resumen de 5 números, valores atípicos

Acceda a los complementos tecnológicos, videos y conjuntos de datos en www.pearsonenespañol.com/triola

StatCrunch

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

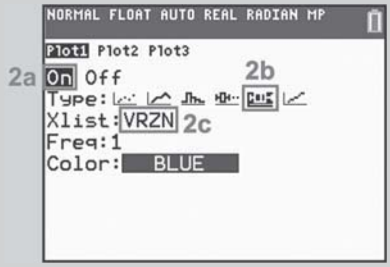
1. Haga clic en **Graph** del menú superior.
2. Seleccione **Boxplot** en el menú desplegable.
3. Seleccione la columna de datos deseada. Para una gráfica de caja modificada, marque la casilla *Use fences*.
4. Haga clic en **Compute!**

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la siguiente manera:

1. Haga clic en **Data** del menú superior.
2. Seleccione **Sort** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Compute!**
5. Examine los valores mínimo y máximo para determinar si están muy alejados de los demás valores.

Calculadora TI-83/84 Plus



Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

1. Abra el menú **STAT PLOTS** pulsando 2ND, Y=
2. Presione ENTER para acceder a la pantalla de configuración *Plot 1* como se muestra:
   a. Seleccione **ON** y pulse ENTER.
   b. Seleccione el segundo ícono de gráfica de caja, presione ENTER. Seleccione el primer ícono de gráfica de caja para una gráfica modificada.
   c. Introduzca el nombre de la lista que contiene los datos.
3. Presione ZOOM luego 9 (ZoomStat) para mostrar la gráfica de caja.
4. Presione TRACE y use ◀ ▶ para ver los valores.

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la manera siguiente:

1. Pulse STAT, seleccione **SortA** (orden ascendente) en el menú y pulse ENTER.
2. Introduzca el nombre de la lista que desea ordenar y pulse ENTER.
3. Para ver la lista ordenada, presione STAT, seleccione **Edit** y pulse ENTER.
4. Resalte la celda superior en una columna vacía, escriba el nombre de la lista y presione ENTER.
5. Utilice ▲ ▼ para examinar los valores mínimo y máximo y determinar si están muy alejados de los demás valores.

Excel

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Complemento XLSTAT

- Después del paso 3 en el procedimiento de estadística descriptiva, haga clic en la ficha **Outputs** y seleccione **Minimum**, **Maximum**, **1st Quartile**, **Median**, **3rd Quartile**. Haga clic en **OK**.

Excel

El complemento para análisis de datos proporciona sólo el mínimo, el máximo y la mediana. Para obtener los cuartiles utilice el siguiente procedimiento:

1. Haga clic en **Insert función** f_x, seleccione la categoría **Statistical** y seleccione la función **QUARTILE.INC**.
2. Introduzca el intervalo de valores de datos en el cuadro *Array*.
3. En el cuadro *Quart*, ingrese **0** para encontrar el mínimo, **1** para encontrar el primer cuartil y **2**, **3**, **4** para encontrar los valores restantes.

Gráficas de caja

Complemento XLSTAT (Requerido)

1. Haga clic en la ficha **XLSTAT** en la barra de opciones y, después haga clic en **Describing Data**.
2. Seleccione **Descriptive Statistics** en el menú desplegable.
3. Marque la casilla de **Quantitative Data** e introduzca el intervalo de datos deseado. Si selecciona dos o más columnas, se generarán varias gráficas. Si la primera fila de datos contiene una etiqueta, también marque la casilla **Sample labels**.
4. Haga clic en la ficha **Options** y confirme que la casilla **Charts** esté marcada.
5. Haga clic en la pestaña **Charts (1)** y marque la casilla **Box plots** en *Quantitative Data*.
6. Haga clic en **OK**.

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento de XLSTAT anterior u ordénela de la siguiente manera:

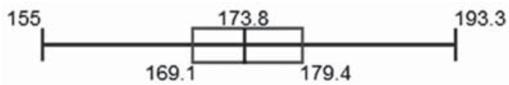
1. Haga clic en la ficha **Data** del menú superior y seleccione el intervalo de valores de datos deseado.
2. Haga clic en el botón **Sort Smallest to Largest** (A → Z) en la barra.
3. Examine los valores mínimo y máximo para determinar si están muy alejados de los demás valores.

3-3 Habilidades y conceptos básicos

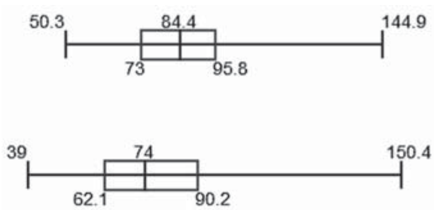
Conocimiento estadístico y pensamiento crítico

1. Puntuaciones z LeBron James, uno de los jugadores de baloncesto más exitosos de todos los tiempos, tiene una estatura de 6 pies y 8 pulgadas, o 203 cm. Con base en estadísticos del conjunto de datos 1 "Datos corporales" en el apéndice B, su estatura se convierte en la puntuación z de 4.07. ¿Cuántas desviaciones estándar está su estatura por arriba de la media?

2. Estaturas El diagrama de caja que se muestra a continuación es el resultado de las estaturas (en cm) de los varones listados en el conjunto de datos 1 "Datos corporales" del apéndice B. ¿Qué nos dicen los números en esa gráfica de caja?



3. Comparación de gráficas de caja Consulte las gráficas de caja que se muestran a continuación, las cuales están dibujadas en la misma escala. Una gráfica representa el peso de los hombres y la otra el peso de las mujeres. ¿Qué gráfica de caja representa el peso de las mujeres? Explique.



4. Puntuaciones z Si su calificación en su próximo examen de estadística se convierte en una puntuación z, ¿cuál de las siguientes puntuaciones prefiere: -2.00, -1.00, 0, 1.00, 2.00? ¿Por qué?

Puntuaciones z. ***En los ejercicios 5 a 8, exprese todas las puntuaciones z con dos decimales.***

5. Velocidades de datos en ATL Para las velocidades de datos en aeropuertos de Verizon (Mbps) listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B, se midió la velocidad más alta de 77.8 Mbps en el aeropuerto internacional de Atlanta (ATL). La lista completa de 50 velocidades de datos para Verizon tiene una media de $\bar{x} = 17.60$ Mbps y una desviación estándar de $s = 16.02$ Mbps.

a. ¿Cuál es la diferencia entre la velocidad de datos para Verizon en el aeropuerto internacional de Atlanta y la media de todas las velocidades de datos para Verizon?

b. ¿Cuántas desviaciones estándar representa esto [la diferencia encontrada en el inciso (a)]?

c. Convierta la velocidad de datos para Verizon en el aeropuerto internacional de Atlanta a una puntuación z.

d. Si consideramos que las velocidades de datos que se convierten a puntuaciones z entre -2 y 2 no son significativamente bajas ni significativamente altas, ¿es la velocidad de Verizon en Atlanta significativa?

6. Velocidades de datos en PHL Repita el ejercicio anterior usando la velocidad de datos para Verizon, de 0.8 Mbps en el Aeropuerto Internacional de Filadelfia (PHL).

7. Pulsos de mujeres En el conjunto de datos 1 "Datos corporales" del apéndice B, se listan pulsos de mujeres. El pulso más bajo es de 36 latidos por minuto, la media de los pulsos listados es $\bar{x} = 74.0$ latidos por minuto y su desviación estándar es $s = 12.5$ latidos por minuto.

a. ¿Cuál es la diferencia entre el pulso de 36 latidos por minuto y el pulso medio de las mujeres?

b. ¿Cuántas desviaciones estándar representa esto [la diferencia encontrada en el inciso (a)]?

c. Convierta el pulso de 36 latidos por minuto en una puntuación z.

d. Si consideramos que los pulsos que se convierten en puntuaciones z entre -2 y 2 no son significativamente bajos ni significativamente altos, ¿es significativo el pulso de 36 latidos por minuto?

8. Desechos de plástico El conjunto de datos 31 "Peso de la basura" en el apéndice B contiene los pesos (en lb) del plástico desechado por los hogares. El peso más alto es 5.28 libras, la media de todos los pesos es $\bar{x} = 1.911$ libras y la desviación estándar de los pesos es $s = 1.065$ lb.

a. ¿Cuál es la diferencia entre el peso de 5.28 lb y la media de los pesos?

b. ¿Cuántas desviaciones estándar representa esto [la diferencia encontrada en el inciso (a)]?

c. Convierta el peso de 5.28 lb en una puntuación z.

d. Si consideramos que los pesos que se convierten en puntuaciones z entre -2 y 2 no son significativamente bajos ni significativamente altos, ¿es significativo el peso de 5.28 lb?

Valores significativos. ***En los ejercicios 9 a 12, considere un valor significativamente bajo si su puntuación z es menor o igual a −2 o significativamente alto si su puntuación z es mayor o igual a 2.***

9. ACT El examen ACT se utiliza con el fin de evaluar la preparación para la universidad. En un año reciente, la media del ACT fue de 21.1 y la desviación estándar fue de 5.1. Identifique las puntuaciones del ACT que son significativamente bajas o significativamente altas.

10. MCAT En un año reciente, las puntuaciones en el Medical College Admission Test (MCAT) tuvieron una media de 25.2 y una desviación estándar de 6.4. Identifique las puntuaciones del MCAT que son significativamente bajas o significativamente altas.

11. Monedas de 25 centavos El conjunto de datos 29 "Pesos de monedas" lista pesos (en gramos) de monedas de 25 centavos fabricados después de 1964. Estos pesos tienen una media de 5.63930 g y una desviación estándar de 0.06194 g. Identifique los pesos que son significativamente bajos o significativamente altos.

12. Diseño de asientos de avión En el proceso de diseño de asientos de avión se encontró que los hombres tienen un ancho de cadera con una media de 36.6 cm y una desviación estándar de 2.5 cm (con base en datos de la encuesta antropométrica de Gordon Clauser *et al.*). Identifique el ancho de cadera de los hombres que son significativamente bajos o significativamente altos.

Comparación de valores. ***En los ejercicios 13 a 16, use puntuaciones z para comparar los valores dados.***

13. Los hombres más altos y más bajos El hombre vivo más alto en el momento en que se escribió esto es Sultan Kosen, que tenía una estatura de 251 cm. El hombre vivo más bajo era Chandra Bahadur Dangi, con una estatura de 54.6 cm. Las estaturas de los hombres tienen una media de 174.12 cm y una desviación estándar de 7.10 cm. ¿Cuál de estos dos hombres tiene la estatura más extrema?

14. Conteo de glóbulos rojos Según el conjunto de datos 1 "Datos corporales" del apéndice B, los varones tienen conteos de glóbulos rojos con una media de 4.719 y una desviación estándar de 0.490, mientras que las mujeres tienen un conteo de glóbulos rojos con media de 4.349 y una desviación estándar de 0.402. ¿Quién tiene el conteo más alto en relación con la muestra de la cual provino: un hombre con un conteo de 5.58 o una mujer con un conteo de 5.23? Explique.

15. Pesos al nacer Según el conjunto de datos 4 "Nacimientos" del apéndice B, los varones recién nacidos tienen pesos con una media de 3272.8 g y una desviación estándar de 660.2 g. Las bebés recién nacidas tienen pesos con una media de 3037.1 g y una desviación estándar de 706.3 g. ¿Quién tiene el peso más extremo en relación con el grupo de donde provienen: un varón que pesó 1500 g o una bebé que pesó 1500 g?

16. Premios Oscar En la 87ª entrega de Premios de la Academia, Eddie Redmayne ganó el Oscar al mejor actor a la edad de 33 años y Julianne Moore el de mejor actriz a la edad de 54. Para todos los mejores actores, la edad media es 44.1 años y la desviación estándar es 8.9 años. Para todas las mejores actrices, la edad media es de 36.2 años y la desviación estándar es de 11.5 años. (Todas las edades se refieren al momento de la ceremonia de entrega de premios). En relación con sus géneros, ¿quién tenía la edad más extrema al ganar el Oscar: Eddie Redmayne o Julianne Moore? Explique.

Percentiles. ***En los ejercicios 17 a 20, utilice las siguientes velocidades de datos en aeropuertos para teléfonos celulares (Mbps) de Sprint. Encuentre el percentil correspondiente a la velocidad de datos dada.***

0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4
0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.7	0.8	1.0
1.1	1.1	1.2	1.2	1.6	1.6	2.1	2.1	2.3	2.4
2.5	2.7	2.7	2.7	3.2	3.4	3.6	3.8	4.0	4.0
5.0	5.6	8.2	9.6	10.6	13.0	14.1	15.1	15.2	30.4

17. 2.4 Mbps **18.** 13.0 Mbps **19.** 0.7 Mbps **20.** 9.6 Mbps

En los ejercicios 21 a 28, use la misma lista de velocidades de datos en aeropuertos para Sprint (Mbps) dada para los ejercicios 17 a 20. Encuentre el percentil o cuartil indicado.

21. P_{60} **22.** Q_1 **23.** Q_3 **24.** P_{40} **25.** P_{50} **26.** P_{75} **27.** P_{25} **28.** P_{85}

Gráficas de caja. ***En los ejercicios 29 a 32, utilice los datos dados para construir una gráfica de caja e identifique el resumen de 5 números.***

29. Citas rápidas Las siguientes son las calificaciones que han dado los varones a las mujeres en un experimento que implica citas rápidas.

2.0 3.0 4.0 5.0 6.0 6.0 7.0 7.0 7.0 7.0 7.0 7.0 8.0 8.0 8.0 8.0 9.0 9.5 10.0 10.

30. Radiación del teléfono celular A continuación se listan las tasas de absorción de radiación medidas (en W/kg) correspondientes a los siguientes teléfonos celulares: iPhone 5S, BlackBerry Z30, Sanyo Vero, Optimus V, Droid Razr, Nokia N97, Samsung Vibrant, Sony Z750a, Kyocera Kona, LG G2 y Virgin Mobile Supreme. Los datos son de la Comisión Federal de Comunicaciones.

1.18 1.41 1.49 1.04 1.45 0.74 0.89 1.42 1.45 0.51 1.38

31. Radiación en dientes de bebés A continuación se listan las cantidades de estroncio-90 (en milibecquerelios o mBq) en una muestra aleatoria simple de dientes de leche obtenida de residentes de Pensilvania nacidos después de 1979 (con base en datos de “An Unexpected Rise in Strontium-90 in U.S. Deciduous Teeth in the 1990s”, de Mangano *et al*., *Science of the Total Environment*).

128 130 133 137 138 142 142 144 147 149 151 151 151 155
156 161 163 163 166 172

32. Medidas de presión arterial Catorce estudiantes diferentes de segundo año de medicina en el Hospital Bellevue midieron la presión arterial de la misma persona. A continuación se indican las lecturas sistólicas (mm Hg).

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

Gráficas de cajas de los conjuntos de datos grandes en el apéndice B. ***En los ejercicios 33 a 36, utilice los conjuntos de datos dados en el apéndice B. Utilice las gráficas de caja para comparar los dos conjuntos de datos.***

33. Pulsos Utilice la misma escala para construir gráficas de caja para los pulsos de hombres y mujeres del conjunto de datos 1 “Datos corporales” en el apéndice B.

34. Edades de los ganadores del Oscar Use la misma escala para construir gráficas de cajas para las edades de las mejores actrices y mejores actores del conjunto de datos 14 “Edades de ganadores del Oscar” en el apéndice B.

35. IMC Utilice los índices de masa corporal (IMC) para hombres y mujeres que se listan en el conjunto de datos 1 “Datos corporales”.

36. Plomo e IQ Utilice la misma escala para construir las gráficas de caja para las puntuaciones de IQ completas (IQF) para el grupo de nivel bajo de plomo (grupo 1) y el grupo de nivel alto de plomo (grupo 3) en el conjunto de datos 1 “IQ y plomo” en el apéndice B.

3-3 Más allá de lo básico

37. Valores atípicos y gráficas de caja modificadas Repita el ejercicio 33 “Pulsos” utilizando gráficas de caja modificadas. Identifique los valores atípicos definidos en la parte 2 de esta sección.

Examen rápido del capítulo

1. Media de sueño Como parte de la Encuesta Nacional de Exámenes de Salud y Nutrición, se preguntó a los sujetos cuánto tiempo durmieron la noche anterior, y se reportaron los siguientes tiempos (horas): 8, 7, 5, 7, 4, 7, 6, 7, 8, 8, 8, 6. Encuentre la media.

2. Mediana de sueño ¿Cuál es la mediana de los valores muestrales listados en el ejercicio 1?

3. Moda de sueño ¿Cuál es la moda de los valores muestrales listados en el ejercicio 1?

4. Varianza de sueño La desviación estándar de los valores muestrales en el ejercicio 1 es de 1.3 horas. ¿Cuál es la varianza (incluyendo las unidades)?

5. Valores atípicos de sueño Si un tiempo de sueño de 0 horas se incluye en los datos muestrales dados en el ejercicio 1, ¿es un valor atípico? ¿Por qué sí o por qué no?

6. Puntuación z de sueño Una muestra mayor de 50 tiempos de sueño (horas) tiene una media de 6.3 horas y una desviación estándar de 1.4 horas. ¿Cuál es la puntuación z para un tiempo de sueño de 5 horas?

7. Q_3 de sueño Para una muestra de 80 tiempos de sueño, ¿aproximadamente cuántos de esos tiempos son menores que Q_3?

8. Resumen de 5 números de sueño Para una muestra de 100 tiempos de sueño, dé los nombres de los valores que constituyen el resumen de 5 números. (Los valores reales no se pueden identificar).

9. Estimación de s Una gran muestra de tiempos de sueño incluye valores que van desde un mínimo de 4 horas hasta un máximo de 10 horas. Utilice la regla práctica del rango para calcular la desviación estándar.

10. Notación de sueño Considere una muestra de los tiempos de sueño tomadas de la población adulta que vive en Alaska. Identifique los símbolos utilizados para la media de la muestra, la media de la población, la desviación estándar de la muestra, la desviación estándar de la población, la varianza de la muestra y la varianza de la población.

Ejercicios de repaso

1. Géiser Old Faithful A continuación se listan los errores de predicción (minutos) que son las diferencias entre los tiempos reales de erupción y los tiempos de erupción previstos. Los números positivos corresponden a erupciones que ocurrieron más tarde de lo previsto, y los números negativos corresponden a erupciones que ocurrieron antes del pronóstico. (Los datos provienen del conjunto de datos 23 “Old Faithful” en el apéndice B). Encuentre (a) la media; (b) la mediana; (c) la modo; (d) la mitad del rango; (e) el rango; (f) la desviación estándar; (g) la varianza; (h) Q_1; (i) Q_3

$$4 \quad -7 \quad 0 \quad 1 \quad -1 \quad 1 \quad -4 \quad -7 \quad 22 \quad 7 \quad -5 \quad 1$$

2. Puntuación z Usando los datos muestrales del ejercicio 1, encuentre la puntuación z correspondiente al error de predicción de 0 min. ¿Es ese error de predicción significativamente bajo o alto? ¿Por qué sí o por qué no?

3. Gráfica de caja Utilizando los mismos errores de predicción que se listan en el ejercicio 1, construya un diagrama de caja e incluya los valores del resumen de 5 números.

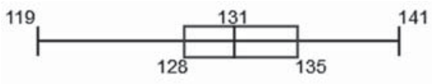
4. Códigos de emergencia En un análisis de las actividades que resultaron en lesiones cerebrales presentadas en las salas de emergencia de hospital, se identificaron las siguientes actividades mediante los códigos que se muestran entre paréntesis: ciclismo (12); fútbol (14); parque infantil (22); baloncesto (27); natación (40). Encuentre la media de 12, 14, 22, 27 y 40. ¿Qué tiene de erróneo este resultado?

5. Comparación de pesos al nacer El peso al nacer de una muestra de varones tiene una media de 3272.8 g y una desviación estándar de 660.2 g. El peso al nacer de una muestra de niñas tiene una media de 3037.1 g y una desviación estándar de 706.3 g (con base en el conjunto de datos 4 “Nacimientos” en el apéndice B). Cuando se considera entre los miembros del mismo sexo, ¿cuál bebé tiene el peso al nacer relativamente mayor: un varón con un peso al nacer de 3400 g o una niña con un peso al nacer de 3200 g? ¿Por qué?

6. Efectos de un valor atípico A continuación se listan los conteos de plaquetas (1000 células/μl) de los sujetos incluidos en el conjunto de datos 1 "Datos corporales". Identifique el valor atípico y luego comente el efecto que tiene sobre la media y la desviación estándar, mediante la determinación de los valores de esos estadísticos incluyendo el valor atípico y sin incluirlo.

263 206 185 246 188 191 308 262 198 253 646

7. Interpretación de una gráfica de caja A continuación se muestra una gráfica de caja de una muestra de 30 anchos máximos de cráneo (mm) medidos en cráneos egipcios de alrededor del año 4000 a.C. ¿Qué representan los números en la gráfica de caja?



8. Estimación de la desviación estándar A continuación se lista una muestra de tiempos de duración (segundos) de erupciones del géiser Old Faithful. Utilice la regla práctica del rango para estimar el valor de la desviación estándar de todos los tiempos de duración y compare el resultado con la desviación estándar de 33.7 segundos obtenida de una muestra de 2634 tiempos de duración.

226 228 247 247 253 256 250 254 229 242 250 241 226 240 117

Ejercicios de repaso acumulado

1. Arsénico en el arroz A continuación se listan las cantidades medidas (μg por porción) de arsénico en una muestra de porciones de arroz integral [datos de la Food and Drug Administration (FDA)]. Construya una distribución de frecuencias. Utilice una anchura de clase de 2 μg y use 0 μg como límite inferior de la primera clase.

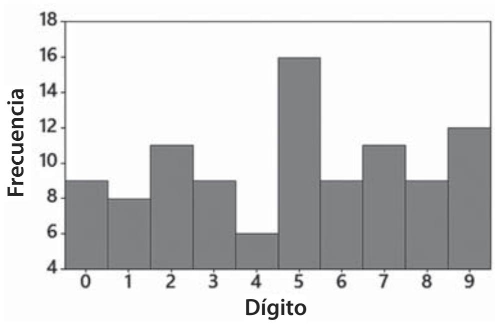
6.1 5.4 6.9 4.9 6.6 6.3 6.7 8.2 7.8 1.5 5.4 7.3

2. Histograma Utilice la distribución de frecuencias del ejercicio 1 para construir un histograma. Utilice los valores medios de clase para la escala horizontal.

3. Diagrama de tallo y hojas Utilice las cantidades de arsénico del ejercicio 1 para construir un diagrama de tallo y hojas.

4. Estadísticos descriptivos Utilice cantidades de arsénico en el ejercicio 1 y encuentre lo siguiente: (a) media, (b) mediana, (c) desviación estándar, (d) varianza y (e) rango. Incluya las unidades de medida apropiadas.

5. Histograma El histograma adjunto muestra los resultados de los dígitos de la lotería Florida Play 4. ¿Cuál es el error más importante en este histograma?



6. Distribución normal Examine la distribución mostrada en el histograma del ejercicio 5. ¿Parece que los datos muestrales provienen de una población con una distribución normal? ¿Por qué sí o por qué no?

Proyecto de tecnología

Palabras pronunciadas por hombres y mujeres Considere el conjunto de datos 24 "Conteos de palabras" en el apéndice B, que incluye los conteos de palabras pronunciadas por hombres y mujeres. Ese conjunto de datos incluye 12 columnas de datos, pero primero apile todos los conteos de palabras masculinas en una columna y todos los conteos de palabras femeninas en otra columna. A continuación, proceda a generar histogramas, o cualquier otra gráfica adecuada, y encuentre los estadísticos que le permiten comparar los dos conjuntos de datos. ¿Existen valores atípicos? ¿Ambos conjuntos de datos tienen propiedades que son básicamente iguales? ¿Hay diferencias significativas? ¿Cuál sería una consecuencia de tener diferencias significativas? Escriba un breve informe incluyendo sus conclusiones y gráficas de apoyo.

DE LOS DATOS A LA DECISIÓN

Humo de segunda mano

El conjunto de datos 12 "Humo pasivo y activo" en el apéndice B contiene medidas de cotinina en tres grupos de sujetos: (1) fumadores, (2) no fumadores expuestos al humo ambiental del tabaco y (3) no fumadores no expuestos al humo ambiental del tabaco. La cotinina es un indicador de la absorción de nicotina.

Pensamiento crítico

Utilice los métodos de este capítulo para explorar y comparar las medidas de cotinina en los tres grupos. ¿Hay diferencias notables? ¿Existen valores atípicos? ¿Qué concluye usted acerca de los efectos que los fumadores tienen sobre los no fumadores? Escriba un informe breve de sus conclusiones, y proporcione evidencia estadística de apoyo.

Actividades en equipo

1. Actividad en clase En clase, cada estudiante debe registrar dos pulsos contando el número de latidos en 1 minuto. El primer pulso se debe medir mientras el estudiante está sentado, y el segundo pulso debe medirse mientras el estudiante está de pie. Utilice los métodos de este capítulo para comparar los resultados. ¿Los hombres y las mujeres parecen tener diversos pulsos? ¿Los pulsos medidos mientras están sentados parecen ser diferentes de los pulsos medidos mientras están de pie?

2. Actividades fuera de clase En el artículo "Weighing Anchors" en la revista *Omni*, el autor John Rubin observó que cuando la gente estima un valor, su estimación suele estar "anclada" (o influenciada por) un número precedente, incluso si ese número anterior no está relacionado con la cantidad que se estima. Para demostrar esto, pidió a las personas que dieran una estimación rápida del valor de $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$. La media de las respuestas dadas fue 2250, pero cuando el orden de los números fue invertido, la media se convirtió en 512. Rubin explicó que cuando comenzamos los cálculos con números más grandes (como en $8 \times 7 \times 6$), nuestras estimaciones tienden a ser más grandes. Señaló que tanto 2250 como 512 están muy por debajo del producto correcto, 40,320. El artículo sugiere que hay cifras irrelevantes que pueden influir en las evaluaciones de bienes raíces, las estimaciones de los valores de los automóviles y las estimaciones de la probabilidad de una guerra nuclear.

Realice un experimento para probar esta teoría. Seleccione algunos sujetos y pídales que estimen rápidamente el valor de

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

Luego 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 las estimaciones junto con el orden particular utilizado. Diseñe cuidadosamente el experimento para que las condiciones sean uniformes y los dos grupos muestrales se seleccionen de manera que se minimice cualquier sesgo. No describa la teoría a los sujetos hasta después de haber proporcionado sus estimaciones. Compare los dos conjuntos de resultados de la muestra utilizando los métodos de este capítulo. Proporcione un informe impreso que incluya los datos recopilados, los métodos detallados que se usaron, el método de análisis, las gráficas y/o estadísticos pertinentes y una declaración de conclusiones. Incluya una crítica del experimento, con razones por las cuales los resultados pudieron no ser correctos, y describa maneras en las cuales el experimento podría ser mejorado.

3. Actividad en clase Complete la siguiente tabla y compare la variación en las velocidades de datos para teléfono inteligente. Consulte el conjunto de datos 32 "Velocidades de datos en aeropuertos".

	Verizon	Sprint	AT&T	T-Mobile
Desviación estándar				
Rango				

4. Actividad fuera de clase Registre los tiempos que los automóviles están estacionados en una bomba de gasolina, y describa las características importantes de esos tiempos.

5. Actividad fuera de clase Varios sitios web, como www.gasbuddy.com, están diseñados para proporcionar una lista de los precios locales de la gasolina. Obtenga una lista de tales precios y explore los datos usando los métodos de este capítulo y el capítulo 2.

6. Actividad fuera de clase El conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B incluye conteos de chispas de chocolate en cinco diferentes marcas de galletas. Obtenga su propia muestra de galletas de chispas de chocolate y proceda a contar el número de chispas en cada galleta. Utilice los datos para generar un histograma y cualquier otra gráfica adecuada. Encuentre los estadísticos descriptivos. Compare sus conteos de chispas de chocolate con los dados en el conjunto de datos 28. ¿Hay alguna diferencia? Explique.

7. Actividad fuera de clase El apéndice B incluye muchos conjuntos de datos reales e interesantes. En grupos de tres o cuatro estudiantes, seleccionen un conjunto de datos del apéndice B y analícenlos usando los métodos discutidos hasta ahora en este libro. Escriban un breve informe resumiendo las conclusiones clave.

8. Actividad fuera de clase En grupos de tres o cuatro estudiantes, recolecten un conjunto de valores de datos originales en el intervalo o nivel de relación de medida. Proporcionen lo siguiente: (1) una lista de los valores muestrales; (2) los resultados de un software para estadísticos descriptivos y gráficas; y (3) una descripción escrita de la naturaleza de los datos, el método de recolección y las características importantes.

4

PROBABILIDAD

PROBLEMA DEL CAPÍTULO

Prueba de drogas a solicitantes de empleo

Aproximadamente 85% de las empresas de Estados Unidos evalúan a sus empleados y/o a los solicitantes de empleo en cuanto al consumo de drogas. Una prueba de orina común y barata (con un costo aproximado de $50) es la prueba EMIT (Técnica de inmunoensayo multiplicado por enzimas), que prueba la presencia de alguna de las siguientes drogas: marihuana, cocaína, anfetaminas, opiáceos o fenciclidina. La mayoría de las empresas exigen que los resultados de las pruebas positivas sean confirmados por una prueba más confiable llamada GC-MS (Espectrometría de masa de cromatografía gaseosa).

Al igual que casi todos los ensayos médicos, las pruebas de detección de drogas pueden resultar erróneas. Los resultados erróneos son de dos tipos: (1) resultados falsos positivos y (2) resultados falsos negativos. En la sociedad actual, estos términos deben entenderse con claridad. Un solicitante de empleo o empleado que obtenga un resultado falso positivo es alguien que se evalúa incorrectamente como usuario de drogas cuando no lo es. Este tipo de error puede resultar injustamente en la negación del trabajo o la terminación del empleo.

Análisis de los resultados

La tabla 4-1 incluye los resultados de 555 adultos en Estados Unidos. Si uno de los sujetos de la tabla 4-1 se selecciona aleatoriamente entre aquellos que no usan drogas, ¿cuál es la probabilidad de un resultado falso positivo? Si uno de los sujetos de la tabla 4-1 es seleccionado aleatoriamente entre aquellos que no usan drogas, ¿cuál es la probabilidad de un resultado verdadero negativo? Abordaremos estas preguntas en el presente capítulo.

- **Prevalencia:** Se considera la proporción de la población que tiene la condición (como el uso o la adicción a alguna droga).
- **Falso positivo:** Resultado erróneo de la prueba que indica incorrectamente que el sujeto posee una condición cuando no la tiene.
- **Falso negativo:** Resultado erróneo de la prueba que indica incorrectamente que el sujeto no posee una condición cuando sí la tiene.
- **Verdadero positivo:** Resultado correcto de la prueba que indica que un sujeto posee una condición cuando el sujeto la tiene.
- **Verdadero negativo:** Resultado correcto de la prueba que indica que un sujeto no posee una condición cuando no la tiene.
- **Sensibilidad de la prueba:** La probabilidad de un resultado verdadero positivo de la prueba, cuando el sujeto realmente tiene la condición que se está probando.
- **Especificidad de la prueba:** La probabilidad de un resultado verdadero negativo de la prueba, cuando el sujeto no tiene la condición que está siendo probada.
- **Valor predictivo positivo:** Probabilidad de que un sujeto realmente tenga la condición, dado que la prueba produce un resultado positivo (indicando que la condición está presente).
- **Valor predictivo negativo:** Probabilidad de que el sujeto no tenga realmente la condición, dado que la prueba produce un resultado negativo (indicando que la condición no está presente).

TABLA 4-1 Resultados de las pruebas de drogas a solicitantes de empleo

	Resultado positivo de la prueba (El examen muestra el uso de drogas)	**Resultado negativo de la prueba** (El examen no muestra uso de drogas)
El sujeto usa drogas	45 (Verdadero positivo)	5 (Falso negativo)
El sujeto no usa drogas	25 (Falso positivo)	480 (Verdadero negativo)

OBJETIVOS DEL CAPÍTULO

El objetivo principal de este capítulo es desarrollar una sólida comprensión de los valores de probabilidad, debido a que éstos constituyen el fundamento subyacente en el que se construyen los métodos de estadística inferencial. Los métodos importantes de la prueba de hipótesis usan comúnmente valores de P, que son valores de probabilidad expresados como números entre 0 y 1, inclusive. Los valores de probabilidad más pequeños, como 0.01, corresponden a eventos que son muy poco probables. Los valores de probabilidad más grandes, como 0.99, corresponden a eventos que son muy probables. Los objetivos del capítulo son:

4-1 Conceptos básicos de probabilidad

- Identificar las probabilidades como valores entre 0 y 1, e interpretar esos valores como expresiones de la posibilidad de ocurrencia de los eventos.
- Desarrollar la capacidad de calcular probabilidades de eventos.
- Definir el complemento de un evento y calcular la probabilidad de ese complemento.

4-2 Regla de la suma y regla de la multiplicación

- Desarrollar la capacidad de calcular la probabilidad de que en un ensayo sencillo ocurra algún evento *A* o algún evento *B*, o que sucedan ambos. Aplicar la regla de la suma ajustando correctamente los eventos que no son disjuntos (o se superponen).
- Desarrollar la capacidad de calcular la probabilidad de que un evento *A* ocurra en un primer ensayo y un evento *B* ocurra en un segundo ensayo. Aplicar la regla de la multiplicación ajustando los eventos que no son independientes.
- Distinguir entre eventos independientes y eventos dependientes.

4-3 Complementos, probabilidad condicional y teorema de Bayes

- Calcular la probabilidad de "al menos una" ocurrencia de un evento *A*.
- Aplicar la regla de la multiplicación para calcular la probabilidad de algún evento, dado que otro evento ya ha ocurrido.

4-4 Conteo

- Desarrollar la capacidad de aplicar la regla fundamental del conteo, la regla del factorial, la regla de las permutaciones y la regla de las combinaciones.
- Distinguir entre las circunstancias que requieren la regla de las permutaciones y las que requieren la regla de las combinaciones.

4-1 Conceptos básicos de probabilidad

Concepto clave La parte 1 de esta sección incluye conceptos básicos de probabilidad. El objetivo más importante de esta sección es aprender a interpretar las probabilidades, que se expresan como valores entre 0 y 1. Una probabilidad pequeña, como 0.001, corresponde a un evento que ocurre en pocas ocasiones.

La parte 2 de esta sección incluye la *ventaja comparativa* y cómo se relaciona con la probabilidad. Los conceptos relacionados con la ventaja comparativa no son necesarios para los temas de los siguientes capítulos, pero se utilizan comúnmente en situaciones como loterías y apuestas.

PARTE 1 Conceptos básicos de probabilidad

Papel de la probabilidad en estadística

La probabilidad tiene un papel central en el importante método estadístico de *prueba de hipótesis* que se presenta más adelante, en el capítulo 8. Los estadísticos toman decisiones usando datos que rechazan explicaciones (como la casualidad) basadas en *probabilidades* muy bajas. Vea el siguiente ejemplo que ilustra el papel de la probabilidad y la manera fundamental en la que los estadísticos piensan.

EJEMPLO 1 Análisis de una afirmación

Algunos investigadores han afirmado lo siguiente (realmente lo han afirmado):

Afirmación: "Hemos desarrollado un método de selección de género que aumenta considerablemente la probabilidad de que un bebé sea niña".

Hipótesis utilizada al probar la afirmación anterior: El método de selección de género *no tiene ningún efecto*, de modo que para las parejas que usan este método, alrededor del 50% de los recién nacidos resultan ser niñas.

continúa

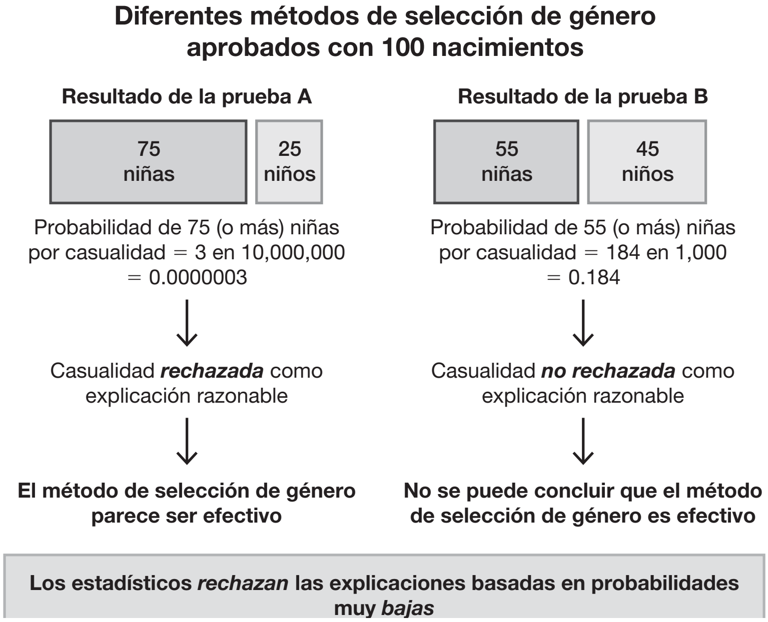


FIGURA 4-1 Datos y conclusiones de pruebas al método de selección de género

La figura 4-1 presenta los datos muestrales de dos pruebas de 100 parejas usando el método de selección de género y la conclusión alcanzada para cada prueba.

INTERPRETACIÓN

Entre los 100 bebés, tanto 75 como 55 niñas es más que las 50 niñas que normalmente esperaríamos, pero sólo el caso de las 75 niñas nos lleva a creer que el método de selección de género es efectivo. A pesar de que hay una probabilidad de obtener 75 niñas en 100 nacimientos sin tratamiento especial, la probabilidad de que eso ocurra es tan pequeña (0.0000003) que debemos rechazar el azar como una explicación razonable. En cambio, se reconocería generalmente que los resultados proporcionan un fuerte apoyo a la afirmación de que el método de selección de género es efectivo. Así es exactamente como piensan los estadísticos: Rechazan las explicaciones (como la casualidad) basadas en *probabilidades* muy bajas.

SU TURNO Resuelva el ejercicio 37 “Predicción del género”.

Fundamentos de probabilidad

En probabilidad se analizan procedimientos (como la generación de nacimientos de hombres/mujeres o la respuesta a una pregunta de un examen de opción múltiple) que producen resultados.

DEFINICIONES

Un **evento** es cualquier colección de resultados o consecuencias de un procedimiento.

Un **evento simple** es un resultado o un evento que no se puede descomponer en componentes más simples.

El **espacio muestral** para un procedimiento consiste en todos los posibles eventos *simples*. Es decir, el espacio muestral consta de todos los resultados que no se pueden descomponer más.

En el ejemplo 2 se ilustran los conceptos definidos anteriormente.

EJEMPLO 2 Eventos simples y espacios muestrales

En la tabla siguiente, se utiliza “h” para expresar un hombre y “m” para expresar una mujer.

Procedimiento	Ejemplo de evento	Espacio muestral: lista completa de eventos simples
Un nacimiento	1 mujer (evento simple)	{h, m}
3 nacimientos	2 hombres y 1 mujer (hhm, hmh y mhh son todos eventos simples que resultan en 2 hombres y 1 mujer).	{hhh, hhm, hmh, hmm, mhh, mhm, mmh, mmm}

Eventos simples:

- Con un nacimiento, el resultado de 1 mujer es un *evento simple* y el resultado de 1 hombre es otro evento simple. Son eventos simples individuales porque no se pueden descomponer más.
- Con tres nacimientos, el resultado de 2 mujeres seguidas por un hombre (mmh) es un evento simple.
- Cuando se lanza un solo dado, el resultado de 5 es un evento simple, pero el resultado de un número par no es un evento simple.

Eventos que no son simples: Con tres nacimientos, el evento de “2 mujeres y 1 hombre” *no es un evento simple* porque puede ocurrir con los siguientes eventos simples diferentes: mmh, mhm, hmm.

Espacio muestral: Con tres nacimientos, el espacio muestral consta de los ocho eventos simples listados en la tabla anterior.

SU TURNO Resuelva el ejercicio 35 “Cuatro niños”.

Probabilidades que desafían la intuición

En ciertos casos, nuestras estimaciones subjetivas de valores de probabilidad difieren en forma drástica de las probabilidades reales. He aquí un ejemplo clásico: si usted inhala profundamente, hay una probabilidad mayor al 99% de que inhale una molécula que haya sido exhalada en el último aliento de Julio César al morir. Con este mismo ánimo morboso y poco intuitivo, si la taza con cicuta que mató a Sócrates contenía agua en su mayor parte, el siguiente vaso de agua que usted beba muy probablemente contendrá una de esas mismas moléculas. He aquí otro ejemplo, pero menos morboso, que puede verificarse: en grupos de 25 estudiantes, la probabilidad de que al menos 2 de ellos cumplan años el mismo día (y mes) es mayor que 50%.

Tres métodos comunes para encontrar la probabilidad de un evento

Primero introducimos una notación básica, luego presentamos tres métodos comunes para encontrar la probabilidad de un evento.

Notación para probabilidades

P expresa una probabilidad.

A, *B* y *C* expresan eventos específicos.

$P(A)$ expresan la “probabilidad de que ocurra el evento *A*”.

Los siguientes tres métodos para encontrar probabilidades dan como resultado valores entre 0 y 1: $0 \leq P(A) \leq 1$. La figura 4-2 muestra los posibles valores de las probabilidades y las expresiones de verosimilitud más familiares y comunes.

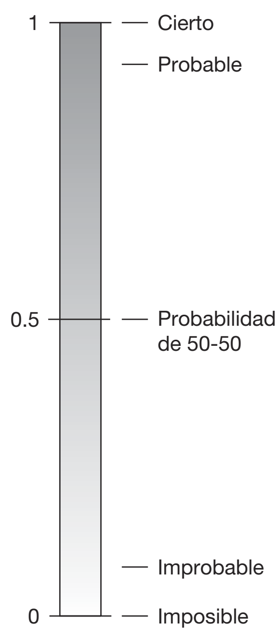


FIGURA 4-2 Valores posibles para las probabilidades

1. **Aproximación de la probabilidad por frecuencias relativas** Realice (u observe) un procedimiento y cuente el número de veces que ocurre el evento *A*. Entonces, $P(A)$ se *aproxima* de la siguiente manera:

$$P(A) = \frac{\text{número de veces que ocurrió } A}{\text{número de veces que se repitió el procedimiento}}$$

Al referirse a las aproximaciones de probabilidades por frecuencias relativas, este texto no distinguirá entre los resultados que son probabilidades exactas y las que son aproximaciones; por ello, una instrucción para "encontrar la probabilidad" podría significar "*estimar* la probabilidad".

2. **Método clásico de la probabilidad (requiere resultados igualmente probables)** Si un procedimiento tiene n eventos simples diferentes que son *igualmente probables*, y si el evento A puede ocurrir de s diferentes maneras, entonces

$$P(A) = \frac{\text{número de formas en que ocurre } A}{\text{número de eventos simples diferentes}} = \frac{s}{n}$$

PRECAUCIÓN Cuando use el método clásico, siempre confirme que los resultados son *igualmente probables*.

3. **Probabilidades subjetivas** $P(A)$ La probabilidad del evento A, se *estima* utilizando el conocimiento de las circunstancias relevantes.

La figura 4-3 ilustra los métodos de las tres definiciones anteriores.

1. Método de la frecuencia relativa: Al tratar de determinar la probabilidad de que un automóvil choque en un año dado, debemos examinar los resultados pasados para determinar el número de autos que se usaron en un año y el número de ellos que chocaron; entonces encontramos la proporción del número de automóviles que chocaron sobre el número total de autos. Para un año reciente, el resultado es una probabilidad de 0.0480. (Vea el ejemplo 3).

2. Método clásico: Al tratar de determinar la probabilidad de ganar el premio mayor en una lotería seleccionando seis números diferentes entre 1 y 60, cada combinación tiene la misma probabilidad de ocurrir. La probabilidad de ganar es 0.0000000200, que se puede encontrar usando los métodos presentados en la sección 4-4. (Vea el ejemplo 4).

3. Probabilidad subjetiva: Al tratar de estimar la probabilidad de quedar atrapado en el siguiente ascensor al que subamos, por experiencia personal sabemos que la probabilidad es bastante pequeña. Vamos a estimar que sea, por ejemplo, 0.001 (equivalente a 1 oportunidad en 1000). (Vea el ejemplo 5).

FIGURA 4-3 Tres métodos para encontrar una probabilidad

Simulaciones En ocasiones no es posible usar ninguno de los tres métodos anteriores. Una *simulación* de un procedimiento es un proceso que se comporta de la misma manera que el procedimiento en sí para que se produzcan resultados similares. Las probabilidades se pueden encontrar a veces mediante una simulación. Vea el proyecto de tecnología, casi al final de este capítulo.

REDONDEO DE PROBABILIDADES

Es difícil proporcionar una regla universal para redondear los valores de probabilidad, pero esta guía se puede aplicar a la mayoría de los problemas de este texto: al expresar el valor de una probabilidad, dé la fracción exacta o decimal o redondee los resultados decimales finales a *tres* dígitos significativos. (*Sugerencia:* cuando una probabilidad no sea una fracción simple como 2/3 o 5/9, exprésela como decimal para que el número pueda entenderse de mejor manera). Todos los dígitos de un número son significativos excepto los ceros que se incluyen para la colocación correcta del punto decimal. Vea los siguientes ejemplos.

- La probabilidad de 0.4450323339 (del ejemplo 6) tiene diez dígitos significativos (4450323339), y puede redondearse a tres dígitos significativos como 0.445.
- La probabilidad de 1/3 se puede dejar como una fracción o redondearse a 0.333. (*No* redondee a 0.3).
- La probabilidad de 2/8 se puede expresar como 1/4 o 0.25. (Debido a que 0.25 es exacto, no hay necesidad de expresarlo con tres dígitos significativos como 0.250).

¿Probabilidades expresadas como porcentajes? Matemáticamente, una probabilidad de 0.25 es equivalente al 25%, pero hay buenas razones para apegarse a las fracciones y a los decimales, y no utilizar porcentajes. Casi universalmente, las revistas profesionales expresan las probabilidades con decimales, no como porcentajes. Más adelante en este libro, usaremos valores de probabilidad generados a partir de software estadístico, y siempre estarán en forma de decimales.

Al encontrar probabilidades con el método de frecuencias relativas, obtenemos una *aproximación* en lugar de un valor exacto. A medida que aumenta el número total de observaciones, las aproximaciones correspondientes tienden a acercarse a la probabilidad real. Esta propiedad se conoce comúnmente como la *ley de los grandes números*.

LEY DE LOS GRANDES NÚMEROS

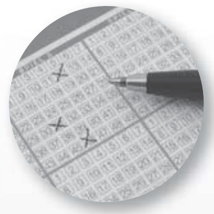
Cuando un procedimiento se repite una y otra vez, la probabilidad por frecuencia relativa de un evento tiende a acercarse a la probabilidad real.

La ley de los grandes números indica que las aproximaciones por frecuencia relativa tienden a mejorar con más observaciones. Esta ley refleja una noción sencilla apoyada por el sentido común: una estimación de la probabilidad basada en sólo unos cuantos ensayos puede estar equivocada en una cantidad sustancial, pero con un número muy grande de ensayos, la estimación tiende a ser mucho más precisa.

PRECAUCIONES

1. La ley de los grandes números se aplica al comportamiento en un gran número de ensayos y no a algún resultado individual. En ocasiones, los apostadores pierden ingenuamente grandes sumas de dinero por pensar incorrectamente que una cadena de pérdidas aumenta las posibilidades de ganar en la próxima apuesta, o que es probable que una serie de ganancias continúe.
2. Si no sabemos nada sobre la verosimilitud de diferentes resultados posibles, no debemos suponer que son igualmente probables. Por ejemplo, no debemos pensar que la probabilidad de pasar el siguiente examen de estadística es 1/2, o 0.5 (porque pasamos el examen o porque no lo hicimos). La probabilidad real depende de factores como la preparación y la dificultad del examen.

Comprensión de las posibilidades de ganar la lotería



En el juego Mega Millions de la Lotería del Estado de Nueva York, se deben elegir cinco números diferentes del 1 al 75, y también es necesario seleccionar otro número llamado “Mega Ball” del 1 al 15. Para ganar el premio, se deben obtener los cinco números correctos y el número correcto de “Mega Ball”. La posibilidad de ganar el premio con un billete es de 1/258,890,850. Los anuncios para esta lotería dicen “todo lo que usted necesita es un poco de suerte”, pero en realidad se requiere una enorme cantidad de suerte. La probabilidad de 1/258,890,850 no es tan fácil de entender, así que consideremos una útil analogía sugerida por el Hermano Donald Kelly del Colegio Marista. Una pila de 258,890,850 monedas de 25 centavos tiene unas 282 millas de altura. Los aviones comerciales suelen volar aproximadamente a 7 millas de altura, por lo que la altura de esta pila de monedas es aproximadamente 40 veces más grande que la altitud crucero de un jet comercial. La posibilidad de ganar el juego de lotería Mega Millions es equivalente a la posibilidad de seleccionar al azar una moneda específica de esa pila que tiene 282 millas de altura. “Cualquiera de nosotros que gaste dinero en esta lotería debe entender que la posibilidad de ganar el premio es muy, muy, muy cercana a cero”.

¿Qué tan probable es?

¿De qué manera interpretamos términos como *probable*, *improbable* o *extremadamente improbable*? La Administración Federal de Aviación (FAA, por sus siglas en inglés) interpreta estos términos de la siguiente forma:

- *Probable*: Una probabilidad del orden de 0.00001 o mayor por cada hora de vuelo. Se espera que estos eventos ocurran varias veces durante la vida útil de un avión.
- *Improbable*: Una probabilidad del orden de 0.00001 o menos. Se espera que estos eventos no ocurran durante la vida útil total de un avión de un tipo particular, pero pueden ocurrir durante la vida útil total de todos los aviones de un tipo particular.
- *Extremadamente improbable*: Una probabilidad del orden de 0.000000001 o menos. Estos eventos son tan poco probables que se debe considerar que nunca ocurrirán.

EJEMPLO 3 Probabilidad por frecuencia relativa: Paracaidismo

Encuentre la probabilidad de morir al hacer un salto en paracaídas.

SOLUCIÓN

En un año reciente, hubo alrededor de 3,000,000 saltos en paracaídas y 21 de ellos resultaron en muertes. Utilizamos el método de la frecuencia relativa de la siguiente manera:

$$P(\text{muerte por paracaidismo}) = \frac{\text{número de muertes por paracaidismo}}{\text{número total de saltos en paracaídas}} = \frac{21}{3{,}000{,}000} = 0.000007$$

Aquí no se puede usar el método clásico porque los dos resultados (morir, sobrevivir) no son igualmente probables. Es posible estimar una probabilidad subjetiva en ausencia de datos históricos.

SU TURNO Resuelva el ejercicio 25 “Selección de género XSORT”.

EJEMPLO 4 Probabilidad clásica: tres hijos del mismo sexo

Cuando nacen tres hijos, el espacio muestral de los géneros es como se muestra en el ejemplo 1: {hhh, hhm, hmh, hmm, mhh, mhm, hhm, mmm}. Si los hombres y las mujeres son igualmente probables, entonces esos ocho eventos simples son igualmente probables. Suponiendo que los hombres y las mujeres son igualmente probables, encuentre la probabilidad de tener tres hijos todos del mismo sexo. (En realidad, la probabilidad de tener un hombre es 0.512 en lugar de 0.5).

SOLUCIÓN

El espacio muestral {hhh, hhm, hmh, hmm, mhh, mhm, hhm, mmm} incluye ocho resultados igualmente probables, y hay exactamente dos resultados en los que los tres niños son del mismo sexo: hhh y mmm. Podemos utilizar el método clásico para obtener

$$P(\text{tres hijos del mismo sexo}) = \frac{2}{8} = \frac{1}{4} = \text{o } 0.25$$

SU TURNO Resuelva el ejercicio 33 “Tres hijos”.

EJEMPLO 5 Probabilidad subjetiva: el dinero de Katy Perry

¿Cuál es la probabilidad de que el siguiente billete de dólar que usted gaste haya sido previamente usado por Katy Perry?

SOLUCIÓN

En ausencia de datos sobre el número de billetes de dólar en circulación y los hábitos de gasto de Katy Perry, hacemos una estimación subjetiva. La experiencia sugiere que la probabilidad es muy pequeña. Estimemos que es, digamos, 0.00001 (equivalente a 1 oportunidad en 100,000). Dependiendo de nuestro conocimiento de las circunstancias relevantes, esa estimación subjetiva podría ser razonablemente exacta o podría ser groseramente errónea.

SU TURNO Resuelva el ejercicio 4 “Probabilidad subjetiva”.

EJEMPLO 6 Envío de textos y conducción

En un estudio de los conductores que estudian la escuela preparatoria en Estados Unidos, se encontró que 3785 enviaron mensajes de texto por el teléfono celular mientras conducían durante los últimos 30 días, y 4120 no enviaron mensajes de texto mientras conducían en ese mismo período (basado en datos de "Texting While Driving ..." por Olsen, Shults, Eaton, *Pediatrics*, Vol. 131, núm. 6). Con base en estos resultados, si se selecciona al azar un conductor que estudia la preparatoria, determine la probabilidad de que haya enviado mensajes de texto mientras conducía durante los 30 días anteriores.

SOLUCIÓN

> **PRECAUCIÓN** Un *error* común consiste en utilizar ciegamente los números para obtener la probabilidad incorrecta de 3785/4720 = 0.802. Debemos *pensar* en lo que estamos haciendo, como se muestra a continuación.

En vez de tratar de determinar una respuesta directamente a partir de la afirmación dada, en primer lugar resuma la información en un formato que permita una comprensión clara, por ejemplo de la siguiente manera:

3785	enviaron mensajes de texto mientras conducían
4120	no enviaron mensajes de texto mientras conducían
8505	número total de conductores en la muestra

Ahora podemos usar el método de frecuencia relativa como sigue:

$$P(\text{mensajes de texto al conducir}) = \frac{\begin{array}{c}\text{número de conductores que enviaron}\\ \text{mensajes de texto al conducir}\end{array}}{\text{número total de conductores en la muestra}} = \frac{3785}{8505}$$
$$= 0.445$$

INTERPRETACIÓN

Hay una probabilidad de 0.445 de que, si se elige aleatoriamente un conductor de la escuela preparatoria, él o ella haya enviado mensajes de texto mientras conducía durante los 30 días anteriores.

SU TURNO Resuelva el ejercicio 27 "Genética mendeliana".

> **PRECAUCIÓN** No cometa el error común de encontrar descuidadamente un valor de probabilidad dividiendo el número menor por el número mayor. En vez de esto, *piense* cuidadosamente sobre los números involucrados y lo que representan. Identifique escrupulosamente el número total de elementos que se están considerando, como se ilustra en el ejemplo 6.

EJEMPLO 7 Día de Acción de Gracias

En un año seleccionado al azar, encuentre la probabilidad de que el Día de Acción de Gracias en Estados Unidos sea (a) un miércoles o (b) un jueves.

SOLUCIÓN

a. En Estados Unidos, el Día de Acción de Gracias siempre se festeja el cuarto jueves de noviembre. Por lo tanto, es imposible que el Día de Acción de Gracias sea un miércoles. Cuando un evento es imposible, su probabilidad es 0. *P*(Acción de Gracias en miércoles) = 0.

continúa

En cifras

10^{80}: Número de partículas en el universo observable. La probabilidad de que un mono golpee al azar un teclado y escriba el *Hamlet* de Shakespeare es $10^{-216,159}$.

b. Es cierto que un Día de Acción de Gracias en Estados Unidos será jueves. Cuando es seguro que un evento ocurra, su probabilidad es 1.

P(Acción de Gracias en jueves) = 1.

Debido a que cualquier evento imaginable es imposible, cierto o se encuentra en algún lugar intermedio, se deduce que la probabilidad matemática de cualquier evento A es 0, 1 o un número entre 0 y 1 (como se muestra en la figura 4-2). Esto es, $0 \leq P(A) \leq 1$.

SU TURNO Resuelva los ejercicios 19 "Clavija cuadrada" y 20 "Muerte e impuestos".

Eventos complementarios

En ocasiones necesitamos determinar la probabilidad de que un evento A *no* ocurra.

DEFINICIÓN

El **complemento** del evento A, expresado por $\overline{A}$, consiste en todos los resultados en los cuales el evento A *no* ocurre.

EJEMPLO 8 Complemento de muerte por paracaidismo

El ejemplo 3 muestra que en un año reciente hubo 3,000,000 saltos en paracaídas y 21 de ellos resultaron en una muerte. Determine la probabilidad de *no* morir al hacer un salto en paracaídas.

SOLUCIÓN

Entre 3,000,000 de saltos hubo 21 muertes, por lo que se deduce que en los otros 2,999,979 saltos la persona sobrevivió. Obtenemos

$$P(\text{no morir al saltar en paracaídas}) = \frac{2{,}999{,}979}{3{,}000{,}000} = 0.999993$$

INTERPRETACIÓN

La probabilidad de *no* morir al hacer un salto en paracaídas es 0.999993.

SU TURNO Resuelva el ejercicio 29 "Redes sociales".

Relación entre $P(A)$ y $P(\overline{A})$. Si expresamos el evento de morir en un salto en paracaídas por D, en el ejemplo 3 se mostró que $P(D) = 0.000007$ y que $P(\overline{D}) = 0.999993$. La probabilidad de $P(\overline{D})$ se puede encontrar simplemente restando $P(D)$ de 1.

Identificación de resultados significativos con probabilidades: la regla del evento inusual para estadística inferencial

Si, en un supuesto dado, la probabilidad de un evento particular observado es muy pequeña y el evento observado ocurre *significativamente menos* o *significativamente más* de lo que normalmente esperaríamos con ese supuesto, concluimos que el supuesto probablemente no es correcto.

Podemos usar probabilidades para identificar valores que son *significativamente bajos* o *significativamente altos* como sigue.

Uso de probabilidades para determinar cuándo los resultados son significativamente altos o significativamente bajos

- **Número significativamente *alto* de éxitos:** x éxitos entre n ensayos es un número *significativamente alto* de éxitos si la probabilidad de x o más éxitos es improbable con una probabilidad de 0.05 o menos. Es decir, x es un número significativamente alto de éxitos si $P(x \text{ o más}) \leq 0.05$.*
- **Número significativamente *bajo* de éxitos:** x éxitos entre n ensayos son un número *significativamente bajo* de éxitos si la probabilidad de x o menos éxitos es improbable con una probabilidad de 0.05 o menos. Es decir, x es un número significativamente bajo de éxitos si $P(x \text{ o menos}) \leq 0.05$.*

*El valor 0.05 no es absolutamente rígido. Podrían usarse otros valores, como 0.01, para distinguir entre los resultados que pueden ocurrir fácilmente por casualidad y los eventos que es muy poco probable que ocurran de esa manera.

Vea el ejemplo 1 en la página 133, que ilustra lo siguiente:

- Entre 100 nacimientos, 75 mujeres es *significativamente alto* porque la probabilidad de 75 o más mujeres es de 0.0000003, que es menor o igual a 0.05 (por lo que el método de selección de género parece ser eficaz).
- Entre 100 nacimientos, 55 mujeres *no es significativamente alto* porque la probabilidad de 55 o más mujeres es de 0.184, que es superior a 0.05 (por lo que la selección de género no parece ser efectiva).

Repaso de probabilidad

- **La probabilidad de un evento es una fracción o un número decimal entre 0 y 1 inclusive.**
- **La probabilidad de un evento imposible es 0.**
- **La probabilidad de que ocurra un evento seguro o cierto es 1.**
- **Notación: $P(A)$ = la probabilidad del evento A.**
- **Notación: $P(\overline{A})$ = la probabilidad de que el evento A *no* ocurra.**

PARTE 2 Posibilidades

Las expresiones de probabilidad se dan a menudo como *posibilidades*, tales como 50:1 (o "50 a 1"). Las siguientes, son las ventajas de las probabilidades y las posibilidades:

- Las posibilidades facilitan el manejo de las transferencias de dinero asociadas con el juego. (Es por eso que las posibilidades se usan comúnmente en casinos, loterías e hipódromos).
- Las probabilidades facilitan los cálculos, por lo que tienden a ser utilizadas por los estadísticos, matemáticos, científicos e investigadores en todos los campos.

Apostar para ganar

En una lotería estatal típica, la "casa" tiene una ventaja del 65 o 70%, ya que sólo entre el 30 y 35% del dinero apostado se devuelve en forma de premios. La ventaja de la casa en los hipódromos suele ser de alrededor del 15%. En los casinos, la ventaja de la casa es del 5.26% en el caso de la ruleta, 1.4% en el juego de dados, y del 3 al 22% en las máquinas tragamonedas.

La ventaja para la casa en el *Veintiuno* (Blackjack) es del 5.9%, pero algunos jugadores profesionales pueden ganar sistemáticamente, usando complicadas técnicas de conteo de cartas que requieren de muchas horas de práctica. Con un sistema, el jugador memoriza las cartas que se muestran y resta un punto por una carta con una figura o 10 por un as, y suma un punto por las cartas 2, 3, 4, 5, 6. Las cartas 7 y 8 se ignoran. Cuando el conteo es alto y el repartidor ya ha entregado una gran cantidad de cartas, el mazo tiene una proporción elevada de cartas altas y la suerte está del lado del jugador. Si un jugador que cuenta las cartas, de manera repentina, comienza a apostar grandes cantidades, el repartidor podría notar el conteo de cartas y expulsaría al jugador. Los contadores de cartas tratan de burlar esta política trabajando en equipo. Cuando el conteo es lo suficientemente alto, el jugador hace una seña a un cómplice para que entre al juego con apuestas cuantiosas. Se supone que un grupo de estudiantes del MIT ganó millones de dólares contando las cartas en el juego del veintiuno.

En las tres definiciones que siguen, las *posibilidades reales en contra* y las *posibilidades reales a favor* reflejan la verosimilitud real de un evento, pero las *posibilidades de pago* describen las cantidades de pago que son determinadas por los operadores de casinos, loterías e hipódromos. Este tipo de negocios buscan obtener un beneficio, por lo que las posibilidades de pago no serán iguales a las probabilidades reales.

DEFINICIONES

Las **posibilidades reales en contra** de la ocurrencia del evento A son la relación $P(\overline{A})/P(A)$, usualmente expresada en la forma de a:b (o "a a b"), donde a y b son enteros. (La expresión se reduce usando el máximo factor común; si $a = 16$ y $b = 4$, exprese las probabilidades como 4:1 en vez de 16:4).

Las **posibilidades reales a favor** de la ocurrencia del evento A son la relación $P(A)/P(\overline{A})$, que es el recíproco de las posibilidades reales en contra de ese evento. Si las posibilidades en contra de un evento son a:b, entonces las posibilidades a favor son b:a.

Las **posibilidades de pago** en contra de la ocurrencia del evento A son la proporción del beneficio neto (si usted gana) a la cantidad apostada:

Posibilidades de pago en contra del evento A = (beneficio neto):(cantidad apostada)

EJEMPLO 9 Posibilidades reales contra posibilidades de pago

Si usted apuesta \$5 al número 13 de la ruleta, su probabilidad de ganar es 1/38, pero las posibilidades de pago están dadas por el casino como 35:1.

a. Encuentre las posibilidades reales en contra del resultado de 13.

b. ¿Cuál sería su beneficio neto si ganara apostando \$5 al 13?

c. Si el casino no operara con fines de lucro y las posibilidades de pago se cambiaran para coincidir con las posibilidades reales en contra del 13, ¿cuánto ganaría usted si el resultado fuera 13?

SOLUCIÓN

a. Con $P(13) = 1/38$ y $P(\text{no } 13) = 37/38$, se obtiene

$$\text{Posibilidades reales contra 13} = \frac{P(\text{no } 13)}{P(13)} = \frac{37/38}{1/38} = \frac{37}{1} = \text{o } 37{:}1$$

b. Debido a que las probabilidades de pago del casino contra 13 son 35:1, tenemos

$$35{:}1 = (\text{beneficio neto}){:}(\text{cantidad apostada})$$

Así que hay un beneficio de \$35 por cada apuesta de \$1. Para una apuesta de \$5, el beneficio neto es \$175 (que es $5 \times \$35$). El apostador obtendría \$175 más la apuesta original de \$5. Después de ganar, la cantidad total recaudada sería \$180, para una ganancia neta de \$175.

c. Si el casino opera sin fines de lucro, las posibilidades de pago se cambiarían a 37:1, que son las posibilidades reales en contra del resultado de 13. Con las posibilidades de pago de 37:1, hay un beneficio neto de \$37 por cada \$1 apostado. Para una apuesta de \$5, el beneficio neto sería \$185. (El casino logra su beneficio proporcionando una ganancia de solamente \$175 en vez de los \$185 que se pagarían en un juego de ruleta que fuera justo en vez de favorecer al casino).

SU TURNO Resuelva el ejercicio 41 "Kentucky Pick 4".

4-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Lotería de Nueva Jersey Sea A el evento de colocar una apuesta directa de $1 en la lotería Pick 3 de Nueva Jersey y ganar. Hay 1000 formas de seleccionar tres dígitos (con la repetición permitida) en esta lotería, y sólo uno de esos números de tres dígitos será el ganador. ¿Cuál es el valor de $P(A)$? ¿Cuál es el valor de $P(\overline{A})$?

2. Probabilidad Reescriba la siguiente afirmación para que la probabilidad de lluvia se exprese como un valor entre 0 y 1: "La probabilidad de que hoy llueva es 25%".

3. Interpretación del tiempo Mientras se estaba escribiendo este ejercicio, Weather.com indicó que había una probabilidad del 60% de lluvia para la región natal del autor. Con base en ese informe, ¿cuál de las siguientes es la interpretación más razonable?

a. Hoy lloverá en 60% de la región del autor.

b. En la región del autor, lloverá 60% del día.

c. Existe una probabilidad de 0.60 que llueva en alguna parte de la región del autor, en algún momento durante el día.

4. Probabilidad subjetiva Estime la probabilidad de que la próxima vez que encienda un interruptor de luz, descubra que la bombilla funciona.

5. Identificación de valores de probabilidad ¿Cuáles de las siguientes son probabilidades?

$$0 \quad 3/5 \quad 5/3 \quad -0.25 \quad 250\% \quad 7{:}3 \quad 1 \quad 50\text{-}50 \quad 5{:}1 \quad 0.135 \quad 2.017$$

6. Penicilina ¿Quién descubrió la penicilina: Sean Penn, William Penn, Penn Jillette, Alexander Fleming o Louis Pasteur? Si usted elige aleatoriamente la respuesta a esta pregunta, ¿cuál es la probabilidad de que su respuesta sea la correcta: Alexander Fleming?

7. Constante de Avogadro Si se le pide en un examen que dé el primer dígito distinto a la izquierda de la constante de Avogadro y, al no conocer la respuesta, usted la elige aleatoriamente, ¿cuál es la probabilidad de que su respuesta sea la correcta: 6?

8. Nacimientos El ejemplo 2 de esta sección incluye el espacio muestral para los géneros de tres nacimientos. Identifique el espacio muestral para los géneros de dos nacimientos.

En los ejercicios 9 a 12, suponga que se seleccionan al azar 50 nacimientos. Utilice el juicio subjetivo para describir el número dado de niñas como (a) significativamente bajo, (b) significativamente alto, (c) ni significativamente bajo ni significativamente alto.

9. 47 niñas. **10.** 26 niñas. **11.** 23 niñas. **12.** 5 niñas.

En los ejercicios 13 a 20, exprese el grado de verosimilitud indicado como un valor de probabilidad entre 0 y 1.

13. Pruebas Si usted hace una elección aleatoria para la respuesta a una pregunta de examen del tipo verdadero/falso, hay posibilidades de 50-50 de que sea correcta.

14. Prueba SAT Al hacer la elección aleatoria de una respuesta a una pregunta de opción múltiple en la prueba de aptitudes académicas (SAT, por sus siglas en inglés), las posibles respuestas son a, b, c, d, e, así que hay 1 oportunidad en 5 de que sea correcta.

15. Equipaje En la salida de un autobús, si selecciona al azar a un viajero, hay 43% de posibilidad de que su equipaje sea de color negro.

16. Sonambulismo Con base en un informe de la revista *Neurology*, el 29.2% de los encuestados ha experimentado sonambulismo.

17. Aleatoriedad Cuando se utiliza una computadora para generar aleatoriamente el último dígito de un número telefónico que se marcará para realizar una encuesta, hay 1 oportunidad en 10 de que el último dígito sea cero.

18. Errores de solicitantes de empleo Con base en una encuesta de Adecco a los gerentes de recursos humanos, en la que se les pidió identificar los errores más grandes que los solicitantes de empleo cometen durante una entrevista, hay una oportunidad de 50-50 que se les identifique con "vestimenta inapropiada".

19. Clavija cuadrada Sydney Smith escribió en "On the Conduct of Understanding" que es imposible conectar una clavija cuadrada en un agujero redondo.

20. Muerte e impuestos Benjamin Franklin afirmó que la muerte es una certeza de la vida.

En los ejercicios 21 a 24, consulte los datos muestrales en la tabla 4-1, que se incluye con el problema del capítulo. Suponga que 1 de los 555 sujetos incluidos en esa tabla se selecciona al azar.

TABLA 4-1 Resultados de las pruebas de drogas a solicitantes de empleo

	Resultado positivo de la prueba (El examen muestra el uso de drogas)	**Resultado negativo de la prueba** (El examen no muestra uso de drogas)
El sujeto usa drogas	45 (Verdadero positivo)	5 (Falso negativo)
El sujeto no usa drogas	25 (Falso positivo)	480 (Verdadero negativo)

21. Pruebas de drogas a solicitantes de empleo Encuentre la probabilidad de seleccionar a alguien que obtuvo un resultado que es un falso negativo. ¿Quién sufriría con un falso negativo? ¿Por qué?

22. Pruebas de drogas a solicitantes de empleo Encuentre la probabilidad de seleccionar a alguien que obtuvo un resultado que es un falso positivo. ¿Quién sufriría con un falso positivo? ¿Por qué?

23. Pruebas de drogas a solicitantes de empleo Encuentre la probabilidad de seleccionar a alguien que usa drogas. ¿El resultado parece ser razonable como una estimación de la "tasa de prevalencia" descrita en el problema del capítulo?

24. Pruebas de drogas a solicitantes de empleo Encuentra la probabilidad de seleccionar a alguien que no usa drogas. ¿El resultado parece ser razonable como una estimación de la proporción de la población adulta que no usa drogas?

En los ejercicios 25 a 32, encuentre la probabilidad y conteste las preguntas.

25. Selección de género XSORT La técnica de selección de género XSORT de MicroSort fue diseñada para aumentar la probabilidad de que un bebé sea una niña. En un punto antes de que se interrumpieran los ensayos clínicos de la técnica de selección de género XSORT, 945 nacimientos resultaron en 879 niñas y 66 niños (según datos del Genetics & IVF Institute). Con base en estos resultados, ¿cuál es la probabilidad de que una pareja que utilizó el método XSORT de MicroSort tenga una niña? ¿Parece que la técnica es efectiva al aumentar la probabilidad de que un bebé sea una niña?

26. Selección de Género YSORT La técnica de selección de género YSORT de MicroSort está diseñada para aumentar la probabilidad de que un bebé sea un niño. En un punto antes de que se interrumpieran los ensayos clínicos de la técnica de selección de género YSORT, 291 nacimientos resultaron en 239 niños y 52 niñas (según datos del Genetics & IVF Institute). Con base en estos resultados, ¿cuál es la probabilidad de que una pareja que utilizó el método YSORT de MicroSort tenga un niño? ¿Parece que la técnica es efectiva al aumentar la probabilidad de que un bebé sea un niño?

27. Genética mendeliana Cuando Mendel llevó a cabo sus famosos experimentos genéticos con chícharos, una muestra de descendientes constó de 428 chícharos verdes y 152 chícharos amarillos. Con base en esos resultados, estime la probabilidad de obtener un chícharo descendiente que sea verde. ¿Es el resultado razonablemente cercano al valor esperado de 3/4, como afirmó Mendel?

28. Adivinación de cumpleaños En su primera cita, Kelly le pide a Mike que adivine la fecha de su nacimiento, sin incluir el año.

a. ¿Cuál es la probabilidad de que Mike adivine correctamente? (No tome en cuenta los años bisiestos).

b. ¿Sería improbable que adivinara correctamente en su primer intento?

c. Si usted fuera Kelly, y Mike adivinara su cumpleaños en el primer intento, ¿creería en su afirmación de que tuvo suerte al adivinar, o estaría convencida de que ya sabía cuándo nació?

d. Si Kelly le pide a Mike que adivine su edad, y la conjetura de Mike es demasiado alta por una cantidad entre 1 y 5 años, ¿cuál es la probabilidad de que Mike y Kelly tengan una segunda cita?

29. Redes sociales En un estudio del Pew Research Center sobre usuarios de Internet, 3732 encuestados dicen que usan sitios de redes sociales y 1380 encuestados dicen *no utilizarlos*. ¿Cuál es la probabilidad de que una persona seleccionada al azar no use un sitio de redes sociales? ¿El resultado sugiere que es poco probable que alguien no utilice los sitios de redes sociales?

30. Volcaduras de automóvil En un año reciente en Estados Unidos, 83,600 automóviles de pasajeros se volcaron al chocar, y 5,127,400 autos no se volcaron al hacerlo. Encuentre la probabilidad de que un accidente de automóvil de pasajeros seleccionado al azar resulte en una volcadura. ¿Es improbable que un auto se vuelque en un accidente?

31. Genética: color de los ojos Cada uno de los padres tiene el genotipo marrón/azul, que consiste en el par de alelos que determinan el color de los ojos, y cada uno de los padres aporta uno de esos alelos a un niño. Suponga que si el niño tiene al menos un alelo marrón, ese color dominará y los ojos serán marrones. (La determinación real del color de los ojos es más complicada que esto).

a. Liste los diferentes resultados posibles. Suponga que estos resultados son igualmente probables.

b. ¿Cuál es la probabilidad de que un hijo de estos padres tenga el genotipo azul/azul?

c. ¿Cuál es la probabilidad de que el niño tenga los ojos marrones?

32. Enfermedad genética vinculada a X Los hombres tienen cromosomas XY (o YX) y las mujeres tienen cromosomas XX. Las enfermedades genéticas recesivas ligadas al cromosoma X (como la retinosquisis juvenil) ocurren cuando hay un cromosoma X defectuoso que se produce sin un cromosoma X pareado no defectuoso. En lo sucesivo, representaremos un cromosoma X defectuoso con x minúscula, por lo que un niño con el par de cromosomas xY o Yx tendrá la enfermedad y un niño con XX o XY o YX o xX o Xx no la padecerá. Cada padre aporta uno de los cromosomas al niño.

a. Si un padre tiene el cromosoma x defectuoso y la madre tiene buenos cromosomas XX, ¿cuál es la probabilidad de que un hijo herede la enfermedad?

b. Si un padre tiene el cromosoma x defectuoso y la madre tiene buenos cromosomas XX, ¿cuál es la probabilidad de que una hija herede la enfermedad?

c. Si una madre tiene un cromosoma x defectuoso y un buen cromosoma X y el padre tiene buenos cromosomas XY, ¿cuál es la probabilidad de que un hijo herede la enfermedad?

d. Si una madre tiene un cromosoma x defectuoso y un buen cromosoma X y el padre tiene buenos cromosomas XY, ¿cuál es la probabilidad de que una hija herede la enfermedad?

Probabilidad de un espacio muestral. ***En los ejercicios 33 a 36, use el espacio muestral dado o elabore el espacio muestral requerido para encontrar la probabilidad indicada.***

33. Tres hijos Utilice este espacio muestral que lista los ocho eventos simples que son posibles cuando una pareja tiene tres hijos (como en el ejemplo 2 en la página 135): {hhh, hhm, hmh, hmm, mhh, mhm, mmh, mmm). Suponga que los hombres y las mujeres son igualmente probables, por lo que los ocho eventos simples también lo son. Encuentre la probabilidad de que entre los tres hijos de una pareja haya exactamente una niña.

34. Tres hijos Utilizando el mismo espacio muestral y el mismo supuesto del ejercicio 33, encuentre la probabilidad de que entre los tres hijos de una pareja haya exactamente dos niñas.

35. Cuatro niños El ejercicio 33 lista el espacio muestral para una pareja que tiene tres hijos. Después de identificar el espacio muestral para una pareja que tiene cuatro hijos, encuentre la probabilidad de tener tres niñas y un niño (en cualquier orden).

36. Cuatro niños Utilizando el mismo espacio muestral y el mismo supuesto del ejercicio 35, encuentre la probabilidad de que entre los cuatro hijos de una pareja los cuatro sean del mismo sexo.

Uso de la probabilidad para formular conclusiones. ***En los ejercicios 37 a 40, use el valor de probabilidad dado para determinar si los resultados de la muestra podrían ocurrir fácilmente por casualidad, y luego formule una conclusión.***

37. Predicción del género Un estudio abordó la cuestión de si las mujeres embarazadas pueden predecir correctamente el sexo de su bebé. Entre 104 mujeres embarazadas, 57 predijeron correctamente el sexo de su bebé (según los datos de "Are Women Carrying 'Basketballs' ..." de Perry DiPietro, Constigan, *Birth*, vol. 26, núm. 3). Si las mujeres embarazadas no tienen tal habilidad, hay una probabilidad 0.327 de obtener estos resultados de la muestra por casualidad. ¿Qué concluye usted?

38. Cinturón ajustado Un estudio sobre el efecto del uso del cinturón de seguridad en colisiones de automóviles de pasajeros descubrió que los conductores que usaban cinturón de seguridad tenían una tasa de supervivencia de 64.1%, mientras que los conductores que no usaban el cinturón de seguridad tenían una tasa de supervivencia de 41.5%. Si los cinturones de seguridad no tienen efecto sobre la tasa de supervivencia, hay menos de 0.0001 de probabilidad de obtener estos resultados (según los datos de "Mortality Reduction with Air Bag and Seat Belt Use in Head-on Passenger Car Collisions", de Crandall Olson, Sklar, *American Journal of Epidemiology*, vol. 153, núm. 3). ¿Qué concluye usted?

39. Charla de café Un estudio sobre el efecto potenciador del café en la memoria a largo plazo encontró que 35 participantes que recibieron 200 mg de cafeína obtuvieron mejores resultados en una prueba de memoria 24 horas más tarde, en comparación con el grupo placebo que no recibió cafeína.

a. Había una probabilidad de 0.049 de que la diferencia entre el grupo del café y el grupo del placebo se debiera al azar. ¿Qué concluye usted?

b. Un grupo que recibió una dosis más alta de 300 mg se desempeñó mejor que el grupo de 200 mg, con una probabilidad de 0.75 de que esta diferencia se debiera al azar. ¿Qué concluye usted?

40. Celulares y cáncer Un estudio de 420,095 usuarios daneses de teléfonos celulares resultó en que 135 desarrollaron cáncer del cerebro o del sistema nervioso (según datos del *Journal of the National Cancer Institute*). Al comparar este grupo de muestra con otro grupo de personas que no usaron teléfonos celulares, se encontró que hay una probabilidad de 0.512 de obtener tales resultados por casualidad. ¿Qué concluye usted?

4-1 Más allá de lo básico

Posibilidades. ***En los ejercicios 41 a 44, responda las preguntas dadas que implican posibilidades.***

41. Kentucky Pick 4 En la lotería Kentucky Pick 4, puede realizar una apuesta "directa" de $1 seleccionando el orden exacto de cuatro dígitos entre 0 y 9 inclusive (con repetición permitida), por lo que la probabilidad de ganar es 1/10,000. Si los mismos cuatro números salen sorteados en el mismo orden, usted gana $5000, por lo que su ganancia neta es de $4999.

a. Encuentra las posibilidades reales de ganar.

b. Encuentra las posibilidades de pago.

c. El sitio web www.kylottery.com indica posibilidades de 1:10,000 para esta apuesta. ¿Es precisa esa descripción?

42. Determinación de posibilidades en la ruleta Una rueda de la ruleta tiene 38 ranuras. Una ranura es 0, otra es 00, y las demás están numeradas del 1 al 36, respectivamente. Usted coloca una apuesta a que el resultado es un número impar.

a. ¿Cuál es su probabilidad de ganar?

b. ¿Cuáles son las posibilidades reales de ganar?

c. Cuando usted apuesta que el resultado será un número impar, las posibilidades de pago son 1:1. ¿Cuánta ganancia obtiene si apuesta $18 y gana?

d. ¿Cuánta ganancia obtendría usted con la apuesta de $18 si pudiera de alguna manera convencer al casino para cambiar sus posibilidades de pago para que sean iguales a las posibilidades reales de ganar? (*Recomendación:* no trate de convencer a ningún casino de esto, su sentido del humor se ausenta de manera notable cuando se trata de cosas de este tipo).

43. Posibilidades en el Kentucky Derby Cuando el caballo California Chrome ganó el Kentucky Derby número 140, una apuesta de $2 por una victoria de California Chrome resultó en un boleto ganador por valor de $7.

a. ¿Cuánta ganancia neta se obtuvo de una apuesta ganadora de $2 por California Chrome?

b. ¿Cuáles fueron las posibilidades de pago contra una victoria de California Chrome?

c. Con base en apuestas preliminares antes de la carrera, los apostantes colectivamente creían que California Chrome tenía una probabilidad de 0.228 de ganar. Asumiendo que 0.228 era la verdadera probabilidad de una victoria de California Chrome, ¿cuáles eran las posibilidades reales contra su victoria?

d. Si las posibilidades de pago eran las probabilidades reales encontradas en la parte (c), ¿cuál sería el valor de un boleto ganador de $2 después de la victoria de California Chrome?

44. Riesgo relativo y tasa de posibilidades En un ensayo clínico de 2103 sujetos tratados con Nasonex, 26 informaron de dolores de cabeza. En un grupo control de 1671 sujetos que recibieron placebo, 22 reportaron dolores de cabeza. Si se expresa la proporción de dolores de cabeza en el grupo de tratamiento con p_t, y la proporción de dolores de cabeza en el grupo de control (placebo) con p_c, el *riesgo relativo* es p_t/p_c. El riesgo relativo es una medida de la fuerza del efecto del tratamiento con Nasonex. Otra medida de este tipo es la *tasa de posibilidades*, que es la proporción de las posibilidades en favor de un dolor de cabeza para el grupo de tratamiento sobre las posibilidades en favor de un dolor de cabeza para el grupo de control (placebo), que se encuentra al calcular lo siguiente:

$$\frac{p_t/(1-p_t)}{p_c/(1-p_c)}$$

El riesgo relativo y la tasa de posibilidades son de uso común en medicina y estudios epidemiológicos. Encuentre el riesgo relativo y la tasa de posibilidades para los datos del dolor de cabeza. ¿Qué sugieren los resultados sobre el riesgo de un dolor de cabeza por el tratamiento con Nasonex?

4-2 Regla de la suma y regla de la multiplicación

Conceptos clave En esta sección presentamos la *regla de la suma* como una herramienta para encontrar *P*(*A* o *B*), que es la probabilidad de que ocurra el evento *A* u ocurra el evento *B* (o que ambos ocurran) como el único resultado de un procedimiento. Para encontrar *P*(*A* o *B*), comenzamos por sumar el número de formas en que *A* puede ocurrir y el número de formas en que *B* puede ocurrir, pero sumamos sin contar doble. La palabra "o" en la regla de la suma está asociada con la adición de probabilidades.

Esta sección también presenta la *regla de la multiplicación* básica, que se usa para encontrar *P*(*A* y *B*), es decir, la probabilidad de que ocurra el evento *A* y ocurra el evento *B*. Si el resultado del evento *A* afecta de algún modo la probabilidad del evento *B*, es importante ajustar la probabilidad de *B* para reflejar la ocurrencia del evento *A*. La regla para encontrar *P*(*A* y *B*) se denomina *regla de la multiplicación* porque implica la multiplicación de la probabilidad del evento *A* y la probabilidad del evento *B* (donde, si es necesario, la probabilidad del evento *B* se ajusta debido al resultado del evento *A*). La palabra "y" en la regla de la multiplicación está asociada con la multiplicación de probabilidades.

En la sección 4-1 consideramos sólo eventos *simples*, pero en esta sección consideraremos *eventos compuestos*.

DEFINICIÓN

Un **evento compuesto** es cualquier evento que combina dos o más eventos simples.

Regla de la suma

Notación para la regla de la suma

$P(A \text{ o } B) = P$(en un solo ensayo, ocurra el evento A u ocurra el evento B o ambos ocurran)

La palabra "o" usada en la notación anterior es la *o inclusiva*; que significa uno o el otro o ambos. La regla de la suma formal se presenta a menudo como una fórmula, pero el uso ciego de las fórmulas no es recomendable. En su lugar, *entienda* el espíritu de la regla y utilice esa comprensión, como en la regla de la suma intuitiva que se presenta a continuación.

REGLA DE LA SUMA INTUITIVA

Para encontrar $P(A \text{ o } B)$, sume el número de formas en que puede ocurrir el evento A y el número de formas en que puede ocurrir el evento B, pero *hágalo de tal manera que cada resultado se contabilice sólo una vez*. $P(A \text{ o } B)$ es igual a esa suma, dividida por el número total de resultados en el espacio muestral.

REGLA DE LA SUMA FORMAL

$$P(A \text{ o } B) = P(A) + P(B) - P(A \text{ y } B)$$

donde $P(A \text{ y } B)$ expresa la probabilidad de que A y B ocurran al mismo tiempo como resultado en un ensayo de un procedimiento.

Una forma de aplicar la regla de la suma es sumar la probabilidad del evento A y la probabilidad del evento B y, si hay alguna superposición que ocasione conteo doble, se compensa al restar la probabilidad de los resultados que se incluyen dos veces. Este método se refleja en la anterior regla de la suma formal.

EJEMPLO 1 Prueba de drogas a solicitantes de empleo

Consulte la tabla 4-1, reproducida aquí para su comodidad y placer visual. Si el sujeto se selecciona al azar de entre los 555 sujetos que fueron sometidos a una prueba de drogas, encuentre la probabilidad de seleccionar un sujeto con resultado positivo en la prueba *o* que use drogas.

TABLA 4-1 Resultados de las pruebas de drogas a solicitantes de empleo

	Resultado positivo de la prueba (El examen muestra el uso de drogas)	**Resultado negativo de la prueba** (El examen no muestra uso de drogas)
El sujeto usa drogas	**45** (Verdadero positivo)	**5** (Falso negativo)
El sujeto no usa drogas	**25** (Falso positivo)	**480** (Verdadero negativo)

**Los números en gris corresponden a resultados positivos en las pruebas o sujetos que consumen drogas, y el total de esos números es 75.*

SOLUCIÓN

Consulte la tabla 4-1 y cuente cuidadosamente el número de sujetos que dieron positivo (primera columna) o usa drogas (primera fila), pero tenga cuidado de contarlos exactamente una vez, no dos. *Al sumar las frecuencias de la primera columna y la primera fila, incluya la frecuencia de 45 sólo una vez*. En la tabla 4-1, hay $45 + 25 + 5 = 75$ sujetos que tuvieron resultados de prueba positivos o usaron drogas. Obtenemos este resultado:

$$P(\text{resultado positivo en la prueba o utiliza drogas}) = 75/555 = 0.135$$

SU TURNO Resuelva el ejercicio 11 "Exactitud de comida rápida en auto".

Eventos disjuntos y la regla de la suma

La regla de la suma se simplifica cuando los eventos son *disjuntos*.

DEFINICIÓN

Los eventos *A* y *B* son **disjuntos** (o **mutuamente excluyentes**) si no pueden ocurrir al mismo tiempo. (Es decir, los eventos disjuntos no se superponen).

EJEMPLO 2 Eventos disjuntos

Eventos disjuntos:	**Evento A** — Seleccionar aleatoriamente a alguien para un ensayo clínico que sea un hombre **Evento B** — Seleccionar aleatoriamente a alguien para un ensayo clínico que sea una mujer (La persona seleccionada *no puede* ser ambas cosas).
Eventos que *no* son disjuntos:	**Evento A** — Seleccionar aleatoriamente a alguien que tome un curso de estadística **Evento B** — Seleccionar aleatoriamente a alguien que sea una mujer. (La persona seleccionada *puede* ser ambas cosas).

SU TURNO Resuelva el ejercicio 12 "Exactitud de comida rápida en auto"

Siempre que *A* y *B* sean disjuntos, $P(A \text{ y } B)$ se convierte en cero en la regla de la suma formal, por lo que para los eventos disjuntos *A* y *B* tenemos $P(A \text{ o } B) = P(A) + P(B)$. Pero, de nuevo, en lugar de usar ciegamente una fórmula, es mejor *entender* y usar la regla de la suma intuitiva.

A continuación se presenta un resumen de los puntos clave de la regla de la suma:

1. **Para encontrar $P(A \text{ o } B)$, primero asocie la palabra *o* con la suma.**
2. **Para encontrar el valor de $P(A \text{ o } B)$, sume el número de formas en que puede ocurrir *A* y el número de formas en que puede ocurrir *B*, pero tenga cuidado de sumar sin contar doble.**

Eventos complementarios y la regla de la suma

En la sección 4-1 usamos $\overline{A}$ para indicar que el evento *A* no ocurre. El sentido común dicta este principio: tenemos la certeza (con probabilidad 1) de que un evento *A* ocurre *o* no ocurre, por lo que $P(A \text{ o } \overline{A}) = 1$. Debido a que los eventos *A* y $\overline{A}$ deben ser disjuntos, podemos utilizar la regla de la suma para expresar este principio de la siguiente manera:

$$P(A \text{ o } \overline{A}) = P(A) + P(\overline{A}) = 1$$

En cifras

2300%: Aumento en el riesgo de accidente cuando se envían mensajes de texto y se conduce. 600%: Aumento del riesgo de accidente cuando se conduce bajo el efecto de bebidas alcohólicas.

Este resultado de la regla de la suma conduce a las tres expresiones siguientes que son "equivalentes" en el sentido de que sólo son formas diferentes del mismo principio.

REGLA DE EVENTOS COMPLEMENTARIOS

$P(A) + P(\overline{A}) = 1 \qquad P(\overline{A}) = 1 - P(A) \qquad P(A) = 1 - P(\overline{A})$

EJEMPLO 3 **Sonambulismo**

Con base en un artículo, la probabilidad de seleccionar aleatoriamente a alguien que ha experimentado sonambulismo es de 0.292, por lo que P(sonambulismo) = 0.292 (según datos de "Prevalence and Comorbidity of Nocturnal Wandering in the U.S. General Population", de Ohayon *et al*., *Neurology*, vol. 78, núm. 20). Si una persona se selecciona aleatoriamente, encuentre la probabilidad de elegir a alguien que *no* ha experimentado sonambulismo.

SOLUCIÓN

Si se utiliza la regla de eventos complementarios, resulta

$$P(\textit{no}\text{ sonambulismo}) = 1 - P(\text{sonambulismo}) = 1 - 0.292 = 0.708$$

La probabilidad de seleccionar aleatoriamente a alguien que no ha experimentado sonambulismo es 0.708.

SU TURNO Resuelva el ejercicio 5 "LOL".

Regla de la multiplicación

Notación para la regla de la multiplicación

Comenzamos con la notación básica seguida por la regla de la multiplicación. Se recomienda utilizar la regla de la multiplicación *intuitiva*, porque se basa en la comprensión y no en el uso ciego de una fórmula.

Notación

$P(A \text{ y } B) = P$(el evento A ocurra en un primer ensayo y el evento B ocurra en un segundo ensayo)

$P(B \mid A)$ representa la probabilidad de que ocurra el evento B después de suponer que el evento A ya ha ocurrido. (Interprete $B \mid A$ como "el evento B ocurre después de que el evento A ya ha ocurrido").

PRECAUCIÓN La notación $P(A \text{ y } B)$ tiene dos significados, dependiendo de su contexto. Para la regla de la multiplicación, $P(A \text{ y } B)$ expresa que el evento A ocurre en un ensayo y el evento B ocurre en otro; para la regla de la suma, utilizamos $P(A \text{ y } B)$ para indicar que los eventos A y B ocurren ambos en el mismo ensayo.

REGLA INTUITIVA DE LA MULTIPLICACIÓN

Para encontrar la probabilidad de que el evento A ocurra en un ensayo y el evento B ocurra en otro, multiplique la probabilidad del evento A por la probabilidad del evento B, pero *asegúrese de que la probabilidad del evento B se encuentra suponiendo que el evento A ya ha ocurrido.*

REGLA DE LA MULTIPLICACIÓN FORMAL

$P(A \text{ y } B) = P(A) \cdot P(B \mid A)$

Independencia y regla de la multiplicación

Al aplicar la regla de la multiplicación y considerar si la probabilidad del evento *B* debe ajustarse para considerar la ocurrencia previa del evento *A*, nos centramos en si los eventos *A* y *B* son *independientes*.

DEFINICIONES

Dos eventos *A* y *B* son **independientes** si la ocurrencia de uno no afecta la *probabilidad* de que ocurra el otro. (Varios eventos son independientes si la ocurrencia de cualquiera no afecta las probabilidades de ocurrencia de los demás). Si *A* y *B* no son independientes, se dice que son **dependientes**.

PRECAUCIÓN No piense que la *dependencia* de dos eventos significa que uno es la *causa* directa del otro. Tener una luz funcionando en su cocina y otra funcionando en su recámara son eventos dependientes porque comparten la misma fuente de alimentación. Una de las luces puede dejar de funcionar por muchas razones, pero si una luz está apagada, hay una mayor probabilidad de que la otra luz se apague (debido a la fuente de alimentación común).

EJEMPLO 4 Prueba de drogas y la regla de la multiplicación básica

Utilicemos sólo los 50 resultados de las pruebas de los sujetos que consumen drogas (de la tabla 4-1), como se muestra a continuación:

Resultados positivos en la prueba:	45
Resultados negativos en la prueba:	5
Total:	50

a. Se seleccionan aleatoriamente *con reemplazo* 2 de estos 50 sujetos que usan drogas; encuentre la probabilidad de que la primera persona seleccionada tuviera un resultado de prueba positivo y la segunda persona tuviera un resultado de prueba negativo.

b. Repita el inciso (a) suponiendo que los dos sujetos se seleccionan *sin reemplazo*.

SOLUCIÓN

a. *Con reemplazo:* Primera selección (con 45 resultados positivos entre los 50 resultados totales):

$$P(\text{resultado positivo en la prueba}) = \frac{45}{50}$$

Segunda selección (con 5 resultados negativos entre los mismos 50 resultados totales):

$$P(\text{resultado negativo en la prueba}) = \frac{5}{50}$$

Ahora aplicamos la regla de la multiplicación como sigue:

$$P(\text{primera selección positiva y segunda selección negativa}) = \frac{45}{50} \cdot \frac{5}{50} = 0.0900$$

continúa

Motores a reacción independientes

Poco después de despegar de Miami, la aeronave del vuelo 855 de Eastern Airlines tuvo que apagar un motor porque se encendió el indicador de baja presión de aceite. Cuando el jet L-1011 regresaba a Miami para aterrizar, los indicadores de baja presión de los otros dos motores también se encendieron. Entonces falló otro motor y después falló el último motor que estaba funcionando. El jet descendió sin propulsión desde 13,000 hasta 4000 pies; entonces la tripulación pudo poner a funcionar un motor. La aeronave, con 172 personas a bordo, aterrizó con seguridad. Con motores a reacción independientes, la probabilidad de que los tres fallen es de sólo 0.0001^3, 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.

b. *Sin reemplazo:* Sin reemplazo del primer sujeto, los cálculos son iguales que en el inciso (a), excepto que la segunda probabilidad se debe ajustar para reflejar el hecho de que la primera selección fue positiva y no está disponible para la segunda selección. Después de seleccionar el primer resultado positivo, tenemos 49 resultados de prueba restantes, y 5 de ellos son negativos. La segunda probabilidad es, por tanto, 5/49, como se muestra a continuación:

$$P(\text{la primera selección es positiva y la segunda es negativa}) = \frac{45}{50} \cdot \frac{5}{49} = 0.0918$$

SU TURNO Resuelva el ejercicio 13 “Exactitud de comida rápida en auto”.

El punto clave del inciso (b) en el ejemplo 4 es el siguiente: *debemos ajustar la probabilidad de que el segundo evento refleje el resultado del primero*. Debido a que la selección del segundo sujeto se realiza *sin* reemplazo del primer sujeto, la segunda probabilidad debe tener en cuenta el hecho de que la primera selección eliminó a un sujeto que dio positivo, de modo que sólo hay 49 sujetos disponibles para la segunda selección y 5 de ellos tuvieron un resultado de prueba negativo. El inciso (a) del ejemplo 4 implicó muestreo con reemplazo, de modo que los eventos son independientes; la parte (b) del ejemplo 4 implicó el muestreo sin sustitución, de modo que los eventos son dependientes. Vea lo siguiente.

Muestreo

En el maravilloso mundo de la estadística, los métodos de muestreo son cruciales, y las siguientes relaciones se sostienen:

- Muestreo *con reemplazo*: Las selecciones son eventos *independientes*.
- Muestreo *sin reemplazo*: Las selecciones son eventos *dependientes*.

Excepción: tratamiento de eventos dependientes como independientes

Algunos cálculos complejos se pueden simplificar en gran medida usando la práctica común de tratar los eventos como independientes cuando se seleccionan *muestras pequeñas* sin reemplazo de *poblaciones grandes*. (En tales casos, es raro seleccionar el mismo elemento dos veces). A continuación se presenta una directriz que se utiliza rutinariamente en aplicaciones como el análisis de resultados de encuestas:

TRATAMIENTO DE EVENTOS DEPENDIENTES COMO INDEPENDIENTES: DIRECTRIZ DEL 5% PARA CÁLCULOS COMPLEJOS

Cuando se usa un muestreo sin reemplazo y el tamaño de la muestra no supera 5% del tamaño de la población, considere que las selecciones son *independientes* (aunque en realidad sean dependientes).

El ejemplo 5 ilustra el uso de la anterior “directriz del 5% para cálculos complejos” y también ilustra que la regla de la multiplicación básica se extiende fácilmente a tres o más eventos.

EJEMPLO 5 Evaluación de drogas y directriz del 5% para cálculos complejos

Suponga que tres adultos se seleccionan aleatoriamente y *sin reemplazo* entre los 247,436,830 adultos en Estados Unidos. También suponga que 10% de los adultos en ese país usan drogas. Encuentre la probabilidad de que los tres adultos seleccionados usen drogas.

SOLUCIÓN

Debido a que los tres adultos se seleccionan al azar sin reemplazo, los tres eventos son dependientes, pero aquí podemos tratarlos como independientes aplicando la directriz del 5% para cálculos complejos. El tamaño muestral de 3 claramente no es mayor al 5% del tamaño de la población de 247,436,830. Obtenemos

$$\begin{aligned} P(\text{los 3 adultos usan drogas}) &= P(\text{el primero usa drogas } y \text{ el segundo usa drogas} \\ &\qquad y \text{ el tercero usa drogas}) \\ &= P(\text{el primero usa drogas}) \cdot P(\text{el segundo usa drogas}) \cdot \\ &\qquad P(\text{el tercero usa drogas}) \\ &= (0.10)(0.10)(0.10) = 0.00100 \end{aligned}$$

Hay una probabilidad de 0.00100 que los tres adultos seleccionados usen drogas.

SU TURNO Resuelva el ejercicio 29 “Helicópteros médicos”.

En el ejemplo 5, si tratamos los eventos como dependientes sin usar la directriz del 5%, obtendremos el siguiente cálculo engorroso que comienza con 247,436,830 adultos, donde 10% de ellos (o 24,743,683) usan drogas:

$$\begin{aligned} \left(\frac{24{,}743{,}683}{247{,}436{,}830}\right)\left(\frac{24{,}743{,}682}{247{,}436{,}829}\right)\left(\frac{24{,}743{,}681}{247{,}436{,}828)}\right) &= 0.0009999998909 \\ &= 0.00100 \text{ (redondeado)} \end{aligned}$$

Imagine la selección aleatoria de 1000 adultos en lugar de sólo 3, como se hace comúnmente en las encuestas. La extensión del cálculo anterior para incluir 1000 factores en lugar de 3 sería lo que los estadísticos conocen como como cálculo “doloroso”.

PRECAUCIÓN En cualquier cálculo de probabilidad, es sumamente importante identificar cuidadosamente el evento que se está considerando. Vea el ejemplo 6, donde los incisos (a) y (b) pueden parecer bastante similares pero sus soluciones son muy diferentes.

EJEMPLO 6 Cumpleaños

Dos personas son seleccionadas aleatoriamente de entre sus compañeros de clase; encuentre la probabilidad indicada suponiendo que los cumpleaños ocurren en los días de la semana con frecuencias iguales.

a. Encuentre la probabilidad de que las dos personas hayan nacido el *mismo día de la semana*.

b. Encuentre la probabilidad de que las dos personas hayan nacido un *lunes*.

SOLUCIÓN

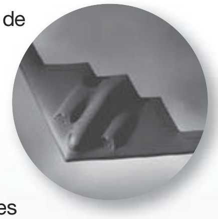
a. Debido a que no se especifica ningún día de la semana, la primera persona puede nacer en cualquiera de los siete días. La probabilidad de que la segunda persona nazca el mismo día que la primera persona es 1/7. La probabilidad de que dos personas nacen el mismo día de la semana es por lo tanto 1/7.

b. La probabilidad de que la primera persona haya nacido un lunes es 1/7 y la probabilidad de que la segunda persona también haya nacido un lunes es 1/7. Debido a que los dos eventos son independientes, la probabilidad de que ambas personas hayan nacido en lunes es

$$\frac{1}{7} \cdot \frac{1}{7} = \frac{1}{49}$$

SU TURNO Resuelva el ejercicio 19 “Exactitud de comida rápida en auto”.

Redundancia



La confiabilidad de sistemas puede mejorarse considerablemente con la redundancia de componentes esenciales. Los automóviles de carreras de las series de la Copa Winston NASCAR tienen dos sistemas de ignición para que, si uno falla, exista otro de reserva. Los aviones poseen dos sistemas eléctricos independientes, y los que se usan para vuelos por instrumentos suelen tener dos radios distintos. La siguiente cita se tomó de un artículo de *Popular Science* acerca de los aviones antirradar: “Un avión construido en buena parte con fibra de carbono fue el Lear Fan 2100, que tenía que llevar dos transpondedores de radar. La razón es que si fallaba una unidad de transpondedor, el avión seguiría siendo casi invisible para el radar”. Tal redundancia es una aplicación de la regla de la multiplicación dentro de la teoría de probabilidad. Si un componente tiene una probabilidad de 0.001 de fallar, la probabilidad de que dos componentes independientes fallen es de sólo 0.000001.

Para ganar, apueste con audacia

El diario *New York Times* publicó una nota de Andrew Pollack en la que se informó que el casino Mirage de Las Vegas tenía menores ganancias que las esperadas. El periodista afirmó que "las ganancias del Mirage resultan particularmente volátiles, ya que se favorece a los grandes apostadores, quienes pueden apostar $100,000 o más en una mano de cartas. La ley de los promedios no funciona con tanta consistencia para unas cuantas apuestas grandes como lo hace para miles de pequeñas...". Esto refleja el principio más fundamental al apostar: para ganar, ¡haga una apuesta grande en vez de muchas apuestas pequeñas! Con el juego adecuado, como puede ser el de los dados, usted tiene poco menos del 50% de probabilidades de duplicar su dinero si realiza una apuesta grande. Al hacer muchas apuestas pequeñas, la probabilidad de duplicar su dinero disminuye sustancialmente.

¡CUIDADO CON SU LENGUAJE! El ejemplo 6 ilustra que encontrar valores de probabilidad correctos o relevantes a menudo requiere más de habilidades lingüísticas que de habilidades de cálculo. En el ejemplo 6, ¿qué significa exactamente "el mismo día de la semana"? Vea cómo los incisos (a) y (b) del ejemplo 6 son muy diferentes.

Redundancia: aplicación importante de la regla de la multiplicación

El principio de *redundancia* se utiliza para aumentar la confiabilidad de muchos sistemas. Nuestros ojos tienen redundancia pasiva en el sentido de que si uno de ellos falla, seguimos viendo. Un hallazgo importante de la biología moderna es que los genes en un organismo a menudo pueden funcionar en lugar de otro. Los ingenieros suelen diseñar componentes redundantes para que no falle todo el sistema debido a la falla de un solo componente, como en el ejemplo siguiente.

EJEMPLO 7 Airbus 310: *Redundancia* para mayor seguridad

En la actualidad, los aviones son altamente confiables, y una característica de diseño que contribuye a esa confiabilidad es el uso de la *redundancia*, por lo que los componentes críticos se duplican de modo que si uno falla, el otro funcione. Por ejemplo, el avión bimotor Airbus 310 tiene tres sistemas hidráulicos independientes, por lo que si falla un sistema, se mantiene el control completo del vuelo con otro sistema en funcionamiento. Para este ejemplo, supondremos que para un vuelo típico, la probabilidad de una falla del sistema hidráulico es 0.002.

a. Si el Airbus 310 tuviera un sistema hidráulico, ¿cuál sería la probabilidad de que el control de vuelo de la aeronave funcionara?

b. Dado que el Airbus 310 tiene en realidad tres sistemas hidráulicos independientes, ¿cuál es la probabilidad de que en un vuelo típico, el control pueda mantenerse con un sistema hidráulico en funcionamiento?

SOLUCIÓN

a. La probabilidad de falla del sistema hidráulico es 0.002, por lo que la probabilidad de que *no* falle es 0.998. Es decir, la probabilidad de que el control del vuelo pueda mantenerse es la siguiente:

$$P(1 \text{ sistema hidráulico } \textit{no falle}) = 1 - P(\text{falla}) = 1 - 0.002\text{: } 0.998$$

b. Con tres sistemas hidráulicos independientes, el control de vuelo se mantendrá si no fallan los tres sistemas. La probabilidad de que los tres sistemas hidráulicos fallen es $0.002 \cdot 0.002 \cdot 0.002 = 0.000000008$. Se deduce que la probabilidad de mantener el control del vuelo es la siguiente:

$$\begin{aligned} &P(\textit{no} \text{ ocurre que los tres sistemas hidráulicos fallen}) \\ &= 1 - 0.000000008 = 0.999999992 \end{aligned}$$

INTERPRETACIÓN

Con un solo sistema hidráulico tenemos una probabilidad de falla de 0.002, pero con tres sistemas hidráulicos independientes, sólo hay una probabilidad de 0.000000008 de que el control del vuelo no pueda mantenerse debido a que los tres sistemas fallaron. Al usar tres sistemas hidráulicos en lugar de uno solo, el riesgo de falla no disminuye en un factor de 1/3, sino en un factor de 1/250,000. Mediante el uso de tres sistemas hidráulicos independientes, el riesgo se reduce drásticamente y la seguridad se incrementa de manera sustancial.

SU TURNO Resuelva el ejercicio 25 "Redundancia en discos duros de computadora".

Justificación de la regla de la multiplicación

Para ver el razonamiento que subyace a la regla de la multiplicación, considere una prueba rápida que consiste en estas dos preguntas:

1. Verdadero o falso: una libra de plumas es más pesada que una libra de oro.
2. ¿Quién dijo: "Por una pequeña muestra, podemos juzgar toda la pieza"? (a) la Juez Judy, b) el Juez Dredd; (c) Miguel de Cervantes, (d) George Gallup, (e) Gandhi

Las respuestas son V (verdadero) y c. (La primera respuesta es verdadero porque los pesos de las plumas están en unidades de *avoirdupois* donde una libra es de 453.59 g, pero los pesos del oro y otros metales preciosos están en unidades *Troy* donde una libra es de 373.24 g. La segunda respuesta es de *Don Quijote* escrito por Cervantes).

A continuación se presenta el espacio muestral para las diferentes respuestas posibles:

Va Vb Vc Vd Ve Fa Fb Fc Fd Fe

Si ambas respuestas son conjeturas aleatorias, entonces los 10 resultados posibles son igualmente probables, por lo que

$$P(\text{ambas correctas}) = P(\text{V y c}) = \frac{1}{10} = 0.1$$

Con $P(T \text{ y } c) = 1/10$, $P(T) = 1/2$ y $P(c) = 1/5$, se observa que

$$\frac{1}{10} = \frac{1}{2} \cdot \frac{1}{5}$$

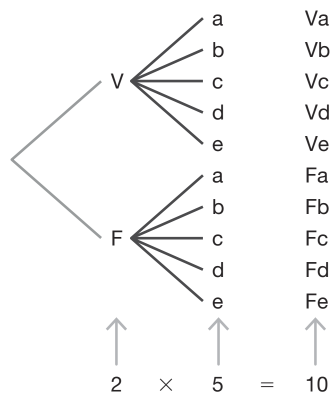


FIGURA 4-4 Diagrama de árbol de las respuestas de la prueba

Un *diagrama de árbol* es una gráfica de los posibles resultados de un procedimiento, como el de la figura 4-4. La figura 4-4 muestra que si ambas respuestas son conjeturas aleatorias, las 10 ramas son igualmente probables y la probabilidad de obtener el par correcto (V, c) es 1/10. Para cada respuesta a la primera pregunta, hay 5 respuestas a la segunda. *El número total de resultados es 5 tomado 2 veces o 10*. El diagrama de árbol de la figura 4-4 proporciona una ilustración visual para el uso de la multiplicación.

Resumen de la regla de la suma y la regla de la multiplicación

Regla de la suma para $P(A \text{ o } B)$: La palabra *o* sugiere adición, y al sumar $P(A)$ y $P(B)$, debemos sumarlas de modo que cada resultado se cuente sólo una vez.

Regla de la multiplicación para $P(\text{A y B})$: La palabra *y* para dos ensayos sugiere multiplicación, y al multiplicar $P(A)$ y $P(B)$, debemos asegurarnos que la probabilidad del evento B considera la ocurrencia previa del evento A.

4-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Notación Al seleccionar aleatoriamente a un adulto, A expresa el evento de seleccionar a alguien con ojos azules. ¿Qué representan $P(A)$ y $P(\overline{A})$?

2. Notación Al seleccionar adultos de manera aleatoria, H expresa el evento de seleccionar aleatoriamente a un hombre y A expresa el evento de seleccionar aleatoriamente a alguien con ojos azules. ¿Qué representa $P(H|A)$? ¿Es $P(H|A)$ lo mismo que $P(A|H)$?

3. Muestra para una encuesta Hay 15,524,971 adultos en Florida. Si la organización Gallup selecciona aleatoriamente a 1068 adultos sin reemplazo, ¿las selecciones son independientes o dependientes? Si las selecciones son dependientes, ¿pueden tratarse como independientes para realizar los cálculos?

4. Regla de los complementos Al seleccionar al azar a un adulto, B representa el evento de seleccionar aleatoriamente a alguien con sangre de tipo B. Redacte una oración que describa lo que dice la regla del complemento: $P(B \text{ o } \overline{B}) = 1$.

Determinación de complementos. ***En los ejercicios 5 a 8, encuentre los complementos indicados.***

5. LOL Una encuesta de U.S. Cellular a los usuarios de teléfonos inteligentes mostró que el 26% de los encuestados respondió "sí" cuando se les preguntó si las abreviaturas (como LOL) son molestas cuando se envían mensajes de texto. ¿Cuál es la probabilidad de seleccionar al azar a un usuario de teléfonos inteligentes y obtener una respuesta distinta de "sí"?

6. Vuelos Según la Oficina de Transporte, el 80.3% de los vuelos de American Airlines llegan a tiempo. ¿Cuál es la probabilidad de seleccionar aleatoriamente un vuelo de American Airlines que no llegue a tiempo?

7. Volar En una encuesta de Harris, se le preguntó a adultos con qué frecuencia viajan en vuelos comerciales, y se encontró que $P(N) = 0.330$, donde N expresa una respuesta de "nunca". ¿Qué representa $P(\overline{N})$ y cuál es su valor?

8. Punto de control de sobriedad Cuando el autor observó un punto de control de la sobriedad conducido por el departamento del comisario del condado de Dutchess, vio que 676 conductores fueron examinados y 6 resultaron arrestados por conducir mientras estaban intoxicados. Con base en esos resultados, podemos estimar que $P(I) = 0.00888$, donde I expresa el evento de seleccionar un conductor y obtener a alguien que está intoxicado. ¿Qué indica $P(\overline{I})$, y cuál es su valor?

En los ejercicios 9 a 20, utilice los datos de la siguiente tabla, que lista la precisión de las órdenes en las cadenas de comida rápida más populares (datos de un estudio de comida rápida ordenada en auto). Suponga que las órdenes se seleccionan aleatoriamente de los incluidos en la tabla.

	McDonald's	Burger King	Wendy's	Taco Bell
Orden exacta	329	264	249	145
Orden inexacta	33	54	31	13

9. Exactitud de comida rápida en auto Si se selecciona una orden, determine la probabilidad de obtener comida que no es de McDonald's.

10. Exactitud de comida rápida en auto Si se selecciona una orden, determine la probabilidad de obtener una orden que no es exacta.

11. Exactitud de comida rápida en auto Si se selecciona una orden, determine la probabilidad de obtener una orden de McDonald's o una orden que sea exacta. ¿Son los eventos de seleccionar una orden de McDonald's y seleccionar una orden exacta de eventos disjuntos?

12. Exactitud de comida rápida en auto Si se selecciona una orden, determine la probabilidad de obtener una orden que no es exacta o es de Wendy's. ¿Son los eventos de seleccionar una orden que no es exacta y seleccionar una orden de Wendy's de eventos disjuntos?

13. Exactitud de comida rápida en auto Si se seleccionan dos órdenes, determine la probabilidad de que ambas sean de Taco Bell.

a. Suponga que las selecciones se hacen con reemplazo. ¿Los eventos son independientes?

b. Suponga que las selecciones se hacen sin reemplazo. ¿Los eventos son independientes?

14. Exactitud de comida rápida en auto Si se seleccionan dos órdenes, determine la probabilidad de que ambas no sean exactas.

a. Suponga que las selecciones se realizan con reemplazo. ¿Los eventos son independientes?

b. Suponga que las selecciones se realizan sin reemplazo. ¿Los eventos son independientes?

15. Exactitud de comida rápida en auto Si se seleccionan dos órdenes, determine la probabilidad de que ambas sean exactas.

a. Suponga que las selecciones se hacen con reemplazo. ¿Los eventos son independientes?

b. Suponga que las selecciones se hacen sin reemplazo. ¿Los eventos son independientes?

16. Exactitud de comida rápida en auto Si se seleccionan dos órdenes, determine la probabilidad de que ambas sean de Burger King.

a. Suponga que las selecciones se hacen con reemplazo. ¿Los eventos son independientes?

b. Suponga que las selecciones se hacen sin reemplazo. ¿Los eventos son independientes?

17. Exactitud de comida rápida en auto Si se selecciona una orden, determine la probabilidad de obtener una orden de McDonald's o Wendy's o una orden que no sea exacta.

18. Exactitud de comida rápida en auto Si se selecciona una orden, determine la probabilidad de obtener una orden de Burger King o Taco Bell o una orden que sea exacta.

19. Exactitud de comida rápida en auto Si se seleccionan tres órdenes diferentes, encuentre la probabilidad de que sean todas de Wendy's.

20. Exactitud de comida rápida en auto Si se seleccionan tres órdenes diferentes, determine la probabilidad de que no sean exactas.

En los ejercicios 21 a 24, utilice estos resultados de la prueba "1-Panel-THC" para el uso de marihuana, que proporciona la compañía Drug Test Success: entre 143 sujetos con resultados positivos, hay 21 resultados falsos positivos; entre 157 resultados negativos, hay 3 resultados falsos negativos. (Sugerencia: elabore una tabla similar a la tabla 4-1, incluida con el problema del capítulo).

21. Pruebas para el uso de marihuana

a. ¿Cuántos sujetos se incluyen en el estudio?

b. ¿Cuántos de los sujetos tuvieron un resultado verdadero negativo?

c. ¿Cuál es la probabilidad de que un sujeto seleccionado al azar tuviera un resultado verdadero negativo?

22. Pruebas para el uso de marihuana Si uno de los sujetos de prueba se selecciona al azar, determine la probabilidad de que sea negativo o use marihuana.

23. Pruebas para el uso de marihuana Si uno de los sujetos de prueba se selecciona al azar, determine la probabilidad de que sea positivo o no use marihuana.

24. Pruebas para el uso de marihuana Si uno de los sujetos de prueba se selecciona al azar, determine la probabilidad de que no haya usado marihuana. ¿Cree usted que el resultado refleja la proporción de la población general que no usa marihuana?

Redundancia. ***Los ejercicios 25 y 26 implican redundancia.***

25. Redundancia en discos duros de computadora Por lo general se reconoce la recomendación de hacer una copia de seguridad de los datos de una computadora. Suponga que hay una tasa del 3% de falla en la unidad de disco en un año (con base en datos de varias fuentes, incluyendo lifehacker.com).

a. Si almacena todos los datos de su computadora en una sola unidad de disco duro, ¿cuál es la probabilidad de que la unidad falle durante un año?

continúa

b. Si todos los datos de su computadora están almacenados en una unidad de disco duro y otra copia almacenada en una segunda unidad de disco duro, ¿cuál es la probabilidad de que ambas unidades fallen durante un año?

c. Si se almacenan copias de todos los datos de su computadora en tres unidades de disco duro independientes, ¿cuál es la probabilidad de que las tres fallen durante un año?

d. Describa la confiabilidad mejorada que se obtiene con las unidades de respaldo.

26. Redundancia en generadores hospitalarios Los hospitales normalmente requieren generadores de respaldo para proporcionar electricidad en caso de un corte de energía. Suponga que los generadores de emergencia fallan 22% de las veces cuando son necesarios (de acuerdo con datos de Arshad Mansoor, vicepresidente senior del Instituto de Investigación en Energía Eléctrica). Un hospital tiene dos generadores de reserva para que la energía esté disponible si uno de ellos falla durante un corte de energía.

a. Determine la probabilidad de que ambos generadores fallen durante un corte de energía.

b. Determine la probabilidad de tener un generador funcionando en caso de un corte de energía. ¿Es esa probabilidad suficientemente alta para el hospital?

Muestreo de aceptación. ***Con un método de un procedimiento llamado muestreo de aceptación, una muestra de artículos se selecciona al azar sin reemplazo y el lote completo se acepta si cada artículo en la muestra está bien. Los ejercicios 27 y 28 incluyen el muestreo de aceptación.***

27. Marcapasos defectuosos Entre 8834 casos de funcionamiento inadecuado de marcapasos cardíacos, 504 resultaron ser causados por el firmware, que es un software programado en el dispositivo (de acuerdo con datos de "Pacemaker and ICD Generator Malfunctions", de Maisel *et al*., *Journal of the American Medical Association*, vol. 295, núm. 16). Si el firmware se prueba en tres marcapasos seleccionados aleatoriamente de este lote de 8834 y se acepta el lote completo si no hay fallas, ¿cuál es la probabilidad de que se acepte el firmware de todo el lote? ¿Es probable que este procedimiento resulte en que se acepte todo el lote?

28. Mariscos La Administración Nacional Oceánica y Atmosférica (NOAA, por sus siglas en inglés) inspecciona los mariscos que se van a consumir. El proceso de inspección consiste en seleccionar muestras de mariscos de un "lote" más grande. Suponga que un lote contiene 2875 contenedores de mariscos y 288 de estos contenedores incluyen mariscos que no cumplen con los requisitos de inspección. ¿Cuál es la probabilidad de que 3 muestras de contenedores seleccionados cumplan con los requisitos y el lote entero se acepte basado en esta muestra? ¿Esta probabilidad parece adecuada?

En los ejercicios 29 y 30, encuentre las probabilidades e indique cuándo se utiliza la "directriz del 5% para cálculos complejos".

29. Helicópteros médicos En un estudio sobre el uso de helicópteros y supervivencia de pacientes, se obtuvieron resultados de 47,637 pacientes transportados en helicóptero y 111,874 pacientes transportados por tierra (basado en datos de "Association Between Helicopter vs Ground Emergency Medical Services and Survival for Adults with Major Trauma", de Galvagno *et al*., *Journal of the American Medical Association*, vol. 307, núm. 15).

a. Si 1 de los 159,511 pacientes en el estudio se selecciona al azar, ¿cuál es la probabilidad de que el sujeto haya sido transportado en helicóptero?

b. Si 5 de los sujetos en el estudio se seleccionan al azar sin reemplazo, ¿cuál es la probabilidad de que todos ellos hayan sido transportados en helicóptero?

30. Helicópteros médicos En el mismo estudio citado en el ejercicio anterior, entre los 47,637 pacientes transportados en helicóptero, 188 de ellos abandonaron el centro de tratamiento contra consejo médico, y los otros 47,449 no abandonaron el centro médico. Si 40 de los sujetos transportados por helicóptero se seleccionan al azar sin reemplazo, ¿cuál es la probabilidad de que ninguno de ellos haya abandonado el centro de tratamiento contra consejo médico?

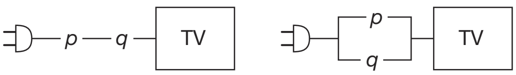
4-2 Más allá de lo básico

31. Protectores contra sobretensión Consulte la figura adjunta que muestra los protectores contra sobretensión p y q utilizados para proteger un televisor caro. Si hay un aumento de voltaje, el protector de sobretensión lo reduce a un nivel seguro. Suponga que cada protector de sobretensión tiene una probabilidad de 0.985 de funcionar correctamente cuando se produce un aumento de voltaje.

a. Si los dos protectores de sobretensión se disponen en serie, ¿cuál es la probabilidad de que una elevación del voltaje no dañe el televisor? (No redondee la respuesta).

b. Si los dos protectores de sobretensión se disponen en paralelo, ¿cuál es la probabilidad de que una elevación del voltaje no dañe el televisor? (No redondee la respuesta).

c. ¿Qué disposición se debe utilizar para una mejor protección?



Configuración en serie Configuración en paralelo

32. Mismos cumpleaños Si 25 personas se seleccionan al azar, determine la probabilidad de que ninguna pareja formada entre ellas tengan el mismo cumpleaños. No tome en cuenta los años bisiestos.

33. O exclusiva La *o exclusiva* significa que ocurre ya sea uno u otro evento, pero no ambos.

a. Para la regla de la suma formal, reescriba la fórmula para $P(A \text{ o } B)$ suponiendo que la regla de la suma utiliza la *o exclusiva* en lugar de la *o inclusiva*.

b. Repita el ejercicio 11 "Exactitud de comida rápida en auto" usando la *o exclusiva* en lugar de la *o inclusiva*.

34. Complementos y la regla de la suma Consulte la tabla utilizada para los ejercicios 9 a 20. Suponga que se selecciona aleatoriamente una orden. Considere que A representa el evento de obtener una orden de McDonald's y B representa el evento de obtener una orden de Burger King. Determine $P(\overline{A \text{ o } B})$, encuentre $P(\overline{A} \text{ o } \overline{B})$; y luego compare los resultados. En general, ¿$P(\overline{A \text{ o } B}) = P(\overline{A} \text{ o } \overline{B})$?

4-3 Complementos, probabilidad condicional y teorema de Bayes

Concepto clave En la parte 1 de esta sección extendemos el uso de la regla de la multiplicación para incluir la probabilidad de que, entre varios ensayos, obtengamos *al menos uno* de algún evento especificado. En la parte 2 consideramos la *probabilidad condicional*: la probabilidad de que ocurra un evento cuando tenemos información adicional de que algún otro evento ya ha ocurrido. En la parte 3 ofrecemos una breve introducción al teorema de Bayes.

PARTE 1 Complementos: La probabilidad de "al menos uno"

Cuando se determina la probabilidad de que ocurra algún evento "al menos una vez", debemos entender lo siguiente:

- "Al menos uno" tiene el mismo significado que "uno o más".
- El *complemento* de obtener "al menos un" evento particular es que no se obtenga ninguna ocurrencia de ese evento.

Sentenciados por probabilidad

Un testigo describió a una asaltante de Los Ángeles como una mujer caucásica de pelo rubio, peinada con cola de caballo, que escapó en un automóvil amarillo conducido por un hombre afroestadounidense que usaba barba y bigote. Janet y Malcolm Collins se ajustaban a esta descripción y se les condenó con fundamento en el testimonio de que existe aproximadamente 1 posibilidad en 12 millones de que cualquier pareja tenga tales características. Se estimó que la probabilidad de tener un automóvil amarillo es de 1/10, en tanto que las demás probabilidades se estimaron en 1/10, 1/3, 1/10 y 1/1000. Más tarde, las condenas se anularon, cuando se señaló que no se presentó evidencia que apoyara las probabilidades estimadas o la independencia de los eventos. Sin embargo, puesto que la pareja no se seleccionó al azar, se cometió un error grave al no considerar la probabilidad de que hubiera otras parejas en la misma región con las mismas características.

Por ejemplo, no conseguir por lo menos 1 niña en 10 nacimientos es lo mismo que no tener niñas, lo que también es igual a conseguir 10 niños.

No obtener por lo menos 1 niña en 10 nacimientos = No obtener niñas = Obtener 10 niños

Los pasos siguientes describen los detalles de la determinación de la probabilidad de obtener un evento al menos una vez:

Determinación de la probabilidad de obtener un evento *al menos una* vez:

1. **Sea A = obtener un evento *al menos una* vez.**
2. **Entonces $\overline{A}$ = no obtener *ninguna* vez el evento considerado.**
3. **Encuentre $P(\overline{A})$ = probabilidad de que el evento A no ocurra.** (Esto es relativamente fácil usando la regla de la multiplicación).
4. **Reste el resultado de 1. Es decir, evalúe esta expresión:**

$$\begin{aligned} &P(\textit{al menos una} \text{ ocurrencia del evento } A) \\ &= 1 - P(\textit{no} \text{ haya ocurrencias del evento } A) \end{aligned}$$

EJEMPLO 1 **Daño accidental de iPad**

Un estudio de SquareTrade encontró que 6% de los iPads dañados fueron afectados por "bolsas/mochilas". Si se seleccionan 20 iPads aleatoriamente, encuentre la probabilidad de obtener *al menos uno* que fue dañado en una bolsa/mochila. ¿Es la probabilidad suficientemente alta para que podamos estar razonablemente seguros de que por lo menos habrá un iPad dañado en una bolsa/mochila?

SOLUCIÓN

Paso 1: Sea A = al menos 1 de los 20 iPads fue dañado en una bolsa/mochila.

Paso 2: Identifique el evento que es el complemento de A.

$$\begin{aligned} \overline{A} &= \textit{no} \text{ obtener al menos 1 iPad dañado en una bolsa/mochila entre 20} \\ &= \text{los 20 iPads dañados de una manera diferente a en bolsa/mochila} \end{aligned}$$

Paso 3: Encuentre la probabilidad del complemento evaluando $P(\overline{A})$.

$$\begin{aligned} P(\overline{A}) &= P(\text{los 20 iPads dañados de una manera diferente a en bolsa/mochila}) \\ &= 0.94 \cdot 0.94 \cdot \ldots \cdot 0.94 \\ &= 0.94^{20} = 0.290 \end{aligned}$$

Paso 4: Encuentre $P(A)$ evaluando $1 - P(\overline{A})$.

$$P(A) = 1 - P(\overline{A}) = 1 - 0.290 = 0.710$$

INTERPRETACIÓN

Para un grupo de 20 iPads dañados, hay una probabilidad de 0.710 de obtener al menos 1 iPad dañado en una bolsa/mochila. Esta probabilidad no es muy alta, por lo que para estar razonablemente seguro de conseguir por lo menos 1 iPad dañado de esta forma, se deben utilizar más de 20 iPads dañados.

SU TURNO Resuelva el ejercicio 7 "Nacimientos en Estados Unidos".

PARTE 2 Probabilidad condicional

Ahora consideramos el principio de que a menudo la probabilidad de un evento es afectada por el conocimiento de que algún otro evento ha ocurrido. Por ejemplo, la probabilidad de que un golfista logre un hoyo en uno es 1/12,000 (con base en resultados anteriores), pero si se tiene el conocimiento adicional de que el golfista seleccionado es un profesional, la probabilidad cambia a 1/2375 (según datos de *USA Today*). En general, una *probabilidad condicional* de un evento se utiliza cuando la probabilidad debe calcularse con algún conocimiento adicional, como la certeza de que algún otro evento ha ocurrido. (Se usaron probabilidades condicionales en la sección 4-2 con situaciones en las que se seleccionaron muestras sin reemplazo).

DEFINICIÓN

Una **probabilidad condicional** de un evento es una probabilidad obtenida con la información adicional de que algún otro evento ya ha ocurrido.

Notación

$P(B \mid A)$ expresa la probabilidad condicional de que ocurra el evento B, dado que el evento A ya ha ocurrido.

MÉTODO INTUITIVO PARA ENCONTRAR $P(B \mid A)$

La probabilidad condicional de que B ocurra dado que A ha ocurrido se puede encontrar *asumiendo que el evento A ha ocurrido* y luego calcular la probabilidad de que ocurra el evento B, como se ilustra en el ejemplo 2.

MÉTODO FORMAL PARA ENCONTRAR $P(B \mid A)$

La probabilidad $P(B \mid A)$ se puede encontrar dividiendo la probabilidad de que los eventos A y B ocurran por la probabilidad del evento A:

$$P(B \mid A) = \frac{P(A \text{ y } B)}{P(A)}$$

La fórmula anterior es una expresión formal de la probabilidad condicional, pero no se recomienda el uso ciego de las fórmulas. En cambio, recomendamos el método intuitivo, como se ilustra en el ejemplo 2.

EJEMPLO 2 Detección de drogas previa al empleo

Consulte la tabla 4-1 en la página siguiente para encontrar lo siguiente:

a. Si 1 de los 555 sujetos de prueba se selecciona al azar, determine la probabilidad de que el sujeto tenga un resultado de prueba positivo, dado que el sujeto realmente usa drogas. Es decir, encuentre $P(\text{resultado positivo en la prueba} \mid \text{el sujeto usa drogas})$.

b. Si 1 de los 555 sujetos de prueba se selecciona al azar, determine la probabilidad de que el sujeto realmente use drogas, dado que tuvo un resultado positivo en la prueba. Es decir, encuentre $P(\text{el sujeto usa drogas} \mid \text{resultado positivo en la prueba})$.

continúa

La falacia del fiscal

La *falacia del fiscal* es una confusión o un malentendido de dos diferentes probabilidades condicionales: (1) la probabilidad de que un acusado sea inocente, dado que la evidencia forense demuestre una coincidencia; (2) la probabilidad de que la evidencia forense demuestre una coincidencia, dado que una persona sea inocente. La falacia del fiscal ha provocado condenas y encarcelamientos de personas inocentes.

Lucia de Berk es una enfermera que fue condenada por asesinato y sentenciada a prisión en los Países Bajos. Los administradores del hospital observaron muertes sospechosas que ocurrieron en los pabellones del hospital en que De Berk había estado presente. Un experto testificó que sólo había una probabilidad en 342 millones de que su presencia fuera una coincidencia. Sin embargo, el matemático Richard Gill calculó que la probabilidad se acercaba más a 1/50, o que posiblemente fuera tan baja como 1/5. El tribunal utilizó la probabilidad de que las muertes sospechosas podrían haber ocurrido con De Berk presente, dado que era inocente. El tribunal debería haber considerado la probabilidad de que De Berk fuera inocente, dado que las muertes sospechosas ocurrieron cuando ella estaba presente. Este error de la falacia del fiscal es útil y puede ser muy difícil de entender y reconocer, aun cuando podría derivar en el encarcelamiento de personas inocentes.

Pruebas compuestas

Durante la Segunda Guerra Mundial, el ejército estadounidense realizó pruebas de sífilis haciendo a cada soldado un análisis de sangre que se estudiaba por separado. Un investigador sugirió mezclar pares de muestras de sangre. Después de probar los pares mixtos, los que tenían sífilis podrían ser identificados reevaluando las pocas muestras de sangre que estaban en las muestras positivas. Dado que el número total de análisis se redujo al emparejar muestras de sangre, ¿por qué no combinarlas en grupos de tres o cuatro o más? Esta técnica de combinar muestras en grupos y volver a probar sólo aquellos grupos que dan positivo se conoce como *pruebas compuestas* o *pruebas en grupo*. El estadístico Christopher Bilder de la Universidad de Nebraska escribió un artículo sobre este tema en la revista *Chance* y citó algunas aplicaciones reales. Señaló que la Cruz Roja estadounidense usa pruebas compuestas para examinar enfermedades específicas, como la hepatitis, y las pruebas en grupo son usadas por veterinarios cuando se analiza el ganado vacuno en busca del virus de la diarrea viral bovina.

TABLA 4-1 Resultados de las pruebas de drogas a solicitantes de empleo

	Resultado positivo de la prueba (El examen muestra el uso de drogas)	**Resultado negativo de la prueba** (El examen no muestra uso de drogas)
El sujeto usa drogas	**45** (Verdadero positivo)	**5** (Falso negativo)
El sujeto no usa drogas	**25** (Falso positivo)	**480** (Verdadero negativo)

SOLUCIÓN

a. Método intuitivo: Deseamos encontrar P(resultado positivo de la prueba | sujeto usa drogas), la probabilidad de obtener a alguien con un resultado de prueba positivo, *dado que el sujeto seleccionado usa drogas*. Aquí está el punto clave: Si asumimos que el sujeto seleccionado realmente usa drogas, estamos tratando solamente con los 50 sujetos en la primera fila de la tabla 4-1. Entre los 50 sujetos, 45 tuvieron resultados positivos, por lo que tenemos el siguiente resultado:

$$P(\text{resultado positivo de la prueba} \mid \text{sujeto usa drogas}) = \frac{45}{50} = 0.900$$

Método formal: El mismo resultado se puede encontrar usando la fórmula para $P(B \mid A)$ dada con el método formal. Usamos la siguiente notación.

$$P(B \mid A) = P(\text{resultado positivo de la prueba} \mid \text{el sujeto usa drogas})$$

donde B = resultado positivo de la prueba y A = el sujeto usa drogas.

En el siguiente cálculo, utilizamos P(el sujeto usa drogas y tuvo un resultado positivo) = 45/555 y P(el sujeto usa drogas) = 50/555 para obtener los siguientes resultados:

$$P(B \mid A) = \frac{P(A \text{ y } B)}{P(A)}$$

se convierte en

$$P(\text{resultado positivo de la prueba} \mid \text{el sujeto usa drogas})$$
$$= \frac{P(\text{el sujeto usa drogas y tuvo un resultado de prueba positivo})}{P(\text{el sujeto usa drogas})}$$
$$= \frac{45/555}{50/555} = 0.900$$

Al comparar el método intuitivo con el método formal, debe quedar claro que el método intuitivo es mucho más fácil de usar, y también es menos probable que resulte en errores. El método intuitivo se basa en la comprensión de la probabilidad condicional, en lugar de la manipulación de una fórmula, y comprender siempre es mucho mejor.

b. Aquí buscamos P(el sujeto usa drogas | resultado positivo de la prueba). Esta es la probabilidad de que el sujeto seleccionado use drogas, dado que el sujeto tuvo un resultado de prueba positivo. Si asumimos que el sujeto tuvo un resultado de prueba positivo, estamos tratando con los 70 sujetos en la primera columna de la tabla 4-1. Entre los 70 sujetos, 45 usan drogas, por lo que

$$P(\text{el sujeto usa drogas} \mid \text{resultado positivo de la prueba}) = \frac{45}{70} = 0.643$$

De nuevo, el mismo resultado se puede encontrar aplicando la fórmula de probabilidad condicional, pero quedará a discreción de aquellas personas con un especial cariño por la manipulación de fórmulas.

INTERPRETACIÓN

El primer resultado de P(resultado positivo de la prueba | el sujeto usa drogas) = 0.900 indica que un sujeto que usa drogas tiene una probabilidad de 0.900 de obtener un resultado positivo de la prueba. El segundo resultado de P(el sujeto usa drogas | resultado positivo de la prueba) = 0.643 indica que para un sujeto que obtiene un resultado positivo de la prueba, existe una probabilidad de 0.643 de que este sujeto realmente use drogas. Observe que P(resultado positivo de la prueba | el sujeto usa drogas) $\neq$ P(el sujeto usa drogas | resultado positivo de la prueba). Vea a continuación "Confusión del inverso".

SU TURNO Resuelva el ejercicio 13 "Efecto de la denominación".

Confusión del inverso

Observe que en el ejemplo 2, P(resultado positivo de la prueba | el sujeto usa drogas) $\neq$ P(el sujeto usa drogas | resultado positivo de la prueba). Este ejemplo demuestra que, en general, $P(B \mid A) \neq P(A \mid B)$. (Podría haber casos individuales donde $P(A \mid B)$ y $P(B \mid A)$ sean iguales, pero generalmente no lo son). Pensar erróneamente que $P(B \mid A)$ y $P(A \mid B)$ son iguales o utilizar incorrectamente un valor en lugar del otro se llama *confusión del inverso*.

EJEMPLO 3 Confusión del inverso

Considere estos eventos:

D: Está oscuro afuera.

M: Es medianoche.

A continuación, convenientemente ignoramos el invierno de Alaska y otras anomalías de este tipo.

$P(D \mid M) = 1$ (Es cierto que está oscuro dado que es medianoche).

$P(M \mid D) = 0$ (La probabilidad de que sea exactamente medianoche dado que está oscuro es casi cero).

Aquí, $P(D \mid M) \neq P(M \mid D)$. La confusión del inverso ocurre cuando intercambiamos incorrectamente esos valores de probabilidad o pensamos que son iguales.

PARTE 3 Teorema de Bayes

En esta sección se extiende el estudio de la probabilidad condicional para incluir aplicaciones del *teorema de Bayes* (o *regla de Bayes*), que usamos para modificar un valor de probabilidad con base en información adicional obtenida posteriormente.

Consideremos un estudio que muestra que los médicos a menudo dan información muy engañosa cuando experimentan la confusión del inverso. Tienden a confundir P(cáncer | resultado positivo de la prueba) con P(resultado positivo de la prueba | cáncer). (Un resultado positivo de la prueba indica que el paciente tiene cáncer, un resultado negativo de la prueba indica que el paciente no tiene cáncer). Aproximadamente 95% de los médicos estimó que P(cáncer | resultado positivo de la prueba positiva) era 10 veces demasiado alto, con el resultado de que algunos pacientes recibieron diagnósticos engañosos y estuvieron innecesariamente angustiados por información incorrecta. Echemos un vistazo a este ejemplo, y esperemos que podamos dar a los médicos la información en un formato mejor que sea fácil de entender.

Probabilidad de un evento que nunca ha ocurrido

tu voto cuenta

Algunos eventos son posibles, pero tan poco probables que nunca han ocurrido. He aquí un problema de este tipo, de gran interés para los científicos políticos: estime la probabilidad de que su voto determine el ganador de una elección presidencial de Estados Unidos. Andrew Gelman, Gary King y John Boscardin escribieron en el *Journal of the American Statistical Association* (vol. 93, núm. 441) que "el valor exacto de esta probabilidad es de interés menor, pero el número tiene implicaciones importantes para la comprensión de la asignación óptima de los recursos de campaña; de esta forma, se podría saber si los estados y los grupos de votantes reciben su parte justa de atención de los candidatos presidenciales y de qué manera los modelos formales de 'elección racional' del comportamiento del votante son capaces de explicar por qué las personas votan". Los autores demuestran cómo se obtiene el valor de probabilidad de 1 en 10 millones para elecciones cerradas.

EJEMPLO 4 Interpretación de resultados de exámenes médicos

Suponga que el cáncer tiene una tasa de prevalencia del 1%, lo que significa que el 1% de la población tiene cáncer. Expresando el evento de tener cáncer por C, tenemos $P(C) = 0.01$ para un sujeto seleccionado aleatoriamente de la población. Este resultado se incluye con las siguientes características de desempeño para la prueba del cáncer (con base en *Probabilistic Reasoning in Clinical Medicine* de David Eddy, Cambridge University Press).

- $P(C) = 0.01$ (Existe una tasa de prevalencia del 1% del cáncer).
- La tasa de falsos positivos es del 10%. Es decir, P(resultado positivo en la prueba dado que el cáncer no está presente) = 0.10.
- La tasa de positivos verdaderos es del 80%. Es decir, P(resultado positivo en la prueba dado que el cáncer está presente) = 0.80.

Determine $P(C\,|$ resultado positivo en la prueba). Es decir, encuentre la probabilidad de que un sujeto realmente tenga cáncer dado que obtuvo un resultado positivo en la prueba.

SOLUCIÓN

Usando la información dada, podemos elaborar una población hipotética con las características anteriores. Es posible encontrar las entradas de la tabla 4-2 de la siguiente página, como sigue.

- Asumir que tenemos 1000 sujetos. Con una tasa de prevalencia del 1%, se espera que 10 de los sujetos tengan cáncer. La suma de las entradas en la primera fila de valores es, por tanto, 10.
- Los otros 990 sujetos no tienen cáncer. La suma de las entradas en la segunda fila de valores es, por tanto, 990.
- Entre los 990 sujetos sin cáncer, el 10% obtiene resultados positivos de la prueba, por lo que el 10% de los 990 sujetos libres de cáncer en la segunda fila obtienen resultados positivos. Vea la entrada de 99 en la segunda fila.
- Para los 990 sujetos en la segunda fila, 99 dieron positivo en la prueba, por lo que los otros 891 deben ser negativos. Vea la entrada de 891 en la segunda fila.
- Entre los 10 sujetos con cáncer en la primera fila, el 80% de los resultados de la prueba son positivos, por lo que 80% de los 10 sujetos de la primera fila resultan positivos. Vea la entrada de 8 en la primera fila.
- Los otros 2 sujetos en la primera fila son negativos. Vea la entrada de 2 en la primera fila.

Para encontrar $P(C\,|$ resultado positivo en la prueba), vea que la primera columna de valores incluye los resultados positivos de la prueba. En esa primera columna, la probabilidad de seleccionar aleatoriamente un sujeto con cáncer es 8/107 o 0.0748, por lo que $P(C\,|$ resultado positivo en la prueba) = 0.0748.

INTERPRETACIÓN

Para los datos dados de este ejemplo, un sujeto seleccionado al azar tiene una probabilidad del 1% de cáncer, pero para un sujeto seleccionado al azar dado que tuvo una prueba con resultado positivo, la probabilidad de cáncer aumenta a 7.48%. Sobre la base de los datos dados en este ejemplo, un resultado de prueba positivo no debe ser una noticia devastadora, porque todavía hay una buena probabilidad de que la prueba sea incorrecta.

TABLA 4-2 Resultados de prueba

	Resultado de prueba positivo (La prueba indica cáncer)	Resultado de prueba negativo (La prueba indica sin cáncer)	Total
Cáncer	8 (verdadero positivo)	2 (falso negativo)	10
Sin cáncer	99 (falso positivo)	891 (verdadero negativo)	990

La solución en el ejemplo 4 no es muy difícil. Otro método es calcular la probabilidad usando la siguiente fórmula que se da comúnmente junto con el teorema de Bayes:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{[P(A) \cdot P(B|A)] + [P(\overline{A}) \cdot P(B|\overline{A})]}$$

Si sustituimos A por C y reemplazamos B por "positivo", obtenemos la solución para el ejemplo 4:

$$P(C|\text{positivo}) = \frac{P(C) \cdot P(\text{positivo}|C)}{P(C) \cdot P(\text{positivo}|C) + P(\overline{C}) \cdot P(\text{positivo}|\overline{C})}$$

$$= \frac{0.01 \cdot 0.80}{(0.01 \cdot 0.80) + (0.99 \cdot 0.10)} = 0.0748$$

Resultados del estudio He aquí un hecho verdaderamente fascinante: cuando a 100 médicos se les dio la información del ejemplo 4, 95 de ellos estimaron que $P(C|\text{positivo})$ estaba alrededor de 0.70 a 0.80, por lo que estaban equivocados por un factor de 10. Los médicos son extremadamente inteligentes, pero aquí probablemente sufrían de la confusión del inverso. La tasa dada de 80% para los resultados de las pruebas positivas entre los que son verdaderos positivos implica que $P(\text{positivo}|C) = 0.80$, pero esto es muy diferente de $P(C|\text{positivo})$. Los médicos hubieran tenido mejores resultados si hubieran visto la información dada, en formato de tabla, en la tabla 4-2.

La importancia y utilidad del teorema de Bayes es que puede usarse con eventos *secuenciales*, donde se obtiene nueva información adicional para un evento subsecuente y esa nueva información se usa para modificar la probabilidad del evento inicial. En este contexto, los términos *probabilidad a priori* y *probabilidad a posteriori* se usan comúnmente.

DEFINICIONES

Una **probabilidad *a priori*** es un valor de probabilidad inicial obtenido originalmente antes de obtener cualquier información adicional.

Una **probabilidad *a posteriori*** es un valor de probabilidad que ha sido modificado con base en información adicional obtenida posteriormente.

En relación con el ejemplo 4, $P(C) = 0.01$, que es la probabilidad de que un sujeto seleccionado aleatoriamente tenga cáncer. $P(C)$ es un ejemplo de una *probabilidad a priori*. Si se utiliza la información adicional de que el sujeto ha obtenido un resultado positivo en la prueba, encontramos que $P(C|\text{resultado positivo en la prueba}) = 0.0748$, y esta es una *probabilidad a posteriori* porque usa la información adicional del resultado positivo en la prueba.

Las coincidencias son más probables de lo que parecen

Evelyn Evans ganó \$3.9 millones en la lotería de Nueva Jersey, luego ganó otros \$1.5 millones sólo 4 meses más tarde. El *New York Times* informó que la posibilidad de que ocurriera esto era sólo de 1 en 17 billones. Pero ese dato es engañoso porque representa la posibilidad de que Evelyn Evans gane con sólo un billete comprado en cada uno de los dos sorteos de lotería específicos. Una mejor pregunta sería la siguiente: ¿cuál es la posibilidad de que alguien en un lugar gane la lotería dos veces? Los estadísticos George McCabe y Steve Samuels descubrieron que en un lapso de 7 años, hay un 53% de posibilidades de que al menos un ganador de la lotería tenga suerte y gane otro sorteo. La posibilidad de "1 en 17 billones" es sensacional, pero la cifra más realista es 53%.

4-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Lenguaje: Complemento de "al menos uno" Sea A = el evento de obtener al menos un iPhone defectuoso cuando se seleccionan de un lote de iPhones 3 unidades al azar con reemplazo. Escriba un enunciado que describa el evento $\overline{A}$.

2. Probabilidad de al menos uno Sea A = el evento de obtener al menos 1 iPhone defectuoso cuando se seleccionan al azar 3 iPhones con reemplazo de un lote. Si el 5% de los iPhones de un lote son defectuosos y el otro 95%, son todos funcionales, ¿cuáles de las siguientes expresiones son correctas?

a. $P(\overline{A}) = (0.95)(0.95)(0.95) = 0.857$

b. $P(A) = 1 - (0.95)(0.95)(0.95) = 0.143$

c. $P(A) = (0.05)(0.05)(0.05) = 0.000125$

3. Notación Al seleccionar a uno de sus amigos de Facebook, sea el evento F = mujer y el evento H = compañero de clase de la escuela preparatoria. Utilice sus propias palabras para convertir la notación $P(H|F)$ en una declaración verbal.

4. Confusión del inverso Si se utilizan los mismos eventos F y H descritos en el ejercicio 3, describa la confusión del inverso.

Al menos uno. ***En los ejercicios 5 a 12, determine la probabilidad.***

5. Tres niñas Determine la probabilidad de que cuando una pareja tiene tres hijos, al menos uno de ellos sea una niña. (Suponga que los niños y las niñas son igualmente probables).

6. Probabilidad de una niña Suponiendo que los niños y las niñas son igualmente probables, determine la probabilidad de que una pareja tenga un niño en el nacimiento de su tercer hijo, dado que sus dos primeros hijos fueron niñas.

7. Nacimientos en Estados Unidos En Estados Unidos, la verdadera probabilidad de que un bebé sea un niño es 0.512 (con base en los datos disponibles en este texto). Entre los siguientes seis nacimientos seleccionados aleatoriamente en Estados Unidos, ¿cuál es la probabilidad de que al menos uno de ellos sea una niña?

8. Nacimientos en China En China, donde a muchas parejas se les permitía tener sólo un hijo, la probabilidad de que un bebé fuera un niño era 0.545. Entre seis nacimientos seleccionados aleatoriamente en China, ¿cuál es la probabilidad de que al menos uno de ellos sea una niña? ¿Podría este sistema seguir funcionando indefinidamente? (La eliminación de esta política inició en 2015).

9. Encuesta telefónica Los sujetos para la próxima encuesta de elecciones presidenciales son contactados usando números telefónicos en los que los últimos cuatro dígitos se seleccionan al azar (con reemplazo). Encuentre la probabilidad de que para uno de estos números telefónicos, los últimos cuatro dígitos incluyan al menos un 0.

10. Al menos una respuesta correcta Si usted hace conjeturas aleatorias para 10 preguntas del examen SAT de opción múltiple (cada una con cinco respuestas posibles), ¿cuál es la probabilidad de conseguir por lo menos 1 correcta? Si estas preguntas son parte de un examen de práctica y un instructor dice que usted debe obtener al menos una respuesta correcta antes de continuar, ¿hay una buena probabilidad de que continúe?

11. Al menos un iPhone defectuoso Se ha informado de que el 20% de los iPhones fabricados por Foxconn para un lanzamiento de producto no cumplen con los estándares de calidad de Apple. Una ingeniera necesita al menos un iPhone defectuoso para intentar identificar el (los) problema(s). Si ella selecciona al azar 15 iPhones de un lote muy grande, ¿cuál es la probabilidad de obtener al menos uno defectuoso? ¿Es esa probabilidad suficientemente alta como para que la ingeniera esté razonablemente segura de tener un defecto sobre el cual realizar su trabajo?

12. Wi-Fi Con base en una encuesta realizada a través de la edición electrónica de *USA Today*, 67% de los usuarios de Internet son más cuidadosos con la información personal cuando usan un hotspot

Wi-Fi público. ¿Cuál es la probabilidad de que, entre cuatro usuarios de Internet seleccionados al azar, al menos uno sea más cuidadoso con la información personal al usar un hotspot Wi-Fi público? ¿Cómo se ve afectado el resultado por la información adicional de que los sujetos encuestados se ofrecieron voluntariamente a responder?

Efecto de la denominación. ***En los ejercicios 13 a 16, utilice los datos de la siguiente tabla. En un experimento para estudiar los efectos de usar cuatro monedas de veinticinco centavos o un billete de $1, a estudiantes universitarios se les dieron las cuatro monedas o el billete y ellos podrían conservar el dinero o gastarlo en goma de mascar. Los resultados se resumen en la tabla (según datos de "The Denomination Effect", de Priya Raghubir y Joydeep Srivastava,*** **Journal of Consumer Research*****, vol. 36).***

	Compraron goma de mascar	Conservaron el dinero
Estudiantes que recibieron cuatro monedas de $0.25	27	16
Estudiantes que recibieron un billete de $1	12	34

13. Efecto de la denominación

a. Encuentre la probabilidad de seleccionar al azar a un estudiante que gastó el dinero, dado que recibió cuatro monedas de $0.25.

b. Determine la probabilidad de seleccionar al azar a un estudiante que conservó el dinero, dado que recibió cuatro monedas de $0.25.

c. ¿Qué sugieren los resultados anteriores?

14. Efecto de la denominación

a. Encuentre la probabilidad de seleccionar al azar a un estudiante que gastó el dinero, dado que recibió un billete de $1.

b. Determine la probabilidad de seleccionar al azar a un estudiante que conservó el dinero, dado que recibió un billete de $1.

c. ¿Qué sugieren los resultados anteriores?

15. Efecto de la denominación

a. Encuentre la probabilidad de seleccionar al azar a un estudiante que gastó el dinero, dado que recibió cuatro monedas de $0.25.

b. Determine la probabilidad de seleccionar al azar a un estudiante que gastó el dinero, dado que recibió un billete de $1.

c. ¿Qué sugieren los resultados anteriores?

16. Efecto de la denominación

a. Encuentre la probabilidad de seleccionar al azar a un estudiante que conservó el dinero, dado que recibió cuatro monedas de $0.25.

b. Determine la probabilidad de seleccionar al azar a un estudiante que conservó el dinero, dado que recibió un billete de $1.

c. ¿Qué sugieren los resultados anteriores?

En los ejercicios 17 a 20, consulte la tabla adjunta que muestra los resultados de una prueba Chembio para la hepatitis C, entre pacientes infectados por el VIH (con base en datos de una variedad de fuentes).

	Resultado positivo en la prueba	Resultado negativo en la prueba
Hepatitis C	335	10
Sin hepatitis C	2	1153

17. Falso positivo Encuentre la probabilidad de seleccionar un sujeto con resultado positivo en la prueba, dado que no tiene hepatitis C. ¿Por qué este caso es problemático para los sujetos de prueba?

18. Falso negativo Determine la probabilidad de seleccionar un sujeto con resultado negativo en la prueba, dado que tiene hepatitis C. ¿Cuál sería una consecuencia desfavorable de este error?

19. Valor predictivo positivo Encuentre el valor predictivo positivo para la prueba. Es decir, determine la probabilidad de que un sujeto tenga hepatitis C, dado que la prueba da un resultado positivo. ¿El resultado hace que la prueba parezca ser efectiva?

20. Valor predictivo negativo Encuentre el valor predictivo negativo para la prueba. Es decir, determine la probabilidad de que un sujeto no tenga hepatitis C, dado que la prueba da un resultado negativo. ¿El resultado hace que la prueba parezca ser efectiva?

21. Redundancia de discos duros de computadora Suponga que hay una tasa de 3% para las fallas en unidades de disco en un año (según datos de varias fuentes, incluyendo lifehacker.com).

a. Si todos los datos de su computadora están almacenados en una unidad de disco duro con una copia almacenada en otra unidad de disco duro, ¿cuál es la probabilidad de que durante un año, pueda evitar una catástrofe con al menos una unidad en funcionamiento? Exprese el resultado con cuatro cifras decimales.

b. Si todos los datos de su computadora se copian y almacenan en tres unidades de disco duro independientes, ¿cuál es la probabilidad de que durante un año, pueda evitar una catástrofe con al menos una unidad en funcionamiento? Exprese el resultado con seis cifras decimales. ¿Qué hay de malo en usar la regla de redondeo usual para las probabilidades en este caso?

22. Redundancia en generadores de estadio Los estadios grandes dependen de generadores de reserva para suministrar electricidad en caso de una falla de energía. Suponga que los generadores de emergencia fallan 22% de las veces cuando son requeridos (según datos de Arshad Mansoor, vicepresidente senior del Instituto de Investigaciones en Energía Eléctrica). Un estadio tiene tres generadores de respaldo para que haya energía disponible si por lo menos uno de ellos funciona durante una falla de energía. Encuentre la probabilidad de que al menos uno de los generadores de respaldo funcione dado que se ha producido una falla de alimentación. ¿El resultado parece ser adecuado para las necesidades del estadio?

23. Pruebas de drogas compuestas Con base en los datos de la tabla 4-1 de la página 162, suponga que la probabilidad de que una persona seleccionada al azar sea positiva para el consumo de drogas es de 0.126. Si se recolectan aleatoriamente las muestras para detección de drogas de 5 sujetos y se combinan, encuentre la probabilidad de que la muestra combinada revele un resultado positivo. ¿Es esa probabilidad suficientemente baja para que las pruebas adicionales de muestras individuales se requieran en muy pocas ocasiones?

24. Muestras de agua compuestas El Departamento de Salud Pública del Condado de Fairfield examina el agua para detectar la presencia de bacterias de *E. coli* (*Escherichia coli*). Para reducir los costos de laboratorio, las muestras de agua de 10 áreas públicas de natación se combinan para una prueba, y se realizan pruebas adicionales sólo si las pruebas de la muestra combinada resultan positivas. Con base en resultados anteriores, hay una probabilidad de 0.005 de encontrar bacterias *E. coli* en un área de natación pública. Determine la probabilidad de que una muestra combinada de 10 áreas de natación pública revele la presencia de bacterias *E. coli*. ¿Es esa probabilidad suficientemente baja para que las pruebas adicionales a muestras individuales se requieran en muy pocas ocasiones?

4-3 Más allá de lo básico

25. Cumpleaños compartidos Encuentre la probabilidad de que entre 25 personas seleccionadas al azar, por lo menos 2 compartan el mismo cumpleaños.

4-4 Conteo

Concepto clave Por lo general, los problemas de probabilidad requieren que conozcamos el número total de eventos simples, pero para encontrar ese número a menudo se debe usar una de las cinco reglas presentadas en esta sección. En la sección 4-2, con la regla de la suma, la regla de la multiplicación y la probabilidad condicional, fomentamos reglas intuitivas basadas en la comprensión y desalentamos el uso ciego de fórmulas, pero en esta sección se necesita un uso mucho mayor de fórmulas al considerar cinco métodos para contar el número de posibles resultados en una variedad de situaciones. No todos los problemas de conteo pueden resolverse con estos cinco métodos, pero proporcionan una base sólida para las aplicaciones reales más comunes.

La secretaria aleatoria

Un problema clásico de probabilidad dice así: una secretaria prepara 50 cartas distintas y las dirige a 50 personas, pero las revuelve al azar antes de meterlas en los sobres. ¿Qué probabilidad hay de que al menos una carta quede en el sobre que le corresponde? Aunque tal vez parezca 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 lejos del alcance de este texto, muy lejos.

1. Regla de conteo por multiplicación

La *regla de conteo por multiplicación* se utiliza para encontrar el número total de posibilidades de una secuencia de eventos.

REGLA DE CONTEO POR MULTIPLICACIÓN

Para una secuencia de eventos en la que el primer evento puede ocurrir de n_1 maneras, el segundo evento puede ocurrir de n_2 maneras, el tercer evento puede ocurrir de n_3 maneras, y así sucesivamente; el número total de resultados es $n_1 \cdot n_2 \cdot n_3 \ldots$.

EJEMPLO 1 **Regla de conteo de multiplicación: Hacker que adivina un código de acceso**

Al hacer suposiciones aleatorias para una contraseña desconocida de cuatro dígitos, cada dígito puede ser 0, 1, . . . , 9. ¿Cuál es el número total de diferentes códigos de acceso posibles? Dado que todas las conjeturas tienen la misma probabilidad de ser correcta, ¿cuál es la probabilidad de adivinar la contraseña correcta en el primer intento?

SOLUCIÓN

Hay 10 posibilidades diferentes para cada dígito, por lo que el número total de posibles códigos de acceso es $n_1 \cdot n_2 \cdot n_3 \cdot n_4 = 10 \cdot 10 \cdot 10 \cdot 10 = 10{,}000$.

Debido a que todos los códigos de acceso son igualmente probables, la probabilidad de obtener el código de acceso correcto en el primer intento es $1/10{,}000$ o 0.0001.

SU TURNO Resuelva el ejercicio 5 "Contraseña de cajero automático".

2. Regla factorial

La regla factorial se utiliza para encontrar el número total de formas en que se pueden reordenar *n* elementos diferentes (el orden de los elementos importa). La regla factorial utiliza la siguiente notación.

NOTACIÓN

El **símbolo factorial (!)** expresa al producto de números enteros positivos. Por ejemplo, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$. Por definición especial, $0! = 1$.

REGLA FACTORIAL

El número de *arreglos* diferentes (el orden importa) de *n* elementos diferentes cuando todos ellos se seleccionan es ***n*!**

La regla factorial se basa en el principio de que el primer elemento puede seleccionarse de diferentes maneras, el segundo elemento puede seleccionarse de $n - 1$ maneras, y así sucesivamente. Esta regla es realmente la regla del conteo por multiplicación modificada para la eliminación de un elemento en cada selección.

EJEMPLO 2 Regla factorial: itinerario de viaje

Un investigador en estadística debe visitar personalmente a los presidentes de las empresas de encuestas Gallup, Nielsen, Harris, Pew y Zogby.

a. ¿Cuántos itinerarios de viaje son posibles?

b. Si el itinerario se selecciona al azar, ¿cuál es la probabilidad de que los presidentes sean visitados en orden de los más jóvenes a los más viejos?

SOLUCIÓN

a. Para los 5 diferentes presidentes, el número de distintos itinerarios de viaje es de $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$. Observe que esta solución podría haberse obtenido aplicando la regla del conteo por multiplicación. La primera persona puede ser cualquiera de los 5 presidentes, la segunda persona puede ser cualquiera de los 4 presidentes restantes, y así sucesivamente. El resultado es de nuevo $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$. El uso de la regla factorial tiene la ventaja de incluir el símbolo factorial, que seguramente impresionará.

b. Sólo hay un itinerario con los presidentes visitados por orden de edad, por lo que la probabilidad es 1/120.

SU TURNO Resuelva el ejercicio 9 "Calificación de exámenes".

¿Cuántas veces hay que barajar?

Después de realizar extensas investigaciones, el matemático de Harvard, Persi Diaconis encontró que se necesita barajar siete veces un mazo de naipes para obtener un mezclado completo. La mezcla es completa en el sentido de que todos los arreglos posibles de los naipes son igualmente probables. Barajar más de siete veces no tendrá un efecto significativo, y menos de siete será insuficiente. Los repartidores de naipes en los casinos rara vez barajan los mazos siete veces o más, así que los mazos no quedan totalmente mezclados. Algunos jugadores expertos han podido aprovechar las mezclas incompletas que resultan de barajar menos de siete veces.

Permutaciones y combinaciones: ¿el *orden* cuenta?

Cuando se usan métodos de conteo diferentes, es esencial saber si las distintas disposiciones de los mismos elementos se cuentan una sola vez o se cuentan por separado. Los términos permutaciones y combinaciones son estándar en este contexto, y se definen como sigue:

DEFINICIONES

Las **permutaciones** de elementos son disposiciones en las que se cuentan por *separado* las diferentes secuencias de los mismos elementos. (Las disposiciones de letras abc, acb, bac, bca, cab y cba se cuentan por separado como seis permutaciones diferentes).

Las **combinaciones** de elementos son disposiciones en las que se cuentan las diferentes secuencias de los mismos elementos como si fueran la *misma*. (Las disposiciones de letras abc, acb, bac, bca, cab y cba son considerados como la misma combinación).

Recursos mnemónicos para permutaciones y combinaciones

- Recuerde "**P**ermutaciones **P**osición", donde la aliteración nos recuerda que con las permutaciones, las posiciones de los elementos hacen una diferencia.
- Recuerde "**C**ombinaciones **C**omité", que nos recuerda que con los miembros de un comité, los reordenamientos de los mismos miembros resultan en el mismo comité, por lo que el orden no cuenta.

3. Regla de las permutaciones (cuando todos los elementos son diferentes)

La regla de las permutaciones se utiliza cuando hay diferentes elementos disponibles para su selección, debemos seleccionar r de ellos sin reemplazo, y la secuencia de los elementos importa. El resultado es el número total de disposiciones (o permutaciones) posibles. (Recuerde, los reordenamientos de los mismos elementos se cuentan como permutaciones diferentes).

REGLA DE LAS PERMUTACIONES

Cuando hay n elementos diferentes disponibles y se seleccionan r de ellos sin reemplazo, el número de permutaciones diferentes (conteos ordenados) está dado por

$$_nP_r = \frac{n!}{(n-r)!}$$

EJEMPLO 3 **Regla de las permutaciones (con diferentes elementos): apuesta trifecta**

En una carrera de caballos, se gana una apuesta *trifecta* al seleccionar correctamente a los caballos que terminan primero, segundo y tercero, y se deben seleccionar en el orden correcto. La edición número 140 del derby de Kentucky tenía 19 caballos en la línea de salida.

a. ¿Cuántas apuestas trifectas diferentes son posibles?

b. Si un apostador selecciona al azar a tres de esos caballos para una apuesta trifecta, ¿cuál es la probabilidad de ganar seleccionando California Chrome como ganador, Commanding Curve para terminar segundo y Danza para terminar tercero, como lo que realmente ocurrió? ¿Todas las diferentes posibles apuestas trifecta tienen la misma posibilidad de ganar? (No considere posibles empates).

SOLUCIÓN

a. Hay $n = 19$ caballos disponibles y debemos seleccionar $r = 3$ sin reemplazo. El número de diferentes disposiciones se encuentra de la siguiente manera:

$$_nP_r = \frac{n!}{(n-r)!} = \frac{19!}{(19-3)!} = 5814$$

b. Hay 5814 diferentes posibles arreglos de 3 caballos seleccionados de los 19 que están disponibles. Si uno de estas disposiciones se selecciona al azar, hay una probabilidad de $1/5814$ de seleccionar el arreglo ganador.

Existen 5814 diferentes posibles apuestas trifecta, pero no todas tienen la misma posibilidad de ganar, porque algunos caballos tienden a ser más rápidos que otros. (Una apuesta trifecta ganadora de \$2 en esta carrera ganó \$3424.60).

SU TURNO Resuelva el ejercicio 11 “Programación de rutas”.

4. Regla de las permutaciones (cuando algunos elementos son idénticos a otros)

Cuando se seleccionan los n elementos sin reemplazo, *pero algunos elementos son idénticos*, el número de permutaciones posibles (el orden importa) se encuentra utilizando la siguiente regla.

Códigos de barras

En 1974, se escaneó el primer código de barras en un paquete de goma de mascar de la marca Juicy Fruit que costó 67¢. En la actualidad, cada día se escanean alrededor de 10 mil millones de códigos de barras o “Códigos Universales de Producto”. Cuando se utiliza para números, el código de barras consta de líneas negras que representan una secuencia de 12 dígitos, por lo que el número total de diferentes secuencias de códigos de barras se puede encontrar mediante la aplicación de la regla de conteo fundamental. El número de secuencias de códigos de barras diferentes es $10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 = 10^{12} = 1{,}000{,}000{,}000{,}000$. La efectividad de los códigos de barras depende del gran número de diferentes productos posibles que pueden identificarse con números únicos.

Cuando se escanea un código de barras, el número detectado no es el precio: es un número que identifica el producto en particular. El escáner utiliza ese número de identificación para buscar el precio en una computadora central. Vea el código de barras adjunto que representa el nombre del autor, de modo que se usen letras en lugar de dígitos. No habrá ningún precio que corresponda a este código de barras, porque esta persona no tiene precio, al menos según la mayoría de los miembros de su familia inmediata.

Elección de códigos de seguridad

Todos utilizamos códigos de seguridad personales para tener acceso a cajeros automáticos, cuentas de Internet y sistemas de seguridad para casas. La seguridad de estos códigos depende del gran número de posibilidades, pero ahora los piratas informáticos cuentan con complejas herramientas que pueden superar este obstáculo con creces. Los investigadores encontraron que usando variaciones del nombre y apellidos del usuario, además de otros 1800 nombres, podrían identificar del 10 al 20% de las contraseñas de sistemas de cómputo típicos. Cuando elija una contraseña, no use variaciones de ningún nombre, ni una palabra del diccionario, ni una secuencia con menos de siete caracteres, ni números telefónicos, ni números del sistema de seguridad social. Incluya caracteres no alfabéticos, como números o signos de puntuación.

REGLA DE LAS PERMUTACIONES (CUANDO ALGUNOS ELEMENTOS SON IDÉNTICOS ENTRE SÍ)

El número de diferentes permutaciones (conteos ordenados) cuando n elementos están disponibles y los n elementos se seleccionan *sin reemplazo*, pero algunos de los elementos son idénticos entre sí, se encuentra de la siguiente manera:

$$\frac{n!}{n_1!n_2!\ldots n_k!} \quad \text{donde } n_1, n_2, \ldots, \text{ y } n_k \text{ son iguales.}$$

EJEMPLO 4 Regla de las permutaciones (con algunos elementos idénticos): Diseño de encuestas

Al diseñar encuestas, en ocasiones los encuestadores repiten una pregunta para ver si un sujeto está dando respuestas sin pensar, sólo para terminar rápidamente. Para una encuesta en particular con 10 preguntas, 2 de las preguntas son idénticas entre sí, y otras 3 preguntas también lo son entre sí. Para esta encuesta, ¿cuántos arreglos diferentes son posibles? ¿Es práctico encuestar suficientes temas para que se utilice cada arreglo diferente posible?

SOLUCIÓN

Tenemos 10 preguntas con 2 que son idénticas entre sí y 3 que también lo son entre sí, y queremos conocer el número de permutaciones. Usando la regla para las permutaciones con algunos elementos idénticos entre sí, obtenemos

$$\frac{n!}{n_1!n_2!\ldots n_k!} = \frac{10!}{2!3!} = \frac{3{,}628{,}800}{2 \cdot 6} = 302{,}400$$

INTERPRETACIÓN

Hay 302,400 arreglos diferentes posibles de las 10 preguntas. No es práctico acomodar todas las permutaciones posibles. Para las encuestas típicas, el número de encuestados es de alrededor de 1000.

SU TURNO Resuelva el ejercicio 12 "Confiabilidad de encuesta".

5. Regla de las combinaciones

La regla de las combinaciones se utiliza cuando hay n elementos diferentes disponibles para su selección, sólo r de ellos se seleccionan sin reemplazo, y el orden no importa. El resultado es el número total de combinaciones posibles. (*Recuerde:* los diferentes arreglos de los mismos elementos se consideran la misma combinación).

REGLA DE LAS COMBINACIONES

Cuando n elementos diferentes están disponibles, pero sólo r de ellos se seleccionan *sin reemplazo*, el número de combinaciones diferentes (el orden no importa) se encuentra de la siguiente manera:

$${}_nC_r = \frac{n!}{(n-r)!r!}$$

EJEMPLO 5 Regla de las combinaciones: lotería

En el juego de lotería Fantasy 5 de California, ganar el premio mayor requiere la selección de 5 números diferentes del 1 al 39, y los mismos 5 números deben salir sorteados en la lotería. Los números ganadores se pueden escoger en cualquier disposición, por lo que el orden no marca una diferencia.

a. ¿Cuántos billetes de lotería diferentes son posibles?

b. Encuentre la probabilidad de ganar el premio mayor al comprar un boleto.

SOLUCIÓN

a. Hay $n = 39$ números diferentes disponibles, y debemos seleccionar $r = 5$ sin reemplazo (porque los números seleccionados deben ser diferentes). Debido a que el orden no cuenta, necesitamos encontrar el número de diferentes *combinaciones* posibles. Obtenemos

$$_nC_r = \frac{n!}{(n-r)!r!} = \frac{39!}{(39-5)!5!} = \frac{39!}{34! \cdot 5!} = 575{,}757$$

b. Si usted selecciona una combinación de 5 números, su probabilidad de ganar es $1/575{,}757$. Las loterías típicas se basan en el hecho de que las personas rara vez conocen el valor de esta probabilidad y no tienen un sentido realista de cuán pequeña es esa probabilidad. Esta es la razón por la que la lotería se considera a menudo un "impuesto a las personas que no saben matemáticas".

SU TURNO Resuelva el ejercicio 29 "Mega Millions".

¿Permutaciones o combinaciones? Debido a que elegir entre permutaciones y combinaciones puede ser complicado, proporcionamos el siguiente ejemplo que enfatiza la diferencia entre ellas.

EJEMPLO 6 Permutaciones y combinaciones: funcionarios y comités empresariales

La compañía Google debe nombrar a tres funcionarios corporativos: director general (CEO), presidente ejecutivo y director de operaciones (COO). También debe nombrar un Comité de Planificación con tres miembros diferentes. Hay ocho candidatos calificados, y los funcionarios también pueden participar en el Comité de Planificación.

a. ¿De cuántas maneras diferentes pueden ser nombrados los funcionarios?

b. ¿De cuántas maneras diferentes se puede nombrar al comité?

SOLUCIÓN

Note que en el inciso (a), el orden es importante porque los funcionarios tienen funciones muy diferentes. Sin embargo, en la parte (b), el orden de selección es irrelevante porque todos los miembros del comité cumplen la misma función.

a. Debido a que el orden *cuenta*, se desea conocer el número de *permutaciones* de $r = 3$ personas seleccionadas entre las $n = 8$ personas disponibles. Obtenemos

$$_nP_r = \frac{n!}{(n-r)!} = \frac{8!}{(8-3)!} = 336$$

continúa

Cómo elegir los números de la lotería

Muchos libros y proveedores de programas informáticos afirman ser útiles para predecir números de lotería ganadores. Algunos utilizan la teoría de que los números particulares son "esperados" (y deben seleccionarse) porque no han aparecido a menudo; otros utilizan la teoría de que algunos números son "fríos" (y deben evitarse) porque no han aparecido con frecuencia, y otros utilizan la astrología, la numerología o los sueños. Debido a que las selecciones de combinaciones ganadoras de números de lotería son eventos independientes, tales teorías carecen de valor. Un método válido es elegir números que son "raros" en el sentido de que no son seleccionados por otras personas, de modo que si usted gana, no necesitará compartir su premio con muchas otras personas. La combinación de 1, 2, 3, 4, 5, 6 es una mala elección porque muchas personas tienden a seleccionarlo. En una lotería de la Florida con 105 millones de dólares en premios, 52,000 boletos tenían 1, 2, 3, 4, 5, 6: si esa combinación hubiera ganado, el premio máximo habría sido sólo de $1000. Es sabio elegir combinaciones no seleccionadas por muchos otros. Evite las combinaciones que forman un patrón en la tarjeta de entrada.

Lo que la mediana no es

43,252,003,274,489,856,000: Número de posiciones posibles en un cubo de Rubik.

b. Debido a que el orden *no cuenta*, se desea conocer el número de combinaciones de $r = 3$ personas seleccionadas entre las $n = 8$ personas disponibles. Obtenemos

$$_nC_r = \frac{n!}{(n-r)!r!} = \frac{8!}{(8-3)!3!} = 56$$

Si se toma en cuenta el orden, hay 336 maneras en que los funcionarios pueden ser nombrados, pero sin considerar el orden, hay 56 diversas comisiones posibles.

SU TURNO Resuelva el ejercicio 23 "Funcionarios y comités corporativos".

4-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Notación ¿Qué representa el símbolo !? Seis personas pueden estar en una línea de 6! maneras diferentes, así que ¿cuál es el número real de maneras en que seis personas pueden estar en una línea?

2. New Jersey Pick 6 En el juego de lotería New Jersey Pick 6, un apostador selecciona seis números *diferentes*, cada uno entre 1 y 49. Ganar el premio mayor requiere que los números seleccionados coincidan con los que salen sorteados, pero el orden no importa. ¿Los cálculos para ganar esta lotería implican permutaciones o combinaciones? ¿Por qué?

3. Oregon Pick 4 En el juego de lotería Oregon Pick 4, el apostador selecciona cuatro números entre 0 y 9, y cualquier número seleccionado puede usarse más de una vez. Ganar el primer premio requiere que los números seleccionados coincidan con los que salen sorteados, en el mismo orden. ¿Los cálculos para esta lotería implican la regla de combinaciones o cualquiera de las dos reglas de permutación presentadas en esta sección? ¿Por qué sí o por qué no? Si la respuesta es no, ¿qué regla se aplica?

4. Combinación de cerradura La combinación de cerradura típica utiliza tres números, cada uno entre 0 y 49. Abrir la cerradura requiere la entrada de los tres números en el orden correcto. ¿Es apropiado el nombre "combinación" de cerradura? ¿Por qué sí o por qué no?

En los ejercicios 5 a 36, exprese todas las probabilidades como fracciones.

5. Contraseña de cajero automático Un ladrón roba una tarjeta bancaria y debe, al azar, adivinar la contraseña correcta que consta de cuatro dígitos (cada uno del 0 al 9), los cuales deben introducirse en el orden correcto. Se permite la repetición de dígitos. ¿Cuál es la probabilidad de una conjetura correcta en el primer intento?

6. Números de seguro social Un número de seguro social consta de nueve dígitos en un orden particular, y se permite la repetición de dígitos. Después de ver los últimos cuatro dígitos impresos en un recibo, si selecciona aleatoriamente los demás dígitos, ¿cuál es la probabilidad de obtener el número de seguro social correcto de la persona a quien se le entregó el recibo?

7. Quiniela En una carrera de caballos, se gana la apuesta en una quiniela si se seleccionan los dos caballos que terminan primero y segundo, y se pueden seleccionar en cualquier orden. En la edición 140 del Derby de Kentucky había 19 caballos en la línea de salida. ¿Cuál es la probabilidad de ganar una quiniela si se hacen selecciones aleatorias de los caballos?

8. Penaltis en fútbol En el fútbol, un empate al final del tiempo regular conduce a tiros de penaltis por tres miembros de cada equipo. ¿De cuántas maneras pueden seleccionarse 3 jugadores entre 11? Para 3 jugadores seleccionados, ¿de cuántas maneras pueden ser designados como primero, segundo y tercero?

9. Calificación de exámenes Su profesor acaba de recoger ocho diferentes exámenes de estadística. Si estos exámenes se disponen en orden aleatorio, ¿cuál es la probabilidad de calificar en orden alfabético los estudiantes que realizaron el examen?

10. Letras en la clave de una estación de radio Si las letras en la clave de una estación de radio deben comenzar con K o W y deben incluir dos o tres letras adicionales, ¿cuántas posibilidades diferentes hay?

11. Programación de rutas Una candidata presidencial planea comenzar su campaña visitando las capitales de 5 de los 50 estados. Si las cinco capitales se seleccionan al azar sin reemplazo, ¿cuál es la probabilidad de que la ruta sea Sacramento, Albany, Juneau, Hartford y Bismarck, en ese orden?

12. Confiabilidad de encuesta Una encuesta con 12 preguntas está diseñada para que 3 de las preguntas sean idénticas entre sí y 4 otras también lo sean entre sí (excepto por cambios menores en la redacción). ¿De cuántas maneras diferentes se pueden disponer las 12 preguntas?

13. Seguridad con números El autor posee una caja fuerte en la que guarda todas sus grandes ideas para la próxima edición de este libro. La "combinación" de seguridad consiste en cuatro números entre 0 y 99, y la caja fuerte está diseñada para que los números puedan repetirse. Si otro autor trata de robar estas ideas, ¿cuál es su probabilidad de obtener la combinación correcta en el primer intento? Suponga que los números se seleccionan al azar. Dado el número de posibilidades, ¿parece factible intentar abrir la caja fuerte haciendo conjeturas al azar para adivinar la combinación?

14. Electricidad Al probar la corriente en un cable con cinco alambres codificados por colores, el autor usó un medidor para probar dos cables a la vez. ¿Cuántas pruebas diferentes se requieren para todos los emparejamientos posibles de dos alambres?

15. Sombrero seleccionador En la Escuela Hogwarts de Brujería y Hechicería, el Sombrero seleccionador elige una de cuatro casas para cada estudiante de primer año. Si 4 estudiantes son seleccionados al azar entre 16 estudiantes disponibles (incluyendo a Harry Potter), ¿cuál es la probabilidad de que sean los cuatro estudiantes más jóvenes?

16. Compañía móvil La empresa de mudanzas United Van Lines tiene un camión lleno para entregas a cinco sitios diferentes. Si el orden de las entregas se selecciona al azar, ¿cuál es la probabilidad de que sea la ruta más corta?

17. Powerball Al momento de escribir esto, la lotería Powerball se jugaba en 44 estados. Ganar el primer premio requiere que usted seleccione los cinco números diferentes correctos entre 1 y 69 y, en una selección separada, también debe escoger el número individual correcto entre 1 y 26. Encuentre la probabilidad de ganar el primer premio.

18. Tiro desde la tee Cuando cuatro golfistas están a punto de comenzar un juego, a menudo lanzan una tee para seleccionar al azar el orden en el que saldrán. ¿Cuál es la probabilidad de que empiecen en el orden alfabético de sus apellidos?

19. Código postal Si usted selecciona aleatoriamente cinco dígitos, cada uno entre 0 y 9, con repetición permitida, ¿cuál es la probabilidad de que obtenga el código postal del autor?

20. Entregas de FedEx Con poco tiempo restante en el día, un conductor de FedEx tiene tiempo para realizar entregas en 6 ubicaciones entre los 9 sitios restantes. ¿Cuántas rutas diferentes son posibles?

21. Números de teléfono Las reglas actuales para códigos de área telefónica permiten el uso de los dígitos del 2 al 9 para el primer dígito y del 0 al 9 para el segundo y tercer dígitos. ¿Cuántos códigos de área diferentes son posibles con estas reglas? Esa misma regla se aplica a los números de *intercambio*, que son los tres dígitos inmediatamente anteriores a los últimos cuatro dígitos de un número de teléfono. Dadas las dos reglas, ¿cuántos números de teléfono de 10 dígitos son posibles? Dado que estas reglas se aplican a Estados Unidos, Canadá y algunas islas, ¿hay suficientes números de teléfono posibles? (Suponga que la población combinada es de unos 400,000,000).

22. Problema clásico de conteo Un problema de conteo clásico es determinar el número de maneras diferentes en que las letras de "Mississippi" se pueden disponer. Encuentre ese número.

23. Funcionarios y comités corporativos Recientemente, Digital Pet Rock Company fue financiada con éxito por Kickstarter y ahora debe nombrar a un presidente, a un consejero delegado, un director de operaciones (COO) y un director financiero (CFO). También debe nombrar un comité de planificación estratégica con cuatro miembros. Hay 10 candidatos calificados y los funcionarios también pueden servir en el comité.

a. ¿De cuántas maneras diferentes pueden ser nombrados los cuatro funcionarios?

b. ¿De cuántas maneras se puede nombrar un comité de cuatro?

c. ¿Cuál es la probabilidad de seleccionar al azar a los miembros del comité y obtener los cuatro candidatos calificados más jóvenes?

24. Cajero automático Usted desea retirar efectivo usando un cajero automático, pero está oscuro y no puede ver su tarjeta cuando la inserta. La tarjeta se debe insertar con el lado frontal hacia arriba y la impresión colocada de modo que el inicio de su nombre entre primero.

a. ¿Cuál es la probabilidad de seleccionar una posición aleatoria e introducir la tarjeta de modo que esté insertada correctamente?

b. ¿Cuál es la probabilidad de seleccionar aleatoriamente la posición de la tarjeta y encontrar en el primer intento que la insertó incorrectamente, pero insertarla correctamente en el segundo intento? (Suponga que la misma posición utilizada para el primer intento también podría usarse para el segundo intento).

c. ¿Cuántas selecciones al azar se requieren para estar absolutamente seguro de que la tarjeta funcionará porque se insertó correctamente?

25. Mezcla en la fiesta DJ Marty T es anfitrión de una fiesta esta noche y ha elegido 8 canciones para su ensamble final (incluyendo "Daydream Believer" de los Monkees). ¿Cuántas listas de 8 canciones son posibles (el orden de las canciones importa)? Si las 8 canciones se seleccionan al azar, ¿cuál es la probabilidad de que estén en orden alfabético por el título de la canción?

26. Robo de identidad con tarjetas de crédito Los números de las tarjetas de crédito suelen tener 16 dígitos, pero no todos son aleatorios.

a. ¿Cuál es la probabilidad de generar aleatoriamente 16 dígitos y obtener *su* número de MasterCard?

b. Los recibos suelen mostrar los últimos cuatro dígitos de un número de tarjeta de crédito. Si sólo se conocen los cuatro últimos dígitos, ¿cuál es la probabilidad de generar aleatoriamente los demás dígitos de su número de MasterCard?

c. Las tarjetas Discover comienzan con los dígitos 6011. Si usted sabe que los cuatro primeros dígitos son 6011 y también conoce los últimos cuatro dígitos de una tarjeta Discover, ¿cuál es la probabilidad de generar aleatoriamente los otros dígitos y conseguir que todos ellos sean correctos? ¿Es esto algo preocupante?

27. ¡Qué palabra! Una de las palabras más largas en la terminología de la estadística estándar es "homocedasticidad", ¿De cuántas maneras pueden disponerse las letras en esa palabra?

28. Fase I de un ensayo clínico Una prueba clínica en seres humanos de un nuevo fármaco se hace normalmente en tres fases. La fase I se lleva a cabo con un número relativamente pequeño de voluntarios sanos. Por ejemplo, una prueba de fase I de bexaroteno sólo involucró a 14 sujetos. Supongamos que queremos tratar a 14 seres humanos sanos con este nuevo medicamento y tenemos disponibles 16 voluntarios adecuados.

a. Si los sujetos son seleccionados y tratados uno a uno en secuencia, ¿cuántos arreglos secuenciales diferentes son posibles si se seleccionan 14 personas de las 16 que están disponibles?

b. Si se seleccionan 14 sujetos entre los 16 que están disponibles y los 14 sujetos seleccionados se tratan al mismo tiempo, ¿cuántos grupos de tratamiento diferentes son posibles?

c. Si 14 sujetos son seleccionados al azar y tratados al mismo tiempo, ¿cuál es la probabilidad de seleccionar los 14 sujetos más jóvenes?

29. Lotería Mega Millions Al momento de escribir esto, la lotería Mega Millions se jugaba en 44 estados. Ganar el primer premio requiere que usted seleccione los cinco números diferentes correctos entre 1 y 75, y en una selección separada, también debe elegir el número correcto entre 1 y 15. Encuentre la probabilidad de ganar el primer premio. ¿Cómo se compara el resultado con la probabilidad de ser alcanzado por un rayo en un año, que el Servicio Meteorológico Nacional calcula en 1/960,000?

30. Diseño experimental Los ensayos clínicos de Nasonex involucraron a un grupo que recibió placebos y otro grupo que recibió tratamientos de Nasonex. Suponga que un ensayo preliminar de la fase I se llevará a cabo con 12 sujetos, incluidos 6 hombres y 6 mujeres. Si 6 de los 12 sujetos son seleccionados al azar para el grupo de tratamiento, encuentre la probabilidad de obtener 6 sujetos del mismo sexo. ¿Habría algún problema con que los miembros del grupo de tratamiento fueran todos del mismo sexo?

31. Clave Morse La clave Morse Internacional es una forma de transmitir texto codificado utilizando secuencias de tonos de encendido/apagado. Cada caracter tiene 1 o 2 o 3 o 4 o 5 segmentos de longitud, y cada segmento es un punto o un guión. Por ejemplo, la letra *G* se transmite como dos guiones seguidos por un punto: — — •. ¿Cuántos caracteres diferentes son posibles con este esquema? ¿Hay suficientes caracteres para el alfabeto y los números?

32. Guisantes de Mendel Mendel realizó algunos de sus famosos experimentos con guisantes que eran amarillos lisos o verdes rugosos. Si cuatro guisantes se seleccionan al azar de un lote consistente en cuatro lisas amarillos y cuatro verdes rugosos, encuentre la probabilidad de que los cuatro guisantes seleccionados sean del mismo tipo.

33. Blackjack En el juego de blackjack jugado con un mazo de cartas, un jugador recibe inicialmente 2 cartas diferentes de las 52 distintas cartas en la baraja. Una mano ganadora de "blackjack" se obtiene con 1 de los 4 ases y 1 de 16 cartas con valor de 10 puntos. Las dos cartas pueden estar en cualquier orden. Encuentre la probabilidad de obtener una mano ganadora de blackjack. ¿Qué porcentaje aproximado de manos son manos ganadoras en este juego?

34. Conteo con los dedos ¿De cuántas maneras diferentes pueden tocarse dos o más dedos entre sí en una mano?

35. Cambio por un cuarto de dólar ¿De cuántas maneras diferentes se puede cambiar una moneda de un cuarto de dólar? (Los diferentes arreglos de las mismas monedas no se cuentan por separado).

36. Gane mil millones Quicken Loans ofreció un premio de mil millones de dólares a cualquiera que pudiera predecir correctamente el ganador de cada partido en el torneo de baloncesto de la NCAA. Después de los juegos del torneo "regular", hay 64 equipos en el torneo.

a. ¿Cuántos juegos se requieren para conseguir 1 equipo campeón entre los 64 equipos?

b. Si usted realiza conjeturas aleatorias para cada juego del torneo, encuentre la probabilidad de elegir al ganador en cada partido.

4-4 Más allá de lo básico

37. Nombres de variables computacionales Una regla común de la programación de computadoras es que los nombres de variables deben tener entre uno y ocho caracteres. El primer carácter puede ser cualquiera de las 26 letras, mientras que los caracteres sucesivos pueden ser cualquiera de las 26 letras o cualquiera de los 10 dígitos. Por ejemplo, los nombres de variables permisibles incluyen A, BBB y M3477K. ¿Cuántos nombres de variables diferentes son posibles? (No tome en cuenta la diferencia entre letras mayúsculas y minúsculas).

38. "Chocadas"

a. Cinco "mateatletas" celebran después de resolver un problema particularmente desafiante durante una competición. Si cada mateatleta "choca" su mano con los demás mateatletas exactamente una vez, ¿cuál es el número total de "chocadas"?

b. Si n mateatletas se dan la mano entre sí exactamente una vez, ¿cuál es el número total de apretones de manos?

c. ¿De cuántas maneras diferentes se pueden sentar cinco mateatletas en una mesa redonda? (Suponga que si todos se mueven hacia la derecha, el arreglo de asientos es el mismo).

d. ¿De cuántas maneras diferentes se pueden sentar n mateatletas en una mesa redonda?

4-5 Probabilidades mediante simulación (disponible en inglés en www.pearsonenespañol.com/triola)

El sitio **www.pearsonenespañol.com/triola** incluye una sección descargable (en inglés) que aborda el uso de métodos de simulación para encontrar probabilidades. Las simulaciones también se analizan en el Proyecto de Tecnología cerca del final de este capítulo.

Examen rápido del capítulo

1. Pruebas estándar Las pruebas estándar, como SAT, ACT o MCAT, tienden a hacer un uso extensivo de preguntas de opción múltiple porque son fáciles de calificar usando software. Si una de estas preguntas de opción múltiple tiene las posibles respuestas correctas a, b, c, d, e, ¿cuál es la probabilidad de una respuesta equivocada si la respuesta es una conjetura aleatoria?

2. Lluvia Mientras el autor escribía este ejercicio, un meteorólogo dijo que hay 20% de probabilidad de lluvia mañana. ¿Cuál es la probabilidad de que no llueva mañana?

3. Meses Si un mes se selecciona al azar después de mezclar las páginas de un calendario en inglés, ¿cuál es la probabilidad de que sea un mes que contenga la letra *y*?

4. Redes sociales Con base en datos del Pew Internet Project, 74% de los usuarios adultos de Internet utilizan sitios de redes sociales. Si dos usuarios adultos de Internet son seleccionados al azar, ¿cuál es la probabilidad de que ambos usen sitios de redes sociales?

5. Probabilidad subjetiva Estime la probabilidad de que la próxima vez que viaje en automóvil, a éste se le desinfle un neumático.

En los ejercicios 6 a 10, utilice los siguientes resultados de las pruebas de un experimento para probar la efectividad de una vacuna experimental para niños (basada en datos de* USA Today*). Exprese todas las probabilidades en forma decimal.

	Desarrolló gripe	No desarrolló gripe
Tratamiento con vacuna	14	1056
Placebo	95	437

6. Si 1 de los 1602 sujetos es seleccionado al azar, determine la probabilidad de obtener alguien que desarrolló gripe.

7. Si 1 de los 1602 sujetos es seleccionado al azar, determine la probabilidad de obtener alguien que tuvo el tratamiento con la vacuna o desarrolló gripe.

8. Si 1 de los 1602 sujetos es seleccionado al azar, determine la probabilidad de elegir a alguien que tuvo el tratamiento con la vacuna y desarrolló gripe.

9. Encuentre la probabilidad de seleccionar aleatoriamente 2 sujetos sin reemplazo y descubrir que ambos desarrollaron gripe.

10. Encuentre la probabilidad de seleccionar aleatoriamente a 1 de los sujetos y encontrar que desarrolló gripe, dado que recibió el tratamiento con la vacuna.

Ejercicios de repaso

En los ejercicios 1 a 10, utilice los datos de la tabla adjunta y exprese todos los resultados en forma decimal. (Los datos provienen de "Reducción de la mortalidad con bolsa de aire y uso del cinturón de seguridad en colisiones de automóviles", Crandall, Olson y Sklar,* American Journal of Epidemiology*, vol. 153, núm. 3).

Conductores involucrados en colisiones frontales de automóviles de pasajeros

	El conductor murió	El conductor no murió
Usó cinturón de seguridad	3655	7005
No usó cinturón de seguridad	4402	3040

1. Uso del cinturón de seguridad Si un conductor es seleccionado al azar, encuentre la probabilidad de que estuviera usando el cinturón de seguridad.

2. Uso del cinturón de seguridad Encuentre la probabilidad de seleccionar al azar un conductor y obtener alguien que no murió dado que el conductor estaba usando el cinturón de seguridad.

3. Sin cinturón de seguridad Encuentre la probabilidad de seleccionar al azar un conductor y obtener alguien que murió dado que el conductor no estaba usando el cinturón de seguridad.

4. Uso del cinturón de seguridad o conductor no murió Si uno de los conductores es seleccionado al azar, encuentre la probabilidad de obtener un conductor que utilizó el cinturón de seguridad o murió.

5. Sin cinturón de seguridad o conductor murió Si uno de los conductores es seleccionado al azar, encuentre la probabilidad de obtener a alguien que no usó el cinturón de seguridad o no murió.

6. Ambos con cinturón de seguridad Si 2 conductores son seleccionados al azar *sin reemplazo*, encuentre la probabilidad de que ambos hayan usado cinturones de seguridad.

7. Ambos conductores murieron Si 2 conductores son seleccionados al azar *con reemplazo*, encuentre la probabilidad de que ambos hayan muerto.

8. Complemento Si A representa el evento de seleccionar al azar un conductor incluido en la tabla y obtener a alguien que usó el cinturón de seguridad, ¿qué representa $\overline{A}$? Encuentre el valor de $P(\overline{A})$.

9. Complemento Si A representa el evento de seleccionar al azar un conductor incluido en la tabla y obtener a alguien que no murió, ¿qué representa $\overline{A}$? Determine el valor de $P(\overline{A})$.

10. Los tres exitosos Si se eligen al azar a 3 conductores sin reemplazo, encuentre la probabilidad de que ninguno de ellos haya muerto.

11. Autos negros Utilice la probabilidad subjetiva para estimar la probabilidad de seleccionar al azar un auto y elegir uno que sea negro.

12. Corrección de la visión Alrededor de 75% de la población de Estados Unidos utiliza algún tipo de corrección de la visión (como gafas o lentes de contacto).

a. Si alguien es seleccionado al azar, ¿cuál es la probabilidad de que no utilice alguna corrección de la visión?

b. Si cuatro personas diferentes son seleccionadas al azar, ¿cuál es la probabilidad de que todos usen alguna corrección de la visión?

c. ¿Sería improbable seleccionar al azar a cuatro personas y descubrir que todas ellas usan alguna corrección de la visión? ¿Por qué sí o por qué no?

13. Día nacional de la estadística

a. Si una persona se selecciona al azar, encuentre la probabilidad de que su cumpleaños sea el 18 de octubre, que es el día nacional de la estadística en Japón. No considere los años bisiestos.

b. Si una persona se selecciona al azar, encuentre la probabilidad de que su cumpleaños sea en octubre. No considere los años bisiestos.

c. Estime una probabilidad subjetiva para el evento de seleccionar al azar a un adulto estadounidense y obtener a alguien que sabe que el 18 de octubre es el día nacional de la estadística en Japón.

d. ¿Es improbable seleccionar al azar a un estadounidense adulto y obtener a alguien que sepa que el 18 de octubre es el día nacional de la estadística en Japón?

14. Muestreo compuesto para la diabetes Actualmente, la tasa de nuevos casos de diabetes en un año es de 3.4 por 1000 (según datos de los Centros para el Control y la Prevención de Enfermedades). Al evaluar la presencia de diabetes, el Portland Diagnostics Laboratory ahorra dinero al combinar muestras de sangre para las pruebas. Las pruebas combinadas de la muestra son positivas si al menos una persona tiene diabetes. Si las pruebas combinadas de la muestra son positivas, se realizan los análisis de sangre individuales. En una prueba de diabetes, se combinan muestras de sangre de 10 sujetos seleccionados aleatoriamente. Determine la probabilidad de que la muestra combinada sea positiva con al menos 1 de las 10 personas con diabetes. ¿Es probable que estas muestras combinadas sean positivas?

15. Lotería Wild Card La lotería Wild Card se juega en los estados de Idaho, Montana, Dakota del Norte y Dakota del Sur. Usted debe seleccionar cinco números diferentes entre 1 y 33, después debe seleccionar una de las cartas con figura (As, Rey, Reina, Joto), y luego seleccionar uno de los cuatro palos (bastos, corazones, diamantes, tréboles). Exprese todas las respuestas como fracciones.

a. ¿Cuál es la probabilidad de seleccionar los cinco números correctos entre 1 y 33?

b. ¿Cuál es la probabilidad de seleccionar la carta con figura correcta?

c. ¿Cuál es la probabilidad de seleccionar el palo correcto?

d. ¿Cuál es la probabilidad de seleccionar los cinco números correctos, la carta con figura correcta y el palo correcto?

16. Pensilvania Cash 5 En la lotería Pensilvania Cash 5, ganar el premio máximo requiere que usted seleccione los cinco números diferentes correctos entre 1 y 43 (en cualquier orden). ¿Cuál es la probabilidad de ganar el primer premio? Exprese la respuesta como una fracción.

17. Redundancia Utilizando despertadores de baterías Braun, el autor estima que la probabilidad de falla en cualquier día dado es 1/1000. (a) ¿Cuál es la probabilidad de que el despertador funcione para un evento importante? (b) Si utiliza dos despertadores para los eventos importantes, ¿cuál es la probabilidad de que al menos uno de ellos funcione?

18. Exacta En una carrera de caballos, se gana una apuesta exacta al seleccionar correctamente los caballos que terminan primero y segundo, y deben estar en el orden en que terminaron. La edición 140 del Derby de Kentucky contó con la participación de 19 caballos. Si usted realiza selecciones aleatorias para una apuesta exacta, ¿cuál es su probabilidad de ganar?

Ejercicios de repaso acumulado

1. Conducción con alcohol fatal A continuación se listan las concentraciones de alcohol en la sangre (g/dL) de los conductores condenados por manejar en estado de ebriedad en accidentes automovilísticos (según datos de la Administración Nacional de Seguridad en el Tráfico por Carretera).

0.09 0.11 0.11 0.13 0.14 0.15 0.17 0.17 0.18 0.18 0.23 0.35

Determine el valor de los siguientes estadísticos e incluya las unidades apropiadas.

a. media **b.** mediana **c.** mitad del rango **d.** rango

e. desviación estándar **f.** varianza

2. Conducción con alcohol fatal Utilice los mismos datos dados en el ejercicio 1.

a. Identifique el resumen de 5 números y los valores que parecen ser atípicos.

b. Elabore un diagrama de caja y bigotes.

c. Elabore una gráfica de tallo y hojas.

3. Donantes de órganos *USA Today* proporcionó información sobre una encuesta (realizada para Donate Life America) de 5100 usuarios adultos de Internet. De los encuestados, 2346 dijeron que están dispuestos a donar órganos después de su muerte. En esta encuesta, se incluyó a 100 adultos en cada estado y el Distrito de Columbia, y los resultados se ponderaron para tomar en cuenta los tamaños de población estatales.

a. ¿Qué porcentaje de encuestados dijo que está dispuesto a donar órganos después de su muerte?

b. Con base en los resultados de la encuesta, ¿cuál es la probabilidad de seleccionar al azar a un adulto que esté dispuesto a donar órganos después de su muerte?

c. ¿Qué término se utiliza para describir el método de muestreo consistente en seleccionar aleatoriamente 100 adultos de cada estado y el Distrito de Columbia?

4. Muestreo del color de los ojos Con base en un estudio realizado por el doctor P. Sorita Soni en la Universidad de Indiana, suponemos que los colores de los ojos en Estados Unidos se distribuyen de la siguiente manera: 40% marrón, 35% azul, 12% verde, 7% gris, 6% avellana.

a. Un profesor de estadística recolecta datos del color de ojos de sus estudiantes. ¿Cuál es el nombre de este tipo de muestra?

b. Identifique un factor que podría hacer que la muestra del inciso (a) estuviera sesgada y no fuera representativa de la población general de Estados Unidos.

c. Si una persona se selecciona al azar, ¿cuál es la probabilidad de que tenga los ojos de color marrón o azul?

d. Si dos personas son seleccionadas al azar, ¿cuál es la probabilidad de que al menos una de ellas tenga ojos marrones?

5. Presión arterial y plaquetas A continuación se presentan las medidas de la presión arterial sistólica (mm Hq) y los conteos de plaquetas (1000 células/μL) de los primeros sujetos incluidos en el conjunto de datos 1 "Datos corporales" del apéndice B. Elabore una gráfica adecuada para explorar una asociación entre la presión arterial sistólica y el conteo de plaquetas ¿Qué sugiere la gráfica sobre esa asociación?

Sistólica	100	112	134	126	114	134	118	138	114	124
Plaquetas	319	187	297	170	140	192	191	286	263	193

6. Nuevo juego de lotería En el juego de lotería Monopoly Millionaires' Club, usted paga $5 y selecciona cinco números diferentes entre 1 y 52, y luego un sexto número entre 1 y 28 es asignado al azar. Ganar requiere que sus cinco números coincidan con los sorteados (en cualquier orden), y luego su sexto número asignado también debe coincidir.

a. ¿Cuál es la probabilidad de seleccionar los cinco números correctos entre 1 y 52?

b. ¿Cuál es la probabilidad de obtener el sexto número correcto que se le asigna?

c. ¿Cuál es la probabilidad de seleccionar los cinco números correctos y obtener el sexto número correcto?

Proyecto de tecnología

Simulaciones En ocasiones, el cálculo de probabilidades es extremadamente difícil, pero las simulaciones nos proporcionan una alternativa muy práctica a los cálculos basados en reglas formales. Una simulación de un procedimiento es un proceso que se comporta de la misma manera que el procedimiento para que se produzcan resultados similares. En lugar de calcular la probabilidad de obtener exactamente 5 niños en 10 nacimientos, se puede tirar varias veces 10 monedas y contar el número de veces que ocurren exactamente 5 caras (o "niños" simulados). Mejor aún, se podría hacer la simulación con un generador de números aleatorios en una computadora o calculadora para generar aleatoriamente 1s (o "niños" simulados) y 0s (o "niñas" simuladas). Considere el siguiente ejercicio de probabilidad:

Encuentre la probabilidad de que entre 50 personas seleccionadas al azar, al menos 3 tengan el mismo día de cumpleaños.

Para el problema anterior, una simulación comienza por representar los cumpleaños con números enteros del 1 al 365, donde 1 representa un cumpleaños el 1 de enero, 2 el 2 de enero, y así sucesivamente. Podemos simular 50 cumpleaños usando una calculadora o computadora para generar 50 números aleatorios (con repetición permitida) entre 1 y 365. Después, estos números se pueden ordenar, con lo que se facilita examinar la lista para determinar si 3 de las fechas de nacimiento simuladas son las mismas. Podemos repetir el proceso tantas veces como queramos, hasta que nos demos cuenta de que tenemos una buena estimación de la probabilidad. Utilice la tecnología para simular 20 grupos diferentes de 50 cumpleaños. Utilice los resultados para estimar la probabilidad de que entre 50 personas seleccionadas al azar, al menos 3 tengan el mismo día de cumpleaños.

continúa

Resumen de funciones de simulación:

Statdisk: Seleccione **Data** en el menú superior, seleccione **Uniform Generator** del menú desplegable.

Excel: Haga clic en **Insert function** f_x, seleccione **Math & Trig**, seleccione **RANDBETWEEN**. Copie a las celdas adicionales.

TI-83/84 Plus: Presione **MATH**, seleccione **PROB** en el menú superior, seleccione **randInt** en el menú.

StatCrunch: Seleccione **Data** en el menú superior, seleccione **Simulate** en el menú desplegable y seleccione **Discrete Uniform** en el submenú.

Minitab: Seleccione **Calc** en el menú superior, seleccione **Random Data** en el menú desplegable, seleccione **Integer** en el submenú.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: Interpretación de resultados a partir de una prueba sobre tabaquismo

Se estima que aproximadamente la mitad de los fumadores mienten cuando se les pregunta acerca de su involucramiento con el hábito de fumar. Los CO-oxímetros de pulso pueden ser una forma de obtener información sobre el hábito de fumar sin depender de las declaraciones de los pacientes. Los CO-oxímetros de pulso usan una luz que brilla a través de una uña y mide el monóxido de carbono en la sangre. Estos dispositivos son utilizados por los bomberos y departamentos de emergencia para detectar intoxicación por monóxido de carbono, pero también se puede usar para identificar a los fumadores. La tabla adjunta muestra los resultados de personas de 18 a 44 años de edad a quienes se les aplicó el CO-oxímetro de pulso configurado para detectar un nivel de 6% o más de carboxihemoglobina (según datos de "El monóxido de carbono puede utilizarse para identificar fumadores", de Patrice Wendling, *Internal Medicine News*, vol. 40, núm. 1, y los Centros para el Control y la Prevención de Enfermedades).

Prueba de CO-oximetría para posibles fumadores

	Resultado positivo de la prueba	Resultado negativo de la prueba
Fumador	49	57
No fumador	24	370

Análisis de los resultados

1. Falso positivo Con base en los resultados de la tabla, determine la probabilidad de que un sujeto no fume, dado que el resultado de la prueba es positivo.

2. Verdadero positivo Con base en los resultados de la tabla, encuentre la probabilidad de que un sujeto fume, dado que el resultado de la prueba es positivo.

3. Falso negativo Con base en los resultados de la tabla, determine la probabilidad de que un sujeto fume, dado que el resultado de la prueba es negativo.

4. Verdadero negativo Con base en los resultados de la tabla, determine la probabilidad de que un sujeto no fume, dado que el resultado de la prueba es negativo.

5. Sensibilidad Encuentre la *sensibilidad* de la prueba determinando la probabilidad de un verdadero positivo, dado que el sujeto realmente fuma.

6. Especificidad Encuentre la *especificidad* de la prueba determinando la probabilidad de un verdadero negativo, dado que el sujeto no fuma.

7. Valor predictivo positivo Encuentre el *valor predictivo positivo* de la prueba determinando la probabilidad de que el sujeto fume, dado que la prueba produce un resultado positivo.

8. Valor predictivo negativo Encuentre el *valor predictivo negativo* de la prueba determinando la probabilidad de que el sujeto no fume, dado que la prueba da un resultado negativo.

9. Confusión del inverso Encuentre los siguientes valores y luego compare. En este caso, ¿cuál es la confusión del inverso?

- $P(\text{fumador} \mid \text{resultado positivo de la prueba})$
- $P(\text{resultado positivo de la prueba} \mid \text{fumador})$

Actividades de cooperación en equipo

1. Actividades fuera de clase Divídanse en grupos de tres o cuatro y creen un nuevo juego de carnaval. Determinen la probabilidad de ganar. Determinen cuánto dinero puede esperar ganar el operador del juego cada vez que se juega.

2. Actividad en clase Divídanse en grupos de tres o cuatro y utilicen el lanzamiento de monedas para desarrollar una simulación que emule al reino que cumple con este decreto: Después de que una madre da a luz a un niño, no tendrá más hijos. Si se sigue este decreto, ¿aumenta la proporción de niñas?

3. Actividad en clase Divídanse en grupos de tres o cuatro y utilicen tachuelas reales o dulces Kisses de Hershey's o conos de papel para estimar la probabilidad de que al caer, aterrizarán con la punta hacia arriba. ¿Cuántos ensayos son necesarios para obtener un resultado que parezca ser razonablemente exacto cuando se redondea al primer decimal?

4. Actividad fuera de clase Los biólogos marinos a menudo usan el *método de captura-recaptura* como una forma de estimar el tamaño de una población, como el número de peces en un lago. Este método implica capturar una muestra de la población, etiquetar a cada miembro de la misma y luego devolverla a la población. Una segunda muestra se captura más tarde, y los miembros marcados se cuentan junto con el tamaño total de esta segunda muestra. Los resultados se pueden utilizar para estimar el tamaño de la población.

En lugar de capturar peces reales, simule el procedimiento utilizando una colección uniforme de elementos tales como cuentas de colores, M&Ms o tarjetas indexadas. Comience con una colección grande de por lo menos 200 artículos. Recoja una muestra de 50 y utilice un marcador para "etiquetar" cada uno. Reemplace los elementos marcados, mezcle toda la población, luego seleccione una segunda muestra y proceda a estimar el tamaño de la población. Compare el resultado con el tamaño real de la población obtenido contando todos los artículos.

5. Actividades fuera de clase Divídanse en grupos de tres o cuatro. Primero, utilicen estimaciones subjetivas para la probabilidad de seleccionar al azar un auto y obtener cada uno de los siguientes colores de auto: negro, blanco, azul, rojo, plata, otros. A continuación, diseñen un plan de muestreo para obtener los colores de auto a través de la observación. Ejecuten el plan de muestreo y obtengan las probabilidades modificadas con base en los resultados observados. Escriban un informe breve de los resultados.

6. Actividad en clase El proceso de fabricación de un nuevo circuito integrado de computadora tiene un rendimiento de 1/6, lo que significa que 1/6 de los circuitos son buenos y los restantes 5/6 están defectuosos. Utilice un dado para simular este proceso de fabricación y considere un resultado de 1 como un buen circuito integrado, mientras que los resultados de 2, 3, 4, 5 o 6 representan circuitos integrados defectuosos. Encuentre el número medio de circuitos que deben fabricarse para obtener uno que sea bueno.

7. Actividad en clase El *problema de Monty Hall* se basa en el antiguo programa de televisión *Let's Make a Deal*, presentado por Monty Hall. Suponga que usted es un competidor que ha seleccionado una de tres puertas después de que se le dijo que dos de ellas no ocultan nada, pero que detrás de una de las tres puertas hay un Corvette rojo nuevo. A continuación, el anfitrión abre una de las puertas que no seleccionó y muestra que no hay nada detrás de ella. Después, le ofrece la opción de quedarse con su primera elección o cambiar a la otra puerta sin abrir. ¿Debería usted seguir con su primera opción o debería cambiar? Divídanse en grupos de dos y simulen este juego para determinar si deben quedarse con la elección inicial o cambiar. (Según la revista *Chance*, las escuelas de negocios de instituciones como Harvard y Stanford usaron este problema para ayudar a los estudiantes a enfrentar la toma de decisiones).

8. Actividades fuera de clase En el ejercicio de repaso acumulado 4, se observó que los colores de ojos en Estados Unidos se distribuyen de la siguiente manera: 40% marrón, 35% azul, 12% verde, 7% gris, 6% avellana. Esa distribución puede constituir la base de las probabilidades. Realice una encuesta pidiendo a sus compañeros que identifiquen el color de sus ojos. ¿La probabilidad de 0.4 para los ojos marrones parece ser consistente con sus resultados? ¿Por qué se requiere una muestra grande para confirmar que $P(\text{ojos avellana}) = 0.06$?

5

DISTRIBUCIONES DE PROBABILIDAD DISCRETA

PROBLEMA DEL CAPÍTULO

En los juegos de fútbol americano que se van a tiempo extra, el equipo que gana el lanzamiento de la moneda al inicio del tiempo extra, ¿tiene alguna ventaja?

Antes de 2012, los partidos de fútbol americano de la NFL que resultaban empatados al final del tiempo regular continuaban después de un lanzamiento de moneda. El equipo que ganaba dicho lanzamiento podía elegir si recibir la pelota o patearla. Un enorme 98.1% de los equipos que ganaron el lanzamiento eligieron recibir el balón. Entre 1974 y 2011, hubo 477 juegos con tiempo extra y 17 de ellos terminaron empatados después de jugar el tiempo extra. Aquí, consideramos los 460 juegos restantes que no terminaron en empate. De los 460 juegos que se decidieron en tiempo extra, 252 fueron ganados por el equipo que ganó el lanzamiento de moneda; los equipos que perdieron el lanzamiento de la moneda ganaron 208 juegos. Vea en la figura 5-1 una gráfica de barras que ilustra estos resultados.

Un análisis de la figura 5-1 podría sugerir que no hay mucha diferencia en victorias entre los ganadores y los perdedores del lanzamiento de moneda, pero podría no ser la mejor herramienta para este análisis. El presente capítulo proporcionará métodos relevantes y críticos para abordar este asunto. A partir de 2012 las reglas de los tiempos extra en la NFL cambiaron, pero ¿por qué ocurrió esto? La pregunta clave es: *¿Bajo las viejas reglas para el tiempo extra, ganar el lanzamiento de la moneda se convierte en una ventaja?* El resultado de 252 victorias en 460 juegos es una proporción ganadora del 54.8% para los equipos que ganaron el lanzamiento. ¿Es eso casi lo mismo que una oportunidad al azar, o es 54.8% significativamente mayor al 50%, de manera que los equipos que ganan el lanzamiento de la moneda tienen una ventaja? Tales preguntas pueden contestarse encontrando un valor de *probabilidad* relevante, como se muestra en este capítulo.

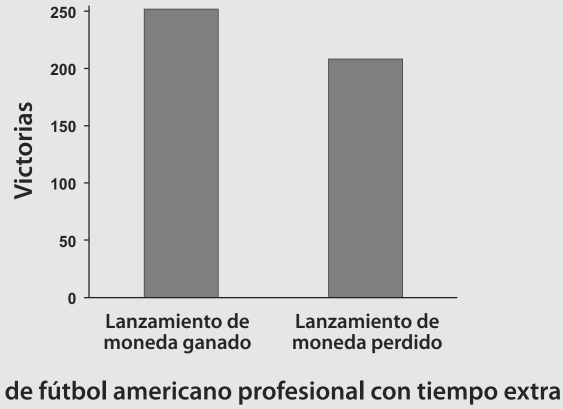


FIGURA 5-1 Juegos de fútbol americano de la NFL ganados y perdidos en tiempo extra

OBJETIVOS DEL CAPÍTULO

La figura 5-2 proporciona una ilustración visual de lo que se logra con este capítulo. Al investigar el número de caras en dos lanzamientos de moneda, podemos utilizar los siguientes dos métodos:

- **Utilizar datos muestrales reales para encontrar resultados *reales*** Registrar las cantidades de caras en lanzamientos de dos monedas, resumirlas en una distribución de frecuencias y encontrar la media $\bar{x}$ y la desviación estándar s (como en los capítulos 2 y 3).
- **Utilizar probabilidades para encontrar resultados *esperados*** Encontrar la probabilidad de cada cantidad posible de caras en dos lanzamientos (como en el capítulo 4), luego resumir los resultados en una tabla que representa una distribución de probabilidad y después encontrar la media μ y la desviación estándar σ.

En este capítulo combinamos los dos métodos anteriores al elaborar una tabla que describe lo que esperamos que suceda (en lugar de lo que sucedió), después encontramos la media μ y la desviación estándar σ de la población. La tabla de la extrema derecha de la figura 5-2 es una *distribución de probabilidad*, porque describe la distribución usando *probabilidades* en vez de conteos de frecuencia. El resto del presente libro y el núcleo de la estadística inferencial se basan en cierto conocimiento de las distribuciones de probabilidad. En este capítulo nos centramos en las distribuciones de probabilidad *discreta*.

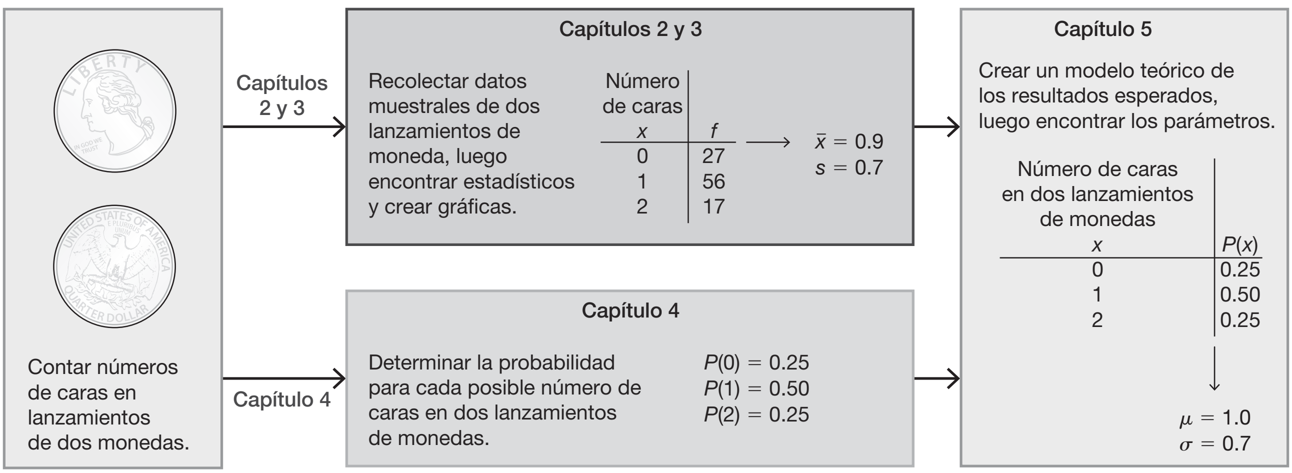


FIGURA 5-2

Los objetivos del capítulo son:

5-1 Distribuciones de probabilidad

- Definir *variable aleatoria* y *distribución de probabilidad*.
- Determinar cuándo una distribución de probabilidad potencial satisface realmente los requisitos necesarios.
- Dada una distribución de probabilidad, calcular la media y la desviación estándar, luego utilizar esos resultados para determinar si son *significativamente bajos* o *significativamente altos*.

5-2 Distribuciones de probabilidad binomial

- Describir una distribución de probabilidad binomial y encontrar valores de probabilidad para una distribución binomial.
- Calcular la media y la desviación estándar para una distribución binomial, luego usar esos resultados para determinar si son *significativamente bajos* o *significativamente altos*.

5-3 Distribuciones de probabilidad de Poisson

- Describir una distribución de probabilidad de Poisson y encontrar valores de probabilidad para una distribución de ese tipo.

5-1 Distribuciones de probabilidad

Concepto clave En esta sección se presentan los conceptos de *variable aleatoria* y *distribución de probabilidad*. Ilustramos cómo un histograma de probabilidad es una gráfica que representa visualmente una distribución de probabilidad. Se muestra cómo determinar los importantes parámetros de la media, la desviación estándar y la varianza para una distribución de probabilidad. Y algo más importante: describimos cómo se determina si los resultados son *significativos* (significativamente bajos o significativamente altos). Iniciamos con los conceptos relacionados de *variable aleatoria* y *distribución de probabilidad*.

PARTE 1 Conceptos básicos de una distribución de probabilidad

DEFINICIONES

Una **variable aleatoria** es una variable (generalmente representada por x) que tiene un único valor numérico, determinado aleatoriamente, para cada resultado de un procedimiento.

Una **distribución de probabilidad** es una descripción que da la probabilidad para cada valor de la variable aleatoria. A menudo se expresa en el formato de una tabla, fórmula o gráfica.

En la sección 1-2 hicimos una distinción entre datos discretos y continuos. Las variables aleatorias también pueden ser discretas o continuas, y las siguientes dos definiciones son consistentes con las que se dieron en la sección 1-2.

DEFINICIONES

Una **variable aleatoria discreta** tiene una colección de valores que es finita o contable. (Si hay una cantidad infinita de valores, tal cantidad es contable si es posible contarlos individualmente, como el número de lanzamientos de una moneda antes de obtener una cara).

Una **variable aleatoria continua** tiene una cantidad infinita de valores, y la colección de valores no es contable. (Es decir, es imposible contar los elementos individuales porque al menos algunos de ellos están en una escala continua, como la temperatura corporal).

Este capítulo trata exclusivamente de variables aleatorias discretas, pero en los siguientes capítulos se estudiarán variables aleatorias continuas.

Distribución de probabilidad: requisitos

Cada distribución de probabilidad debe satisfacer los siguientes tres requisitos.

1. Hay una variable aleatoria *numérica* (no categórica) x, y sus valores numéricos están asociados con probabilidades correspondientes.
2. $\Sigma P(x) = 1$ donde x asume todos los valores posibles. (La suma de todas las probabilidades debe ser 1, pero las sumas como 0.999 o 1.001 son aceptables porque resultan de errores de redondeo).
3. $0 \leq P(x) \leq 1$ para cada valor individual de la variable aleatoria x. (Es decir, cada valor de probabilidad debe estar entre 0 y 1 inclusive).

El segundo requisito proviene del simple hecho de que la variable aleatoria x representa todos los eventos posibles en el espacio muestral completo, por lo que estamos seguros (con probabilidad 1) de que uno de los eventos ocurrirá. El tercer requisito proviene del principio básico de que cualquier valor de probabilidad debe ser 0 o 1, o un valor entre 0 y 1.

EJEMPLO 1 Lanzamiento de monedas

Considere el lanzamiento de dos monedas, con la siguiente variable aleatoria:

$$x = \text{número de caras cuando se lanzan dos monedas}$$

La x anterior es una variable aleatoria porque sus valores numéricos dependen del azar. Al lanzar dos monedas, el número de caras puede ser 0, 1 o 2, y la tabla 5-1 es una distribución de probabilidad porque proporciona la probabilidad para cada valor de la variable aleatoria x y satisface los tres requisitos listados anteriormente:

1. La variable x es una variable aleatoria *numérica*, y sus valores están asociados con probabilidades, como en la tabla 5-1.
2. $\Sigma P(x) = 0.25 + 0.50 + 0.25 = 1$
3. Cada valor de $P(x)$ está entre 0 y 1. (Específicamente 0.25, 0.50 y 0.25 están cada uno entre 0 y 1 inclusive).

La variable aleatoria x de la tabla 5-1 es una variable aleatoria *discreta*, porque tiene tres valores posibles (0, 1, 2), y tres es un número finito, por lo que satisface el requisito de ser finito o contable.

SU TURNO Resuelva el ejercicio 7 "Trastorno genético".

TABLA 5-1 Distribución de probabilidad para el número de caras en dos lanzamientos de monedas

x: Número de caras cuando se lanzan dos monedas	$P(x)$
0	0.25
1	0.50
2	0.25

Notación para 0+

En tablas como la tabla 5-1 o las probabilidades binomiales listadas en la tabla A-1 del apéndice A, a veces usamos 0+ para representar un valor de probabilidad que es positivo pero muy pequeño, como 0.000000123. (Al redondear un valor de probabilidad para su inclusión en dicha tabla, redondear a 0 sería engañoso porque sugeriría incorrectamente que el evento es imposible).

Histograma de probabilidad: gráfica de una distribución de probabilidad

Hay varias maneras de graficar una distribución de probabilidad, pero por ahora consideraremos solamente el **histograma de probabilidad**. La figura 5-3 es un histograma de probabilidad correspondiente a la tabla 5-1. Observe que es similar a un histograma de frecuencias relativas (descrito en la sección 2-2), pero la escala vertical muestra las probabilidades en vez de las frecuencias relativas basadas en los resultados reales de la muestra.

En la figura 5-3, se observa que los valores de 0, 1, 2 a lo largo del eje horizontal están situados en los centros de los rectángulos. Esto implica que los rectángulos son cada uno de 1 unidad de ancho, por lo que las áreas de los rectángulos son 0.25, 0.50 y 0.25. Las *áreas* de estos rectángulos son lo mismo que las *probabilidades* en la tabla 5-1. En el capítulo 6 y subsecuentes se verá que esa correspondencia entre áreas y probabilidades resulta muy útil.

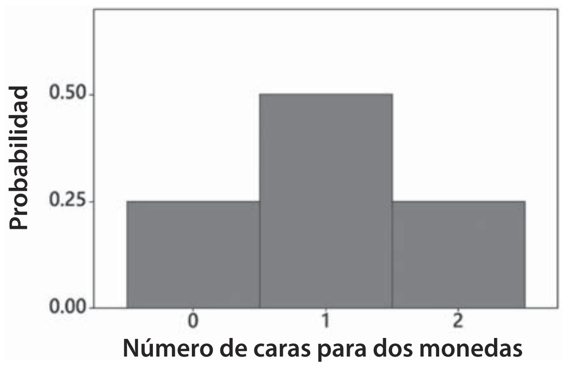


FIGURA 5-3 Histograma de probabilidad para el número de caras cuando se lanzan dos monedas

Fórmula de probabilidad El ejemplo 1 implica una tabla, pero una distribución de probabilidad también podría tener la forma de una fórmula. Considere la fórmula $P(x) = \dfrac{1}{2(2-x)!x!}$ (donde x puede ser 0, 1 o 2). A partir de esa fórmula, encontramos que $P(0) = 0.25$, $P(1) = 0.50$ y $P(2) = 0.25$. Las probabilidades que se encuentran usando esta fórmula son iguales que las de la tabla 5-1. La fórmula describe una distribución de probabilidad porque se cumplen los tres requisitos, como se muestra en el ejemplo 1.

EJEMPLO 2 Errores en una entrevista de trabajo

Se pidió a los gerentes de recursos humanos que identificaran los errores más grandes que los solicitantes de empleo cometían durante una entrevista, y la tabla 5-2 se basa en sus respuestas (según datos de una encuesta de Adecco). ¿La tabla 5-2 describe una distribución de probabilidad?

TABLA 5-2 Errores en una entrevista de trabajo

x	$P(x)$
Vestimenta inapropiada	0.50
Impuntualidad	0.44
Falta de contacto visual	0.33
Revisar su teléfono o enviar mensajes de texto	0.30
Total	**1.57**

SOLUCIÓN

La tabla 5-2 viola el primer requisito porque x no es una variable aleatoria *numérica*. En cambio, los "valores" de x son datos categóricos, no números. La tabla 5-2 también viola el segundo requisito porque la suma de las probabilidades es 1.57, pero esa suma debe ser 1. Debido a que no se satisfacen los tres requisitos, concluimos que la tabla 5-2 *no* describe una distribución de probabilidad.

SU TURNO Resuelva el ejercicio 9 "Frase de conquista".

¿Hay alguien en casa?

Los encuestadores no pueden ignorar simplemente a quienes no estaban en casa cuando acudieron por primera vez. Una solución implica regresar varias veces hasta localizar a la persona. Alfred Politz y Willard Simmons describen una forma para compensar los resultados faltantes, sin tener que regresar varias veces. Ellos sugieren ponderar los resultados con base en la frecuencia con que las personas no se encuentran en su casa. Por ejemplo, una persona que está en su casa sólo dos de seis días a la semana tendrá una probabilidad de 2/6 o 1/3 de estar allí en la primera visita. Cuando se localiza a esa persona por primera vez, sus resultados se ponderan de manera que se cuenten tres veces, respecto a un individuo que siempre está en su casa. Esta ponderación compensa a los demás individuos similares que permanecen en casa dos de seis días a la semana y que no respondieron cuando se les buscó por primera vez. Esta inteligente solución se presentó inicialmente en 1949.

Parámetros de una distribución de probabilidad

Recuerde que con una distribución de probabilidad, tenemos una descripción de una *población* en vez de una muestra, por lo que los valores de la media, la desviación estándar y la varianza son *parámetros*, no datos estadísticos. La media, la varianza y la desviación estándar de una distribución discreta de probabilidad se pueden encontrar con las siguientes fórmulas:

FÓRMULA 5-1 Media μ para una distribución de probabilidad

$$\mu = \Sigma[x \cdot P(x)]$$

FÓRMULA 5-2 Varianza σ^2 para una distribución de probabilidad

$$\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)] \text{ (Este formato es más fácil de entender)}$$

FÓRMULA 5-3 Varianza σ^2 para una distribución de probabilidad

$$\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2 \text{ (Este formato es más fácil para los cálculos mentales)}$$

FÓRMULA 5-4 Desviación estándar σ para una distribución de probabilidad

$$\sigma = \sqrt{\Sigma[x^2 \cdot P(x)] - \mu^2}$$

Al aplicar las fórmulas 5-1 a 5-4, utilice la siguiente regla para redondear los resultados.

Regla de redondeo para μ, σ y σ^2 de una distribución de probabilidad

Redondee los resultados *con un decimal más* que el número de decimales utilizados para la variable aleatoria x. Si los valores de x son enteros, redondee μ, σ y σ^2 a un decimal.

Excepciones a la regla de redondeo En algunos casos especiales, la regla de redondeo anterior resulta en valores que son engañosos o inapropiados. Por ejemplo, en los aviones a propulsión de cuatro motores, el número medio de motores que funcionan exitosamente durante un vuelo es 3.999714286, que se redondea a 4.0; pero eso es engañoso porque sugiere que todos los motores a reacción funcionan siempre con éxito. Aquí necesitamos más precisión para reflejar correctamente la verdadera media, por ejemplo al especificar el valor de 3.999714.

Meta-análisis

El término *meta-análisis* se refiere a una técnica para la realización de un estudio que en esencia combina los resultados de otros estudios. Tiene la ventaja de que las muestras más pequeñas se pueden combinar en una gran muestra, haciendo que los resultados colectivos sean más significativos. También tiene la ventaja de utilizar el trabajo que ya se ha hecho. El meta-análisis tiene la desventaja de ser tan bueno como los estudios que se utilizan, pero si los estudios previos son defectuosos, puede suceder que al "entrar basura, salga basura". En la actualidad, el uso del meta-análisis es popular en la investigación médica y psicológica. Como ejemplo, un estudio de los tratamientos para migraña se basó en datos de otros 46 estudios. (Vea "Meta-Analysis of Migraine Headache Treatments: Combining Information from Heterogeneous Designs", de Dominici *et al*., *Journal of the American Statistical Association*, vol. 94, núm. 445).

Valor esperado

La media de una variable aleatoria discreta x es el resultado teórico medio para infinidad de ensayos. Podemos pensar en esa media como el *valor esperado* en el sentido de que es el valor promedio que esperaríamos obtener si los ensayos pudieran continuar indefinidamente.

DEFINICIÓN

El **valor esperado** de una variable aleatoria discreta x se expresa con E, y es el valor medio de los resultados, por lo que $E = \mu$ y E también se puede encontrar evaluando $\Sigma[x \cdot P(x)]$, como en la fórmula 5-1.

PRECAUCIÓN Un valor esperado no tiene que ser un número entero, aunque los diferentes valores posibles de x lo sean. El número esperado de niñas en cinco nacimientos es de 2.5, aunque cinco nacimientos en particular nunca pueden resultar en 2.5 niñas. Si decidiéramos encuestar a muchas parejas con cinco hijos, esperaríamos que el número medio de niñas fuera de 2.5.

EJEMPLO 3 Determinación de la media, la varianza y la desviación estándar

La tabla 5-1 de la página 187 describe la distribución de probabilidad para el número de caras cuando se lanzan dos monedas. Encuentre la media, la varianza y la desviación estándar para la distribución de probabilidad descrita en la tabla 5-1 del ejemplo 1.

SOLUCIÓN

En la tabla 5-3, las dos columnas de la izquierda describen la distribución de probabilidades que se dio anteriormente en la tabla 5-1. Creamos las dos columnas a la derecha con el propósito de realizar los cálculos requeridos.

Si usamos las fórmulas 5-1 y 5-2 y los resultados de la tabla, obtenemos

$$\text{Media: } \mu = \Sigma[x \cdot P(x)] = 1.0$$

$$\text{Varianza: } \sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)] = 0.5$$

La desviación estándar es la raíz cuadrada de la varianza, por lo que

$$\text{Desviación estándar: } \sigma = \sqrt{0.5} = 0.707107 = 0.7 \text{ (redondeado)}$$

Redondeo: En la tabla 5-3, usamos $\mu = 1.0$. Si μ hubiera tenido el valor de 1.23456, podríamos redondear μ a 1.2, pero debemos usar su valor *no redondeado* de 1.23456 en los cálculos de la tabla 5-3. El redondeo en medio de los cálculos puede dar lugar a resultados con errores demasiado grandes.

TABLA 5-3 Cálculo de μ y σ para una distribución de probabilidad

x	$P(x)$	$x \cdot P(x)$	$(x - \mu)^2 \cdot P(x)$
0	0.25	$0 \cdot 0.25 = 0.00$	$(0 - 1.0)^2 \cdot 0.25 = 0.25$
1	0.50	$1 \cdot 0.50 = 0.50$	$(1 - 1.0)^2 \cdot 0.50 = 0.00$
2	0.25	$2 \cdot 0.25 = 0.50$	$(2 - 1.0)^2 \cdot 0.25 = 0.25$
Total		1.00 ↑ $\mu = \Sigma[x \cdot P(x)]$	0.50 ↑ $\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)]$

INTERPRETACIÓN

Al lanzar dos monedas, el número medio de caras es 1.0, la varianza es 0.5 caras^2, y la desviación estándar es 0.7 caras. Además, el valor esperado para el número de caras cuando se lanzan dos monedas es 1.0, que es el mismo valor de la media. Si tuviéramos que recopilar datos sobre un gran número de ensayos con dos monedas lanzadas en cada ensayo, esperaríamos obtener una media de 1.0 caras.

SU TURNO Resuelva el ejercicio 15 “Media y desviación estándar”.

Sentido de los resultados: valores significativos

Presentamos los siguientes dos métodos para determinar si un valor de una variable aleatoria *x* es significativamente bajo o alto.

Identificación de resultados significativos con la regla práctica del rango

La regla práctica del rango (presentada en la sección 3-2) puede ser útil para interpretar el valor de una desviación estándar. De acuerdo con la regla práctica del rango, la gran mayoría de los valores debe estar dentro de 2 desviaciones estándar de la media, por lo que podemos considerar un valor *significativo* si está por lo menos 2 desviaciones estándar alejado de la media. Entonces, podemos identificar los valores “significativos” como sigue:

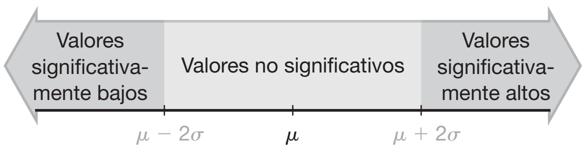
Regla práctica para la identificación de valores significativos

Los valores ***significativamente bajos*** son $(\mu - 2\sigma)$ o inferiores.

Los valores ***significativamente altos*** son $(\mu + 2\sigma)$ o superiores.

Los valores ***no significativos*** están entre $(\mu - 2\sigma)$ y $(\mu + 2\sigma)$.

La figura 3-3 de la sección 3-2 ilustra los criterios anteriores:



PRECAUCIÓN Recuerde que el uso del número 2 en la regla práctica del rango es algo arbitrario y se trata de una directriz, no de una regla absolutamente rígida.

EJEMPLO 4 Identificación de resultados significativos con la regla práctica del rango

En el ejemplo 3 se encontró que cuando se lanzan dos monedas, el número medio de caras es $\mu = 1.0$ y la desviación estándar es $\sigma = 0.7$ caras. Utilice esos resultados y la regla práctica del rango para determinar si 2 es un número significativamente alto de caras.

SOLUCIÓN

De acuerdo con la regla práctica del rango, el resultado de 2 caras es significativamente alto si es mayor o igual que $\mu + 2\sigma$. Con $\mu = 1.0$ caras y $\sigma = 0.7$ caras, se obtiene

$$\mu + 2\sigma = 1 + 2(0.7) = 2.4 \text{ caras}$$

Las cantidades significativamente altas de caras son 2.4 y superiores.

continúa

INTERPRETACIÓN

Con base en estos resultados, concluimos que 2 caras no es un número significativamente alto de caras (porque 2 no es mayor o igual que 2.4).

SU TURNO Resuelva el ejercicio 17 "Regla práctica del rango para eventos significativos".

Identificación de resultados significativos con probabilidades:

- **Número *significativamente alto* de éxitos:** x éxitos entre n ensayos es un número *significativamente alto* de éxitos si la probabilidad de x o más éxitos es 0.05 o menos. Es decir, x es un número significativamente alto de éxitos si $P(x \text{ o más}) \leq 0.05$.*
- **Número *significativamente bajo* de éxitos:** x éxitos entre n ensayos es un número *significativamente bajo* de éxitos si la probabilidad de x o menos éxitos es 0.05 o menos. Es decir, x es un número significativamente bajo de éxitos si $P(x \text{ o menos}) \leq 0.05$.*

*El valor de 0.05 no es absolutamente rígido. Se podrían usar otros, como 0.01, para distinguir entre los resultados que son significativos y los que no lo son.

En ocasiones se utiliza la identificación de números significativamente bajos o significativamente altos de éxitos con el propósito de rechazar supuestos, como se establece en la siguiente regla de eventos raros.

Regla de eventos raros para estadística inferencial

Si, bajo un supuesto dado, la probabilidad de un resultado particular es muy pequeña y el resultado es *significativamente menor que*, o *significativamente mayor que*, lo esperado con ese supuesto, se concluye que el supuesto probablemente es incorrecto.

Por ejemplo, si se prueba que los niños y las niñas son igualmente probables, el resultado de 20 niñas en 100 nacimientos es significativamente bajo y sería una base para rechazar ese supuesto.

EJEMPLO 5 Identificación de resultados significativos con probabilidades

¿Son 252 caras en 460 lanzamientos de monedas un número significativamente alto de caras? ¿Qué sugiere el resultado sobre el problema del capítulo, que incluye los resultados de 460 juegos con tiempos extra? (Entre los 460 equipos que ganaron el lanzamiento de moneda, 252 de ellos ganaron el juego. ¿Son 252 victorias en esos 460 juegos significativamente altas?)

SOLUCIÓN

Un resultado de 252 caras en 460 lanzamientos de monedas es mayor de lo que esperaríamos con una probabilidad aleatoria, pero necesitamos determinar si 252 caras son *significativamente altas*. Aquí, la probabilidad relevante es la probabilidad de obtener 252 *o más* caras en 460 lanzamientos de monedas. Con base en métodos que se estudian más adelante en la sección 5-2, podemos encontrar que P(252 o más caras en 460 lanzamientos de monedas) = 0.0224 (redondeado). Debido a que la probabilidad de obtener 252 o más caras es menor o igual que 0.05, concluimos que 252 caras en 460 lanzamientos de monedas son un número *significativamente alto*. Vea la figura 5-4, que es un histograma con la probabilidad de los diferentes números de caras.

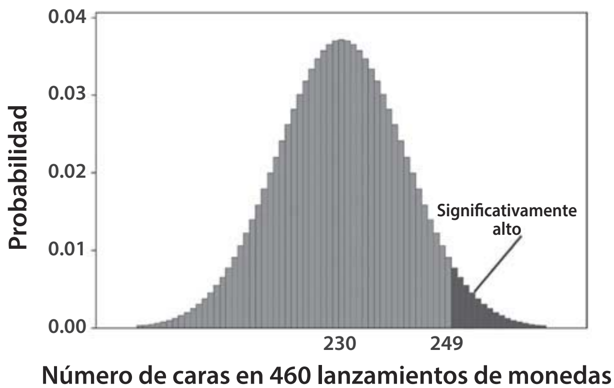


FIGURA 5-4 Histograma de probabilidades de caras en 460 lanzamientos de monedas

INTERPRETACIÓN

Es poco probable que obtengamos 252 o más caras en 460 lanzamientos de monedas por casualidad. Se deduce que 252 victorias de los equipos que ganaron el lanzamiento de moneda para los tiempos extra son significativamente altas, por lo que ganar el lanzamiento de monedas es una ventaja. Esta es una justificación para cambiar las reglas del tiempo extra, como se hizo en 2012.

SU TURNO Resuelva el ejercicio 19 "Uso de probabilidades para eventos significativos".

No *exactamente*, pero "Al menos tan extremo"

Debe ser obvio que entre 1000 lanzamientos de una moneda, 502 caras no son significativamente altas, mientras que 900 caras sí lo son. ¿Qué hace que 900 caras sean significativas pero que 502 caras no lo sean? No son las probabilidades *exactas* de 900 y 502 caras (ambas menores que 0.026). Es el hecho de que la probabilidad de 502 *o más* caras (0.2162) no es baja, pero la probabilidad de 900 *o más* caras (0+) es muy baja.

PARTE 2 Valor esperado y justificación de las fórmulas

Valor esperado

En la parte 1 de esta sección se observó que el valor esperado de una variable aleatoria x es igual a la media μ. Por lo tanto, podemos encontrar el valor esperado calculando $\Sigma[x \cdot P(x)]$, tal como lo hacemos para encontrar el valor de μ. También observamos que el concepto de valor esperado se utiliza en la *teoría de decisiones*. En el ejemplo 6 ilustramos este uso del valor esperado con una situación en la que debemos elegir entre dos apuestas diferentes. El ejemplo 6 implica una decisión real y práctica.

EJEMPLO 6 Cómo ser un mejor apostador

Usted tiene \$5 para colocar en una apuesta en el casino Golden Nugget en Las Vegas. Ha reducido su elección a una de dos apuestas:

Ruleta: Apostar al número 7 de la ruleta.

Dados: Apostar a la "línea de paso" en el juego de dados.

a. Si usted apuesta \$5 al número 7 de la ruleta, la probabilidad de perder \$5 es 37/38 y la probabilidad de hacer una ganancia neta de \$175 es 1/38. (El premio es de \$180, incluyendo su apuesta de \$5, por lo que la ganancia neta es de \$175). Encuentre su valor esperado si apuesta \$5 al número 7 de la ruleta.

continúa

b. Si apuesta \$5 a la línea de paso en el juego de los dados, la probabilidad de perder \$5 es 251/495 y la probabilidad de tener una ganancia neta de \$5 es 244/495. (Si apuesta \$5 a la línea de paso y gana, recibe \$10 que incluye su apuesta, por lo que la ganancia neta es de \$5). Encuentre el valor esperado si usted apuesta \$5 a la línea de paso.

¿Cuál de las dos apuestas es mejor en el sentido de producir un mayor valor esperado?

SOLUCIÓN

a. Ruleta Las probabilidades y ganancias de las apuestas de \$5 al número 7 de la ruleta se resumen en la tabla 5-4. Esa tabla también muestra que el valor esperado es $\Sigma[x \cdot P(x)] = -26¢$. Es decir, para cada apuesta de \$5 al número 7, se puede esperar *perder* un promedio de 26 centavos.

TABLA 5-4 Ruleta

Evento	x	$P(x)$	$x \cdot P(x)$
Pérdida	−\$5	37/38	−\$4.868421
Ganancia (neta)	\$175	1/38	\$4.605263
Total			−\$0.26 (redondeado) (o −26¢)

b. Juego de dados Las probabilidades y ganancias por apostar \$5 a la línea de paso en los dados se resumen en la tabla 5-5. Esta tabla también muestra que el valor esperado es $\Sigma[x \cdot P(x)] = -7¢$. Es decir, por cada apuesta de \$5 en la línea de paso, usted puede esperar perder un promedio de 7¢.

TABLA 5-5 Juego de dados

Evento	x	$P(x)$	$x \cdot P(x)$
Pérdida	−\$5	251/495	−\$2.535353
Ganancia (neta)	\$5	244/495	\$2.464646
Total			−\$0.07 (redondeado) (o −7¢)

INTERPRETACIÓN

La apuesta de \$5 en la ruleta resulta en un valor esperado de −26¢ y la apuesta de \$5 en los dados resulta en un valor esperado de −7¢. Puesto que es mejor perder 7¢ que perder 26¢, el juego de dados es mejor a largo plazo, a pesar de que el juego de la ruleta ofrece la oportunidad de una mayor ganancia cuando se juega una sola vez.

SU TURNO Resuelva el ejercicio 27 “Valor esperado en el juego Pick 3 de Virginia”.

Justificación de las fórmulas 5-1 a 5-4

En lugar de aceptar y utilizar ciegamente las fórmulas, es mucho mejor tener cierta comprensión de por qué funcionan. Cuando se calcula la media a partir de una distribución de frecuencias, f representa la frecuencia de clase y N representa el tamaño de la población. En la expresión que sigue, reescribimos la fórmula para la media de una tabla de frecuencias de modo que sea aplicable a una población. En la fracción f/N, el valor de f es la frecuencia con la que se produce el valor x y N es el tamaño de la población, por lo que f/N es la probabilidad para el valor de x. Cuando reemplazamos f/N por $P(x)$, hacemos la transición de la frecuencia relativa basada en un número limitado de observaciones a la probabilidad basada en infinitos ensayos. Este resultado muestra por qué la fórmula 5-1 es como se ha dado anteriormente en esta sección.

$$\mu = \frac{\Sigma(f \cdot x)}{N} = \sum\left[\frac{f \cdot x}{N}\right] = \sum\left[x \cdot \frac{f}{N}\right] = \Sigma[x \cdot P(x)]$$

Un razonamiento similar permite tomar la fórmula de la varianza del capítulo 3 y aplicarla a una variable aleatoria para una distribución de probabilidad; el resultado es la fórmula 5-2. La fórmula 5-3 es una versión rápida que siempre producirá el mismo resultado que la fórmula 5-2. Aunque la fórmula 5-3 suele ser más fácil de trabajar, la fórmula 5-2 es más fácil de entender directamente. Con base en la fórmula 5-2, podemos expresar la desviación estándar como

$$\sigma = \sqrt{\Sigma[(x - \mu)^2 \cdot P(x)]}$$

o como la forma equivalente dada en la fórmula 5-4.

5.1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Variable aleatoria La tabla adjunta lista las probabilidades para el número correspondiente de niñas en cuatro nacimientos. ¿Cuál es la variable aleatoria, cuáles son sus posibles valores y cuáles sus valores numéricos?

2. ¿Discreta o continua? ¿La variable aleatoria dada en la tabla adjunta es discreta o continua? Explique.

3. Distribución de probabilidad Para la tabla adjunta, ¿es la suma de los valores de $P(x)$ igual a 1, como se requiere para una distribución de probabilidad? ¿La tabla describe una distribución de probabilidad?

4. Significativo Para 100 nacimientos, P(exactamente 56 niñas) = 0.0390 y P(56 niñas o más) = 0.136. ¿Es 56 niñas en 100 nacimientos un número significativamente alto de niñas? ¿Qué probabilidad es relevante para responder esa pregunta?

Número de niñas en cuatro nacimientos

Número de niñas x	$P(x)$
0	0.063
1	0.250
2	0.375
3	0.250
4	0.063

Identificación de variables aleatorias discretas y continuas. ***En los ejercicios 5 y 6, consulte los valores dados, luego identifique cuál de las siguientes expresiones es la más apropiada:* variable aleatoria discreta, variable aleatoria continua, *o* una variable no aleatoria.**

5. a. Pesos exactos de los próximos 100 bebés nacidos en Estados Unidos

b. Respuestas a la pregunta de encuesta "¿Qué partido político prefiere?"

c. Cantidad de vueltas a la ruleta necesarias para obtener el número 7

d. Longitudes exactas de los pies de los seres humanos

e. Tamaños de zapato (como 8 u 8½) para los seres humanos

6. a. Calificaciones (A, B, C, D, F) obtenidas en las clases de estadística

b. Estatura de estudiantes en las clases de estadística

c. Cantidad de estudiantes en las clases de estadística

d. Colores de los ojos de los estudiantes de estadística

e. Número de veces que los estudiantes de estadística deben lanzar una moneda antes de obtener caras

Identificación de distribuciones de probabilidad. ***En los ejercicios 7 a 14, determine si se da una distribución de probabilidad. Si se tiene una distribución de probabilidad, encuentre su media y desviación estándar. Si no se tiene una distribución de probabilidad, identifique los requisitos que no se satisfacen.***

7. Trastorno genético Cinco varones con un trastorno genético ligado al cromosoma X tienen un hijo cada uno. La variable aleatoria x es el número de niños entre los cinco que heredan el trastorno genético ligado al cromosoma X.

x	$P(x)$
0	0.031
1	0.156
2	0.313
3	0.313
4	0.156
5	0.031

8. Varones con daltonismo Al realizar una investigación sobre el daltonismo en varones, un investigador forma grupos al azar con cinco varones en cada grupo. La variable aleatoria x es el número de varones en el grupo que tienen una forma de daltonismo (según datos de los Institutos Nacionales de Salud).

x	$P(x)$
0	0.659
1	0.287
2	0.050
3	0.004
4	0.001
5	0+

9. Frase de conquista Ted no es particularmente creativo. Utiliza la frase de conquista "¿Acaba de salir el sol o me has sonreído?" La variable aleatoria x es el número de chicas a las que Ted se acerca antes de encontrar a alguien que reaccione positivamente.

x	$P(x)$
1	0.001
2	0.009
3	0.030
4	0.060

10. Maneras divertidas de coquetear En una encuesta de la mensajería instantánea de Microsoft, se pidió a los encuestados que eligieran la forma más divertida de coquetear y la tabla adjunta se basa en los resultados.

	$P(x)$
Correo electrónico	0.06
En persona	0.55
Mensaje instantáneo	0.24
Mensaje de texto	0.15

11. Maneras divertidas de coquetear Un sociólogo selecciona aleatoriamente a adultos solteros en distintos grupos de tres, y la variable aleatoria x es la cantidad de personas en el grupo que dice que la manera más divertida de coquetear es en persona (según una encuesta de la mensajería instantánea de Microsoft).

x	$P(x)$
0	0.091
1	0.334
2	0.408
3	0.166

12. Vehículo autoconducido Se seleccionan grupos de adultos al azar y se organizan de tres en tres. La variable aleatoria x es la cantidad de personas en el grupo que dice sentirse cómodo en un vehículo autoconducido (según una encuesta de TE Connectivity).

x	$P(x)$
0	0.358
1	0.439
2	0.179
3	0.024

13. Uso del teléfono celular En una encuesta, a los usuarios de teléfonos celulares se les preguntó qué oído utilizan para escuchar su teléfono y la tabla se basa en sus respuestas (según datos de "Hemispheric Dominance and Cell Phone Use", Seidman *et al*., *JAMA Otolaryngology-Head & Neck Surgery*, vol. 139, núm. 5).

	$P(x)$
Izquierdo	0.636
Derecho	0.304
Sin preferencia	0.060

14. Juegos de casino Cuando se apuesta a la línea de paso en el juego de dados del casino Mohegan Sun en Connecticut, la tabla lista las probabilidades del número de apuestas que se deben hacer para ganar una vez.

x	$P(x)$
1	0.493
2	0.250
3	0.127
4	0.064

Genética. ***En los ejercicios 15 a 20, consulte la tabla adjunta, que describe los resultados de grupos de 8 nacimientos de 8 grupos de padres. La variable aleatoria x representa el número de niñas entre 8 nacimientos.***

Número de niñas x	$P(x)$
0	0.004
1	0.031
2	0.109
3	0.219
4	0.273
5	0.219
6	0.109
7	0.031
8	0.004

15. Media y desviación estándar Encuentre la media y la desviación estándar para el número de niñas en 8 nacimientos.

16. Regla práctica del rango para eventos significativos Utilice la regla práctica del rango para determinar si una niña en 8 nacimientos es un número significativamente bajo de niñas.

17. Regla práctica del rango para eventos significativos Utilice la regla práctica del rango para determinar si 6 niñas en 8 nacimientos es un número significativamente alto de niñas.

18. Uso de probabilidades para eventos significativos

a. Encuentre la probabilidad de obtener exactamente 7 niñas en 8 nacimientos.

b. Encuentre la probabilidad de obtener 7 o más niñas en 8 nacimientos.

c. ¿Qué probabilidad es relevante para determinar si 7 es un número significativamente alto de niñas en 10 nacimientos: el resultado del inciso (a) o del inciso (b)?

d. ¿Es 7 un número significativamente alto de niñas en 8 nacimientos? ¿Por qué sí o por qué no?

19. Uso de probabilidades para eventos significativos

a. Encuentre la probabilidad de tener exactamente 6 niñas en 8 nacimientos.

b. Encuentre la probabilidad de tener 6 o más niñas en 8 nacimientos.

c. ¿Qué probabilidad es relevante para determinar si 6 es un número significativamente alto de niñas en 8 nacimientos: el resultado del inciso (a) o del inciso (b)?

d. ¿Es 6 un número significativamente alto de niñas en 8 nacimientos? ¿Por qué sí o por qué no?

20. Uso de probabilidades para eventos significativos

a. Encuentre la probabilidad de tener exactamente 1 niña en 8 nacimientos.

b. Encuentre la probabilidad de tener 1 o menos niñas en 8 nacimientos.

c. ¿Qué probabilidad es relevante para determinar si 1 es un número significativamente bajo de niñas en 8 nacimientos: el resultado del inciso (a) o del inciso (b)?

d. ¿Es 1 un número significativamente bajo de niñas en 8 nacimientos? ¿Por qué sí o por qué no?

Sonambulismo. ***En los ejercicios 21 a 25, consulte la tabla adjunta, que describe el número de adultos en grupos de cinco que reportaron sonambulismo (según datos de "Prevalence and Comorbidity of Nocturnal Wandering in the U.S. Adult General Population", de Ohayon et al., Neurology, vol. 78, núm. 20).***

x	$P(x)$
0	0.172
1	0.363
2	0.306
3	0.129
4	0.027
5	0.002

21. Media y desviación estándar Encuentre la media y la desviación estándar para el número de sonámbulos en grupos de cinco.

22. Regla práctica del rango para eventos significativos Utilice la regla práctica del rango para determinar si 4 es un número significativamente alto de sonámbulos en un grupo de 5 adultos.

23. Regla práctica del rango para eventos significativos Utilice la regla práctica del rango para determinar si 3 es un número significativamente alto de sonámbulos en un grupo de 5 adultos.

24. Uso de probabilidades para identificar eventos significativos

a. Encuentre la probabilidad de tener exactamente 4 sonámbulos entre 5 adultos.

b. Encuentre la probabilidad de tener 4 o más sonámbulos entre 5 adultos.

c. ¿Qué probabilidad es relevante para determinar si 4 es un número significativamente alto de sonámbulos entre 5 adultos: el resultado del inciso (a) o del inciso (b)?

d. ¿Es 4 un número significativamente alto de sonámbulos entre 5 adultos? ¿Por qué sí o por qué no?

25. Uso de probabilidades para identificar eventos significativos

a. Encuentre la probabilidad de tener exactamente 1 sonámbulo entre 5 adultos.

b. Encuentre la probabilidad de tener 1 o menos sonámbulos entre 5 adultos.

c. ¿Qué probabilidad es relevante para determinar si 1 es un número significativamente bajo de sonámbulos entre 5 adultos: el resultado del inciso (a) o del inciso (b)?

d. ¿Es 1 un número significativamente bajo de sonámbulos entre 5 adultos? ¿Por qué sí o por qué no?

5-1 Más allá de lo básico

26. Valor esperado para la Lotería Pick 4 de Ohio En la Lotería Pick 4 de Ohio, usted puede apostar $1 seleccionando cuatro dígitos, cada uno entre 0 y 9 inclusive. Si los mismos cuatro números salen sorteados en el mismo orden, usted gana y cobra $5000.

a. ¿Cuántas selecciones diferentes son posibles?

b. ¿Cuál es la probabilidad de ganar?

c. Si usted gana, ¿cuál es su ganancia neta?

d. Encuentre el valor esperado por una apuesta de $1.

e. Si apuesta $1 a la línea de paso en el juego de dados del casino, el valor esperado es −1.4¢. ¿Qué apuesta es mejor en el sentido de producir un mayor valor esperado: una apuesta de $1 en la Lotería Pick 4 de Ohio o una apuesta de $1 a la línea de paso en los dados?

27. Valor esperado en el juego Pick 3 de Virginia En el juego de lotería Pick 3 de Virginia, usted puede pagar $1 para seleccionar un número de tres dígitos de 000 a 999. Si selecciona la misma secuencia de tres dígitos que sale sorteada, usted gana y recibe $500.

a. ¿Cuántas selecciones diferentes son posibles?

b. ¿Cuál es la probabilidad de ganar?

c. Si usted gana, ¿cuál es su ganancia neta?

d. Encuentre el valor esperado.

e. Si usted apuesta \$1 en el juego Pick 4 de Virginia, el valor esperado es de $-50¢$. ¿Qué apuesta es mejor en el sentido de producir un mayor valor esperado: una apuesta de \$1 en el juego Pick 3 de Virginia o una apuesta de \$1 en el juego Pick 4 de Virginia?

28. Valor esperado en la ruleta Mientras juega a la ruleta en el casino Venetian de Las Vegas, un jugador está tratando de decidir si apostará \$5 al número 27 o \$5 a que el resultado es cualquiera de estas cinco posibilidades: 0, 00, 1, 2, 3. Del ejemplo 6, sabemos que el valor esperado de la apuesta de \$5 para un solo número es $-26¢$. Para la apuesta de \$5 a que el resultado es 0, 00, 1, 2 o 3, hay una probabilidad de 5/38 de obtener una ganancia neta de \$30 y una probabilidad de 33/38 de perder \$5.

a. Encuentre el valor esperado para la apuesta de \$5 a que el resultado será 0, 00, 1, 2 o 3.

b. ¿Qué apuesta es mejor: una apuesta de \$5 al número 27 o una apuesta de \$5 que el resultado es cualquiera de los números 0, 00, 1, 2 o 3? ¿Por qué?

29. Valor esperado para un seguro de vida Existe una probabilidad de 0.9986 de que un varón de 30 años, seleccionado al azar, viva durante todo el año (según datos del Departamento de Salud y los Servicios Humanos de Estados Unidos). Una compañía de seguros de vida llamada Fidelity cobra \$161 por asegurar que el hombre vivirá durante todo el año. Si el hombre no sobrevive al año, la póliza paga \$100,000 como beneficio por su muerte.

a. Desde la perspectiva del hombre de 30 años, ¿cuáles son los valores monetarios correspondientes a los dos eventos de sobrevivir al año y no hacerlo?

b. Si un hombre de 30 años compra la póliza, ¿cuál es su valor esperado?

c. ¿Puede la compañía de seguros esperar obtener beneficios de muchas de estas pólizas? ¿Por qué?

30. Valor esperado para un seguro de vida Existe una probabilidad de 0.9968 de que una mujer de 50 años, seleccionada al azar, viva durante todo el año (según datos del Departamento de Salud y los Servicios Humanos de Estados Unidos). Una compañía de seguros de vida llamada Fidelity cobra \$226 por asegurar que la mujer vivirá durante todo el año. Si la mujer no sobrevive al año, la póliza paga \$50,000 como beneficio por su muerte.

a. Desde la perspectiva de la mujer de 50 años, ¿cuáles son los valores correspondientes a los dos eventos de sobrevivir al año y no hacerlo.

b. Si una mujer de 50 años compra la póliza, ¿cuál es su valor esperado?

c. ¿Puede la compañía de seguros esperar obtener beneficios de muchas de estas pólizas? ¿Por qué?

5-2 Distribuciones de probabilidad binomial

Concepto clave En la sección 5-1 se introdujo el importante concepto de una distribución de probabilidad discreta. Entre las diferentes distribuciones de probabilidad discreta que existen, el foco de esta sección está en la *distribución de probabilidad binomial*. La parte 1 de la sección introduce la distribución de probabilidad binomial junto con métodos para encontrar las probabilidades. La parte 2 presenta métodos sencillos para encontrar la media y la desviación estándar de una distribución binomial. Como en otras secciones, destacamos la importancia de *interpretar* los valores de probabilidad para determinar si los eventos son *significativamente bajos* o *significativamente altos*.

PARTE 1 Fundamentos de la distribución de probabilidad binomial

Las distribuciones de probabilidad binomial permiten tratar con circunstancias en las que los resultados pertenecen a *dos* categorías, como cara/cruz o aceptable/defectuoso o sobreviviente/fallecido.

En cifras

9,000,000: Cantidad de personas que nacieron el mismo día que usted.

DEFINICIÓN

Una **distribución de probabilidad binomial** resulta de un procedimiento que cumple con estos cuatro requisitos:

1. El procedimiento tiene un *número fijo de ensayos*. (Un ensayo es una sola observación).
2. Los ensayos son *independientes*, lo que significa que el resultado de cualquier ensayo individual no afecta las probabilidades en los otros ensayos.
3. Cada ensayo debe tener todos los resultados clasificados en exactamente *dos* categorías, comúnmente llamadas *éxito* y *fracaso*.
4. La probabilidad de un éxito se conserva igual en todos los ensayos.

Notación para distribuciones de probabilidad binomial

E y F (éxito y fracaso) indican las dos posibles categorías de todos los resultados.

$P(\text{E}) = p$	(p = probabilidad de un éxito)
$P(\text{F}) = 1 - p = q$	(q = probabilidad de un fracaso)
n	número fijo de ensayos
x	un número específico de éxitos en n ensayos, por lo que x puede ser cualquier número entero entre 0 y n, inclusive
p	probabilidad de *éxito* en *uno* de los n ensayos
q	probabilidad de *fracaso* en *uno* de los n ensayos
$P(x)$	probabilidad de obtener exactamente x éxitos entre los n ensayos

La palabra *éxito* como se utiliza aquí es arbitraria y no necesariamente representa algo bueno. Cualquiera de las dos categorías posibles puede denominarse éxito E siempre que su probabilidad se identifique como p. (El valor de q se puede encontrar siempre a partir de $q = 1 - p$. Si $p = 0.95$, entonces $q = 1 - 0.95 = 0.05$).

PRECAUCIÓN Cuando se utiliza una distribución de probabilidad binomial, siempre asegúrese de que x y p sean *consistentes*, en el sentido de que ambas se refieran a la *misma* categoría denominada éxito.

EJEMPLO 1 Twitter

Cuando se selecciona un adulto al azar (con reemplazo), hay una probabilidad de 0.85 de que esta persona sepa lo que es Twitter (según los resultados de una encuesta de Pew Research Center). Suponga que se desea encontrar la probabilidad de que exactamente tres de cinco adultos seleccionados al azar sepan lo que es Twitter.

a. ¿Este procedimiento resulta en una distribución binomial?

b. Si este procedimiento resulta en una distribución binomial, identifique los valores de n, x, p y q.

SOLUCIÓN

a. Este procedimiento satisface los requisitos para una distribución binomial, como se muestra a continuación.

1. El número de ensayos (5) es fijo.
2. Los 5 ensayos son independientes porque la probabilidad de que cualquier adulto conozca Twitter no se ve afectada por los resultados de otros adultos seleccionados.

3. Cada uno de los 5 ensayos tiene dos categorías de resultados: la persona seleccionada sabe lo que es Twitter o no lo sabe.
4. Para cada adulto seleccionado al azar, existe una probabilidad de 0.85 de que esta persona sepa qué es Twitter y esa probabilidad sigue siendo la misma para cada una de las cinco personas seleccionadas.

b. Después de concluir que el procedimiento dado genera una distribución binomial, ahora procedemos a identificar los valores de n, x, p y q.

1. Con cinco adultos seleccionados aleatoriamente, tenemos $n = 5$.
2. Queremos conocer la probabilidad de que exactamente tres sepan lo que es Twitter, por lo que $x = 3$.
3. La probabilidad de éxito (obtener una persona que sepa qué es Twitter) para una selección es 0.85, por lo que $p = 0.85$.
4. La probabilidad de fracaso (no obtener a alguien que sepa lo que es Twitter) es de 0.15, por lo que $q = 0.15$.

Una vez más, es muy importante estar seguros de que x y p se refieran al mismo concepto de "éxito". En este ejemplo, usamos x para contar el número de personas que saben lo que es Twitter, por lo que p debe ser la probabilidad de que la persona seleccionada sepa qué es Twitter. Por lo tanto, x y p utilizan el mismo concepto de éxito: saber lo que es Twitter.

SU TURNO Resuelva el ejercicio 5 "Ensayo clínico de YSORT".

Tratamiento de eventos dependientes como independientes

Cuando se selecciona una muestra (como en una encuesta), generalmente tomamos muestras sin reemplazo. El muestreo sin reemplazo produce eventos dependientes, lo que viola un requisito de una distribución binomial. Sin embargo, es frecuente dar tratamiento a los eventos como si fueran independientes aplicando la siguiente directriz del 5% introducida en la sección 4-2:

Directriz del 5% para cálculos complejos

Cuando se toman muestras sin reemplazo y el tamaño de la muestra no supera el 5% del tamaño de la población, se considera que las selecciones son *independientes* (aunque en realidad sean dependientes).

Métodos para encontrar probabilidades binomiales

Ahora procedemos a presentar tres métodos para encontrar probabilidades correspondientes a la variable aleatoria x en una distribución binomial. El primer método implica cálculos que utilizan la *fórmula de probabilidad binomial* y es la base para los otros dos métodos. El segundo método involucra el uso de software o una calculadora, y el tercer método implica el uso de la tabla A-1 del apéndice A. (Dado que la tecnología tiene un uso tan extenso, las tablas se están volviendo obsoletas). Si se utiliza tecnología que automáticamente produce probabilidades binomiales, recomendamos que se resuelvan uno o dos ejercicios con el método 1 para comprender mejor la base de los cálculos.

"Cómo la estadística puede ayudar a salvar corazones"

Un artículo del *New York Times* de David Leonhardt presentó el título "Cómo la estadística puede ayudar a salvar corazones". Leonhardt escribió que los pacientes tienen mayores posibilidades de recuperación si sus arterias obstruidas se abren en un periodo de dos horas después de un ataque cardíaco. En 2005, el Departamento de Salud y Servicios Humanos de Estados Unidos comenzó a publicar datos hospitalarios en su sitio web www.hospitalcompare.hhs.gov, e incluyó el porcentaje de pacientes con ataque cardíaco que recibieron tratamiento en sus arterias bloqueadas dentro de las dos horas posteriores a su llegada al hospital. Al no querer sentirse avergonzados por datos negativos, los médicos y hospitales están reduciendo el tiempo que les toma desbloquear las arterias obstruidas. Leonhardt ha escrito sobre el San Francisco Medical Center de la Universidad de California, que redujo su tiempo a la mitad, de casi tres horas a unos 90 minutos. El uso efectivo de la estadística simple puede salvar vidas.

Método 1: Uso de la fórmula de probabilidad binomial En una distribución de probabilidad binomial, las probabilidades se pueden calcular usando la fórmula 5-5.

FÓRMULA 5-5 Fórmula de probabilidad binomial

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x} \qquad \text{para } x = 0, 1, 2, \ldots, n$$

donde

n = número de ensayos

x = número de éxitos entre n ensayos

p = probabilidad de éxito en cualquier ensayo

q = probabilidad de fracaso en cualquier ensayo ($q = 1 - p$)

La fórmula 5-5 también se puede expresar como $P(x) = {}_nC_x p^x q^{n-x}$. Con x elementos idénticos entre sí, y otros $n - x$ elementos idénticos entre sí, el número de permutaciones es ${}_nC_x = n!/[(n-x)!x!]$. De modo que los dos lados de esta ecuación son intercambiables. El símbolo factorial, presentado en la sección 4-4, expresa el producto de factores decrecientes. Dos ejemplos de factoriales son $3! = 3 \cdot 2 \cdot 1 = 6$ y $0! = 1$ (por definición).

EJEMPLO 2 Twitter

Dado que hay una probabilidad de 0.85 de que un adulto seleccionado al azar sepa lo que es Twitter, use la fórmula de probabilidad binomial para encontrar la probabilidad de que, cuando cinco adultos se seleccionan al azar, exactamente tres sepan qué es Twitter. Es decir, aplique la fórmula 5-5 para encontrar $P(3)$ dado que $n = 5$, $x = 3$, $p = 0.85$ y $q = 0.15$.

SOLUCIÓN

Si usamos los valores dados de n, x, p y q en la fórmula de probabilidad binomial (fórmula 5-5), obtenemos

$$\begin{aligned} P(3) &= \frac{5!}{(5-3)!3!} \cdot 0.85^3 \cdot 0.15^{5-3} \\ &= \frac{5!}{2!3!} \cdot 0.614125 \cdot 0.0225 \\ &= (10)(0.614125)(0.0225) = 0.138178 \\ &= 0.138 \text{ (redondeada a tres dígitos significativos)} \end{aligned}$$

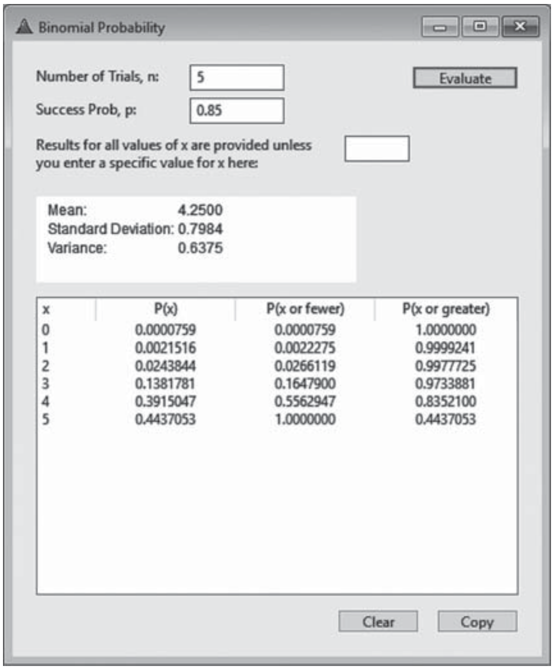
La probabilidad de obtener exactamente tres adultos que conozcan Twitter entre los cinco adultos seleccionados al azar es 0.138.

SU TURNO Resuelva el ejercicio 13 "Respuestas adivinadas".

Sugerencia de cálculo: Cuando se calcule una probabilidad con la fórmula de probabilidad binomial, es útil obtener un número único para $n!/[(n - x)!x!]$ o ${}_nC_x$, un número único para p^x y un único número para q^{n-x}. Simplemente multiplique los tres factores como se muestra en la tercera línea del cálculo en el ejemplo anterior. No redondee cuando encuentre esos tres factores; redondee solamente al final, hasta tres dígitos significativos.

Método 2: Uso de la tecnología Las probabilidades binomiales se pueden determinar mediante el uso de tecnologías de cálculo. La pantalla muestra las probabilidades binomiales listadas para $n = 5$ y $p = 0.85$, como en el ejemplo 2. Observe que en cada pantalla, la distribución de probabilidad se da como una tabla.

Statdisk



x	P(x)	P(x or fewer)	P(x or greater)
0	0.0000759	0.0000759	1.0000000
1	0.0021516	0.0022275	0.9999241
2	0.0243844	0.0266119	0.9977725
3	0.1381781	0.1647900	0.9733881
4	0.3915047	0.5562947	0.8352100
5	0.4437053	1.0000000	0.4437053

Minitab

x	P(X = x)
0	0.000076
1	0.002152
2	0.024384
3	0.138178
4	0.391505
5	0.443705

Excel

	A	B
1	x	P(x)
2	0	7.594E-05
3	1	0.0021516
4	2	0.0243844
5	3	0.1381781
6	4	0.3915047
7	5	0.4437053

TI-83/84 Plus CE

NORMAL FLOAT AUTO REAL RADIAN MP

L1	L2	L3	L4	L5
0	7.6E-5	------	------	------
1	.00215			
2	.02438			
3	.13818			
4	.3915			
5	.44371			
------	------			

EJEMPLO 3 Regla del tiempo extra en el fútbol americano

En el problema del capítulo, observamos que entre 1974 y 2011, hubo 460 partidos de fútbol americano de la NFL que se decidieron en tiempo extra, y 252 de ellos fueron ganados por el equipo que ganó el lanzamiento de moneda previo al tiempo extra. ¿Es el resultado de 252 victorias en 460 juegos equivalente a una oportunidad aleatoria, o 252 victorias son *significativamente altas*? Podemos responder a esa pregunta encontrando la probabilidad de 252 victorias o más en 460 juegos, suponiendo que las victorias y las derrotas son igualmente probables.

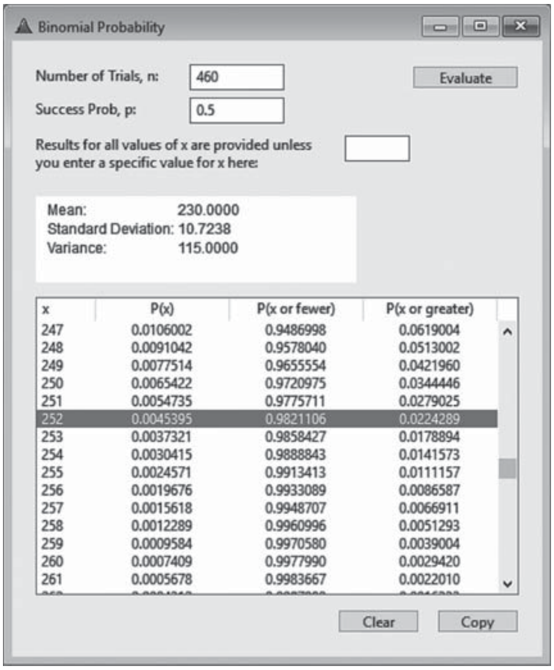
SOLUCIÓN

Si usamos la notación para las probabilidades binomiales, tenemos $n = 460$, $p = 0.5$, $q = 0.5$, y queremos encontrar la suma de todas las probabilidades para cada valor de x de 252 a 460. La fórmula no es práctica aquí, porque tendríamos que aplicarla 209 veces… y es algo que no queremos hacer. La tabla A-1 (probabilidades binomiales) no es aplicable porque $n = 460$, lo que está muy por encima del alcance de esa tabla. En cambio, elegimos usar la tecnología.

La pantalla de Statdisk en la página siguiente muestra que la probabilidad de 252 o más victorias en 460 juegos con tiempo extra es 0.0224 (redondeada), que es baja (menos de 0.05). Esto muestra que es poco probable que obtengamos 252 o más victorias por casualidad. Si descartamos efectivamente la casualidad, nos quedamos con la explicación más razonable de que el equipo que gana el lanzamiento de moneda previo al tiempo extra tiene una mayor oportunidad de ganar el juego.

continúa

Statdisk



SU TURNO Resuelva el ejercicio 23 "Teléfonos inteligentes".

El ejemplo 3 ilustra de buena manera la potencia y facilidad de uso de la tecnología. El ejemplo 3 también demuestra la regla del pensamiento estadístico del evento raro: si bajo un supuesto dado (como el de que ganar el lanzamiento de moneda previo al tiempo extra no tiene ningún efecto), la probabilidad de un evento particular observado (como 252 o más victorias en 460 juegos) es extremadamente pequeña (0.05 o menos), concluimos que la suposición probablemente no es correcta.

Método 3: Uso de la tabla A-1 en el apéndice A Este método puede omitirse si hay tecnología disponible. La tabla A-1 del apéndice A lista las probabilidades binomiales para valores seleccionados de n y p. No se puede utilizar si $n > 8$ o si la probabilidad p no es uno de los 13 valores incluidos en la tabla.

Para usar la tabla de probabilidades binomiales, primero debemos localizar n y el valor correspondiente deseado de x. En esta etapa, se debe aislar una fila de números. Luego alineamos esa fila con la probabilidad deseada de p usando la columna en la parte superior. El número aislado representa la probabilidad deseada. Una probabilidad muy pequeña, como 0.000064, está indicada por 0+.

EJEMPLO 4 Diablo de problema

Con base en una encuesta de Harris, 60% de los adultos creen en el diablo. Suponga que seleccionamos al azar a cinco adultos, utilice la tabla A-1 para encontrar lo siguiente:

a. La probabilidad de que exactamente tres de los cinco adultos crean en el diablo

b. La probabilidad de que el número de adultos que creen en el diablo sea de al menos dos

SOLUCIÓN

a. El siguiente fragmento de la tabla muestra que cuando $n = 5$ y $p = 0.6$, la probabilidad para $x = 3$ está dada por $P(3) = 0.346$.

TABLA A-1 Probabilidades binomiales

				p	
n	*x*	.01	.50	.60	.70
5	0	.951	.031	.010	.002
	1	.048	.156	.077	.028
	2	.001	.313	.230	.132
	3	0+	.313	.346	.309
	4	0+	.156	.259	.360
	5	0+	.031	.078	.168

x	*P*(*x*)
0	.010
1	.077
2	.230
3	.346
4	.259
5	.078

b. La frase “al menos dos” éxitos significa que el número de éxitos es 2, 3, 4 o 5.

$$\begin{aligned} P(\text{al menos 2 creen en el diablo}) &= P(2, 3, 4 \text{ o } 5) \\ &= P(2) + P(3) + P(4) + P(5) \\ &= 0.230 + 0.346 + 0.259 + 0.078 \\ &= 0.913 \end{aligned}$$

SU TURNO Resuelva el ejercicio 15 “Prueba SAT” usando la tabla A-1.

Proporciones de hombres y mujeres

Es bien sabido que cuando un bebé nace, no es igualmente probable el nacimiento de una niña que el de un niño. Actualmente se cree que hay 105 niños por cada 100 niñas, por lo que la probabilidad de un niño es de 0.512. Kristen Navara, de la Universidad de Georgia, llevó a cabo un estudio que muestra que alrededor del mundo nacen más niños que niñas, pero la diferencia se hace más pequeña a medida que la gente se encuentra más cerca del ecuador. Utilizó latitudes, temperaturas, tasas de desempleo, productos brutos y nacionales de 200 países y realizó un análisis estadístico que muestra que las proporciones de niños parecen estar afectadas sólo por la latitud y su clima. Hasta el momento, nadie ha identificado una explicación razonable para este fenómeno.

PARTE 2 Uso de la media y la desviación estándar para el pensamiento crítico

En la sección 5-1 se incluyeron fórmulas para hallar la media, la varianza y la desviación estándar de cualquier distribución discreta de probabilidad. Una distribución binomial es un tipo particular de distribución de probabilidad discreta, por lo que podríamos usar las mismas fórmulas, pero si conocemos los valores de n y p, es mucho más fácil usar lo siguiente:

Para las distribuciones binomiales

Fórmula 5-6 Media: $\mu = np$

Fórmula 5-7 Variación: $\sigma^2 = npq$

Fórmula 5-8 Desviación estándar: $\sigma = \sqrt{npq}$

Como en las secciones anteriores, encontrar valores para μ y σ puede ser muy divertido, pero es especialmente importante *interpretar* y *entender* esos valores, por lo que la regla práctica del rango y la regla de los eventos raros para la estadística inferencial pueden ser muy útiles. A continuación se presenta un breve resumen de la regla práctica del rango: los valores son significativamente bajos o altos si difieren de la media por más de 2 desviaciones estándar, como se describe a continuación:

Regla práctica del rango

Valores ***significativamente bajos*** $\leq (\mu - 2\sigma)$

Valores ***significativamente altos*** $\geq (\mu + 2\sigma)$

Valores ***no significativos***: Entre $(\mu - 2\sigma)$ y $(\mu + 2\sigma)$

EJEMPLO 5 Uso de parámetros para determinar la significancia

El problema del capítulo y el ejemplo 3 implican $n = 460$ partidos de fútbol americano de la NFL con tiempo extra. Obtenemos $p = 0.5$ y $q = 0.5$, suponiendo que ganar el lanzamiento de moneda antes del tiempo extra no proporciona una ventaja, por lo que ambos equipos tienen la misma oportunidad de 0.5 de ganar el juego en tiempo extra.

a. Encuentre la media y la desviación estándar para el número de victorias en grupos de 460 juegos.

b. Utilice la regla práctica del rango para encontrar los valores que separan el número de victorias que son significativamente bajas o significativamente altas.

c. ¿El resultado de 252 victorias en tiempo extra en 460 juegos es significativamente alto?

SOLUCIÓN

a. Con $n = 460$, $p = 0.5$, y $q = 0.5$, las fórmulas 5-6 y 5-8 se pueden aplicar como sigue:

$$\mu = np = (460)(0.5) = 230.0 \text{ juegos}$$

$$\sigma = \sqrt{npq} = \sqrt{(460)(0.5)(0.5)} = 10.7 \text{ (redondeado)}$$

Para grupos aleatorios de 460 juegos con tiempo extra, el número medio de victorias es de 230.0 juegos, y la desviación estándar es de 10.7 juegos.

b. Los valores que separan el número de victorias que son significativamente bajas o significativamente altas son los valores que indican dos desviaciones estándar desde la media. Con $\mu = 230.0$ juegos y $\sigma = 10.7$ juegos, obtenemos

$$\mu - 2\sigma = 230.0 - 2(10.7) = 208.6 \text{ juegos}$$

$$\mu + 2\sigma = 230.0 + 2(10.7) = 251.4 \text{ juegos}$$

Un número significativamente bajo de victorias son 208.6 juegos o menos, un número significativamente alto de victorias son 251.4 juegos o más, y los valores no significativos están entre 208.6 y 251.4 juegos.

c. El resultado de 252 victorias es significativamente alto porque es mayor que el valor de 251.4 juegos que se encontró en el inciso (b).

SU TURNO Resuelva el ejercicio 29 “Selección de género”.

En lugar de la regla práctica del rango, también podríamos usar probabilidades para determinar cuándo los valores son significativamente altos o bajos.

Uso de probabilidades para determinar cuándo los resultados son significativamente altos o bajos

- **Número significativamente *alto* de éxitos:** x éxitos entre n ensayos es *significativamente alto* si la probabilidad de x o más éxitos es 0.05 o menos. Es decir, x es un número significativamente alto de éxitos si $P(x \text{ o más}) \leq 0.05$.*
- **Número significativamente *bajo* de éxitos:** x éxitos entre n ensayos es *significativamente bajo* si la probabilidad de x o menos éxitos es de 0.05 o menos. Es decir, x es un número significativamente bajo de éxitos si $P(x \text{ o menos}) \leq 0.05$.

*El valor de 0.05 no es absolutamente rígido. Se podrían usar otros valores, como 0.01, para distinguir entre los resultados que son significativos y los que no lo son.

Justificación de la fórmula de probabilidad binomial

La fórmula de probabilidad binomial es la base para los tres métodos presentados en esta sección. En lugar de aceptar y usar esa fórmula a ciegas, veamos por qué funciona.

En el ejemplo 2, utilizamos la fórmula de probabilidad binomial para encontrar la probabilidad de obtener exactamente tres adultos que conozcan Twitter cuando se seleccionan cinco adultos al azar. Con $P(\text{conoce Twitter}) = 0.85$, podemos usar la regla de la multiplicación de la sección 4-2 para encontrar la probabilidad de que los tres primeros adultos conozcan Twitter y los dos últimos adultos no lo conozcan. Obtenemos el siguiente resultado:

$$\begin{aligned} &P(\text{3 adultos conocen Twitter seguidos por 2 adultos que no lo conocen}) \\ &= 0.85 \cdot 0.85 \cdot 0.85 \cdot 0.15 \cdot 0.15 \\ &= 0.85^3 \cdot 0.15^2 \\ &= 0.0138 \end{aligned}$$

Este resultado da una probabilidad de seleccionar al azar cinco adultos y encontrar que los tres primeros conocen Twitter y los dos últimos no. Sin embargo, la probabilidad de 0.0138 no es la probabilidad de obtener exactamente tres adultos que conozcan Twitter porque se encontró la probabilidad de una secuencia en particular. Existen otras secuencias posibles.

En la sección 4-4 vimos que con tres sujetos idénticos entre sí (como los adultos que conocen Twitter) y otros dos sujetos idénticos entre sí (como los adultos que no conocen Twitter), el número total de arreglos, o permutaciones, es $5!/[(5-3)!3!]$ o 10. Cada uno de esos 10 arreglos diferentes tiene una probabilidad de $0.85^3 \cdot 0.15^2$, por lo que la probabilidad total es la siguiente:

$$\begin{aligned} P(\text{3 adultos conocen Twitter entre 5}) &= \frac{5!}{(5-3)!3!} \cdot 0.85^3 \cdot 0.15^2 \\ &= 0.138 \end{aligned}$$

Este resultado particular puede generalizarse como la fórmula de probabilidad binomial (fórmula 5-5). Es decir, la fórmula de probabilidad binomial es una combinación de la regla de la multiplicación y la regla del conteo para el número de disposiciones de n elementos cuando x de ellos son idénticos entre sí y los otros $n - x$ también lo son entre sí.

$$P(x) = \overbrace{\frac{n!}{(n-x)!x!}}^{\substack{\text{El número de resultados} \\ \text{con exactamente } x \text{ éxitos} \\ \text{entre } n \text{ ensayos}}} \cdot \overbrace{p^x \cdot q^{n-x}}^{\substack{\text{La probabilidad de } x \text{ éxitos} \\ \text{entre } n \text{ ensayos para} \\ \text{cualquier orden particular}}}$$

En cifras

\$5: Costo del boleto en el primer vuelo de una aerolínea comercial en Estados Unidos en 1914; el vuelo tuvo una trayectoria de 21 millas.

CENTRO DE TECNOLOGÍA

Distribuciones binomiales

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga **clic** en **Analysis** en el menú superior.
2. Seleccione **Probability Distributions** en el menú desplegable y seleccione **Binomial Distribution** en el submenú.
3. Introduzca los valores de *n*, *p* y haga clic en **Evaluate**.

Sugerencia: Introduzca un valor específico para *x* a fin de obtener una sola probabilidad.

Minitab

1. Introduzca los valores de *x* para los que busca una probabilidad (como 0, 1, 2, 3, 4, 5) en la columna C1.
2. Seleccione **Calc** en el menú superior.
3. Seleccione **Probability Distributions** en el menú desplegable y **Binomial** en el submenú.
4. Seleccione **Probability**, ingrese el número de ensayos, la probabilidad del evento y seleccione C1 para la *columna de entrada*.
5. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Calculators** en el menú desplegable y **Binomial** en el submenú.
3. En el cuadro de diálogo, introduzca los valores deseados para *n*, *p*, *x*. Seleccione = o la desigualdad deseada para *x*.
4. Haga clic en **Compute**.

Calculadora TI-83/84 Plus

1. Pulse las teclas 2ND y VARS para acceder al menú *DISTR* (distribuciones).
2. Seleccione **binompdf** y haga clic en ENTER.
3. Introduzca los valores de los ensayos *n*, la probabilidad *p* y el número de éxitos *x* para completar el comando **binompdf(*n*, *p*, *x*)**. Pulse ENTER.

Sugerencia: Al omitir un valor para *x* se obtiene una lista de todas las probabilidades correspondientes a $x = 0, 1, 2, ..., n$. Pulse STOP luego 2ND y después 2 para guardar las probabilidades como la lista *L2*. A continuación, puede introducir manualmente los valores de *x* en la lista L1 para realizar los cálculos.

Sugerencia: Seleccione **binomcdf** en el paso 2 para obtener probabilidades acumuladas.

Excel

1. Introduzca los valores de *x* para los que desea probabilidades (como 0, 1, 2, 3, 4, 5) en la columna A.
2. Seleccione la celda **B1**, haga clic en **Insert Function** f_x, seleccione la categoría **Statistical**, seleccione la función **BINOM.DIST** y haga clic en **OK**.
3. Introduzca **A1** para *Number_s* y luego ingrese el número de ensayos *n* y la probabilidad *p*.
4. Introduzca **0** en el cuadro *acumulativo*.
5. Haga clic en **OK** y la probabilidad aparecerá en la celda B1.
6. Copie **B1** en la columna para obtener la probabilidad de cada valor de *x* listado en la columna A.

Sugerencia: Introduzca 1 en el paso 4 para la distribución binomial acumulada.

5-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Entregas con drones De acuerdo con una encuesta de Pitney Bowes, suponga que 42% de los consumidores se sienten cómodos con el hecho de que sus compras les sean entregadas con drones. Suponga que se desea encontrar la probabilidad de que cuando se seleccionan cinco consumidores al azar, exactamente dos de ellos se sientan cómodos con los drones. ¿Qué hay de erróneo en usar la regla de la multiplicación para encontrar la probabilidad de que dos consumidores se sientan cómodos con drones, seguidos por tres consumidores que no se sientan cómodos, como en este cálculo: $(0.42)(0.42)(0.58)(0.58)(0.58) = 0.0344$?

2. Notación Suponga que deseamos encontrar la probabilidad de que al seleccionar cinco consumidores al azar, exactamente dos de ellos se sientan cómodos con la entrega por drones. También suponga que 42% de los consumidores están cómodos con los drones (según una encuesta de Pitney Bowes). Identifique los valores de n, x, p y q.

3. Eventos independientes Con base en una encuesta de Pitney Bowes, cuando se les preguntó a 1009 consumidores si se sentían cómodos con que sus compras les fueran entregadas usando drones, 42% dijo que sí. Considere la probabilidad de que entre 30 consumidores diferentes seleccionados al azar de los 1009 encuestados, haya al menos 10 que se sientan cómodos con los drones. Dado que los sujetos encuestados fueron seleccionados sin reemplazo, ¿son independientes las 30 selecciones? ¿Pueden ser tratados como independientes? ¿Es posible encontrar la probabilidad usando la fórmula de la probabilidad binomial? Explique.

4. Notación de 0+ Si se usa la misma encuesta del ejercicio 3, la probabilidad de seleccionar al azar a 30 de los 1009 consumidores y obtener exactamente 24 que se sienten cómodos con los drones, se representa como 0+. ¿Qué indica 0+? ¿0+ indica que es imposible obtener exactamente 24 consumidores que se sientan cómodos con los drones?

Identificación de distribuciones binomiales. ***En los ejercicios 5 a 12, determine si el procedimiento dado resulta en una distribución binomial (o una distribución que pueda tratarse como binomial). Para aquellas que no sean binomiales, identifique al menos un requisito que no se satisfaga.***

5. Ensayo clínico de YSORT El método YSORT para la selección de género, desarrollado por el Instituto de Genética y FIV, fue diseñado para aumentar la probabilidad de que un bebé sea un niño. Cuando 291 parejas usan el método YSORT y tienen 291 bebés, se registran sus pesos.

6. Ensayo clínico de YSORT El método YSORT para la selección de género, desarrollado por el Instituto de Genética y FIV, fue diseñado para aumentar la probabilidad de que un bebé sea un niño. Cuando 291 parejas usan el método YSORT y tienen 291 bebés, se registran los géneros de los bebés.

7. LOL En una encuesta de US Cellular a 500 usuarios de teléfonos inteligentes, se pregunta a los sujetos si encuentran que las abreviaturas (como LOL o BFF) son molestas, y cada respuesta se registra como "sí" u "otro".

8. LOL En una encuesta de US Cellular a 500 usuarios de teléfonos inteligentes, se pregunta a los sujetos si encuentran que las abreviaturas (como LOL o BFF) son molestas, y cada respuesta se registra como "sí", "no" o "no estoy seguro".

9. Sondeo a senadores Los miembros del Senado del Congreso 113 incluyen 80 varones y 20 mujeres. Cuarenta senadores diferentes son seleccionados al azar sin reemplazo, y se registra el género de cada senador seleccionado.

10. Sondeo a senadores Diez senadores diferentes del Congreso 113 son seleccionados aleatoriamente sin reemplazo, y se registra la cantidad de periodos que han servido en el Congreso.

11. Encuesta de tarjetas de crédito En una encuesta de *AARP Bulletin*, 1019 adultos fueron seleccionados al azar sin reemplazo. Se preguntó a los encuestados si tenían una o más tarjetas de crédito, y las respuestas se registraron como "sí" o "no".

12. Investigación de citas En una encuesta patrocinada por TGI Friday's, 1000 participantes adultos fueron seleccionados al azar sin reemplazo, y a cada uno se le preguntó si investigan a sus citas en las redes sociales antes de reunirse con ellas. Las respuestas pueden ser "sí" o "no".

Fórmula binomial de probabilidad. ***En los ejercicios 13 y 14, responda las preguntas diseñadas para ayudar a entender la justificación de la fórmula de probabilidad binomial.***

13. Respuestas adivinadas Los exámenes estándar, como SAT, ACT, o el Examen de Admisión a Profesiones Médicas (MCAT), suelen utilizar preguntas de opción múltiple, cada una con cinco opciones de respuesta (a, b, c, d, e) y sólo una es correcta. Suponga que adivina las respuestas a las tres primeras preguntas.

a. Utilice la regla de la multiplicación para encontrar la probabilidad de que las dos primeras conjeturas estén equivocadas y la tercera sea correcta. Es decir, encuentre P(IIC), donde I expresa una respuesta incorrecta y C, una respuesta correcta.

b. Comenzando con IIC, haga una lista completa de los diferentes arreglos posibles de dos respuestas incorrectas y una respuesta correcta, luego determine la probabilidad de cada entrada en la lista.

c. Con base en los resultados anteriores, ¿cuál es la probabilidad de obtener exactamente una respuesta correcta cuando se hacen tres conjeturas?

14. Fuente de noticias De acuerdo con los datos de una encuesta de Harris Interactive, 40% de los adultos dicen que prefieren recibir sus noticias en línea. Se seleccionan cuatro adultos al azar.

continúa

a. Use la regla de la multiplicación para encontrar la probabilidad de que los tres primeros prefieran recibir sus noticias en línea y el cuarto prefiera una fuente diferente. Es decir, encuentre *P*(LLLD), donde L expresa una preferencia por las noticias en línea y D, una preferencia por una fuente de noticias diferente.

b. Comenzando con LLLD, haga una lista completa de los diferentes arreglos posibles de esas cuatro letras, luego encuentre la probabilidad de cada entrada en la lista.

c. Con base en los resultados anteriores, ¿cuál es la probabilidad de obtener exactamente tres adultos que prefieran recibir sus noticias en línea y un adulto que prefiera una fuente de noticias diferente.

Prueba SAT. ***En los ejercicios 15 a 20, suponga que se hacen conjeturas aleatorias para ocho preguntas de opción múltiple en una prueba SAT, de modo que hay $n = 8$ ensayos, cada uno con probabilidad de éxito (correcta) dada por $p = 0.20$. Encuentre la probabilidad indicada para el número de respuestas correctas.***

15. Encuentre la probabilidad de que el número x de respuestas correctas sea exactamente 7.

16. Encuentre la probabilidad de que el número x de respuestas correctas sea al menos 4.

17. Encuentre la probabilidad de que el número x de respuestas correctas sea menor que 3.

18. Encuentre la probabilidad de que el número x de respuestas correctas no sea mayor que 2.

19. Encuentre la probabilidad de no tener respuestas correctas.

20. Encuentre la probabilidad de que al menos una respuesta sea correcta.

En los ejercicios 21 a 24, suponga que cuando se seleccionan aleatoriamente adultos con teléfonos inteligentes, 54% los usa en reuniones o clases (según datos de una encuesta de LG Smartphone).

21. Si se seleccionan al azar 8 usuarios adultos de teléfonos inteligentes, encuentre la probabilidad de que exactamente 6 de ellos los utilicen en reuniones o clases.

22. Si se seleccionan al azar 20 usuarios adultos de teléfonos inteligentes, encuentre la probabilidad de que exactamente 15 de ellos los utilicen en reuniones o clases.

23. Si se seleccionan al azar 10 usuarios adultos de teléfonos inteligentes, encuentre la probabilidad de que al menos 8 de ellos los utilicen en reuniones o clases.

24. Si se seleccionan al azar 12 usuarios adultos de teléfonos inteligentes, encuentre la probabilidad de que menos de 3 usuarios los utilicen en reuniones o clases.

En los ejercicios 25 a 28, encuentre las probabilidades y responda las preguntas.

25. *Whitus vs Georgia* En el caso clásico de *Whitus vs Georgia*, se suponía que un grupo de jurados de 90 personas se seleccionaba al azar de una población en la que 27% eran minorías. Entre las 90 personas seleccionadas, 7 eran minorías. Encuentre la probabilidad de obtener 7 o menos minorías si el grupo de jurados fue seleccionado aleatoriamente. ¿Es el resultado de 7 minorías significativamente bajo? ¿Qué sugiere el resultado sobre el proceso de selección del jurado?

26. Corrección de la visión Una encuesta patrocinada por el Consejo de la Visión mostró que 79% de los adultos necesitan corrección (gafas, lentes de contacto, cirugía, etc.) para su vista. Si 20 adultos son seleccionados al azar, encuentre la probabilidad de que al menos 19 de ellos necesiten corrección para su vista. ¿Es 19 un número significativamente alto de adultos que necesitan corrección para la vista?

27. Nos vemos después De acuerdo con una encuesta de Harris Interactive, 20% de los adultos creen en la reencarnación. Suponga que seis adultos son seleccionados al azar y encuentre la probabilidad indicada.

a. ¿Cuál es la probabilidad de que exactamente cinco de los adultos seleccionados crean en la reencarnación?

b. ¿Cuál es la probabilidad de que todos los adultos seleccionados crean en la reencarnación?

c. ¿Cuál es la probabilidad de que al menos cinco de los adultos seleccionados crean en la reencarnación?

d. Si seis adultos son seleccionados al azar, ¿cinco personas que creen en la reencarnación son un número significativamente alto?

28. Demasiado joven para tatuarse Con base en una encuesta de Harris, entre los adultos que se arrepienten de hacerse tatuajes, 20% dice que eran demasiado jóvenes cuando se tatuaron. Suponga que cinco adultos arrepentidos de haberse hecho tatuajes son seleccionados al azar y encuentre la probabilidad indicada.

a. Determine la probabilidad de que ninguno de los adultos seleccionados diga que eran demasiado jóvenes para hacerse tatuajes.

b. Determine la probabilidad de que exactamente uno de los adultos seleccionados diga que era demasiado joven para hacerse tatuajes.

c. Determine la probabilidad de que el número de adultos seleccionados que digan que eran demasiado jóvenes sea 0 o 1.

d. Si se seleccionan al azar cinco adultos, ¿uno que dice que era demasiado joven para hacerse tatuajes es un número significativamente bajo?

Significancia con la regla práctica del rango. ***En los ejercicios 29 y 30, suponga que diferentes grupos de parejas usan el método XSORT de selección de género y cada pareja tiene un bebé. El método XSORT está diseñado para aumentar la probabilidad de que un bebé sea niña, pero suponga que el método no tiene ningún efecto, por lo que la probabilidad de tener una niña es 0.5.***

29. Selección de género Suponga que los grupos consisten en 36 parejas.

a. Encuentre la media y la desviación estándar para el número de niñas en grupos de 36 nacimientos.

b. Utilice la regla práctica del rango para encontrar los valores que separan los resultados que son significativamente bajos o altos.

c. ¿El resultado de 26 niñas es un resultado significativamente alto? ¿Qué sugiere sobre la efectividad del método XSORT?

30. Selección de género Suponga que los grupos consisten en 16 parejas.

a. Encuentre la media y la desviación estándar para el número de niñas en grupos de 16 nacimientos.

b. Utilice la regla práctica del rango para encontrar los valores que separan los resultados que son significativamente bajos o altos.

c. ¿Es el resultado de 11 niñas un resultado significativamente alto? ¿Qué sugiere sobre la efectividad del método XSORT?

Significancia con la regla práctica del rango. ***En los ejercicios 31 y 32, suponga que se llevan a cabo experimentos de hibridación con chícharos, los cuales poseen la propiedad de que, para los descendientes, existe una probabilidad de 0.75 de que un chícharo tenga vainas verdes (como en uno de los famosos experimentos de Mendel).***

31. Híbridos Suponga que los chícharos descendientes se seleccionan al azar en grupos de 10.

a. Encuentre la media y la desviación estándar para el número de chícharos con vainas verdes en los grupos de 10.

b. Utilice la regla práctica del rango para encontrar los valores que separan los resultados que son significativamente bajos o altos.

c. ¿Es el resultado de 9 chícharos con vainas verdes significativamente alto? ¿Por qué sí o por qué no?

32. Híbridos Suponga que los chícharos descendientes se seleccionan al azar en grupos de 16.

a. Encuentre la media y la desviación estándar para el número de chícharos con vainas verdes en los grupos de 16.

b. Utilice la regla práctica del rango para encontrar los valores que separan los resultados que son significativamente bajos o altos.

c. ¿Es un resultado de 7 chícharos con vainas verdes significativamente bajo? ¿Por qué sí o por qué no?

Muestreo compuesto. ***Los ejercicios 33 y 34 incluyen el método de muestreo compuesto, mediante el cual un laboratorio de pruebas médicas ahorra tiempo y dinero al combinar muestras de sangre para las pruebas, de modo que sólo se realice una prueba para varias personas. Una muestra***

combinada muestra positivo si al menos una persona tiene la enfermedad. Si una muestra combinada es positiva, se usan análisis de sangre individuales para identificar al individuo con la enfermedad o el trastorno.

33. VIH Se estima que, en todo el mundo, 1% de las personas entre 15 y 49 años está infectado con el virus de la inmunodeficiencia humana (VIH) (según datos de los Institutos Nacionales de Salud). En las pruebas para el VIH, se combinan las muestras de sangre de 36 personas. ¿Cuál es la probabilidad de que una muestra combinada sea positiva para el VIH? ¿Es improbable que una muestra combinada de este tipo sea positiva?

34. Anemia Con base en datos de Bloodjournal.org, 10% de las mujeres de 65 años de edad y mayores tienen anemia, que es una deficiencia de los glóbulos rojos. En las pruebas de anemia, se combinan las muestras de sangre de 8 mujeres mayores de 65 años. ¿Cuál es la probabilidad de que las pruebas a las muestras combinadas sean positivas para la anemia? ¿Es probable que una muestra combinada tenga un resultado positivo?

Muestreo de aceptación. ***Los ejercicios 35 y 36 incluyen el método de muestreo de aceptación, mediante el cual se acepta un envío de un gran número de artículos con base en los resultados de una muestra de tales artículos.***

35. Aspirina La compañía farmacéutica MedAssist recibe grandes envíos de tabletas de aspirina y utiliza este plan de muestreo de aceptación: selecciona y prueba aleatoriamente 40 tabletas, luego acepta el lote completo si sólo hay una o ninguna que no cumpla con las especificaciones requeridas. Si un envío de 5000 tabletas de aspirina realmente tiene una tasa de 3% de defectos, ¿cuál es la probabilidad de que todo este envío sea aceptado? ¿Casi todos estos envíos serán aceptados, o muchos serán rechazados?

36. Baterías AAA Compañías como Duracell, Energizer, Eveready y Panasonic fabrican baterías AAA. Al comprar pedidos a granel de baterías AAA, un fabricante de juguetes utiliza este plan de muestreo de aceptación: selecciona al azar 50 baterías y determina si cada una está dentro de las especificaciones. El envío entero se acepta si a lo más 2 baterías no cumplen con las especificaciones. Un envío contiene 2000 baterías AAA, y 2% de ellas no cumplen con las especificaciones. ¿Cuál es la probabilidad de que todo este envío sea aceptado? ¿Casi todos estos envíos serán aceptados, o muchos serán rechazados?

¡Ejercicios binomiales extremos! ***Los ejercicios 37 a 40 implican hallar probabilidades binomiales, encontrar parámetros y determinar si los valores son significativamente altos o bajos usando la regla práctica del rango y las probabilidades.***

37. M&Ms El conjunto de datos 27 "Pesos de M&M" en el apéndice B incluye datos de 100 caramelos M&M y 19 de ellos son verdes. Mars, Inc. afirma que 16% de sus caramelos M&M son verdes. De aquí en adelante, suponga que la afirmación del 16% es cierta y que una muestra consta de 100 M&Ms.

a. Utilice la regla práctica del rango para identificar los límites que separan los valores que son significativamente bajos y altos. Con base en los resultados, ¿19 M&Ms verdes son significativamente altos?

b. Encuentre la probabilidad de exactamente 19 M&Ms verdes.

c. Encuentre la probabilidad de 19 o más M&Ms verdes.

d. ¿Qué probabilidad es relevante para determinar si el resultado de 19 M&Ms verde es significativamente alto: la probabilidad del inciso (b) o del inciso (c)? Con base en la probabilidad relevante, ¿es el resultado de 19 M&Ms significativamente alto?

e. ¿Qué sugieren los resultados sobre la afirmación del 16% hecha por Mars. Inc.?

38. Política El secretario del Condado de Essex, Nueva Jersey, fue acusado de hacer trampa al no utilizar la aleatoriedad al asignar posiciones en la línea de votación. En 41 votos, los demócratas fueron asignados a la línea superior 40 veces. Suponga que a los demócratas y a los republicanos se les asigna la línea superior usando un método de selección aleatoria para que tengan la misma probabilidad de obtener tal línea.

a. Utilice la regla práctica del rango para identificar los límites que separan los valores significativamente bajos y altos. Con base en los resultados, ¿es el resultado de 40 líneas superiores para los demócratas significativamente alto?

b. Encuentre la probabilidad de exactamente 40 líneas superiores para los demócratas.

c. Encuentre la probabilidad de 40 o más líneas superiores para los demócratas.

d. ¿Qué probabilidad es relevante para determinar si las 40 líneas superiores para los demócratas son significativamente altas: la probabilidad del inciso (b) o del inciso (c)? De acuerdo con la probabilidad relevante, ¿es el resultado de 40 líneas superiores para los demócratas significativamente alto?

e. ¿Qué sugieren los resultados acerca de cómo el secretario cumplió con el requisito de asignar las posiciones de línea usando un método aleatorio?

39. Percepción y realidad En una elección presidencial, 611 votantes seleccionados aleatoriamente fueron encuestados y 308 de ellos dijeron que votaron por el candidato ganador (basado en datos del ICR Survey Research Group). El porcentaje real de votos para el candidato ganador fue del 43%. Suponga que 43% de los votantes realmente votaron por el candidato ganador y que los 611 votantes se seleccionaron al azar.

a. Utilice la regla práctica del rango para identificar los límites que separan los valores que son significativamente bajos y altos. Con base en los resultados, ¿los 308 votantes que dijeron haber votado por el ganador son significativamente altos?

b. Encuentre la probabilidad de que exactamente 308 votantes realmente hayan votado por el ganador.

c. Encuentre la probabilidad de que 308 o más votantes realmente hayan votado por el ganador.

d. ¿Qué probabilidad es relevante para determinar si el valor de 308 votantes es significativamente alto: la probabilidad del inciso (b) o del inciso (c)? De acuerdo con la probabilidad relevante, ¿el resultado de 308 votantes que dijeron haber votado por el ganador es significativamente alto?

e. ¿Cuál es una observación importante sobre los resultados de la encuesta?

40. Híbridos Uno de los famosos experimentos de Mendel con chícharos resultó en 580 descendientes y 152 de ellos eran amarillos. Mendel afirmó que en las mismas condiciones, 25% de los chícharos descendientes serían amarillos. Suponga que la afirmación de Mendel del 25% es verdadera y que una muestra consta de 580 chícharos.

a. Utilice la regla práctica del rango para identificar los límites que separan los valores que son significativamente bajos y altos. Con base en los resultados, ¿152 chícharos amarillos son significativamente bajos o significativamente altos?

b. Encuentre la probabilidad de exactamente 152 chícharos amarillos.

c. Encuentre la probabilidad de 152 o más chícharos amarillos.

d. ¿Cuál probabilidad es relevante para determinar si 152 chícharos son significativamente altos: la probabilidad del inciso (b) o del inciso (c)? De acuerdo con la probabilidad relevante, ¿es el resultado de 152 chícharos amarillos significativamente alto?

e. ¿Qué sugieren los resultados sobre la afirmación de Mendel del 25%?

5-2 Más allá de lo básico

41. Distribución geométrica Si un procedimiento cumple las condiciones de una distribución binomial, excepto que el número de ensayos no es fijo, entonces se puede utilizar la **distribución geométrica**. La probabilidad de obtener el primer éxito en el ensayo x está dada por $P(x) = p(1 - p)^{x-1}$, donde p es la probabilidad de éxito en cualquier ensayo. Los sujetos se seleccionan al azar para la Encuesta Nacional de Exámenes de Salud y Nutrición realizada por el Centro Nacional de Estadística de la Salud, Centros de Control y Prevención de Enfermedades. La probabilidad de que alguien sea un donante universal (con el grupo O y el tipo de sangre Rh negativo) es de 0.06. Encuentre la probabilidad de que el primer sujeto que sea un donante de sangre universal sea la quinta persona seleccionada.

42. Distribución multinomial La distribución binomial se aplica sólo a los casos que implican dos tipos de resultados, mientras que la distribución multinomial implica más de dos categorías. Suponga que tenemos tres tipos de resultados mutuamente excluyentes expresados mediante A, B y C. Sea $P(\mathrm{A}) = p_1$, $P(\mathrm{B}) = p_2$ y $P(\mathrm{C}) = p_3$. En n ensayos independientes, la probabilidad de x_1 resultados del tipo A, x_2 resultados del tipo B, y x_3 resultados del tipo C está dada por

$$\frac{n!}{(x_1)!(x_2)!(x_3)!} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3}$$

Una rueda de la ruleta en el casino Venetian de Las Vegas tiene 18 casillas rojas, 18 casillas negras y 2 casillas verdes. Si la ruleta se juega 15 veces, encuentre la probabilidad de obtener 7 resultados rojos, 6 resultados negros y 2 resultados verdes.

43. Distribución hipergeométrica Si tomamos muestras de una pequeña población finita sin reemplazo, la distribución binomial no debe ser usada porque los eventos no son independientes. Si el muestreo se hace sin reemplazo y los resultados pertenecen a uno de dos tipos, podemos usar la distribución hipergeométrica. Si una población tiene objetos de tipo A (como los números de lotería seleccionados), mientras que los objetos B restantes son del otro tipo (como los números de lotería que no seleccionó) y si se muestrean n objetos sin reemplazo (como los seis números de lotería sorteados), entonces la probabilidad de obtener x objetos de tipo A y $n - x$ objetos de 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 el juego de lotería Pick 6 de Nueva Jersey, un apostador selecciona seis números del 1 al 49 (sin repetición) y posteriormente resulta al azar una combinación ganadora de seis números. Encuentre las probabilidades de obtener exactamente dos números ganadores con un boleto. (*Sugerencia:* Utilice $A = 6$, $B = 43$, $n = 6$ y $x = 2$).

5-3 Distribuciones de probabilidad de Poisson

Concepto clave En la sección 5-1 introdujimos las distribuciones de probabilidad discretas generales y en la sección 5-2 se estudiaron las distribuciones de probabilidad binomial, que es una categoría particular de las distribuciones de probabilidad discretas. En esta sección introducimos las *distribuciones de probabilidad de Poisson*, que son otra categoría de las distribuciones de probabilidad discretas.

La siguiente definición establece que las distribuciones de Poisson se utilizan con ocurrencias de un evento en un intervalo especificado, y a continuación se listan algunas aplicaciones:

- Número de usuarios de Internet que ingresan a un sitio web en un día
- Número de pacientes que llegan a una sala de emergencias en una hora
- Número de huracanes en el Atlántico en un año

DEFINICIÓN

Una **distribución de probabilidad de Poisson** es una distribución de probabilidad discreta que se aplica a las ocurrencias de algún evento *en un intervalo especificado*. La variable aleatoria x es el número de ocurrencias del evento en un intervalo. El intervalo puede ser tiempo, distancia, área, volumen o alguna unidad similar. La probabilidad de que el suceso ocurra x veces en un intervalo está dada por la fórmula 5-9.

FÓRMULA 5-9. Distribución de probabilidad de Poisson

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!}$$

donde $e \approx 2.71828$

μ = número medio de ocurrencias del evento en los intervalos

En cifras

42: Número de años antes de que nos quedemos sin petróleo.

Requisitos para la distribución de probabilidad de Poisson

1. La variable aleatoria x es el número de ocurrencias de un evento *en algún intervalo*.
2. Las ocurrencias deben ser *aleatorias*.
3. Las ocurrencias deben ser *independientes* entre sí.
4. Las ocurrencias deben estar *uniformemente distribuidas* a lo largo del intervalo utilizado.

Parámetros de la distribución de probabilidad de Poisson

- La media es μ.
- La desviación estándar es $\sigma = \sqrt{\mu}$.

Propiedades de la distribución de probabilidad de Poisson

1. Una distribución particular de Poisson está determinada solamente por la media μ.
2. Una distribución de Poisson tiene posibles valores x de 0, 1, 2, . . . , sin límite superior.

EJEMPLO 1 Huracanes en el Atlántico

Para el período de 55 años a partir de 1960, hubo 336 huracanes en el Atlántico. Suponga que la distribución de Poisson es un modelo adecuado.

a. Encuentre μ, el número medio de huracanes por año.

b. Encuentra la probabilidad de que en un año seleccionado al azar, hayan existido exactamente 8 huracanes. Es decir, encuentre $P(8)$, donde $P(x)$ es la probabilidad de x huracanes en el Atlántico en un año.

c. En este período de 55 años, hubo en realidad 5 años con 8 huracanes del Atlántico. ¿Cómo se compara este resultado real con la probabilidad encontrada en el inciso (b)? ¿La distribución de Poisson parece ser un buen modelo en este caso?

SOLUCIÓN

a. La distribución de Poisson se aplica porque estamos tratando con las ocurrencias de un evento (huracanes) durante algún intervalo (un año). El número medio de huracanes por año es

$$\mu = \frac{\text{número de huracanes}}{\text{número de años}} = \frac{336}{55} = 6.1$$

continúa

b. A partir de la fórmula 5-9, la probabilidad de $x = 8$ huracanes en un año es la siguiente (con $x = 8$, $\mu = 6.1$, y $e = 2.71828$):

$$P(8) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{6.1^8 \cdot 2.71828^{-6.1}}{8!}$$

$$= \frac{(1{,}917{,}073.13)(0.0022428769)}{40{,}320} = 0.107$$

La probabilidad de exactamente 8 huracanes en un año es $P(8) = 0.107$.

c. La probabilidad de $P(8) = 0.107$ del inciso (b) es la probabilidad de que se formen 8 huracanes del Atlántico en 1 año. En 55 años, el número esperado de años con 8 huracanes del Atlántico es $55 \times 0.107 = 5.9$ años. El número esperado de años con 8 huracanes es de 5.9, lo que está razonablemente cerca de los 5 años en los que en realidad se formaron 8 huracanes, por lo que en este caso, el modelo de Poisson parece funcionar bastante bien.

SU TURNO Resuelva el ejercicio 5 "Huracanes".

Distribución de Poisson como aproximación a la distribución binomial

La distribución de Poisson se utiliza a veces para aproximar la distribución binomial cuando n es grande y p es pequeña. Una regla general es usar tal aproximación cuando se cumplen los dos requisitos siguientes.

Requisitos para utilizar Poisson como una aproximación a binomial

1. $n \geq 100$
2. $np \leq 10$

Si se satisfacen ambos requisitos y queremos usar la distribución de Poisson como una aproximación a la distribución binomial, necesitamos un valor para p. Ese valor se puede calcular usando la fórmula 5-6 (de la sección 5-2):

FÓRMULA 5-6 Media para Poisson como aproximación a binomial

$\mu = np$

EJEMPLO 2 Maine Pick 4

En el juego Maine Pick 4, usted paga 50¢ para seleccionar una secuencia de cuatro dígitos (del 0 al 9), como 1377. Si juega este juego una vez al día, encuentre la probabilidad de ganar al menos una vez en un año de 365 días.

SOLUCIÓN

El intervalo de tiempo es un día, y al jugar una vez cada día se tiene $n = 365$ juegos. Debido a que hay un conjunto ganador de números entre los 10,000 que son posibles (de 0000 a 9999), la probabilidad de un juego ganado es $p = 1/10{,}000$. Con $n = 365$ y $p = 1/10{,}000$, se satisfacen las condiciones $n \geq 100$ y $np \leq 10$, por lo que podemos usar la distribución de Poisson como una aproximación a la distribución binomial. Primero necesitamos el valor de p, que se encuentra como sigue:

$$\mu = np = 365 \cdot \frac{1}{10{,}000} = 0.0365$$

Después de haber encontrado el valor de μ, podemos proceder a encontrar la probabilidad de valores específicos de x. Debido a que queremos la probabilidad de que x sea "al menos 1", usaremos la estrategia inteligente de encontrar primero $P(0)$, la probabilidad de no ganar en 365 días. La probabilidad de al menos un juego ganado se puede encontrar rescatando ese resultado de 1. Encontramos $P(0)$ usando $x = 0$, $\mu = 0.0365$ y $e = 2.71828$, como se muestra aquí:

$$P(0) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{0.0365^0 \cdot 2.71828^{-0.0365}}{0!} = \frac{1 \cdot 0.9642}{1} = 0.9642$$

Si se usa la distribución de Poisson como una aproximación a la distribución binomial, encontramos que hay una probabilidad 0.9642 de no ganar un juego, por lo que la probabilidad de al menos un juego ganado es $1 - 09642 = 0.0358$. Si usamos la distribución binomial, obtenemos una probabilidad de 0.0358, por lo que la distribución de Poisson funciona muy bien aquí.

SU TURNO Resuelva el ejercicio 17 "Powerball: aproximación de Poisson a binomial".

CENTRO DE TECNOLOGÍA

Distribuciones de Poisson

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Probability Distributions** en el menú desplegable y seleccione **Poisson Distribution** en el submenú.
3. Introduzca el valor de la media y haga clic en **Evaluate**.

Minitab

1. Introduzca los valores de *x* para los que desea probabilidades (como 0, 1, 2, 3, 4, 5) en la columna C1.
2. Seleccione **Calc** en el menú superior.
3. Seleccione **Probability Distributions** del menú desplegable y **Poisson** en el submenú.
4. Seleccione **Probability**, ingrese la media y seleccione **C1** para la *columna de entrada*.
5. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Calculators** en el menú desplegable y **Poisson** en el submenú.
3. En el cuadro de diálogo introduzca el valor de la media y el valor de *x*. Seleccione = o la desigualdad deseada para *x*.
4. Haga clic en **Compute**.

Calculadora TI-83/84 Plus

1. Pulse las teclas **2ND** y **VARS** para acceder al menú *DISTR* (distribuciones).
2. Seleccione **poissonpdf** y pulse **ENTER**.
3. Introduzca los valores de la media (μ) y *x* para completar el comando **poissonpdf(μ, *x*)**. Pulse **ENTER**.

Sugerencia: Seleccione **poissoncdf** en el paso 2 para la probabilidad *acumulada*.

Excel

1. Introduzca los valores de *x* para los que desea probabilidades (como 0, 1, 2, 3, 4, 5) en la columna A.
2. Seleccione la celda **B1**, haga clic en **Insert Function f_x**, seleccione la categoría **Statistical**, seleccione la función **POISSON.DIST** y haga clic en **OK**.
3. Introduzca **A1** para *X* y después ingrese el valor de la media.
4. Introduzca **0** en el cuadro *acumulativo*.
5. Haga clic en **OK** y la probabilidad aparecerá en la celda B1.
6. Copie B1 hacia abajo en la columna para obtener la probabilidad de cada valor de *x* listado en la columna A.

Sugerencia: Introduzca **1** en el paso 4 para la distribución de Poisson *acumulada*.

5-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Notación Al analizar los impactos de bombas V-1 en la Segunda Guerra Mundial, el sur de Londres se dividió en 576 regiones, cada una con un área de 0.25 km^2. Un total de 535 bombas impactaron en el área combinada de 576 regiones. Suponga que deseamos encontrar la probabilidad de que una región seleccionada al azar tenga exactamente dos impactos. Al aplicar la fórmula 5-9, identifique los valores de μ, x y e. Además, describa brevemente lo que representa cada uno de esos símbolos.

2. Tornados Durante un período reciente de 64 años, Nuevo México tuvo un total de 153 tornados con intensidad de 1 o más en la escala de Fujita. Considere que la variable aleatoria x representa el número de estos tornados que impactaron Nuevo México en un año, y suponga que tiene una distribución de Poisson. ¿Cuál es el número medio de estos tornados en Nuevo México en un año? ¿Cuál es la desviación estándar? ¿Cuál es la varianza?

3. Probabilidad de la distribución de Poisson La variable aleatoria x representa el número de llamadas telefónicas que el autor recibe en un día, y tiene una distribución de Poisson con una media de 7.2 llamadas. ¿Cuáles son los posibles valores de x? ¿Es posible un valor de $x = 2.3$? ¿Es x una variable aleatoria discreta o una variable aleatoria continua?

4. Probabilidad si 0 Para la fórmula 5-9, ¿qué representa $P(0)$? Simplifique la fórmula 5-9 para el caso en que $x = 0$.

Huracanes. ***En los ejercicios 5 a 8, suponga que se aplica la distribución de Poisson; también asuma que el número medio de huracanes del Atlántico en Estados Unidos es de 6.1 por año, como en el ejemplo 1, y proceda a encontrar la probabilidad indicada.***

5. Huracanes

a. Encuentre la probabilidad de que, en un año, haya 5 huracanes.

b. En un período de 55 años, ¿en cuántos años se espera que haya 5 huracanes?

c. ¿Cómo se compara el resultado del inciso (b) con el período reciente de 55 años en el que 8 años tuvieron 5 huracanes? ¿Funciona bien la distribución de Poisson aquí?

6. Huracanes

a. Encuentre la probabilidad de que en un año no se formen huracanes.

b. En un período de 55 años, ¿en cuántos años se espera que no haya huracanes?

c. ¿Cómo se compara el resultado del inciso (b) con el período reciente de 55 años en el que no hubo años sin huracanes? ¿Funciona bien la distribución de Poisson aquí?

7. Huracanes

a. Encuentre la probabilidad de que en un año haya 7 huracanes.

b. En un período de 55 años, ¿en cuántos años se espera que haya 7 huracanes?

c. ¿Cómo se compara el resultado del inciso (b) con el período reciente de 55 años en el que en 7 años se formaron 7 huracanes? ¿Funciona bien la distribución de Poisson aquí?

8. Huracanes

a. Encuentre la probabilidad de que en un año se formen 4 huracanes.

b. En un período de 55 años, ¿en cuántos años se espera que haya 4 huracanes?

c. ¿Cómo se compara el resultado del inciso (b) con el período reciente de 55 años en el que 10 años tuvieron 4 huracanes? ¿Funciona bien la distribución de Poisson aquí?

En los ejercicios 9 a 16, use la distribución de Poisson para encontrar las probabilidades indicadas.

9. Nacimientos En un año reciente, NYU-Langone Medical Center tuvo 4221 nacimientos. Encuentre el número medio de nacimientos por día, luego use ese resultado para encontrar la probabilidad de que en un día haya 15 nacimientos. ¿Parece probable que en un día dado haya exactamente 15 nacimientos?

10. Asesinatos En un año reciente, hubo 333 asesinatos en la ciudad de Nueva York. Encuentre el número medio de asesinatos por día, luego use ese resultado para encontrar la probabilidad de que en un día no haya asesinatos. ¿Puede esperarse que haya muchos días sin asesinatos?

11. Decadencia radiactiva Los átomos radiactivos son inestables porque tienen demasiada energía. Cuando liberan su energía extra, se descomponen. Al estudiar el cesio-l37, un ingeniero nuclear descubrió que durante los 365 días, 1,000,000 de átomos radiactivos decaían a 97,728 átomos radiactivos: por lo tanto, 22,713 átomos decaían durante 365 días.

a. Calcule el número medio de átomos radiactivos que se descomponen en un día.

b. Encuentre la probabilidad de que, en un día dado, exactamente 50 átomos radiactivos decaigan.

12. Muertes por patadas de caballo Un ejemplo clásico de la distribución de Poisson implica el número de muertes causadas por patadas de caballo a hombres del ejército prusiano entre 1875 y 1894. Los datos de 14 cuerpos fueron combinados para el período de 20 años y los 280 cuerpos-año incluyeron un total de 196 muertes. Después de hallar el número medio de muertes por cuerpo-año, determine la probabilidad de que un cuerpo-año seleccionado al azar tenga el siguiente número de muertes: (**a**) 0, (**b**) 1, (**c**) 2, (**d**) 3, (**e**) 4. Los resultados reales consistieron en las siguientes frecuencias: 0 muertes (en 144 cuerpo-años); 1 muerte (en 91 cuerpo-años); 2 muertes (en 32 cuerpo-años); 3 muertes (en 11 cuerpo-años); 4 muertes (en 2 cuerpo-años). Compare los resultados reales con los esperados usando probabilidades de Poisson. ¿La distribución Poisson sirve como una buena herramienta para predecir los resultados reales?

13. Bombas de la Segunda Guerra Mundial En el ejercicio 1 "Notación", observamos que al analizar los impactos de bombas V-1 en la Segunda Guerra Mundial, el sur de Londres se dividió en 576 regiones, cada una con un área de 0.25 km^2. Un total de 535 bombas impactaron el área combinada de 576 regiones.

a. Encuentre la probabilidad de que una región seleccionada al azar tenga exactamente 2 impactos.

b. Entre las 576 regiones, encuentre el número esperado de regiones con exactamente 2 impactos.

c. ¿Cómo se compara el resultado del inciso (b) con este resultado real: Hubo 93 regiones que tuvieron exactamente 2 impactos?

14. Enfermedades en grupos El neuroblastoma, una forma rara de cáncer, se produce en 11 niños de cada millón, por lo que su probabilidad es 0.000011. Cuatro casos de neuroblastoma ocurrieron en Oak Park, Illinois, donde había 12,419 niños.

a. Suponiendo que el neuroblastoma ocurre como es usual, encuentre el número medio de casos en grupos de 12,429 niños.

b. Usando la media no redondeada del inciso (a), determine la probabilidad de que el número de casos de neuroblastoma en un grupo de 12,429 niños sea 0 o 1.

c. ¿Cuál es la probabilidad de más de un caso de neuroblastoma?

d. ¿El grupo con cuatro casos parece atribuible a la casualidad? ¿Por qué sí o por qué no?

15. Muertes en automóvil La tasa reciente de muertes en automóvil es de 33,561 muertes por 2969 mil millones de millas recorridas (según datos de la Administración Nacional de Seguridad del Tráfico en Carreteras). Encuentre la probabilidad de que para los próximos mil millones de millas recorridas haya por lo menos una muerte. ¿Qué indica el resultado sobre la probabilidad de al menos una muerte?

16. Cheques En un año reciente, el autor escribió 181 cheques. Encuentre la probabilidad de que en un día seleccionado aleatoriamente, escriba al menos un cheque.

5-3 Más allá de lo básico

17. Powerball: aproximación de Poisson a binomial Hay una probabilidad de 1/292,201,338 de ganar el premio mayor de la lotería Powerball con un solo boleto. Suponga que usted compra un boleto en cada uno de los 5200 juegos de Powerball diferentes que se realizarán en los próximos 50 años. Encuentre la probabilidad de ganar el premio con al menos uno de esos boletos. ¿Hay una buena probabilidad de que gane el premio al menos una vez en 50 años?

Examen rápido del capítulo

1. ¿Se define una distribución de probabilidad si los únicos valores posibles de una variable aleatoria son 0, 1, 2, 3 y $P(0) = P(1) = P(2) = P(3) = 1/3$?

2. Hay 80 preguntas en una prueba SAT, y todas son de opción múltiple con posibles respuestas de a, b, c, d, e. Para cada pregunta, sólo una respuesta es correcta. Encuentre la media y la desviación estándar para el número de respuestas correctas para aquellos que hacen conjeturas aleatorias en las 80 preguntas.

3. ¿Los valores encontrados en el ejercicio 2 son estadísticos o parámetros? ¿Por qué?

4. Usando las mismas preguntas del SAT descritas en el ejercicio 2, ¿es 20 un número significativamente alto de respuestas correctas para alguien que hace conjeturas al azar?

5. Usando las mismas preguntas del SAT descritas en el ejercicio 2, ¿es 8 un número significativamente bajo de respuestas correctas para alguien que hace conjeturas al azar?

x	$P(x)$
0	0+
1	0.006
2	0.051
3	0.205
4	0.409
5	0.328

En los ejercicios 6 a 10, utilice lo siguiente: cinco vuelos de American Airlines se seleccionan al azar y la tabla al margen muestra las probabilidades para el número de vuelos que llegan a tiempo (según datos del Departamento de Transporte). Suponga que cinco vuelos se seleccionan aleatoriamente.

6. ¿La tabla describe una distribución de probabilidad?

7. Encuentre la media del número de vuelos entre cinco que llegan a tiempo.

8. Con base en la tabla, la desviación estándar es 0.9 vuelos. ¿Cuál es la varianza? Incluya las unidades apropiadas.

9. ¿Qué indica la probabilidad de 0+? ¿Indica que entre cinco vuelos seleccionados aleatoriamente, es imposible que ninguno de ellos llegue a tiempo?

10. ¿Cuál es la probabilidad de que menos de tres de los cinco vuelos lleguen a tiempo?

Ejercicios de repaso

En los ejercicios 1 a 5, suponga que 74% de los adultos seleccionados al azar tienen una tarjeta de crédito (según los resultados de una encuesta del* AARP Bulletin*). Suponga que un grupo de cinco adultos se selecciona aleatoriamente.

1. Tarjetas de crédito Encuentre la probabilidad de que exactamente tres de los cinco adultos tengan tarjetas de crédito.

2. Tarjetas de crédito Encuentre la probabilidad de que al menos uno de los cinco adultos tenga una tarjeta de crédito. ¿El resultado es aplicable a cinco amigos adultos que están de vacaciones juntos? ¿Por qué sí o por qué no?

3. Tarjetas de crédito Encuentre la media y la desviación estándar para el número de adultos en grupos de cinco que tienen tarjetas de crédito.

4. Tarjetas de crédito Si los cinco adultos tienen tarjetas de crédito, ¿cinco es un número significativamente alto? ¿Por qué sí o por qué no?

5. Tarjetas de crédito Si el grupo de cinco adultos incluye exactamente 1 con tarjeta de crédito, ¿ese valor de 1 es significativamente bajo?

6. Encuesta de Seguridad En una encuesta de *USA Today*, se preguntó a los sujetos si las contraseñas deberían ser reemplazadas por seguridad biométrica, como las huellas dactilares. Los resultados de esa encuesta se han utilizado para crear la tabla adjunta. ¿Esta tabla describe una distribución de probabilidad? ¿Por qué sí o por qué no?

Respuesta	$P(x)$
Sí	0.53
No	0.17
No estoy seguro	0.30

7. Reconocimiento de marca En un estudio de reconocimiento de marca de Sony, se entrevistó a grupos de cuatro consumidores. Si x es el número de personas en el grupo que reconocen la marca Sony, entonces x puede ser 0, 1, 2, 3 o 4 y las probabilidades correspondientes son 0.0016, 0.0250, 0.1432, 0.3892 y 0.4096. ¿La información dada describe una distribución de probabilidad? ¿Por qué sí o por qué no?

8. Familia/Pareja Grupos de personas entre 15 y 65 años se seleccionan al azar y se organizan en grupos de seis. La variable aleatoria x es el número de personas en el grupo que dicen que su familia y/o pareja contribuyen más a su felicidad (con base en una encuesta de Coca-Cola). La tabla adjunta lista los valores de x junto con sus correspondientes probabilidades. ¿La tabla describe una distribución de probabilidad? Si es así, encuentre la media y la desviación estándar.

x	$P(x)$
0	0+
1	0.003
2	0.025
3	0.111
4	0.279
5	0.373
6	0.208

9. Detección de fraude La oficina del fiscal del distrito de Brooklyn analizó los dígitos principales (a la izquierda) de las cantidades en los cheques para identificar fraudes. Se espera que el dígito principal 1 ocurra 30.1% de las veces, de acuerdo con la "ley de Benford", que se aplica en este caso. Entre 784 cheques emitidos por una empresa sospechosa, no había ninguno con cantidades que tuvieran un dígito principal de 1.

a. Si hay un 30.1% de probabilidad de que el dígito principal de la cantidad en el cheque sea 1, ¿cuál es el número esperado de cheques entre 784 que debería tener un dígito principal de 1?

b. Suponga que se seleccionan al azar grupos de 784 cheques. Encuentre la media y la desviación estándar para el número de cheques con cantidades que tengan el dígito principal de 1.

c. Utilice los resultados del inciso (b) y la regla práctica del rango para identificar los valores que son significativamente bajos.

d. Dado que las 784 cantidades en los cheques reales no tuvieron dígitos principales de 1, ¿hay evidencia contundente de que los cheques sospechosos son muy diferentes de los resultados esperados? ¿Por qué sí o por qué no?

10. Poisson: Muertes En la actualidad, un promedio de 7 residentes del pueblo de Westport (760 habitantes) mueren cada año (según datos del Centro Nacional Estadounidense de Estadísticas de la Salud).

a. Calcule el número medio de muertes por día.

b. Encuentre la probabilidad de que en un día dado, no haya muertes.

c. Encuentra la probabilidad de que en un día dado, haya más de una muerte.

d. Con base en los resultados anteriores, ¿debe Westport tener un plan de contingencia para manejar más de una muerte al día? ¿Por qué sí o por qué no?

Ejercicios de repaso acumulado

1. Planetas Los planetas del sistema solar, ordenados de acuerdo con su distancia desde el Sol, tienen el número de lunas que se lista a continuación. (Plutón no está incluido porque desde 2006 ya no es invitado a la fiesta del Sistema Solar). Incluya las unidades apropiadas cuando sea relevante.

0 0 1 2 17 28 21 8

a. Encuentre la media.

b. Encuentre la mediana.

c. Encuentra la moda.

continúa

d. Encuentre el rango.

e. Encuentre la desviación estándar.

f. Encuentre la varianza.

g. Utilice la regla práctica del rango para identificar los valores que separan los resultados significativos de los no significativos.

h. Con base en el resultado del inciso (g), ¿tiene alguno de los planetas un número de lunas significativamente bajo o significativamente alto? ¿Por qué sí o por qué no?

i. ¿Cuál es el nivel de medición de los datos: nominal, ordinal, de intervalo o de razón?

j. ¿Los datos son discretos o continuos?

2. South Carolina Pick 3 En el juego de lotería South Carolina Pick 3, usted puede pagar $1 para seleccionar una secuencia de tres dígitos, como 221. Si compra solamente un boleto y gana, su premio es de $500 y su ganancia neta es de $499.

a. Si compra un billete, ¿cuál es la probabilidad de ganar?

b. Encuentre el número medio de juegos ganados en los años con exactamente 365 días si participara en este juego una vez al día.

c. Encuentre la probabilidad de ganar exactamente una vez en 365 días si participara en este juego una vez al día.

d. Encuentre el valor esperado para la compra de un boleto.

3. Desafío de tenis En un reciente Torneo Abierto de Tenis en Estados Unidos, hubo 879 desafíos realizados por jugadores individuales, y 231 de ellos resultaron en decisiones de los árbitros que fueron derrotadas. La tabla adjunta muestra los resultados por género.

	Desafío con decisión revocada	Desafío rechazado sin cambio
Desafíos de hombres	152	412
Desafíos de mujeres	79	236

a. Si 1 de los 879 desafíos se selecciona al azar, ¿cuál es la probabilidad de que resulte en una decisión revocada?

b. Si una de las decisiones revocadas se selecciona al azar, ¿cuál es la probabilidad de que el desafío haya sido hecho por una mujer?

c. Si dos desafíos diferentes son seleccionados al azar sin reemplazo, encuentre la probabilidad de que ambos resulten en una decisión revocada.

d. Si 1 de los 879 desafíos es seleccionado al azar, encuentre la probabilidad de que haya sido hecho por un hombre o haya resultado en una decisión revocada.

e. Si uno de los desafíos es seleccionado al azar, encuentre la probabilidad de que haya sido hecho por un hombre, dado que el reto resultó en una decisión revocada.

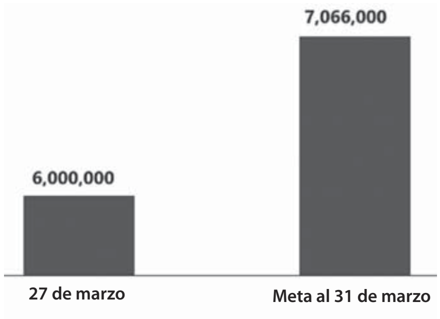
4. Solicitantes de empleo La Sociedad para la Administración de Recursos Humanos realizó una encuesta a 347 profesionales de los recursos humanos y encontró que 73% informó que sus empresas hacen verificaciones de antecedentes penales de todos los solicitantes de empleo.

a. Encuentre el número de encuestados que informaron que sus empresas realizan verificaciones de antecedentes penales de todos los solicitantes de empleo.

b. Identifique la muestra y la población.

c. ¿El valor de 73% es un estadístico o un parámetro? Explique.

5. Gráfica de barras Fox News transmitió una gráfica similar a la mostrada aquí. La gráfica pretende comparar el número de personas realmente inscritas en un plan de salud gubernamental (barra izquierda) y la meta para el número de inscritos (barra derecha). ¿La gráfica representa los datos correctamente o es de alguna manera engañosa? Explique.



Fuente: HHS

6. Lavado de manos Con base en los resultados de una encuesta de Bradley Corporation, suponga que 70% de los adultos siempre se lavan las manos después de usar un baño público.

a. Encuentre la probabilidad de que entre 8 adultos seleccionados al azar, exactamente 5 siempre se laven las manos después de usar un baño público.

b. Encuentre la probabilidad de que entre 8 adultos seleccionados al azar, al menos 7 siempre se laven las manos después de usar un baño público.

c. Para grupos de 8 adultos seleccionados al azar, encuentre la media y la desviación estándar de las cantidades de personas en los grupos que siempre se lavan las manos después de usar un baño público.

d. Si se seleccionan al azar 8 adultos y se descubre que exactamente uno de ellos se lava las manos después de usar un baño público, ¿es éste un número significativamente bajo?

Proyecto de tecnología

Sobreventa de vuelos El vuelo 171 de American Airlines del aeropuerto JFK de Nueva York al aeropuerto LAX de Los Ángeles utiliza un avión Airbus A321 con 189 asientos disponibles para los pasajeros. American Airlines puede sobrevender aceptando más reservaciones que los asientos disponibles. Si el vuelo no se sobrevende, la aerolínea pierde los ingresos de los asientos vacíos; si se venden demasiados asientos, la aerolínea pierde dinero por la compensación que debe pagar a los pasajeros que no alcanzan asiento. Suponga que hay una probabilidad de 0.0995 de que un pasajero con reservación no llegue para tomar el vuelo (según datos del documento de investigación de IBM "Passenger-Based Predictive Modeling of Airline No-Show Rates", de Lawrence, Hong y Cherrier). También suponga que American Airlines acepta 205 reservaciones para los 189 asientos disponibles.

• Encuentre la probabilidad de que al aceptar 205 reservaciones para el vuelo 171 de American Airlines, haya más pasajeros que asientos disponibles. ¿Es la probabilidad de sobreventa suficientemente pequeña para que no suceda muy a menudo, o parece suficientemente alta para que deban hacerse cambios con el fin de reducirla?

• Utilice prueba y error para encontrar el número máximo de reservaciones que se podrían aceptar de modo que la probabilidad de tener más pasajeros que asientos sea de 0.05 o menos.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿Los resultados de los experimentos de Mendel sobre la hibridación de plantas contradicen su teoría?

Gregor Mendel realizó experimentos originales para estudiar los rasgos genéticos de las plantas de chícharos. En 1865 escribió "Experiments in Plant Hybridization", que fue publicado en *Proceedings of the Natural History Society*. Mendel presentó una teoría de que cuando hay dos rasgos heredables, uno de ellos será dominante y el otro será recesivo. Cada uno de los padres aporta un gen a la descendencia y, dependiendo de la combinación de genes, esa descendencia podría heredar el rasgo dominante o el rasgo recesivo. Mendel llevó a cabo un experimento utilizando plantas de chícharos. Las vainas de las plantas de chícharo pueden ser verdes o amarillas. Cuando un chícharo que lleva un gen verde dominante y un gen amarillo recesivo se cruza con otro chícharo que lleva los mismos genes verdes/amarillos, el descendiente puede heredar cualquiera de estas cuatro combinaciones de genes: (1) verde/verde; (2) verde/amarillo; (3) amarillo/verde; (4) amarillo/amarillo. Debido a que el verde es dominante y el amarillo es recesivo, la vaina de la descendencia será verde si cualquiera de los dos genes heredados es verde. La descendencia puede tener una vaina amarilla solamente si hereda el gen amarillo de cada uno de los dos padres. Dadas estas condiciones, esperamos que 3/4 de los chícharos descendientes tengan vainas verdes; es decir, $P(\text{vaina verde}) = 3/4$.

Cuando Mendel llevó a cabo sus famosos experimentos de hibridación utilizando plantas de chícharos con la combinación verde/amarillo en los genes, obtuvo 580 descendientes. Según la teoría de Mendel, 3/4 de la descendencia debía tener vainas verdes, pero el número real de plantas con vainas verdes fue de 428. Por lo tanto, la proporción de descendientes con vainas verdes sobre el número total de descendientes es $428/580 = 0.738$. Mendel *esperaba* una proporción de 3/4 o 0.75, pero su *resultado real t*uvo una proporción de 0.738.

a. Suponiendo que $P(\text{vaina verde}) = 3/4$, encuentre la probabilidad de que entre 580 descendientes, el número de chícharos con vainas verdes sea *exactamente* 428.

b. Suponiendo que $P(\text{vaina verde}) = 3/4$, encuentre la probabilidad de que entre 580 descendientes, el número de chícharos con vainas verdes sea 428 *o menos*.

c. ¿Cuál de las dos probabilidades anteriores debería usarse para determinar si 428 es un número significativamente bajo de chícharos con vainas verdes?

d. Utilice las probabilidades para determinar si 428 chícharos con vainas verdes es un número significativamente bajo. (*Sugerencia:* Consulte "Identificación de resultados significativos con probabilidades" en la sección 5-1).

Actividades en equipo

1. Actividad en clase ¡Gane $1,000,000! La Fundación Educativa James Randi ofrece un premio de $1,000,000 a cualquier persona que pueda mostrar "bajo condiciones apropiadas de observación, evidencia de cualquier poder o evento paranormal, sobrenatural u oculto". Divídanse en grupos de tres. Seleccione una persona que será probada en cuanto a su percepción extrasensorial (PES) tratando de identificar correctamente un dígito (de 0 a 9) seleccionado al azar por otro miembro del equipo. Realice al menos 20 ensayos. Otro miembro del equipo debe registrar el dígito elegido al azar, el dígito adivinado por el sujeto y si la conjetura fue correcta o incorrecta. Construya una tabla de la distribución de probabilidad para los dígitos generados al azar, construya la tabla de frecuencias relativas para los dígitos aleatorios que se obtuvieron realmente y construya una tabla de frecuencias relativas para las conjeturas que se hicieron. Después de comparar las tres tablas, ¿qué concluye? ¿Qué proporción de conjeturas es correcta? ¿Parece que el sujeto tiene la capacidad de seleccionar el dígito correcto significativamente más a menudo de lo que se esperaría por casualidad?

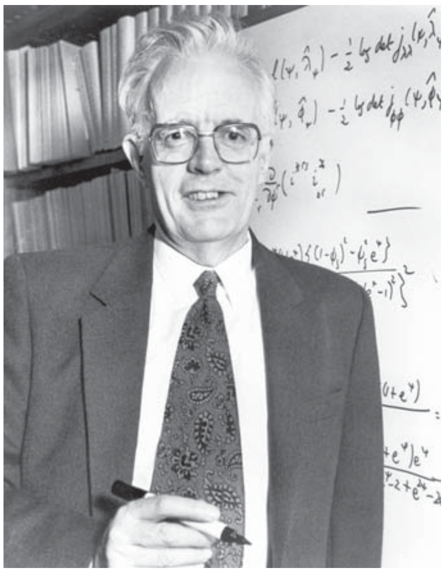
2. Actividad en clase Vea la actividad anterior y *diseñe un experimento* que sería efectivo para probar la afirmación de que alguien tiene la capacidad de identificar el color de una carta seleccionada de un paquete estándar de naipes. Describa el experimento con gran detalle. Debido a que el premio de $1,000,000 está en juego, queremos tener cuidado de evitar el grave error de concluir que la persona tiene un poder paranormal cuando tal poder no existe en realidad. Probablemente habrá alguna posibilidad de que el sujeto pueda hacer conjeturas aleatorias y adivinar en cada ocasión, así que identifique una probabilidad que sea razonable para el evento de que el sujeto pase la prueba con suposiciones al azar. Asegúrese de que la prueba esté diseñada para que esta probabilidad sea igual o menor que el valor de probabilidad seleccionado.

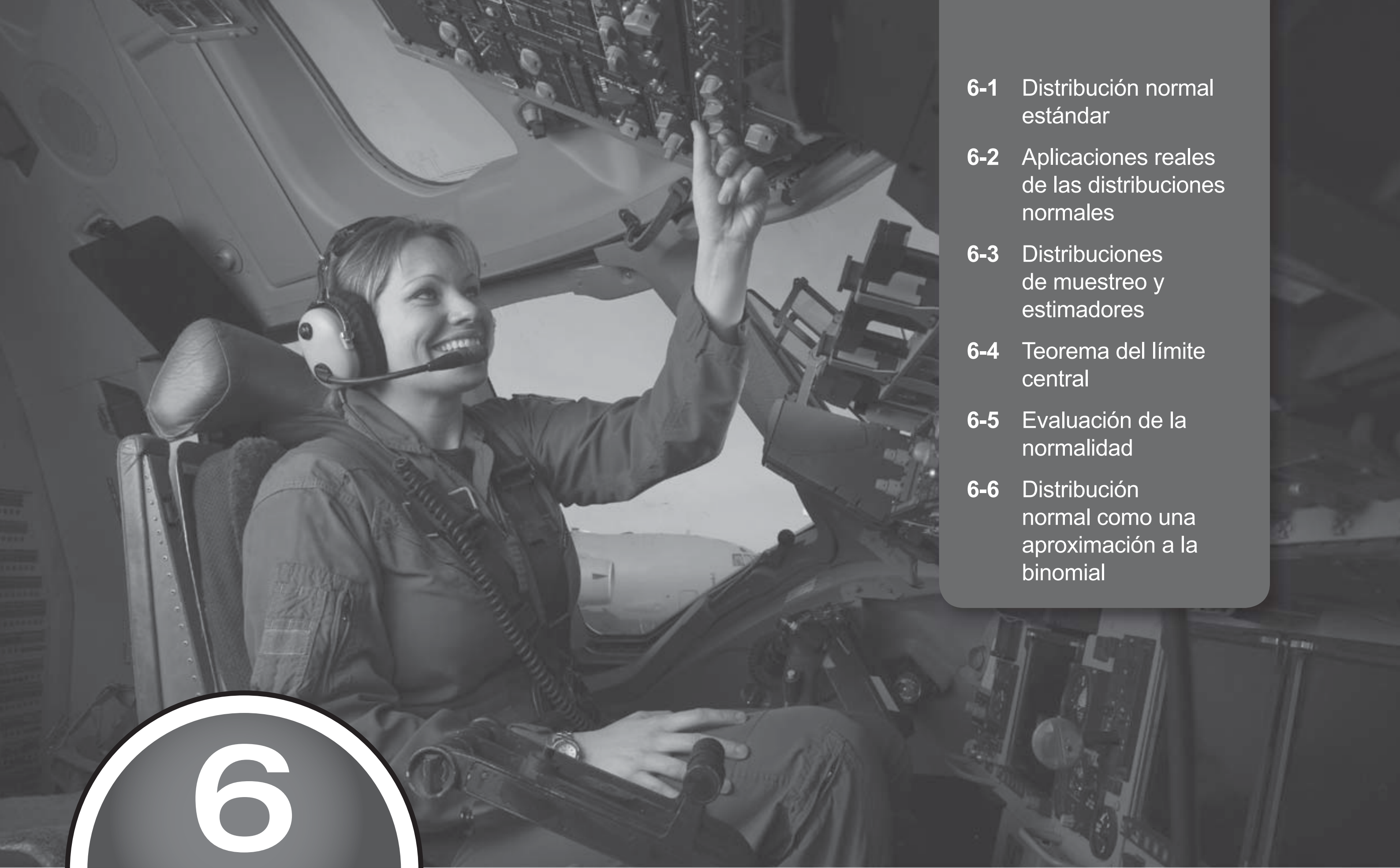
3. Actividad en clase Suponga que queremos identificar la distribución de probabilidad del número de niños en familias con al menos un hijo. Para cada estudiante en la clase, encuentre el número de hermanos y hermanas y registre el número total de niños (incluyendo el estudiante) en cada familia. Construya la tabla de frecuencias relativas para el resultado obtenido. (Los valores de la variable aleatoria *x* serán 1, 2, 3, . . .). ¿Qué hay de erróneo en usar esta tabla de frecuencias relativas como una estimación de la distribución de probabilidad para el número de niños en familias seleccionadas al azar?

4. Actividad fuera de clase En ocasiones, el análisis de los últimos dígitos de los datos puede revelar si los datos han sido recolectados a través de mediciones reales o si han sido reportados por los sujetos. Consulte un almanaque en Internet y encuentre una colección de datos (como las longitudes de los ríos en el mundo), luego analice la distribución de los últimos dígitos para determinar si los valores se obtuvieron a través de mediciones reales.

5. Actividad fuera de la clase En el pasado, los dígitos principales (a la izquierda) de los importes en los cheques se han analizado por fraude. Para los cheques que no involucran fraude, el dígito principal de 1 se espera en alrededor de 30.1% de los cheques. Obtenga una muestra aleatoria de cantidades reales de cheques y anote los dígitos iniciales. Compare el número real de cheques con cantidades que tienen un dígito inicial de 1 y la proporción de 30.1% esperada. ¿Los cheques reales se ajustan a la proporción esperada, o hay una discrepancia sustancial? Explique.

6. Actividad fuera de clase Las fotos que se muestran a continuación muestran a famosos estadísticos con nombres de David y John, no necesariamente en ese orden. Haga una encuesta con la siguiente pregunta: "¿Quién se llama David y quién se llama John?" ¿Los encuestados parecen dar resultados significativamente diferentes de lo que se espera con suposiciones aleatorias? (Vea "Who Do You Look Like? Evidence of Facial Stereotypes for Male Names", de Lea, Thomas, Lamkin y Bell, *Psychonomic Bulletin & Review*, vol. 14, artículo 5.





6 DISTRIBUCIONES DE PROBABILIDAD NORMAL

PROBLEMA DEL CAPÍTULO

Ergonomía 101

Este capítulo introduce herramientas estadísticas con muchas aplicaciones importantes del mundo real, incluyendo la ergonomía. La *ergonomía* es una disciplina centrada en el diseño de herramientas y equipos para que puedan utilizarse de forma segura, cómoda y eficiente. A continuación se muestran ejemplos de problemas ergonómicos:

- La sección 4.4.2 de la Ley para Estadounidenses con Discapacidades se refiere a los espacios libres verticales con la siguiente declaración: “Los pasillos, pasadizos, corredores, pasajes, u otros espacios de circulación tendrán un espacio libre mínimo de 80 pulgadas (2030 mm)”. ¿Qué porcentaje de adultos varones son más altos de 80 pulgadas?
- British Airways y muchas otras compañías aéreas tienen el requisito de que la tripulación en cabina debe tener alturas entre 62 y 73 pulgadas. ¿Qué porcentaje de mujeres adultas tienen una estatura menor a 62 pulgadas?

- El ejército de Estados Unidos requiere que las mujeres tengan entre 58 y 80 pulgadas de estatura. ¿Qué porcentaje de mujeres cumple este requisito?
- El ascensor en la instalación para la renta de autos en el aeropuerto de San Francisco tiene un cartel que indica una carga máxima de 4000 libras o 27 pasajeros. ¿Qué tan probable es que alrededor de 27 pasajeros superen la carga máxima de 4000 libras?

Con frecuencia los problemas ergonómicos implican cuestiones de seguridad extremadamente importantes. Estos son algunos casos reales que resultaron mortales:

- "Tenemos una emergencia en el vuelo 54-80 de Air Midwest", dijo la piloto Katie Leslie, justo antes de que su avión Beech se estrellara en Charlotte, Carolina del Norte, con el resultado de la muerte de los 21 tripulantes y pasajeros. El presunto exceso de peso de los pasajeros en su conjunto fue un factor que contribuyó al accidente.
- Después de que 20 pasajeros perecieron cuando el bote de paseo *Ethan Allen* se volcó en el lago George de Nueva York, una investigación mostró que aunque el número de pasajeros era inferior al máximo permitido, la embarcación debió haber sido certificada para un número mucho menor de pasajeros.
- Un taxi acuático se hundió en el Muelle Interior de Baltimore, matando a cinco de las 25 personas a bordo. El bote estaba certificado para transportar a 25 pasajeros, pero su peso total superaba la carga segura de 3500 libras, por lo que el número de pasajeros debió haberse limitado a 20.

OBJETIVOS DEL CAPÍTULO

El capítulo 5 introdujo distribuciones de probabilidad *discretas*; en este capítulo incluimos distribuciones de probabilidad *continuas*, aunque la mayor parte se centra en las distribuciones *normales*. Los objetivos del capítulo son:

6-1 La distribución normal estándar

- Describir las características de una distribución normal estándar.
- Encontrar la probabilidad de un rango de valores z en una distribución normal estándar.
- Determinar las puntuaciones z correspondientes a las regiones bajo la curva que representa una distribución normal estándar.

6-2 Aplicaciones reales de distribuciones normales

- Desarrollar la capacidad de describir una distribución normal (no necesariamente una distribución normal estándar).
- Encontrar la probabilidad de algún rango de valores en una distribución normal.
- Determinar las puntuaciones x correspondientes a las regiones bajo la curva que representa una distribución normal.

6-3 Distribuciones de muestreo y estimadores

- Desarrollar la capacidad de describir una *distribución muestral de un estadístico*.
- Determinar si un estadístico sirve como un buen estimador del parámetro de población correspondiente.

6-4 Teorema del límite central

- Describir lo que establece el teorema del límite central.
- Aplicar el teorema del límite central determinando la probabilidad de que la media de la muestra se encuentre dentro de un rango específico de valores.
- Identificar las condiciones para las cuales es apropiado utilizar una distribución normal para la distribución de las medias muestrales.

>>>

6-5 Evaluación de la normalidad

- Desarrollar la capacidad de examinar histogramas, valores atípicos y gráficas cuantilares normales para determinar si los datos muestrales parecen ser de una población que tiene una distribución que es aproximadamente normal.

6-6 Distribución normal como aproximación a una distribución binomial

- Identificar las condiciones para las cuales es apropiado utilizar una distribución normal como una aproximación a una distribución de probabilidad binomial.
- Utilizar la distribución normal para aproximar probabilidades de una distribución binomial.

6-1 Distribución normal estándar

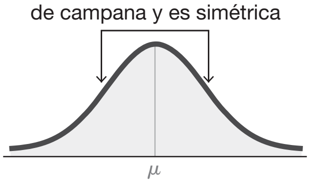


FIGURA 6-1 La distribución normal

Concepto clave En esta sección presentamos la *distribución normal estándar*, que es una distribución normal específica con tres propiedades:

1. Forma de campana: La gráfica de la distribución normal estándar tiene forma de campana (como en la figura 6-1).
2. $\mu = 0$: La distribución normal estándar tiene una media igual a 0.
3. $\sigma = 1$: La distribución normal estándar tiene una desviación estándar igual a 1.

En esta sección desarrollamos la habilidad de encontrar áreas (o probabilidades o frecuencias relativas) correspondientes a diferentes regiones bajo la gráfica de la distribución normal estándar. Además, determinamos puntuaciones z que corresponden a áreas bajo la curva. Tales habilidades se vuelven importantes en la siguiente sección, donde estudiaremos las distribuciones normales no estándar, las aplicaciones reales y su importancia.

Distribuciones normales

Hay una cantidad infinita de distribuciones normales diferentes, dependiendo de los valores utilizados para la media y la desviación estándar. Comenzaremos con una breve introducción a la familia general de distribuciones normales.

DEFINICIÓN

Si una variable aleatoria continua tiene una distribución con gráfica simétrica y forma de campana, como en la figura 6-1, y puede describirse mediante la ecuación dada en la fórmula 6-1, se dice que tiene una **distribución normal**.

FÓRMULA 6-1

$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Por fortuna, no será necesario usar la fórmula 6-1, aunque hay que destacar que el lado derecho de la ecuación revela que cualquier distribución normal particular está determinada por dos parámetros: la media poblacional, μ, y la desviación estándar de la población, σ. (En la fórmula 6-1, x es una variable que puede cambiar, $\pi = 3.14159\ldots$ y $e = 2.71828\ldots$). Una vez que se seleccionan los valores específicos para μ y σ, la fórmula 6-1 es una ecuación que relaciona a x y a y , y podemos graficar esa ecuación para obtener un resultado que se verá como en la figura 6-1. ¡Eso es todo lo que necesitamos saber sobre la fórmula 6-1!

Distribuciones uniformes

Este capítulo se centra en el concepto de una distribución de probabilidad normal; comenzaremos con una *distribución uniforme* de modo que podamos revisar estas dos propiedades importantes:

1. El área bajo la gráfica de una distribución de probabilidad continua es igual a 1.
2. Existe una correspondencia entre área y probabilidad, por lo que es posible determinar *probabilidades* identificando las *áreas* correspondientes en la gráfica; esto se consigue por medio de la fórmula para el área de un rectángulo:

$$\text{Área} = \text{altura} \times \text{anchura}$$

DEFINICIÓN

Una variable aleatoria continua tiene una **distribución uniforme** si sus valores se distribuyen *equitativamente* en el rango de posibilidades. La gráfica de una distribución uniforme da como resultado una forma rectangular.

Curva de densidad La gráfica de cualquier distribución de probabilidad continua se denomina **curva de densidad** y cualquier curva de densidad debe satisfacer el requisito de que el área total bajo la curva es exactamente 1. Esto simplifica los problemas de probabilidad, por lo que la siguiente aseveración es realmente importante:

Debido a que el área total bajo cualquier curva de densidad es igual a 1, existe una correspondencia entre *área* y *probabilidad*.

EJEMPLO 1 Tiempos de espera en la seguridad aeroportuaria

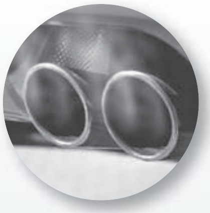
Durante ciertos periodos de tiempo en el aeropuerto JFK de la ciudad de Nueva York, los pasajeros que llegan al puesto de control de seguridad tienen tiempos de espera que están uniformemente distribuidos entre 0 y 5 minutos, como se ilustra en la figura 6-2 de la página siguiente.

Consulte la figura 6-2 y observe las siguientes propiedades:

- Todos los tiempos de espera posibles son *igualmente probables*.
- Los tiempos de espera pueden tener cualquier valor entre 0 y 5 minutos, por lo que es posible tener un tiempo de espera de 1.234567 minutos.
- Al asignar la probabilidad de 0.2 a la altura de la línea vertical en la figura 6-2, el *área cerrada es exactamente 1*. (En general, se debe hacer que la altura de la línea vertical en una distribución uniforme sea igual a 1/rango).

continúa

El poder de las muestras pequeñas



La Agencia de Protección Ambiental de Estados Unidos (EPA, por sus siglas en inglés) descubrió que los automóviles de Chrysler tenían un mal funcionamiento en sus carburadores, lo que ocasionaba que las emisiones de monóxido de carbono fueran demasiado altas. Los automóviles involucrados tenían desplazamientos de 360 y 400 pulgadas cúbicas y carburadores de dos barriles. La EPA ordenó a Chrysler solucionar el problema, pero ante la negativa de la compañía, prosiguió el caso de *Chrysler Corporation contra la Agencia de Protección Ambiental*. El desahogo del caso referido condujo a la conclusión de que había "pruebas sustanciales" de que esta marca de autos producía niveles excesivos de monóxido de carbono. La EPA ganó el caso y Chrysler se vio obligada a retirar y reparar 208,000 vehículos. Al analizar este caso en un artículo de *AMSTAT News*, el estadístico en jefe de la EPA, Barry Nussbaum, escribió: "El muestreo es costoso y el muestreo ambiental suele ser muy costoso. En la EPA, tenemos que hacer lo mejor que podamos con muestras pequeñas o desarrollar modelos... ¿Cuál fue el tamaño de muestra requerido para llegar a tal recogida de autos (los 208,000 Chryslers)? La respuesta es sólo 10. Esto representa una afirmación del poder de la estadística inferencial, pero también un reto para explicar cómo una muestra pequeña puede resultar suficiente".

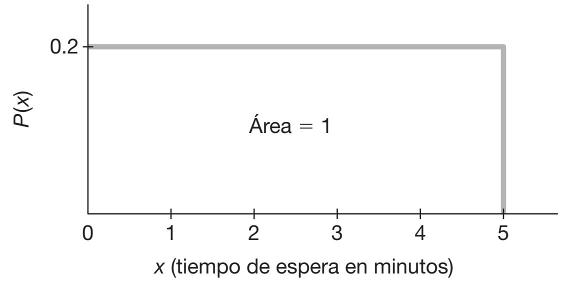


FIGURA 6-2 Distribución uniforme del tiempo de espera

SU TURNO Resuelva el ejercicio 5 "Distribución uniforme continua".

EJEMPLO 2 Tiempos de espera en la seguridad aeroportuaria

Dada la distribución uniforme ilustrada en la figura 6-2, determine la probabilidad de que un pasajero seleccionado al azar tenga un tiempo de espera de al menos 2 minutos.

SOLUCIÓN

El área sombreada en la figura 6-3 representa los tiempos de espera de al menos 2 minutos. Debido a que el área total bajo la curva de densidad es igual a 1, existe una correspondencia entre el área y la probabilidad. Es posible encontrar fácilmente la *probabilidad* deseada usando las *áreas* de la siguiente manera:

$$\begin{aligned} P(\text{tiempo de espera de al menos 2 min}) &= \text{altura} \times \text{anchura del área sombreada} \\ &\quad\ \text{en la figura 6-3} \\ &= 0.2 \times 3 \\ &= 0.6 \end{aligned}$$

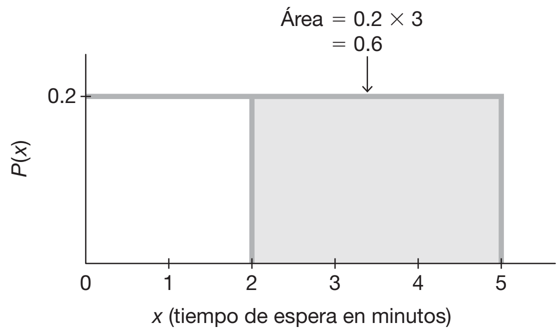


FIGURA 6-3 Uso del área para encontrar la probabilidad

INTERPRETACIÓN

La probabilidad de seleccionar aleatoriamente a un pasajero con un tiempo de espera de al menos 2 minutos es de 0.6.

SU TURNO Resuelva el ejercicio 7 "Distribución uniforme continua".

Distribución normal estándar

La curva de densidad de una distribución uniforme es una recta horizontal, por lo que podemos encontrar el área de cualquier región rectangular mediante la aplicación de la siguiente fórmula:

$$\text{Área} = \text{altura} \times \text{anchura}$$

Debido a que la curva de densidad de una distribución normal tiene una forma de campana más complicada, como se muestra en la figura 6-1, es más difícil encontrar áreas. Sin embargo, el principio básico es el mismo: *hay una correspondencia entre el área y la probabilidad*. En la figura 6-4 se muestra que para una distribución normal estándar, el área bajo la curva de densidad es igual a 1. En la figura 6-4, usamos la etiqueta "Puntuación z" para el eje horizontal, lo cual es común para la distribución normal estándar, definida como sigue.

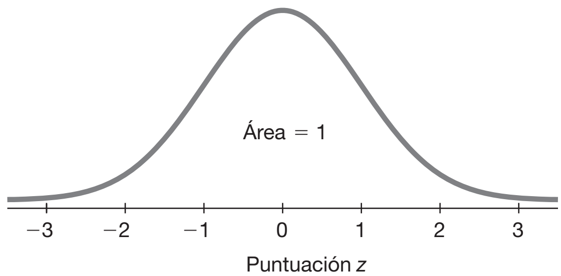


FIGURA 6-4 Distribución normal estándar

DEFINICIÓN

La **distribución normal estándar** es una distribución normal con los parámetros $\mu = 0$ y $\sigma = 1$. El área total bajo su curva de densidad es igual a 1 (como se muestra en la figura 6-4).

Determinación de probabilidades cuando se tienen puntuaciones *z*

No es fácil encontrar áreas en la figura 6-4, pero sí podemos encontrar áreas (o probabilidades) para muchas regiones diferentes usando la tecnología; también es posible utilizar la tabla A-2 (del apéndice A y de la inserción de *Fórmulas y tablas*). Las características clave de los diferentes métodos se resumen en la tabla 6-1. (StatCrunch ofrece opciones para una región izquierda acumulada, una región derecha acumulada o la región entre dos límites). Debido a que las calculadoras y el software generalmente dan resultados más precisos que la tabla A-2, se recomienda *enfáticamente* utilizar la tecnología. (Cuando haya discrepancias, las respuestas del apéndice D incluirán generalmente resultados basados en tecnología, así como las respuestas basadas en la tabla A-2).

Si se utiliza la tabla A-2, resulta esencial entender los siguientes puntos:

1. La tabla A-2 está diseñada sólo para la distribución normal *estándar*, que es una distribución normal con media 0 y desviación estándar 1.
2. La tabla A-2 abarca dos páginas: la página izquierda incluye las puntuaciones *z negativas* y la página derecha las puntuaciones *z positivas*.
3. Cada valor en el cuerpo de la tabla es un *área acumulada desde la izquierda* hasta un límite vertical sobre una puntuación z específica.
4. Cuando trabaje con una gráfica, evite caer en confusión entre las puntuaciones z y las áreas.

   Puntuación *z*: ***Distancia* a lo largo de la escala horizontal de la distribución normal estándar (correspondiente al número de desviaciones estándar por encima o por debajo de la media); consulte la columna extrema izquierda y la fila superior de la tabla A-2.**

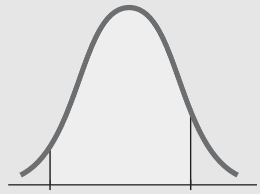
   Área: ***Región* bajo la curva; consulte los valores en el *cuerpo* de la tabla A-2.**

5. La parte de la puntuación z que denota centésimas se encuentra en la fila superior de la tabla A-2.

En cifras

134: Número de veces que la gente revisa diariamente sus teléfonos inteligentes, según una encuesta de Dignity Health a 2000 usuarios de teléfonos inteligentes.

TABLA 6-1 Formatos utilizados para determinar áreas de la distribución normal

Área acumulada desde la izquierda Las siguientes opciones proporcionan el *área acumulada desde la izquierda* hasta una línea vertical sobre un valor específico de *z*: • **Tabla A-2** • **Statdisk** • **Minitab** • **Excel** • **StatCrunch**	 **Región izquierda acumulada**
Área entre dos límites Las siguientes opciones proporcionan el área delimitada a la izquierda y delimitada a la derecha por líneas verticales sobre valores específicos. • **Calculadora TI-83/84 Plus** • **StatCrunch**	 **Área entre dos límites**

PRECAUCIÓN Cuando trabaje con una distribución normal, tenga cuidado de no confundir las puntuaciones *z* y las áreas.

Los siguientes ejemplos ilustran procedimientos que pueden usarse con aplicaciones reales y cruciales, como las que se presentan en las secciones siguientes.

EJEMPLO 3 Prueba de densidad ósea

Una prueba de densidad mineral ósea puede ser útil para identificar la presencia o la probabilidad de existencia de osteoporosis, enfermedad que causa que los huesos se vuelvan más frágiles y más propensos a romperse. El resultado de una prueba de densidad ósea se mide comúnmente como una puntuación z. La población de puntuaciones z se distribuye normalmente con una media de 0 y una desviación estándar de 1, por lo que tales resultados de las pruebas cumplen los requisitos de una distribución normal estándar, a la vez que la gráfica de los resultados de la prueba de densidad ósea es como se muestra en la figura 6-5.

Un adulto seleccionado al azar se somete a una prueba de densidad ósea. Determine la probabilidad de que esta persona tenga una puntuación en su prueba de densidad ósea inferior a 1.27.

SOLUCIÓN

Note que los siguientes valores son *iguales* (debido a la correspondencia entre probabilidad y área):

- *Probabilidad* de que la puntuación en la prueba de densidad ósea sea inferior a 1.27
- *Área sombreada* mostrada en la figura 6-5

Por lo tanto, necesitamos encontrar el área de la figura 6-5 por debajo de $z = 1.27$. Si usa la tecnología, consulte las instrucciones del centro de tecnología que se incluye al final de esta sección. Si usa la tabla A-2, comience con la puntuación $z = 1.27$ localizando 1.2 en la columna izquierda; después, encuentre el valor en la fila contigua de probabilidades que está directamente por debajo de 0.07, como se muestra en el extracto anexo. La tabla A-2 muestra que hay un área de 0.8980 correspondiente a $z = 1.27$. Queremos el área por *debajo* de 1.27, y la tabla A-2 da el área acumulada desde la izquierda, por lo que el área deseada es 0.8980. Debido a la correspondencia entre área y probabilidad, sabemos que la probabilidad de una puntuación z por debajo de 1.27 es 0.8980.

INTERPRETACIÓN

La *probabilidad* de que una persona seleccionada al azar tenga un resultado menor a 1.27 en la prueba de densidad ósea es 0.8980, la cual se muestra como la región sombreada de la figura 6-5. Otra forma de interpretar este resultado es concluir que 89.80% de las personas tienen niveles de densidad ósea por debajo de 1.27.

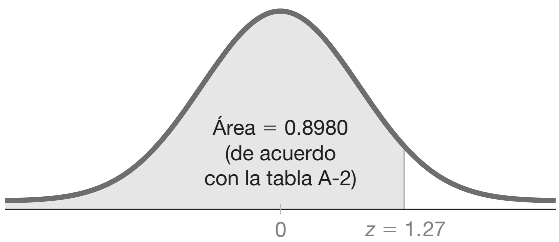


FIGURA 6-5 Determinación del área a la izquierda de z = 1.27

TABLA A-2 *(continuación)* Área acumulada desde la IZQUIERDA

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
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

SU TURNO Resuelva el ejercicio 9 “Distribución normal estándar”.

EJEMPLO 4 Prueba de densidad ósea: Determinación del área a la derecha de un valor

Utilice la misma prueba de densidad ósea del ejemplo 3 y encuentre la probabilidad de que una persona seleccionada al azar tenga un resultado por encima de −1.00. Se considera que un valor superior a −1.00 está en el rango “normal” de las lecturas de densidad ósea.

SOLUCIÓN

Encontramos nuevamente la *probabilidad* deseada determinando un *área* correspondiente. Buscamos el área de la región a la derecha de $z = -1.00$ que está sombreada en la figura 6-6 de la página siguiente. La pantalla de Statdisk muestra que el área a la derecha de $z = -1.00$ es 0.841345.

Si utilizamos la tabla A-2, debemos saber que está diseñada para aplicarse sólo a las áreas acumuladas de la *izquierda*. En relación con la página de las puntuaciones *z negativas*, encontramos que el área acumulada desde la izquierda hasta $z = -1{,}00$ es 0.1587, como se muestra en la figura 6-6. Debido a que el área total bajo la curva es 1, podemos

continúa

encontrar el área sombreada restando 0.1587 de 1. El resultado es 0.8413. Aunque la tabla A-2 está diseñada sólo para áreas acumuladas desde la izquierda, podemos usarla para encontrar áreas a la derecha, como se muestra en la figura 6-6.

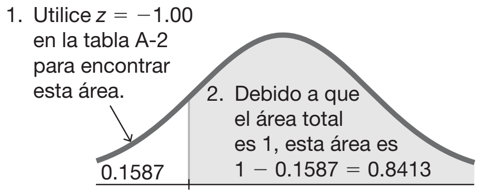
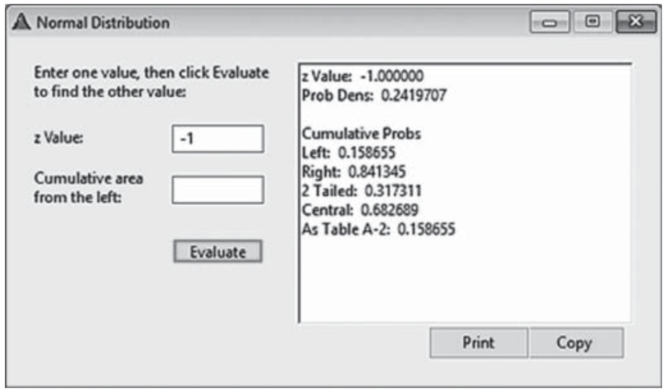


FIGURA 6-6 Determinación del área a la derecha de $z = -1$

Statdisk



INTERPRETACIÓN

Debido a la correspondencia entre probabilidad y área, concluimos que la *probabilidad* de seleccionar aleatoriamente a alguien con una densidad ósea por encima de -1 es 0.8413 (que es el *área* a la derecha de $z = -1.00$). También podríamos decir que 84.13% de las personas tienen niveles de densidad ósea por encima de -1.00.

El ejemplo 4 ilustra una manera en que la tabla A-2 puede usarse indirectamente para encontrar un área acumulada desde la derecha. El siguiente ejemplo ilustra otra forma de encontrar indirectamente un área usando la tabla A-2.

EJEMPLO 5 Prueba de densidad ósea: Determinación del área entre dos valores

Una prueba de densidad ósea entre -1.00 y -2.50 indica que el sujeto tiene osteopenia, que es una forma de pérdida ósea. Encuentre la probabilidad de que un sujeto seleccionado al azar tenga una lectura entre -1.00 y -2.50.

SOLUCIÓN

Estamos tratando de nuevo con valores normalmente distribuidos que tienen una media de 0 y una desviación estándar de 1. Los valores entre -1.00 y -2.50 corresponden a la región sombreada en la tercera gráfica de la figura 6-7. La tabla A-2 no puede usarse para encontrar esa área de manera directa, pero la podemos utilizar para determinar lo siguiente:

- El área a la izquierda de $z = -1.00$ es 0.1587.
- El área a la izquierda de $z = -2.50$ es 0.0062.

- El *área* entre $z = -2.50$ y $z = -1.00$ (el área sombreada en el extremo derecho de la figura 6-7) es la diferencia entre las áreas encontradas en los dos pasos anteriores:

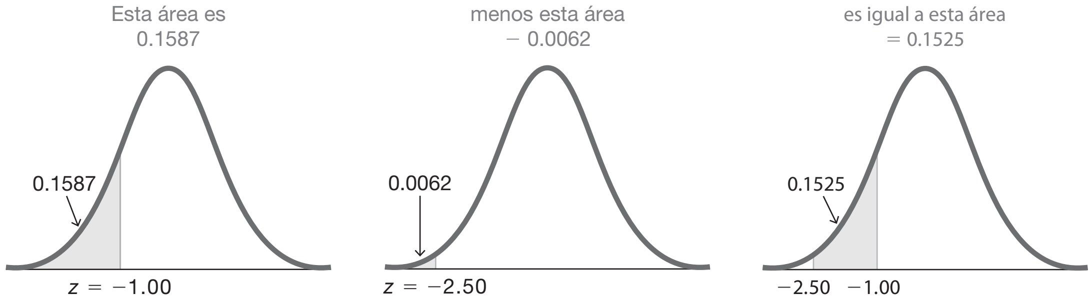


FIGURA 6-7 Determinación del área entre dos puntuaciones *z*

INTERPRETACIÓN

Si se utiliza la correspondencia entre probabilidad y área, concluimos que existe una probabilidad de 0.1525 de que un sujeto seleccionado al azar tenga una densidad ósea entre -1.00 y -2.50. Otra forma de interpretar este resultado es afirmar que 15.25% de las personas tienen osteopenia, con lecturas de densidad ósea entre -1.00 y -2.50.

SU TURNO Resuelva el ejercicio 11 "Distribución normal estándar".

El ejemplo 5 puede generalizarse como la siguiente regla:

El área correspondiente a la región *entre* dos puntuaciones *z* se puede encontrar determinando la diferencia entre las dos áreas que se encuentran en la tabla A-2.

La figura 6-8 ilustra esta regla general. Se puede encontrar la región sombreada *B* calculando la *diferencia* entre las dos áreas localizadas en la tabla A-2.

SUGERENCIA No trate de memorizar una regla o fórmula para este caso. Concéntrese en su *comprensión* usando una gráfica. Dibuje la gráfica, sombree el área deseada y luego piense creativamente en una manera de encontrar el área deseada trabajando con áreas acumuladas desde la izquierda.

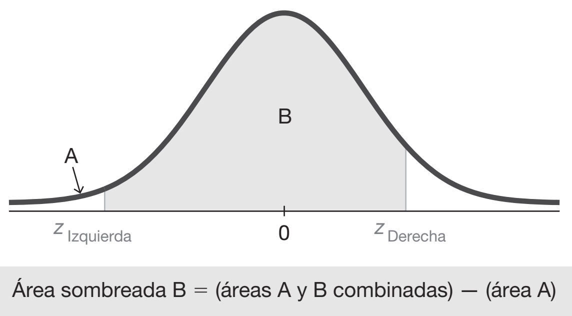


FIGURA 6-8 Determinación del área entre dos puntuaciones *z*

Las probabilidades como las de los ejemplos anteriores también se pueden expresar con la siguiente notación.

Notación

$P(a < z < b)$ denota la probabilidad de que la puntuación z esté entre a y b.

$P(z > a)$ denota la probabilidad de que la puntuación z sea mayor que a.

$P(z < a)$ denota la probabilidad de que la puntuación z sea menor que a.

Con esta notación, $P(-2.50 < z < -1.00) = 0.1525$, indica en símbolos que la probabilidad de una puntuación z que cae entre -2.50 y -1.00 es 0.1525 (como en el ejemplo 5).

Determinación de puntuaciones *z* a partir de áreas conocidas

Los ejemplos 3, 4 y 5 involucraron la distribución normal estándar, y todos fueron ejemplos con este mismo formato: dadas las puntuaciones z, encuentre áreas (o probabilidades). En muchos casos, necesitamos un método para invertir el formato: dada un área conocida (o probabilidad), encuentre la puntuación z correspondiente. En tales casos, es realmente importante evitar la confusión entre las puntuaciones z y las áreas. Recuerde que las puntuaciones z son *distancias* a lo largo de la escala horizontal, y que las áreas (o probabilidades) son regiones bajo la curva de densidad. (La tabla A-2 lista las puntuaciones z en la columna izquierda y en la fila superior, mientras que las áreas se encuentran en el *cuerpo* de la tabla). También debemos recordar que las puntuaciones z situadas en la mitad izquierda de la curva son siempre negativas. Si ya conocemos una probabilidad y queremos encontrar la puntuación z correspondiente, usamos el siguiente procedimiento.

Procedimiento para encontrar una puntuación *z* a partir de un área conocida

1. Dibuje una curva en forma de campana e identifique la región bajo la curva que corresponde a la probabilidad dada. Si esa región no es una región acumulada desde la izquierda, trabaje en su lugar con una región conocida que lo sea.
2. Use la tecnología o la tabla A-2 para encontrar la puntuación z. Con la tabla A-2, use el área acumulada desde la izquierda, localice la probabilidad más cercana en el *cuerpo* de la tabla e identifique la puntuación z correspondiente.

Casos especiales en la tabla A-2

Puntuación z	Área acumulada desde la izquierda
1.645	0.9500
-1.645	0.0500
2.575	0.9950
-2.575	0.0050
Arriba de 3.49	0.9999
Debajo de -3.49	0.0001

Casos especiales En la siguiente solución del ejemplo 6, la tabla A-2 conduce a una puntuación z de 1.645, que está a la mitad entre 1.64 y 1.65. Cuando se usa la tabla A-2, normalmente podemos evitar la interpolación con sólo seleccionar el valor más cercano. La tabla adjunta muestra los casos especiales generalmente utilizados en una amplia variedad de aplicaciones. (Para uno de esos casos especiales, el valor de $z = 2.576$ da un área ligeramente más cercana al área de 0.9950, pero $z = 2.575$ tiene la ventaja de ser el valor exactamente a la mitad entre $z = 2.57$ y $z = 2.58$). A excepción de estos casos especiales, normalmente podemos seleccionar el valor más cercano en la tabla. (Si un valor deseado está a la mitad entre dos valores de la tabla, seleccione el valor mayor). Para puntuaciones z por encima de 3.49, podemos usar 0.9999 como una aproximación del área acumulada desde la izquierda; para puntuaciones z por debajo de -3.49, podemos usar 0.0001 como una aproximación del área acumulada desde la izquierda.

EJEMPLO 6 **Prueba de densidad ósea: Determinación de una puntuación de prueba**

Utilice los mismos resultados de las pruebas de densidad ósea usados en los ejemplos anteriores. Estas puntuaciones se distribuyen normalmente con una media de 0 y una desviación estándar de 1, por lo que cumplen los requisitos de una distribución normal estándar. Encuentre la densidad ósea correspondiente a P_{95}, el percentil 95. Es decir, determine la puntuación de densidad ósea que separa el 95% inferior del 5% superior. Vea la figura 6-9.

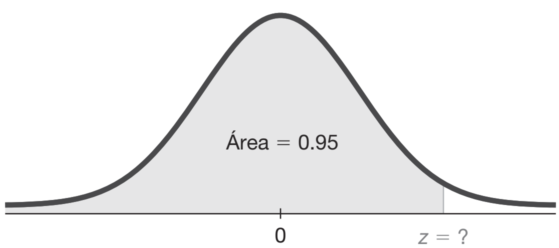


FIGURA 6-9 **Determinación del percentil 95**

SOLUCIÓN

La figura 6-9 muestra la puntuación z que es el percentil 95, con 95% del área (o 0.95) por debajo de ella.

Tecnología: También podemos encontrar la puntuación z usando la tecnología. La pantalla de Excel muestra que la puntuación z con un área de 0.95 a su izquierda es $z = 1.644853627$, o 1.645 al redondearla.

Excel

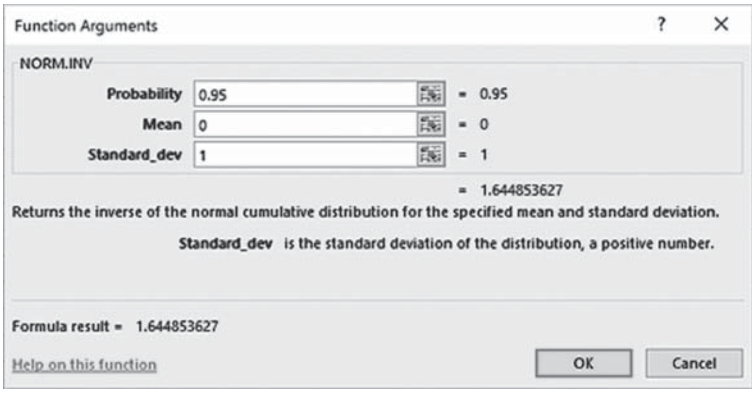


Tabla A-2: Si usa la tabla A-2, busque el área de 0.95 *en el cuerpo* de la tabla y luego encuentre la puntuación z correspondiente. En la tabla A-2 encontramos las áreas de 0.9495 y 0.9505, pero hay un asterisco con una nota especial que indica que 0.9500 corresponde a una puntuación z de 1.645. Ahora podemos concluir que la puntuación z en la figura 6-9 es 1.645, por lo que el percentil 95 es $z = 1.645$.

INTERPRETACIÓN

Para las puntuaciones de la prueba de densidad ósea, el 95% de las puntuaciones son menores o iguales a 1.645, y 5% de ellas son mayores o iguales a 1.645.

SU TURNO Resuelva el ejercicio 37 "Determinación de puntuaciones de densidad ósea".

EJEMPLO 7 Prueba de densidad ósea

Si utilizamos la misma prueba de densidad ósea descrita en el ejemplo 3, tenemos una distribución normal estándar con una media de 0 y una desviación estándar de 1. Encuentre la puntuación de la prueba de densidad ósea que separa el 2.5% inferior y determine la puntuación que separa el 2.5% superior.

SOLUCIÓN

Las puntuaciones z requeridas se muestran en la figura 6-10 de la página siguiente. Se pueden encontrar dichas puntuaciones z usando la tecnología. Si se utiliza la tabla A-2 para encontrar la puntuación z situada a la izquierda, se busca en el *cuerpo de la tabla* un área de 0.025. El resultado es $z = -1.96$. Para encontrar la puntuación z ubicada a la derecha, buscamos en el *cuerpo de* la tabla A-2 un área de 0.975. (Recuerde que la tabla A-2 siempre da áreas acumuladas desde la *izquierda*). El resultado es $z = 1.96$. Los valores de $z = -1.96$ y $z = 1.96$ separan respectivamente el 2.5% inferior y el 2.5% superior, como se muestra en la figura 6-10.

continúa

¿Los zurdos mueren más rápido?

Un estudio realizado por los psicólogos Diane Halpern y Stanley Coren atrajo la atención de los medios de comunicación y generó mucho interés al concluir que los zurdos no viven tanto tiempo como las personas diestras. Según su estudio, pareciera que los zurdos viven en promedio nueve años menos que los diestros. El estudio de Halpern/Coren ha sido criticado por usar datos imprecisos. Utilizaron datos de "segunda mano" al encuestar a familiares sobre personas que habían fallecido recientemente. El mito de que los zurdos mueren más jóvenes ha sobrevivido muchos años. Sin embargo, estudios más recientes muestran que las personas zurdas no tienen una vida más corta que las diestras.

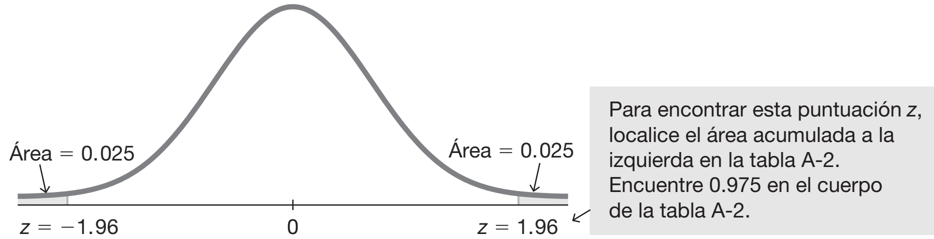


FIGURA 6-10 Determinación de puntuaciones z

INTERPRETACIÓN

Para la población de puntuaciones de la prueba de densidad ósea, el 2.5% de las puntuaciones son iguales o menores que -1.96, mientras que 2.5% de las puntuaciones son iguales o mayores que 1.96. Otra interpretación es que 95% de todos los resultados de la prueba de densidad ósea están entre -1.96 y 1.96.

SU TURNO Resuelva el ejercicio 39 “Determinación de puntuaciones de densidad ósea”.

Valores críticos Para una distribución normal, un *valor crítico* es una puntuación z en el límite que separa las puntuaciones z que son *significativamente bajas* o *significativamente altas*. Los valores críticos comunes son $z = -1.96$ y $z = 1.96$, y se obtienen como se mostró en el ejemplo 7, donde los valores de $z = -1.96$ o menos son significativamente bajos porque sólo el 2.5% de la población tiene puntuaciones iguales o inferiores a -1.96, mientras que los valores iguales o superiores a $z = 1.96$ son significativamente altos porque sólo el 2.5% de la población tiene puntuaciones iguales o superiores a 1.96. Los valores críticos serán muy importantes en los capítulos siguientes. La siguiente notación se utiliza para los valores críticos de z, los cuales se determinan empleando la distribución normal estándar.

DEFINICIÓN

Para la distribución normal estándar, un **valor crítico** es una puntuación z en el límite que separa las puntuaciones z que son *significativamente bajas* o *significativamente altas*.

Notación

La expresión z_α denota la puntuación z con un área de α a su derecha (α es la letra griega alfa).

EJEMPLO 8 Determinación del valor crítico z_α

Encuentre el valor de $z_{0.025}$. (Sea $\alpha = 0.025$ en la expresión z_α).

SOLUCIÓN

La notación de $z_{0.025}$ se utiliza para representar la puntuación z con un área de 0.025 a su *derecha*. Observe la figura 6-10 y note que el valor de $z = 1.96$ tiene un área de 0.025 a su derecha, por lo que $z_{0.025} = 1.96$. Tenga en cuenta que $z_{0.025}$ corresponde a un área acumulada de 0.975 desde la izquierda.

SU TURNO Resuelva el ejercicio 41 “Valores críticos”.

PRECAUCIÓN Cuando trate de determinar un valor de z_α para un valor particular de α, tenga en cuenta que α es el área a la *derecha* de z_α, pero la tabla A-2, y también algunas tecnologías, dan áreas acumuladas a la *izquierda* de determinada puntuación z. Para encontrar el valor de z_α, resuelva esta situación usando el valor de $1 - \alpha$. Por ejemplo, para encontrar $z_{0.1}$, refiérase a la puntuación z con un área de 0.9 a su izquierda.

Nueva tecnología, nuevos datos, nueva visión

Los residentes de la ciudad de Nueva York creían que los taxis escaseaban alrededor de la hora pico de la tarde. Sus quejas no podían atenderse porque no había datos que apoyaran la supuesta escasez. Sin embargo, se instalaron unidades de GPS en los taxis para que los directivos pudieran rastrear sus ubicaciones. Después de analizar los datos de GPS, se encontró que 20% o un porcentaje menor de taxis estaban en servicio entre las 4:00 pm y 5:00 pm que en la hora anterior. Las creencias subjetivas y las historias anecdóticas estaban ahora corroboradas con datos objetivos.

Se descubrió que dos factores eran responsables de la falta de taxis en la tarde. En primer lugar, los cambios de los turnos de 12 horas estaban programados a las 5:00 pm para que los conductores de los dos turnos tuvieran una participación igual en la hora pico. En segundo lugar, el aumento de las rentas en Manhattan obligó a muchas compañías de taxis a alojar sus taxis en Queens, por lo que los conductores tenían que regresar alrededor de las 4:00 pm para lograr llegar a tiempo y evitar multas por retardo. En los últimos años, la escasez de taxis se ha evitado con el crecimiento de empresas como Uber y Lyft.

Los ejemplos 3 a 7 de esta sección se basan en la aplicación real de la prueba de densidad ósea, con puntuaciones que se distribuyen normalmente con media de 0 y desviación estándar de 1, de manera que tales puntuaciones tienen una distribución normal estándar. Además de los resultados de la prueba de densidad ósea, es raro encontrar parámetros convenientes, debido a que las distribuciones normales típicas tienen medias diferentes a 0 y desviaciones estándar diferentes a 1. En la siguiente sección presentamos métodos para trabajar con tales distribuciones normales.

CENTRO TECNOLÓGICO

Determinación de puntuaciones *z*/áreas (normal estándar)

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Probability Distributions** en el menú desplegable y seleccione **Normal Distribution** en el submenú.
3. Introduzca la puntuación *z* deseada o el área acumulada desde la izquierda de la puntuación *z* y haga clic en **Evaluate**.

Minitab

1. Haga clic en **Calc** en el menú superior.
2. Seleccione **Probability Distributions** en el menú desplegable y seleccione **Normal** en el submenú.

Determinación del área acumulada a la izquierda de una puntuación *z*.

- Seleccione **Cumulative probability**, introduzca la media de **0** y la desviación estándar de **1**.
- Seleccione **Input Constant**, ingrese la puntuación *z* deseada y haga clic en **OK**.

Determinación de una puntuación *z* a partir de una probabilidad conocida

- Seleccione **Inverse cumulative probability**, introduzca la media de **0** y la desviación estándar de **1**.
- Seleccione **Input Constant**, ingrese el área total a la izquierda de la puntuación *z* y haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Calculators** en el menú desplegable y **Normal** en el submenú.
3. En la casilla de la calculadora ingrese la media de **0** y la desviación estándar de **1**.
4. Introduzca la puntuación *z* deseada (casilla central) o la probabilidad/área conocida (casilla más a la derecha). Seleccione la desigualdad deseada.
5. Haga clic en **Compute**

Calculadora TI-83/84 Plus

A diferencia de la mayoría de las otras tecnologías, la TI-83/84 Plus basa las áreas en la región entre dos puntuaciones z, en vez de hacerlo en la región acumulada desde la izquierda.

Determinación del área entre dos puntuaciones *z*

1. Pulse las teclas **2ND** y **VARS** para acceder al menú *DISTR* (distribuciones).
2. Seleccione **normalcdf** y pulse **ENTER**.
3. Introduzca la puntuación *z* inferior y la puntuación *z* superior deseadas. Introduzca **0** para μ y 1 para σ a fin de completar el comando **normalcdf(lower *z*,upper *z*,μ,σ)**. Pulse **ENTER**.

SUGERENCIA: Si no hay una puntuación *z* inferior, ingrese −999999; si no hay una puntuación *z* superior, ingrese 999999.

Determinación de una puntuación *z* a partir de una probabilidad conocida

1. Pulse las teclas **2ND** y **VARS** para acceder al menú *DISTR* (distribuciones).
2. Seleccione **invNorm** y pulse **ENTER**.
3. Introduzca el área a la izquierda de la puntuación *z*, **0** para μ y **1** para σ a fin de completar el comando **invNorm (area,μ,σ)**. Pulse **ENTER**.

Excel

Determinación del área acumulada a la izquierda de una puntuación *z*

1. Haga clic en **Insert Function f_x**, seleccione la categoría **Statistical**, seleccione la función **NORM.DIST** y haga clic en **OK**.
2. Para *x* ingrese la puntuación *z*, introduzca **0** para *Mean*, **1** para *Standard_dev* y **1** para *Cumulative*.
3. Haga clic en **OK**.

Determinación de puntuación *z* a partir de una probabilidad conocida

1. Haga clic en **Insert function f_x**, seleccione la categoría **Statistical** y la función **NORM.INV**.
2. Ingrese la probabilidad, introduzca **0** para *Mean* y **1** para *Standard_dev*.
3. Haga clic en **OK**.

6-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Distribución normal ¿Qué es erróneo en la afirmación siguiente? “Debido a que los dígitos 0, 1, 2, . . . , 9 son los resultados normales de los sorteos de lotería, tales números seleccionados al azar tienen una distribución normal”.

2. Distribución normal Una distribución normal se describe de manera informal como una distribución de probabilidad que tiene “forma de campana” cuando se grafica. Trace un bosquejo aproximado de una curva con la forma de campana característica de una distribución normal.

3. Distribución normal estándar Identifique los dos requisitos para que una distribución normal sea una distribución normal *estándar*.

4. Notación ¿Qué indica la notación z_α?

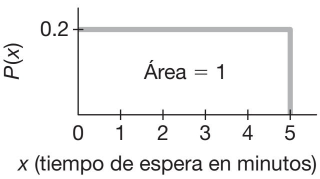


FIGURA 6-2

Distribución uniforme continua. ***En los ejercicios 5 a 8, consulte la distribución uniforme continua representada en la figura 6-2 y descrita en el ejemplo 1. Suponga que se selecciona un pasajero al azar y compruebe la probabilidad de que el tiempo de espera esté dentro del rango dado.***

5. Más de 3.00 minutos

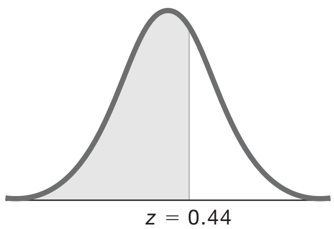
6. Menos de 4.00 minutos

7. Entre 2 minutos y 3 minutos

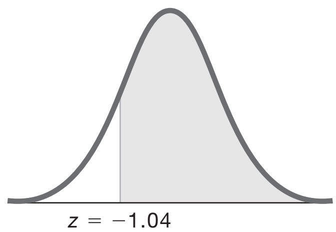
8. Entre 2.5 minutos y 4.5 minutos

Distribución normal estándar. ***En los ejercicios 9 a 12, determine el área de la región sombreada. La gráfica representa la distribución normal estándar de las puntuaciones de densidad ósea con media 0 y desviación estándar 1.***

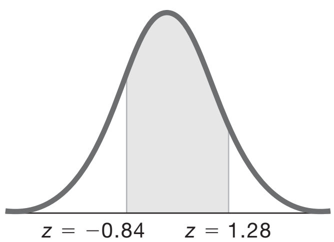
9.



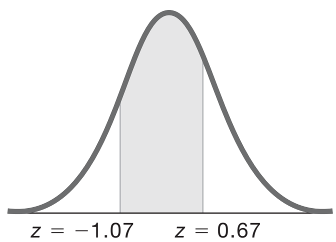
10.



11.

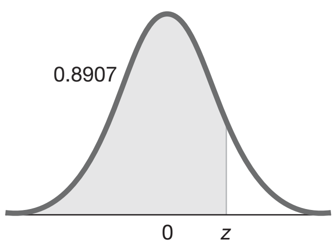


12.

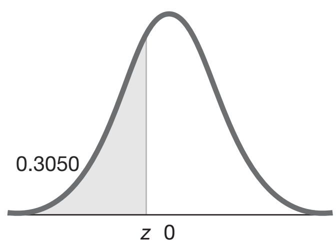


Distribución normal estándar. ***En los ejercicios 13 a 16, determine la puntuación z indicada. La gráfica representa la distribución normal estándar de las puntuaciones de densidad ósea con media 0 y desviación estándar 1.***

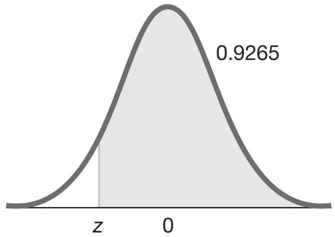
13.



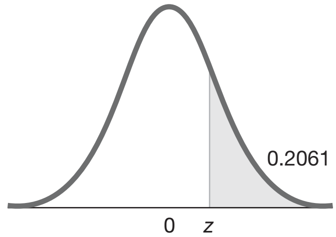
14.



15.



16.



Distribución normal estándar. ***En los ejercicios 17 a 36, suponga que se aplica una prueba de densidad ósea a un sujeto seleccionado al azar. Las puntuaciones de las pruebas se distribuyen normalmente con una media de 0 y una desviación estándar de 1. En cada caso, trace una gráfica y luego determine la probabilidad de las puntuaciones dadas de la prueba de densidad ósea. Si utiliza la tecnología en vez de la tabla A-2, redondee las respuestas a cuatro cifras decimales.***

17. Menor que −1.23

18. Menor que −1.96

19. Menor que 1.28

22. Menor que 2.56

21. Mayor que 0.25

22. Mayor que 0.18

23. Mayor que −2.00

24. Mayor que −3.05

25. Entre 2.00 y 3.00

26. Entre 1.50 y 2.50

27. Entre −2.55 y −2.00

28. Entre −2.75 y −0.75

29. Entre −2.00 y 2.00

30. Entre −3.00 y 3.00

31. Entre −1.00 y 5.00

32. Entre −4.27 y 2.34

33. Menor que 4.55

34. Mayor que −3.75

35. Mayor que 0

36. Menor que 0

Determinación de las puntuaciones de densidad ósea. ***En los ejercicios 37 a 40 suponga que a un sujeto seleccionado al azar se le aplica una prueba de densidad ósea. Las puntuaciones de la prueba de densidad ósea se distribuyen normalmente con una media de 0 y una desviación estándar de 1. En cada caso, trace una gráfica y después encuentre la puntuación de la prueba de densidad ósea correspondiente a la información dada. Redondee los resultados a dos decimales.***

37. Encuentre P_{99}, el percentil 99. Esta es la puntuación de densidad ósea que separa el 99% inferior del 1% superior.

38. Encuentre P_{10}, el décimo percentil. Esta es la puntuación de densidad ósea que separa el 10% inferior del 90% superior.

39. Si las puntuaciones de densidad ósea en el 2% inferior y el 2% superior se usan como puntos de corte para niveles que son demasiado bajos o demasiado altos, encuentre las dos lecturas que representan los valores de corte.

40. Encuentre las puntuaciones de densidad ósea que se pueden usar como valores de corte que separan el 3% más bajo y el 3% más alto.

Valores críticos. ***En los ejercicios 41 a 44, encuentre el valor crítico indicado. Redondee los resultados a dos decimales.***

41. $z_{0.10}$

42. $z_{0.02}$

43. $z_{0.04}$

44. $z_{0.15}$

Bases para la regla práctica del rango y la regla empírica. ***En los ejercicios 45 a 48, encuentre el área indicada bajo la curva de la distribución normal estándar; luego conviértala a un porcentaje y rellene el espacio en blanco. Los resultados forman la base de la regla práctica del rango y la regla empírica que se presentaron en la sección 3-2.***

45. Alrededor del ___% del área está entre $z = -1$ y $z = 1$ (o dentro de 1 desviación estándar de la media).

46. Alrededor del ___% del área está entre $z = -2$ y $z = 2$ (o dentro de 2 desviaciones estándar de la media).

47. Alrededor del ___% del área está entre $z = -3$ y $z = 3$ (o dentro de 3 desviaciones estándar de la media).

48. Alrededor del ___% del área está entre $z = -3.5$ y $z = 3.5$ (o dentro de 3.5 desviaciones estándar de la media).

6-1 Más allá de lo básico

49. Significancia. Para las puntuaciones de densidad ósea que se distribuyen normalmente con una media de 0 y una desviación estándar de 1, determine el *porcentaje* de puntuaciones que:

a. *son significativamente altas* (o que están al menos 2 desviaciones estándar por encima de la media).

b. *son significativamente bajas* (o que están al menos 2 desviaciones estándar por debajo de la media).

c. *no son significativas* (o que están a menos de 2 desviaciones estándar de la media).

50. Distribuciones. En una distribución uniforme continua,

$$\mu = \frac{\text{mínimo} + \text{máximo}}{2} \quad \text{y} \quad \sigma = \frac{\text{rango}}{\sqrt{12}}$$

a. Encuentre la media y la desviación estándar de la distribución de tiempos de espera representados en la figura 6-2, que acompaña a los ejercicios del 5 al 8.

b. Para una distribución uniforme continua con $\mu = 0$ y $\sigma = 1$, el mínimo es $-\sqrt{3}$ y el máximo es $\sqrt{3}$. Para esta distribución uniforme continua, determine la probabilidad de seleccionar aleatoriamente un valor entre -1 y 1 y compárela con el valor que se obtendría tratando incorrectamente la distribución como una distribución normal estándar. ¿La distribución afecta mucho los resultados?

6-2 Aplicaciones reales de las distribuciones normales

Concepto clave Ahora abordaremos temas reales mientras ampliamos los métodos de la sección anterior, con el fin de poder trabajar con cualquier *distribución normal no estándar* (con una media diferente de 0 y/o una desviación estándar distinta de 1). La clave es una conversión simple (fórmula 6-2) que nos permite "estandarizar" cualquier distribución normal de modo que los valores de x se puedan transformar en puntuaciones z; así, será posible utilizar los métodos de la sección anterior.

FÓRMULA 6-2

$$z = \frac{x - \mu}{\sigma} \quad \text{(puntuaciones } z \text{ redondeadas a 2 decimales)}$$

La figura 6-11 ilustra la conversión de una distribución normal no estándar a una normal estándar. El área en *cualquier* distribución normal limitada por una puntuación x (como en la figura 6-11a) es *igual* que el área delimitada por la puntuación z correspondiente en la distribución normal estándar (como en la figura 6-11b).

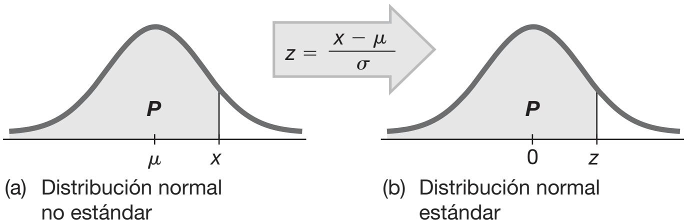


FIGURA 6-11 Conversión de distribuciones

En cifras

183,225,323,712: Cantidad de correos electrónicos enviados el día de hoy.

Algunas calculadoras y software no requieren el uso de la fórmula 6-2 para convertir a puntuaciones z porque las probabilidades se pueden encontrar directamente. Sin embargo, si se usa la tabla A-2, primero debemos convertir los valores en puntuaciones z estándar.

Cuando encuentre áreas con una distribución normal no estándar, use el siguiente procedimiento.

Procedimiento para encontrar áreas con una distribución normal no estándar

1. Trace una curva normal, etiquete la media y cualquier valor específico de x; después sombree la región que representa la probabilidad deseada.
2. Para cada valor relevante de x que sea un límite para la región sombreada, utilice la fórmula 6-2 a fin de convertir ese valor en la puntuación z equivalente. (En muchas tecnologías, puede omitirse este paso).
3. Use la tecnología (software o una calculadora) o la tabla A-2 para encontrar el área de la región sombreada. Esta área es la probabilidad deseada.

El siguiente ejemplo ilustra el procedimiento anterior.

EJEMPLO 1 **¿Qué proporción de hombres tiene una estatura mayor a las 72 pulgadas de altura requeridas para los cabezales de ducha (según la mayoría de los códigos de construcción)?**

Las alturas de los hombres se distribuyen normalmente con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas (de acuerdo con el conjunto de datos 1 "Datos corporales" en el apéndice B). Determine el porcentaje de hombres que son más altos que un cabezal de ducha a 72 pulgadas.

SOLUCIÓN

Paso 1: Vea la figura 6-12, que incorpora esta información: Los hombres tienen alturas que se distribuyen normalmente con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas. La región sombreada representa los hombres que son más altos que la altura del cabezal de una ducha a 72 pulg.

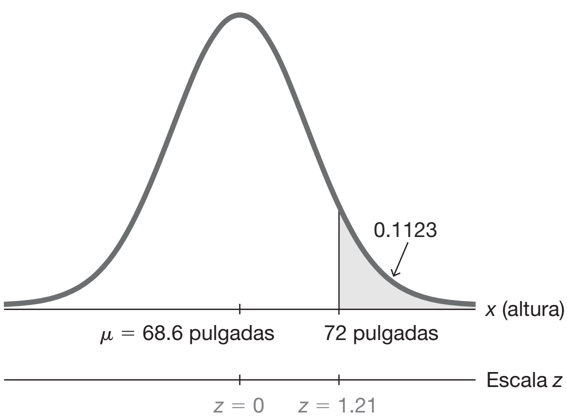


FIGURA 6-12 Alturas de los hombres

continúa

Paso 2: Podemos convertir la altura del cabezal de la ducha de 72 pulgadas a la puntuación z de 1.21, usando la fórmula 6-2 como sigue:

$$z = \frac{x - \mu}{\sigma} = \frac{72 - 68.6}{2.8} = 1.21 \text{ (redondeado a dos decimales)}$$

Paso 3: Tecnología: Se puede usar la tecnología para encontrar que el área redondeada a la derecha de 72 pulgadas en la figura 6-12 es 0.1123. (Con muchas tecnologías, se puede omitir el paso 2. Vea las instrucciones en el centro de tecnología al final de esta sección). El resultado de 0.1123 es más preciso que el resultado de 0.1131 encontrado mediante la tabla A-2.

Tabla A-2: Utilice la tabla A-2 para encontrar que el área acumulada a la *izquierda* de $z = 1.21$ es 0.8869. (Recuerde que la tabla A-2 está diseñada para que todas las áreas sean áreas acumuladas desde la *izquierda*). Debido a que el área total bajo la curva es 1, se sigue que el área sombreada en la figura 6-12 es $1 - 0.8869 = 0.1131$.

INTERPRETACIÓN

La proporción de hombres con mayor estatura que la altura del cabezal de la ducha a 72 pulgadas es 0.1123, o 11.23%. Alrededor del 11% de los hombres pueden asumir que el diseño es inadecuado. (*Nota:* Se ha sabido que algunos equipos de la NBA manejan duchas más bajas en los vestuarios de los equipos de baloncesto visitantes).

SU TURNO Resuelva el ejercicio 13 “Diseños de asiento”.

EJEMPLO 2 Requisito de estatura en la fuerza aérea

La Fuerza Aérea de Estados Unidos requiere que los pilotos tengan estaturas entre 64 y 77 pulgadas. Las alturas de las mujeres se distribuyen normalmente con una media de 63.7 pulgadas y una desviación estándar de 2.9 pulgadas (según el conjunto de datos 1 “Datos corporales” en el apéndice B). ¿Qué porcentaje de mujeres cumple el requisito de estatura?

SOLUCIÓN

En la figura 6-13 se muestra la región sombreada que representa las alturas de las mujeres entre 64 y 77 pulgadas.

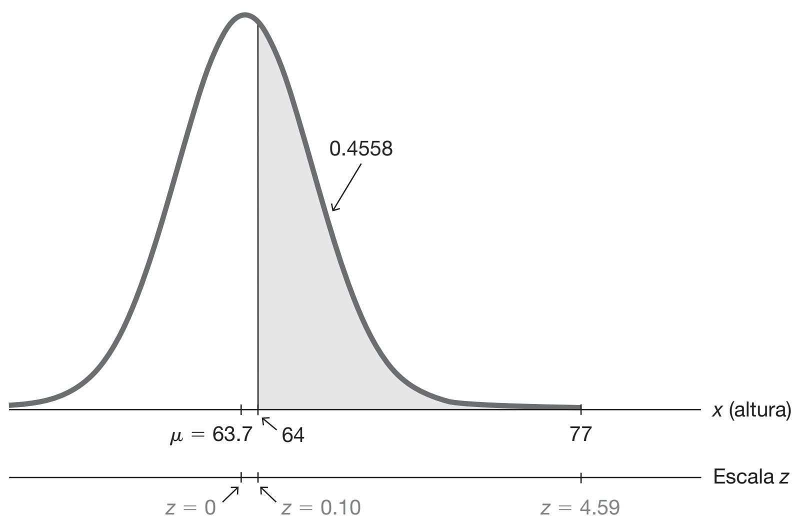


FIGURA 6-13 Alturas de las mujeres

Paso 1: Vea la figura 6-13, que incorpora esta información: Las mujeres tienen alturas que se distribuyen normalmente con una media de 63.7 pulgadas y una desviación estándar de 2.9 pulgadas. La región sombreada representa las mujeres con alturas entre 64 y 77 pulgadas.

Paso 2: Con algunas tecnologías, el área sombreada de la figura 6-13 puede encontrarse directamente y no es necesario convertir las puntuaciones x de 64 y 77 pulgadas en puntuaciones z (vea el paso 3).

Si usamos la tabla A-2, no podemos encontrar el área sombreada directamente, pero podemos determinarla de manera indirecta usando los mismos procedimientos de la sección 6-1, como sigue: (1) Encuentre el área acumulada desde la izquierda hasta 77 pulgadas (o $z = 4.59$); (2) encuentre el área acumulada desde la izquierda hasta 64 pulgadas (o $z = 0.10$); (3) encuentre la diferencia entre ambas áreas. Las alturas de 77 y 64 pulgadas se convierten en puntuaciones z usando la fórmula 6-2 como sigue:

$$\text{Para } x = 77 \text{ pulgadas: } z = \frac{x - \mu}{\sigma} = \frac{77 - 63.7}{2.9} = 4.59$$

($z = 4.59$ produce un área de 0.9999).

$$\text{Para } x = 64 \text{ pulgadas: } z = \frac{x - \mu}{\sigma} = \frac{64 - 63.7}{2.9} = 0.10$$

($z = 0.10$ produce un área de 0.5398).

Paso 3: Tecnología: Para usar la tecnología, consulte las instrucciones del centro de tecnología al final de esta sección. La tecnología mostrará que el área sombreada en la figura 6-13 es 0.4588.

Tabla A-2: Consulte la tabla A-2 con $z = 4.59$ y compruebe que el área acumulada a la *izquierda* de $z = 4.59$ es 0.9999. (Recuerde que la tabla A-2 está diseñada para que todas las áreas sean áreas acumuladas desde la *izquierda).* La tabla A-2 también muestra que $z = 0.10$ corresponde a un área de 0.5398. Debido a que las áreas de 0.9999 y 0.5398 son *áreas acumuladas desde la izquierda*, encontramos el área sombreada de la figura 6-13 como sigue:

$$\text{Área sombreada en la figura 6-13} = 0.9999 - 0.5398 = 0.4601$$

Existe una discrepancia relativamente pequeña entre el área de 0.4588 encontrada con la tecnología y el área de 0.4601 de la tabla A-2. El área obtenida con la tecnología es más precisa porque se basa en resultados no redondeados, mientras que la tabla A-2 requiere puntuaciones z redondeadas a dos decimales.

INTERPRETACIÓN

Al expresar el resultado como un porcentaje, concluimos que alrededor de 46% de las mujeres satisfacen el requisito de tener una altura entre 64 y 77 pulgadas. Alrededor del 54% de las mujeres no cumplen con ese requisito y no son elegibles para ser pilotos en la Fuerza Aérea de Estados Unidos.

SU TURNO Resuelva el ejercicio 15 “Diseños de asiento”.

Determinación de valores a partir de áreas conocidas

A continuación se proporcionan recomendaciones útiles para aquellos casos en que se conoce el área (o la probabilidad o el porcentaje) y debemos encontrar el(los) valor(es) relevante(s):

1. Las gráficas son extremadamente útiles para visualizar, entender y trabajar con éxito con las distribuciones de probabilidad normales, por lo que siempre deben utilizarse.
2. *No confunda las puntuaciones z y las áreas*. Recuerde que las puntuaciones z son *distancias* a lo largo de la escala horizontal y que las áreas son *regiones* bajo la curva normal. La tabla A-2 lista las puntuaciones z en la columna de la izquierda y en la fila superior, y las áreas se encuentran en el cuerpo de la tabla.

3. *Elija el lado correcto (derecho/izquierdo) de la gráfica*. Un valor que separa el 10% *superior* del resto estará situado en el lado derecho de la gráfica, pero un valor que separa el 10% *inferior* se ubicará en el lado izquierdo de la gráfica.
4. Una puntuación z debe ser *negativa* siempre que se encuentre en la mitad *izquierda* de la distribución normal.
5. Las áreas (o probabilidades) están siempre entre 0 y 1, y nunca son negativas.

Procedimiento para encontrar valores a partir de áreas o probabilidades conocidas

1. Trace una curva de distribución normal, escriba la probabilidad dada o el porcentaje en la región apropiada de la gráfica e identifique el valor x buscado.
2. Si utiliza tecnología, consulte las instrucciones del centro de tecnología al final de esta sección. Si usa la tabla A-2, consulte el *cuerpo* de la tabla A-2 para encontrar el área a la izquierda de x, después identifique la puntuación z correspondiente a esa área.
3. Si conoce z y debe convertir al valor x equivalente, utilice la fórmula 6-2 introduciendo los valores de μ, σ y la puntuación z encontrada en el paso 2 y, enseguida, determine x. Con base en la fórmula 6-2, es posible determinar x como sigue:

$$x = \mu + (\underset{\uparrow}{z} \cdot \sigma) \quad \text{(otra forma de la fórmula 6-2)}$$

(Si z se encuentra a la izquierda de la media, asegúrese de que sea un número negativo).

4. Consulte el bosquejo de la curva para verificar que la solución tenga sentido, en el contexto de la gráfica y del problema.

El siguiente ejemplo utiliza este procedimiento para encontrar un valor a partir de un área conocida.

EJEMPLO 3 Diseño de una cabina de avión

Al diseñar equipos, un criterio común es usar un diseño que se ajuste al 95% de la población. En el ejemplo 2 vimos que sólo el 46% de las mujeres satisface los requisitos de estatura para los pilotos de la Fuerza Aérea de Estados Unidos. ¿Cuál sería la altura máxima aceptable de una mujer si se cambiaran los requisitos para permitir que el 95% de las mujeres con menor estatura pudieran ser pilotos? Es decir, encuentre el percentil 95 de las alturas de las mujeres. Suponga que las estaturas de las mujeres se distribuyen normalmente con una media de 63.7 pulgadas y una desviación estándar de 2.9 pulgadas. Además de la estatura máxima permitida, ¿debería existir también una estatura mínima requerida? ¿Por qué?

SOLUCIÓN

Paso 1: La figura 6-14 muestra la distribución normal con la altura x que se desea identificar. El área sombreada representa el 95% de las mujeres con menor estatura.

Paso 2: Tecnología: La tecnología proporcionará el valor de x en la figura 6-14. Por ejemplo, vea la pantalla de Excel adjunta que muestra que $x = 68.47007552$ pulgadas o 68.5 pulgadas al redondear.

Excel

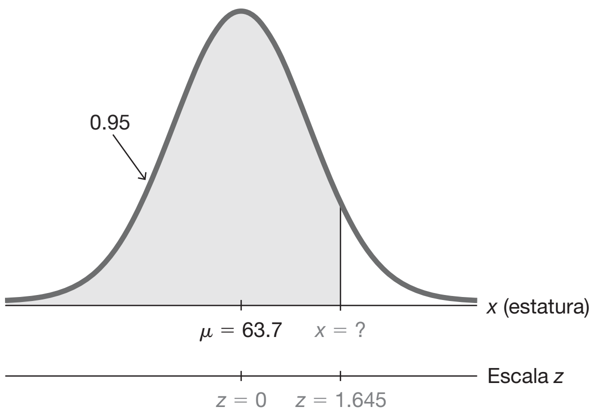


FIGURA 6-14 **Determinación del percentil 95**

Tabla A-2: Si usa la tabla A-2, busque un área de 0.9500 *en el cuerpo* de la tabla. (El área de 0.9500 que aparece en la figura 6-14 es un área acumulada desde la izquierda y es exactamente el tipo de área que se muestra en la tabla A-2). El área de 0.9500 está entre 0.9495 y 0.9505, pero hay un asterisco y nota al pie de página que indica que un área de 0.9500 corresponde a $z = 1.645$.

Paso 3: Con $z = 1.645$, $\mu = 63.7$ pulg. y $\sigma = 2.9$ pulg., podemos determinar x usando la fórmula 6-2:

$$z = \frac{x - \mu}{\sigma} \quad \text{se convierte en} \quad 1.645 = \frac{x - 63.7}{2.9}$$

El resultado de $x = 68.4705$ pulgadas se puede encontrar directamente o usando la siguiente versión de la fórmula 6-2:

$$x = \mu + (z \cdot \sigma) = 63.7 + (1.645 \cdot 2.9) = 68.4705 \text{ pulgadas}$$

Paso 4: La solución de $x = 68.5$ pulgadas (redondeadas) en la figura 6-14 es razonable porque es mayor que la media de 63.7 pulgadas.

INTERPRETACIÓN

Un requisito de una altura inferior a 68.5 pulgadas permitiría que 95% de las mujeres fueran elegibles como pilotos de la Fuerza Aérea de Estados Unidos. También debe haber un requisito de altura *mínima* para que el piloto pueda alcanzar fácilmente todos los controles.

SU TURNO Resuelva el ejercicio 17 “Diseños de asiento”.

Significancia

En el capítulo 4 vimos que las probabilidades se pueden usar para determinar si los valores son *significativamente altos* o *significativamente bajos*. El capítulo 4 se refiere a x éxitos entre n ensayos, pero podemos adaptar esos criterios para aplicarlos a variables continuas de la siguiente manera:

Significativamente alto: El valor de x es *significativamente alto* si $P(x \text{ o mayor}) \leq 0.05$.*

Significativamente bajo: El valor de x es *significativamente bajo* si $P(x \text{ o menos}) \leq 0.05$.*

*El valor de 0.05 no es absolutamente rígido, y en su lugar se pueden usar otros valores como 0.01.

EJEMPLO 4 Pulsos de mujeres significativamente bajos o significativamente altos

Utilice los criterios anteriores para identificar los pulsos de las mujeres que son significativamente bajos o significativamente altos. Con base en el conjunto de datos 1 “Datos corporales” en el apéndice B, suponga que las mujeres tienen pulsos distribuidos normalmente con una media de 74.0 latidos por minuto y una desviación estándar de 12.5 latidos por minuto.

SOLUCIÓN

Paso 1: Comenzamos con la gráfica mostrada en la figura 6-15. Hemos introducido la media de 74.0, y hemos identificado los valores de x que separan el 5% más bajo y el 5% más alto.

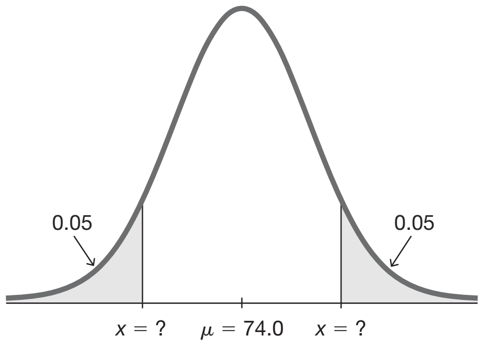


FIGURA 6-15 Pulsos de las mujeres

Paso 2: Tecnología: Para usar la tecnología, consulte las instrucciones del centro de tecnología al final de esta sección. La tecnología demostrará que los valores redondeados de x en la figura 6-15 son 53.4 y 94.6 latidos por minuto.

Tabla A-2: Si se utiliza la tabla A-2, debemos trabajar con áreas acumuladas desde la izquierda. Para el valor más a la izquierda de x, el área acumulada desde la izquierda es 0.05, así que buscamos un área de 0.05 *en el cuerpo* de la tabla para obtener $z = -1.645$ (identificado por el asterisco entre 0.0505 y 0.0495). Para el valor más a la derecha de x, el área acumulada desde la izquierda es 0.95, así que se busca un área de 0.9500 *en el cuerpo* de la tabla para obtener $z = 1.645$ (identificado por el asterisco entre 0.9495 y 0.9505). Una vez encontradas las dos puntuaciones z, procedemos a convertirlas a valores de pulso.

Paso 3: Ahora determinamos dos valores de x usando la fórmula 6-2 directamente, o usando la siguiente versión de la fórmula 6-2:

Valor más a la izquierda de x: $x = \mu + (z \cdot \sigma) = 74.0 + (-1.645 \cdot 12.5) = 53.4$

Valor más a la derecha de x: $x = \mu + (z \cdot \sigma) = 74.0 + (1.645 \cdot 12.5) = 94.6$

Paso 4: Con referencia a la figura 6-15, se observa que el valor más a la izquierda de $x =$ 53.4 es razonable porque es menor que la media de 74.0. Además, el valor más a la derecha de 94.6 es razonable porque está por encima de la media de 74.0.

INTERPRETACIÓN

Los pulsos significativos de las mujeres son:

- Significantemente bajo: 53.4 latidos por minuto o menos
- Significantemente alto: 94.6 latidos por minuto o más

Los médicos podrían usar estos resultados para investigar problemas de salud que pudieran causar que los pulsos sean significativamente bajos o significativamente altos.

SU TURNO Resuelva el ejercicio 19 “Significancia”.

CENTRO DE TECNOLOGÍA

Determinación de valores de *x* / áreas

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Probability Distributions** en el menú desplegable y seleccione **Normal Distribution** en el submenú.
3. Introduzca el valor *z* deseado o el área acumulada a la izquierda de la puntuación *z* y haga clic en **Evaluate**.

SUGERENCIA: Statdisk no funciona directamente con las distribuciones normales no estándar, así que utilice las puntuaciones *z* correspondientes.

Minitab

1. Haga clic en **Calc** en el menú superior.
2. Seleccione **Probability Distributions** en el menú desplegable y seleccione **Normal** en el submenú.

Determinación del área acumulada a la izquierda de un valor *x*

- Seleccione **Cumulative probability**, introduzca la media y la desviación estándar.
- Seleccione **Input Constant**, introduzca el valor *x* deseado y haga clic en **OK**.

Determinación del valor *x* a partir de una probabilidad conocida

- Seleccione **Inverse cumulative probability**, ingrese la media y la desviación estándar.
- Seleccione **Input Constant**, ingrese el área total a la izquierda del valor de *x* y haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Calculators** en el menú desplegable y **Normal** en el submenú.
3. En la casilla de la calculadora introduzca la media y la desviación estándar.
4. Introduzca el valor *x* deseado (casilla media) o la probabilidad (casilla en la extrema derecha).
5. Haga clic en **Compute**

Calculadora TI-83/84 Plus

A diferencia de la mayoría de las otras tecnologías, la TI-83/84 Plus basa las áreas en la región entre dos puntuaciones z, y no en las regiones acumuladas a la izquierda.

Determinación del área entre dos valores *x*

1. Pulse las teclas 2ND y VARS para acceder al menú *DISTR* (distribuciones).
2. Seleccione **normalcdf** y pulse ENTER.
3. Introduzca el valor *inferior* y el valor *superior* de *x* deseados. Introduzca la media (μ) y la desviación estándar (σ) para completar el comando **normalcdf(lower *x*,upper *x*,μ,σ)**. Pulse ENTER.

Sugerencia: Si no hay un valor inferior de *x*, ingrese −999999; si no existe un valor superior de *x*, introduzca 999999.

Determinación del valor *x* correspondiente a un área conocida

1. Pulse las teclas 2ND y VARS para acceder al menú *DISTR* (distribuciones).
2. Seleccione **invNorm** y presione ENTER.
3. Introduzca el área a la izquierda del valor *x*, ingrese la media (μ) y la desviación estándar (σ) para completar el comando **invNorm(area, μ, σ)**. Pulse ENTER.

Excel

Determinación del área acumulada a la izquierda de un valor *x*

1. Haga clic en **Insert Function f_x**, seleccione la categoría **Statistical**, seleccione la función **NORM.DIST** y haga clic en **OK**.
2. Para *x* ingrese el valor *x*, ingrese *Mean*, introduzca *Standard_dev* e ingrese 1 para *Cumulative*.
3. Haga clic en **OK**.

Determinación del valor *x* correspondiente a una probabilidad conocida

1. Haga clic en **Insert Function f_x**, seleccione la categoría **Statistical**, seleccione la función **NORM.INV** y haga clic en **OK**.
2. Introduzca la probabilidad o el área a la izquierda del valor de *x* deseado, ingrese *Mean*, e introduzca *Standard_dev*.
3. Haga clic en **OK**.

6-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Pesos al nacer Con base en el conjunto de datos 4 "Nacimientos" del apéndice B, los pesos al nacer se distribuyen normalmente con una media de 3152.0 g y una desviación estándar de 693.4 g.

a. ¿Cuáles son los valores de la media y la desviación estándar después de convertir todos los pesos al nacer a puntuaciones z usando $z = (x - \mu)/\sigma$?

b. Los pesos al nacer originales se dan en gramos. ¿Cuáles son las unidades de las puntuaciones z correspondientes?

2. Pesos al nacer Con base en el conjunto de datos 4 "Nacimientos" del apéndice B, los pesos al nacer se distribuyen normalmente con una media de 3152.0 g y una desviación estándar de 693.4 g.

a. Para la gráfica con forma de campana, ¿cuál es el área bajo la curva?

b. ¿Cuál es el valor de la mediana?

c. ¿Cuál es el valor de la moda?

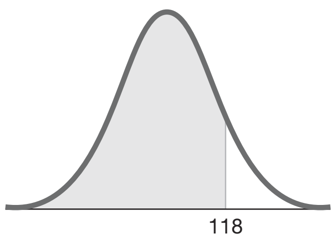
d. ¿Cuál es el valor de la varianza?

3. Distribuciones normales ¿Cuál es la diferencia entre una distribución normal estándar y una distribución normal no estándar?

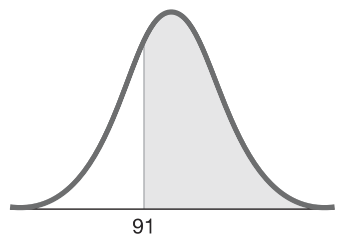
4. Dígitos aleatorios Por lo general, se utilizan computadoras para generar aleatoriamente los dígitos de los números de teléfono a los que se marcará como parte de la realización de una encuesta. ¿Se pueden usar los métodos de esta sección para encontrar la probabilidad de que cuando un dígito se genera aleatoriamente, sea menor a 3? ¿Por qué sí o por qué no? ¿Cuál es la probabilidad de obtener un dígito menor a 3?

Puntuaciones de IQ. ***En los ejercicios 5 a 8, encuentre el área de la región sombreada. Las gráficas representan las puntuaciones de IQ de los adultos, y esas puntuaciones se distribuyen normalmente con una media de 100 y una desviación estándar de 15 (como en la prueba Wechsler IQ).***

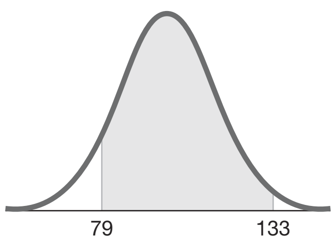
5.



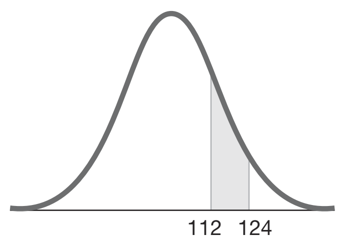
6.



7.

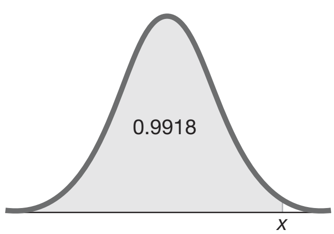


8.

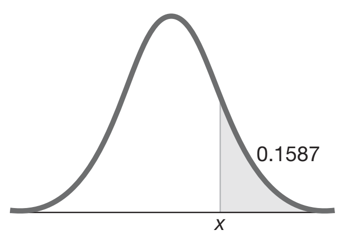


Puntuaciones de IQ. ***En los ejercicios 9 a 12, encuentre la puntuación de IQ indicada y redondee al número entero más cercano. Las gráficas representan las puntuaciones de IQ de los adultos, y esas puntuaciones se distribuyen normalmente con una media de 100 y una desviación estándar de 15 (como en la prueba Wechsler IQ).***

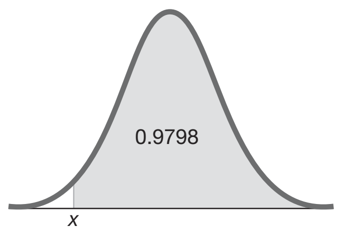
9.



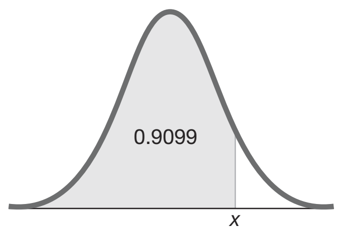
10.



11.



12.



Diseños de asientos. ***En los ejercicios 13 a 20, utilice los datos de la siguiente tabla para hombres y mujeres adultos sentados (según datos de encuestas antropométricas de Gordon, Churchill y otros). Estos datos se usan frecuentemente en el diseño de diferentes asientos, incluyendo asientos de avión, tren, teatro y asientos escolares. (Sugerencia: Trace una gráfica en cada caso).***

Longitud de la espalda a la rodilla (pulgadas) cuando la persona está sentada

	Media	Desv. Estd.	Distribución
Hombres	23.5 pulg.	1.1 pulg.	Normal
Mujeres	22.7 pulg.	1.0 pulg.	Normal

13. Encuentre la probabilidad de que un hombre tenga una longitud de espalda a rodilla menor a 21 pulgadas.

14. Encuentre la probabilidad de que una mujer tenga una longitud de espalda a rodilla mayor a 24.0 pulgadas.

15. Encuentre la probabilidad de que una mujer tenga una longitud de espalda a rodilla entre 22.0 y 24.0 pulgadas.

16. Encuentre la probabilidad de que un hombre tenga una longitud de espalda a rodilla entre 22.0 y 24.0 pulgadas.

17. Para los hombres, encuentre P_{90}, que es la longitud que separa el 90% inferior del 10% superior.

18. Para las mujeres, encuentre el primer cuartil Q_1, que es la longitud que separa el 25% inferior del 75% superior.

19. Significancia En lugar de utilizar 0.05 para identificar valores significativos, utilice el criterio de que un valor x es *significativamente alto* si $P(x$ o mayor$) \leq 0.01$ y un valor es *significativamente bajo* si $P(x$ o menor$) \leq 0.01$. Encuentre las longitudes de espalda a rodilla para los hombres, que separan los valores significativos de los que no son significativos. Utilizando estos criterios, ¿un hombre con una longitud de espalda a rodilla de 26 pulgadas es significativamente alto?

20. Significancia En lugar de usar 0.05 para identificar valores significativos, utilice el criterio de que un valor x es *significativamente alto* si $P(x$ o mayor$) \leq 0.025$ y un valor es *significativamente bajo* si $P(x$ o menor$) \leq 0.025$. Encuentre la longitud femenina de la espalda a la rodilla que separa los valores significativos de los que no son significativos. Usando estos criterios, ¿una mujer con una longitud de espalda a rodilla de 20 pulgadas es significativamente baja?

En los ejercicios 21 a 24, utilice los siguientes parámetros (de acuerdo con el conjunto de datos 1 "Datos corporales" en el apéndice B):

- ***Las estaturas de los hombres se distribuyen normalmente con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas.***
- ***Las estaturas de las mujeres se distribuyen normalmente con una media de 63.7 pulgadas y una desviación estándar de 2.9 pulgadas.***

21. Pilotos de la marina La Marina de Estados Unidos requiere que los pilotos de caza tengan estaturas entre 62 y 78 pulgadas.

a. Encuentre el porcentaje de mujeres que cumplen con los requisitos de estatura. ¿Las mujeres no calificadas son muchas porque son demasiado bajas o demasiado altas?

b. Si la Marina cambia los requisitos de estatura para que todas las mujeres sean elegibles excepto el 3% más bajo y el 3% más alto, ¿cuáles son los nuevos requisitos de altura para las mujeres?

22. Pilotos de la Fuerza Aérea La Fuerza Aérea de Estados Unidos requiere que los pilotos tengan estaturas entre 64 y 77 pulgadas.

a. Encuentre el porcentaje de hombres que cumplen con el requisito de estatura.

b. Si se cambian los requisitos de estatura de la Fuerza Aérea para excluir sólo el 2.5% más alto y el 2.5% más bajo de los hombres, ¿cuáles son los nuevos requisitos de altura?

23. Mickey Mouse Disney World requiere que las personas empleadas como el personaje de Mickey Mouse tengan una estatura entre 56 y 62 pulgadas.

a. Encuentre el porcentaje de hombres que cumplen con el requisito de estatura. ¿Qué sugiere el resultado sobre los géneros de las personas que son empleados como personajes de Mickey Mouse?

b. Si se cambian los requisitos de estatura para excluir al 50% más alto y al 5% más bajo de los hombres, ¿cuáles son los nuevos requisitos de altura?

24. Puerta ejecutiva de un jet El Gulfstream 100 es un jet ejecutivo de seis asientos, y tiene una altura de entrada de 51.6 pulgadas.

a. ¿Qué porcentaje de hombres adultos puede entrar a través de la puerta sin inclinarse?

b. ¿El diseño de la puerta con una altura de 51.6 pulgadas parece ser el adecuado? ¿Por qué los ingenieros no diseñaron una puerta más grande?

c. ¿Qué altura de la puerta permitiría que 40% de los hombres pudieran entrar sin inclinarse?

25. Contacto visual En un estudio de comportamiento facial, las personas en un grupo de control fueron cronometrados en cuanto al contacto visual en un periodo de 5 minutos. Los tiempos se distribuyen normalmente con una media de 184.0 segundos y una desviación estándar de 55.0 segundos (según datos del *Estudio Etológico del Comportamiento Facial en Pacientes Esquizofrénicos No Paranoicos y Paranoides* de Pittman, Olk, Orr y Singh, *Psychiatry*, vol. 144, núm. 1). Para una persona seleccionada aleatoriamente del grupo de control, determine la probabilidad de que el tiempo de contacto visual sea superior a 230.0 segundos, que es la media para los esquizofrénicos paranoicos. Con base en la experiencia personal, ¿el resultado parece ser la proporción de personas que son esquizofrénicos paranoicos?

26. Diseño de una estación de trabajo Un requisito de diseño común es que el entorno debe ajustarse al rango de personas que se encuentran entre el percentil 5 para las mujeres y el percentil 95 para los hombres. Al diseñar una mesa de trabajo de ensamble, debemos considerar la *altura de la rodilla sentada*, que es la distancia desde la parte inferior de los pies hasta la parte superior de la rodilla. Los hombres tienen alturas de rodilla sentada que se distribuyen normalmente con una media de 21.4 pulgadas y una desviación estándar de 1.2 pulgadas; las mujeres tienen alturas de rodilla sentada que se distribuyen normalmente con una media de 19.6 pulgadas y una desviación estándar de 1.1 pulgadas, según datos del Departamento de Transporte).

a. ¿Cuál es el espacio mínimo requerido para satisfacer el requisito de ajustarse al 95% de los hombres? ¿Por qué se ignora el percentil 95 para las mujeres en este caso?

b. El autor está escribiendo este ejercicio en una mesa con un espacio libre de 23.5 pulgadas por encima del piso. ¿Qué porcentaje de hombres se ajusta a esta mesa, y qué porcentaje de mujeres lo hace? ¿La mesa parece ser hecha para que pueda ser usada por la mayoría de las personas?

27. Asientos de eyección en jets La Fuerza Aérea de Estados Unidos utilizó en algún momento los asientos de eyección ACES-II diseñados para hombres que pesan entre 140 y 211 libras. Dado que los pesos de las mujeres normalmente se distribuyen con una media de 171.1 libras y una desviación estándar de 46.1 libras (según datos de la Encuesta Nacional de Salud), ¿qué porcentaje de mujeres tiene pesos que están dentro de esos límites? ¿Se excluyeron muchas mujeres con las especificaciones mencionadas?

28. Monedas de ¢25 Después de 1964, se fabrican monedas de ¢25 de manera que sus pesos tengan una media de 5.67 g y una desviación estándar de 0.06 g. Algunas máquinas expendedoras están diseñadas para que el peso de las monedas aceptables puedan ajustarse. Si se encuentran muchas monedas falsas, es posible reducir el rango de pesos aceptables con el efecto de que la mayoría de las monedas falsas sean rechazadas junto con algunas monedas legítimas.

a. Si usted ajusta sus máquinas expendedoras para aceptar pesos entre 5.60 g y 5.74 g, ¿qué porcentaje de las monedas de ¢25 legales se rechazan? ¿Este porcentaje es demasiado alto?

b. Si ajusta las máquinas expendedoras para aceptar todas las monedas de ¢25 legales excepto aquellas con pesos en el 2.5% superior y el 2.5% inferior, ¿cuáles son los límites de los pesos aceptables?

29. Peso bajo al nacer El Centro Médico de la Universidad de Maryland considera que los menores con "bajo peso al nacer" son aquellos que tienen 5.5 lb o 2495 g. Los pesos al nacer se distribuyen normalmente con una media de 3152.0 g y una desviación estándar de 693.4 g (de acuerdo con el conjunto de datos 1 "Nacimientos" en el apéndice B).

a. Si se selecciona al azar un peso al nacer, ¿cuál es la probabilidad de que sea un "peso bajo"?

b. Encuentre los pesos considerados como *significativamente bajos*, usando el criterio de un peso al nacer con probabilidad de 0.05 o menos.

c. Compare los resultados de los incisos (a) y (b).

30. Temperaturas corporales Con base en los resultados de la muestra del conjunto de datos 3 "Temperaturas corporales" en el apéndice B, suponga que las temperaturas del cuerpo humano se distribuyen normalmente con una media de 98.20 °F y una desviación estándar de 0.62 °F.

a. De acuerdo con emedicinehealth.com, una temperatura corporal de 100.4 °F o superior se considera una fiebre. ¿Qué porcentaje de las personas normales y sanas se considera que tiene fiebre? ¿Este porcentaje sugiere que un punto de corte de 100.4 °F es apropiado?

b. Los médicos desean seleccionar una temperatura mínima para solicitar más pruebas médicas. ¿Cuál debería ser esa temperatura, si queremos que sólo el 2.0% de las personas sanas lo superen? (Este resultado es un *falso positivo*, lo que significa que aunque el resultado de la prueba sea positivo, el sujeto no está realmente enfermo).

31. Duración de los embarazos La duración de los embarazos se distribuye normalmente con una media de 268 días y una desviación estándar de 15 días.

a. En una carta a "Querida Abby", una esposa afirmó haber dado a luz 308 días después de una breve visita de su marido, que estaba trabajando en otro país. Encuentre la probabilidad de un embarazo de 308 días o más. ¿Qué sugiere el resultado?

b. Si estipulamos que un bebé es *prematuro* si la duración del embarazo está en el 3% más bajo, encuentre la duración que separa a los bebés prematuros de aquellos que no lo son. Los bebés prematuros a menudo requieren cuidados especiales, y este resultado podría ser útil para los directores de un hospital en la planificación de esa atención.

32. Seguridad en taxis acuáticos Cuando un taxi acuático se hundió en el muelle interior de Baltimore, una investigación reveló que la carga segura de pasajeros para el taxi acuático era de 3500 libras. También se observó que el peso medio supuesto de un pasajero era de 140 libras. Suponga un escenario del "peor de los casos" en el que todos los pasajeros son hombres adultos. Suponga que los pesos de los hombres se distribuyen normalmente con una media de 188.6 libras y una desviación estándar de 38.9 libras (según el conjunto de datos 1 "Datos corporales" del apéndice B).

a. Si un hombre es seleccionado al azar, encuentre la probabilidad de que pese menos de 174 libras (que es el nuevo valor sugerido por el Consejo Nacional del Transporte y la Seguridad).

b. Con un límite de carga de 3500 libras, ¿cuántos pasajeros masculinos pueden permitirse si suponemos un peso medio de 140 libras?

c. Con un límite de carga de 3500 libras, ¿cuántos pasajeros masculinos pueden permitirse si asumimos el peso promedio actualizado de 188.6 libras?

d. ¿Por qué es necesario revisar y modificar periódicamente el número de pasajeros permitido?

Conjuntos de datos grandes. ***En los ejercicios 33 y 34, consulte los conjuntos de datos del apéndice B y use un software o una calculadora.***

33. Pulso de los hombres Consulte el conjunto de datos 1 "Datos corporales" en el apéndice B y utilice los pulsos de los hombres.

a. Determine la media y la desviación estándar, y verifique que los pulsos tienen una distribución más o menos normal.

b. Si se tratan los valores no redondeados de la media y la desviación estándar como parámetros, y se supone que los pulsos masculinos se distribuyen normalmente, encuentre el pulso que separa al 2.57% más bajo y el pulso que separa al 2.5% más alto. Estos valores podrían ser útiles cuando los médicos tratan de determinar si los pulsos son significativamente bajos o significativamente altos.

34. Ponderaciones de M&Ms Consulte el conjunto de datos 27 "Pesos de M&Ms" en el apéndice B y use los pesos (gramos) de todos los M&Ms listados.

a. Encuentre la media y la desviación estándar, y verifique que los datos tengan una distribución que sea más o menos normal.

b. Si se tratan los valores no redondeados de la media y la desviación estándar como parámetros, y se supone que los pesos están distribuidos normalmente, encuentre el peso que separa el 0.5% más bajo y el peso que separa el 0.5% más alto. Estos valores podrían ser útiles cuando los especialistas en control de calidad tratan de controlar el proceso de fabricación.

6-2 Más allá de lo básico

35. Curvas de calificaciones de exámenes Una profesora aplica un examen y las calificaciones se distribuyen normalmente con una media de 60 y una desviación estándar de 12. Ella planea aplicar curvas a las calificaciones.

a. Si aplica curvas añadiendo 15 a cada calificación, ¿cuáles son la nueva media y la desviación estándar?

b. ¿Es justo añadir 15 a cada calificación? ¿Por qué sí o por qué no?

c. Si se aplican curvas a las calificaciones para que se den calificaciones de B a las puntuaciones por encima del 70% inferior y por debajo del 10% superior, encuentre los límites numéricos para una calificación de B.

d. ¿Qué método de aplicación de curvas a las calificaciones es más justo: agregar 15 a cada calificación original o usar un esquema como el que se da en el inciso (c)? Explique.

36. Valores atípicos Con el propósito de construir bigotes y gráficas de caja modificados, como se describe en la sección 3-3, los valores atípicos se definen como valores de datos que están por encima de Q_3 en una cantidad mayor que 1.5 × IQR, o por debajo de Q_1 en una cantidad mayor que 1.5 × IQR, donde IQR es el rango intercuartil. Utilice esta definición de valores atípicos para determinar la probabilidad de que, cuando un valor se selecciona aleatoriamente de una distribución normal, sea un valor atípico.

6-3 Distribuciones muestrales y estimadores

Concepto clave Ahora consideraremos el concepto de una *distribución muestral de un estadístico*. En lugar de trabajar con valores de la población original, queremos centrarnos en los valores de los *estadísticos* (como las proporciones o las medias muestrales) obtenidos de la población. La figura 6-16 muestra los puntos clave que necesitamos conocer, así que trate de entender realmente la historia que narra la figura 6-16.

Una historia corta Entre la población de adultos, exactamente 70% no se siente cómodo en un vehículo auto-conducido (el autor sólo sabe esto). En una encuesta de TE Connectivity a 1000 adultos, 69% dijo que no se sentían cómodos en un vehículo auto-conducido. Fortalecidos por las visiones de multitudes de automóviles sin conductor, 50,000 personas se volvieron tan entusiastas que cada uno llevó a cabo su propia encuesta individual de 1000 adultos seleccionados al azar sobre el mismo tema. Cada uno de los 50,000 encuestadores novatos informó el porcentaje que encontraron, con resultados como el 68%, 72%, 70%. El autor obtuvo cada uno de los 50,000 porcentajes muestrales, los convirtió en proporciones y luego construyó el histograma que se muestra en la figura 6-17. ¿Observa algo sobre la *forma* del histograma? Es *normal* (a diferencia de los 50,000 topógrafos novatos). ¿Observa algo sobre la media de las proporciones muestrales? Se centran en el valor de 0.70, que es la proporción poblacional. Moraleja, cuando se toman muestras del mismo tamaño de la misma población, se aplican las siguientes dos propiedades:

1. Las proporciones muestrales tienden a distribuirse normalmente.
2. La media de las proporciones muestrales es igual que la media de la población.

Las implicaciones de las propiedades anteriores se extenderán en los capítulos subsecuentes. ¡Qué final tan feliz!

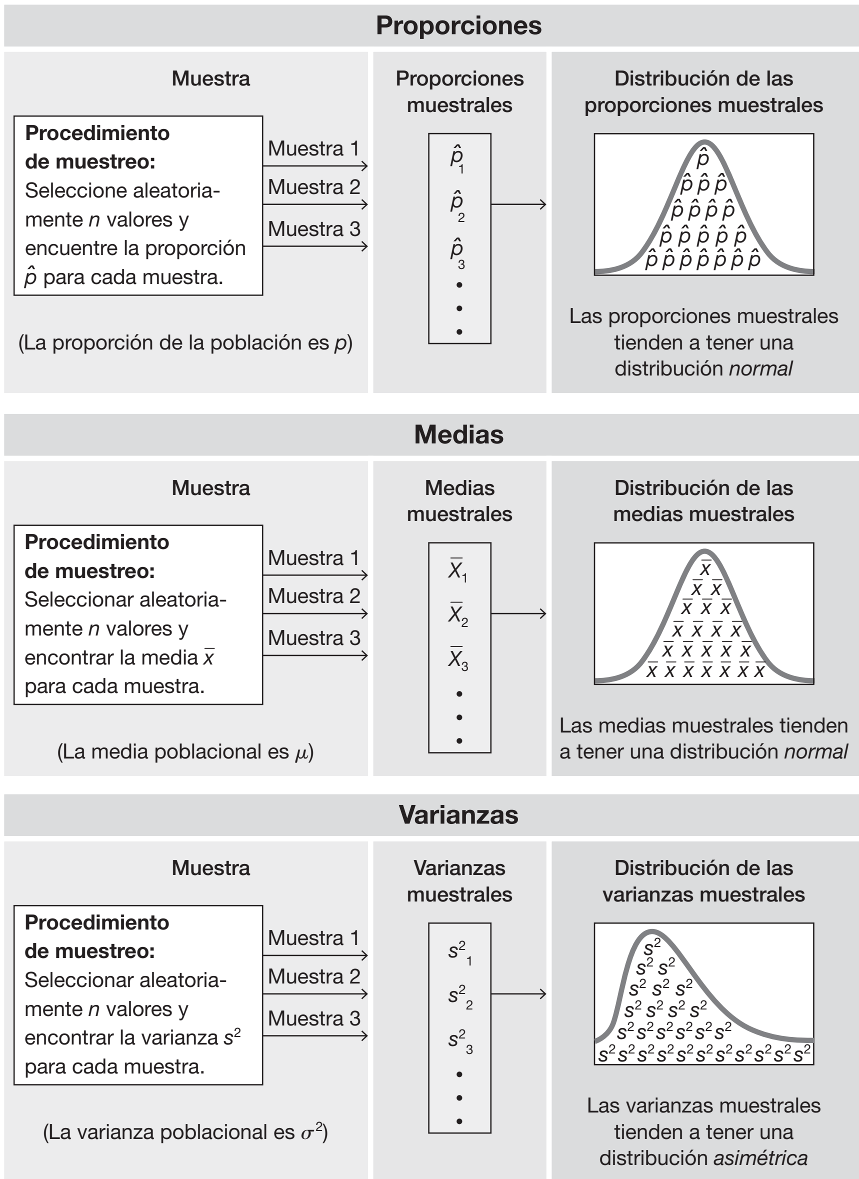


FIGURA 6-16 Comportamiento general de las distribuciones muestrales

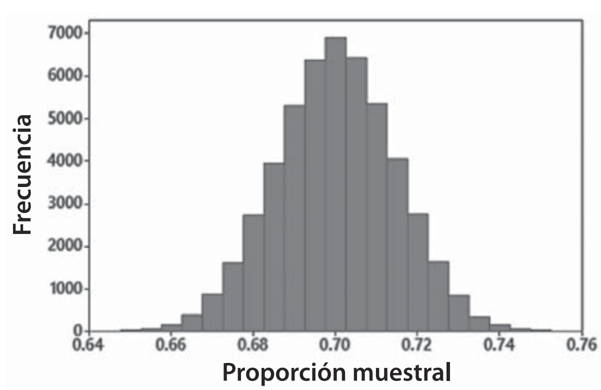


FIGURA 6-17 Histograma de 50,000 proporciones muestrales

Definamos formalmente la *distribución muestral*, el personaje principal del relato anterior.

DEFINICIÓN

La **distribución muestral de un estadístico** (por ejemplo una proporción muestral o una media muestral) es la distribución de todos los valores del estadístico cuando todas las muestras posibles del mismo tamaño n se toman de la misma población. (La distribución muestral de un estadístico se representa típicamente como una distribución de probabilidad en el formato de un histograma de probabilidad, una fórmula o una tabla).

Distribución muestral de la proporción muestral

La definición general anterior de la distribución muestral de un estadístico puede ahora reformularse para el caso específico de una proporción muestral:

DEFINICIÓN

La **distribución muestral de la proporción muestral** es la distribución de las proporciones muestrales (o la distribución de la variable $\hat{p}$), donde todas las muestras tienen el mismo tamaño muestral n tomado de la misma población. (La distribución muestral de la proporción muestral se representa típicamente como una distribución de probabilidad en el formato de un histograma de probabilidad, una fórmula o una tabla de probabilidad).

Es necesario distinguir entre una proporción poblacional p y una proporción muestral, por lo que la siguiente notación común se usará en el resto del libro.

Notación para proporciones

p = proporción *poblacional*

$\hat{p}$ = proporción *muestral*

SUGERENCIA $\hat{p}$ se pronuncia “p-gorro”. Cuando se utilizan símbolos encima de una letra; como en $\bar{x}$ y $\hat{p}$, éstas representan *estadísticos*, no parámetros.

Comportamiento de las proporciones muestrales

1. La distribución de las proporciones muestrales tiende a aproximarse a una distribución normal.
2. Las proporciones muestrales *se dirigen* al valor de la proporción poblacional en el sentido de que la media de todas las proporciones muestrales $\hat{p}$ es igual a la proporción poblacional p; el valor esperado de la proporción muestral es igual a la proporción de la población.

EJEMPLO 1 **Distribución muestral de la proporción muestral**

Considere la repetición del siguiente proceso: Lanzar un dado 5 veces y encontrar la proporción de números *impares* (1 o 3 o 5). ¿Qué sabemos sobre el comportamiento de todas las proporciones muestrales que se generan cuando este proceso continúa indefinidamente?

En cifras

90%: Porcentaje de los datos mundiales que se generaron en los últimos dos años (según *Science Daily*).

SOLUCIÓN

La figura 6-18 ilustra el proceso de lanzar un dado 5 veces y encontrar la proporción de números impares. (La figura 6-18 muestra los resultados de repetir este proceso 10,000 veces, pero la verdadera distribución muestral de la proporción muestral implica repetir el proceso indefinidamente). La figura 6-18 muestra que las proporciones de la muestra se distribuyen aproximadamente en forma normal. (Dado que los valores de 1, 2, 3, 4, 5, 6 son todos igualmente probables, la proporción de números impares en la población es de 0.5 y la figura 6-18 muestra que las proporciones muestrales tienen una media de 0.50).

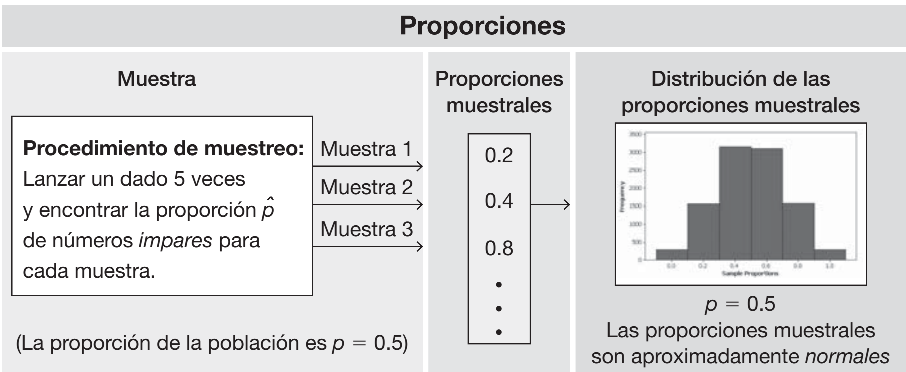


FIGURA 6-18 **Proporciones muestrales de 10,000 ensayos**

SU TURNO Resuelva el ejercicio 10 “Distribución muestral de la proporción muestral”.

Distribución muestral de la media muestral

Ahora vamos a considerar las medias muestrales.

DEFINICIÓN

La **distribución muestral de la media muestral** es la distribución de todas las medias muestrales posibles (o la distribución de la variable $\bar{x}$), donde todas las muestras tienen el mismo tamaño n tomado de la misma población. (La distribución muestral de la media muestral se representa típicamente como una distribución de probabilidad en el formato de un histograma de probabilidad, una fórmula o una tabla).

Comportamiento de las medias muestrales

1. La distribución de las medias muestrales tiende a ser una distribución normal. (Esto se analizará en la siguiente sección, pero la distribución tiende a acercarse a una distribución normal a medida que aumenta el tamaño de la muestra).
2. Las medias muestrales se *dirigen* al valor de la media de la población. (Es decir, la media de las medias muestrales es la media de la población). El valor esperado de la media muestra es igual a la media de la población.

EJEMPLO 2 Distribución muestral de la media muestral

Un amigo del autor tiene tres hijos con edades de 4, 5 y 9 años. Consideremos la población que consiste en {4, 5, 9}. (Por lo general, no conocemos todos los valores de una población y normalmente no trabajamos con una población tan pequeña, pero funciona bien para los propósitos de este ejemplo). Si se seleccionan dos edades al azar con reemplazo de la población {4, 5, 9}, identifique la distribución muestral de la media muestral mediante la creación de una tabla que represente la distribución de probabilidad de las medias muestrales. ¿Los valores de las medias muestrales se dirigen al valor de la media de la población?

SOLUCIÓN

Si se seleccionan dos valores aleatoriamente con reemplazo de la población {4, 5, 9}, la columna de la izquierda en la tabla 6-2 lista las nueve diferentes muestras posibles. La segunda columna lista las medias muestrales correspondientes. Las nueve muestras son igualmente probables con una probabilidad de 1/9. Vimos en la sección 5-1 que una distribución de probabilidad proporciona la probabilidad para cada valor de una variable aleatoria, como en la segunda y tercera columnas de la tabla 6-2. La segunda y tercera columnas de la tabla 6-2 constituyen una distribución de probabilidad para la variable aleatoria que representa las medias muestrales, por lo que la segunda y tercera columnas representan la distribución muestral de la media muestral. En la tabla 6-2, se repiten algunos de los valores de la media muestral, por lo que los combinamos en la tabla 6-3.

TABLA 6-2 Distribución muestral de la media

Muestra	Media muestral $\bar{x}$	Probabilidad
4, 4	4.0	1/9
4, 5	4.5	1/9
4, 9	6.5	1/9
5, 4	4.5	1/9
5, 5	5.0	1/9
5, 9	7.0	1/9
9, 4	6.5	1/9
9, 5	7.0	1/9
9, 9	9.0	1/9

TABLA 6-3 Distribución muestral de la media (condensada)

Media muestral $\bar{x}$	Probabilidad
4.0	1/9
4.5	2/9
5.0	1/9
6.5	2/9
7.0	2/9
9.0	1/9

INTERPRETACIÓN

Debido a que la tabla 6-3 lista los valores posibles de la media muestral junto con sus correspondientes probabilidades, la tabla 6-3 es un ejemplo de una distribución muestral de la media muestral.

El valor de la media de la población {4, 5, 9} es $\mu = 6.0$. Utilizando la tabla 6-2 o 6-3, podríamos calcular la media de los valores muestrales y obtener 6.0. Debido a que la media de las medias muestrales (6.0) es igual a la media de la población (6.0), concluimos que los valores de las medias muestrales se *dirigen* al valor de la media de la población. Desafortunadamente, esto suena como una redundancia, pero ilustra que *la media de las medias muestrales es igual a la media de la población* μ.

SUGERENCIA Lea la última frase del párrafo anterior varias veces hasta que tenga sentido.

Si tuviéramos que crear un histograma de probabilidad a partir de la tabla 6-2, no tendría la forma de campana que es característica de una distribución normal, pero eso se debe a que estamos trabajando con muestras muy pequeñas. Si la población de {4, 5, 9} fuera mucho más grande y si seleccionáramos muestras mucho más grandes que $n = 2$, como sucede en este ejemplo, obtendríamos un histograma de probabilidad mucho más cercano a la forma de campana, lo que indicaría una distribución normal, como en el ejemplo 3.

SU TURNO Resuelva el ejercicio 11 "Distribución muestral de la media muestral".

EJEMPLO 3 Distribución muestral de la media muestral

Considere la repetición del siguiente proceso: Lanzar un dado 5 veces para seleccionar aleatoriamente 5 valores de la población {1, 2, 3, 4, 5, 6}, y después encontrar la media $\bar{x}$ de los resultados. ¿Qué sabemos sobre el comportamiento de todas las medias muestrales que se generan cuando este proceso continúa indefinidamente?

SOLUCIÓN

La figura 6-19 ilustra el proceso de rodar un dado 5 veces y encontrar la media de los resultados. La figura 6-19 muestra los resultados de repetir este proceso 10,000 veces, pero la verdadera distribución muestral de la media implica repetir el proceso indefinidamente. Debido a que los valores de 1, 2, 3, 4, 5, 6 son todos igualmente probables, la población tiene una media de $\mu = 3.5$. Las 10,000 medias muestrales incluidas en la figura 6-19 tienen una media de 3.5. Si el proceso continúa indefinidamente, la media de las medias muestrales será 3.5. Además, la figura 6-19 expresa que la distribución de las medias muestrales es aproximadamente una distribución normal.

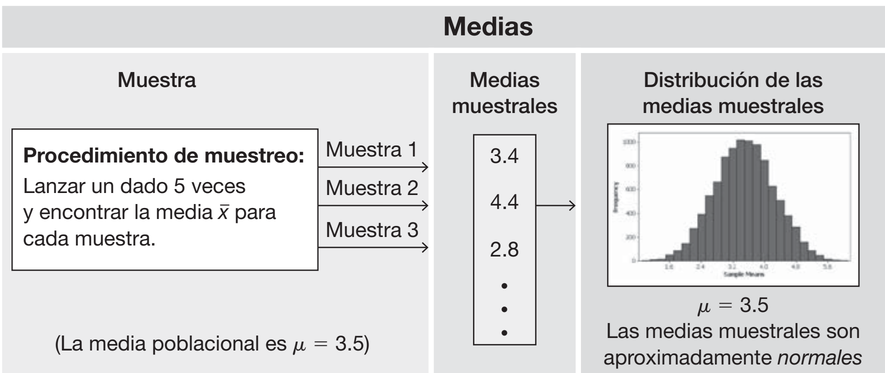


FIGURA 6-19 Medias muestrales de 10,000 ensayos

Distribución muestral de la varianza muestral

Consideremos ahora la distribución muestral de las varianzas muestrales.

DEFINICIÓN

La **distribución muestral de la varianza muestral** es la distribución de las varianzas muestrales (la variable s^2), donde todas las muestras tienen el mismo tamaño n tomado de la misma población. (La distribución muestral de la varianza muestral se representa típicamente como una distribución de probabilidad en el formato de una tabla, un histograma de probabilidad o una fórmula).

PRECAUCIÓN Al trabajar con desviaciones estándar o varianzas poblacionales, asegúrese de evaluarlas correctamente. En la sección 3-2 vimos que los cálculos para las desviaciones estándar o varianzas *poblacionales* implican la división por el tamaño de población N en lugar de $n - 1$, como aquí se muestra.

$$\text{Desviación estándar poblacional:} \quad \sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$

$$\text{Varianza poblacional:} \quad \sigma^2 = \frac{\sum (x - \mu)^2}{N}$$

Debido a que los cálculos se realizan comúnmente con software o calculadoras, tenga cuidado de distinguir correctamente entre la varianza de una muestra y la varianza de una población.

Comportamiento de las varianzas muestrales

1. La distribución de las varianzas muestrales tiende a ser una distribución sesgada a la derecha.
2. Las varianzas muestrales se *dirigen* al valor de la varianza poblacional. (Es decir, la media de las varianzas muestrales es la varianza poblacional. El valor esperado de la varianza muestral es igual a la varianza poblacional).

EJEMPLO 4 Distribución muestral de la varianza muestral

Considere la repetición del siguiente proceso: Lanzar un dado 5 veces y encontrar la varianza s^2 de los resultados. ¿Qué sabemos sobre el comportamiento de todas las varianzas muestrales que se generan cuando este proceso continúa indefinidamente?

SOLUCIÓN

La figura 6-20 ilustra el proceso de lanzar un dado 5 veces y encontrar la varianza de los resultados. La figura 6-20 muestra los resultados de repetir este proceso 10,000 veces, pero la verdadera distribución muestral de la varianza muestral implica repetir el proceso indefinidamente. Debido a que los valores de 1, 2, 3, 4, 5, 6 son todos igualmente probables, la población tiene una varianza de $\sigma^2 = 2.9$ y las 10,000 varianzas muestrales incluidas en la figura 6-20 tienen una media de 2.9. Si el proceso continúa indefinidamente, la media de las varianzas muestrales será 2.9. Además, la figura 6-20 establece que la distribución de las varianzas muestrales es una distribución asimétrica, en vez de una distribución normal con su forma de campana característica.

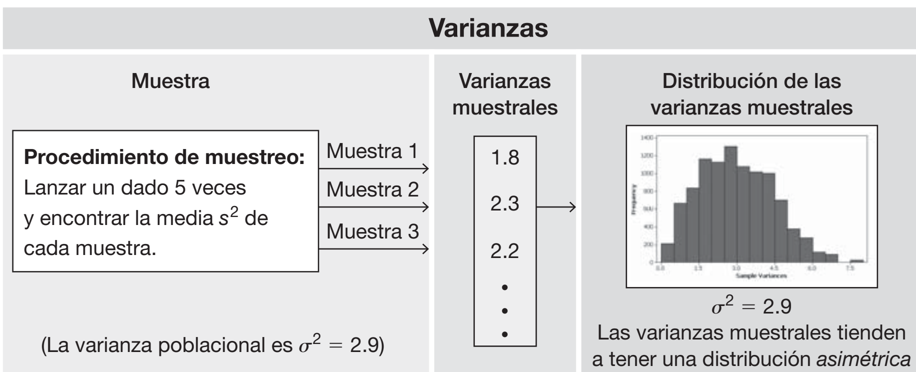


FIGURA 6-20 Varianzas muestrales de 10,000 ensayos

SU TURNO Resuelva el ejercicio 14 “Distribución muestral de la varianza”.

Estimadores: no sesgados y sesgados

Los ejemplos anteriores muestran que las proporciones, las medias y las varianzas muestrales tienden a *dirigirse* a los parámetros poblacionales correspondientes. De manera más formal, decimos que las proporciones, las medias y las varianzas muestrales son *estimadores no sesgados*. Vea las dos definiciones siguientes.

DEFINICIONES

Un **estimador** es un estadístico utilizado para inferir (o estimar) el valor de un parámetro poblacional.

Un **estimador no sesgado** es un estadístico que se dirige al valor del parámetro poblacional correspondiente en el sentido de que la distribución muestral del estadístico tiene una media que es igual al parámetro poblacional correspondiente.

Estimadores no sesgados Estos estadísticos son estimadores no sesgados. Es decir, cada uno se dirige al valor del parámetro poblacional correspondiente (con una distribución muestral que tiene una media igual al parámetro poblacional):

- Proporción $\hat{p}$
- Media $\overline{x}$
- Varianza s^2

Estimadores sesgados Estos estadísticos son estimadores sesgados. Es decir, *no* se dirigen al valor del parámetro poblacional correspondiente:

- Mediana
- Rango
- Desviación estándar s

Nota importante: Las desviaciones estándar muestrales no se dirigen a la desviación estándar poblacional σ, pero el sesgo es relativamente pequeño en muestras grandes, por lo que **a menudo se utiliza s para estimar σ** aunque s sea un estimador sesgado de σ.

EJEMPLO 5 Distribución muestral del rango muestral

Como en el ejemplo 2, considere muestras de tamaño $n = 2$ seleccionadas aleatoriamente de la población $\{4, 5, 9\}$.

a. Liste las diferentes muestras posibles junto con la probabilidad de cada muestra, luego encuentre el rango para cada muestra.

b. Describa la distribución muestral del rango muestral en el formato de una tabla que resuma la distribución de probabilidad.

c. Con base en los resultados, ¿los rangos muestrales se dirigen al rango poblacional, que es $9 - 4 = 5$?

d. ¿Qué indican estos resultados sobre el rango muestral como estimador del rango poblacional?

SOLUCIÓN

a. En la tabla 6-4 se listan las nueve diferentes muestras posibles de tamaño $n = 2$ seleccionadas con reemplazo de la población $\{4, 5, 9\}$. Las nueve muestras son igualmente probables, así que cada una tiene probabilidad de 1/9. La tabla 6-4 también presenta el rango para cada una de las nueve muestras.

continúa

TABLA 6-4 Distribución muestral del rango

Muestra	Rango muestral	Probabilidad
4, 4	0	1/9
4, 5	1	1/9
4, 9	5	1/9
5, 4	1	1/9
5, 5	0	1/9
5, 9	4	1/9
9, 4	5	1/9
9, 5	4	1/9
9, 9	0	1/9

b. Las dos últimas columnas de la tabla 6-4 listan los valores del rango junto con sus probabilidades correspondientes, de manera que las dos últimas columnas constituyen una tabla que resume la distribución de probabilidad. Por lo tanto, la tabla 6-4 describe la *distribución muestral* del rango muestral.

c. La media de los rangos muestrales en la tabla 6-4 es 20/9 o 2.2. La población de {4, 5, 9} tiene un rango de $9 - 4 = 5$. Debido a que la media de los rangos muestrales (2.2) no es igual al rango de la población (5), los rangos muestrales *no* se dirigen al valor del rango poblacional.

d. Debido a que los rangos muestrales no se dirigen al rango poblacional, el rango muestral es un *estimador sesgado* del rango poblacional.

INTERPRETACIÓN

Debido a que el rango muestral es un estimador sesgado del rango poblacional, generalmente no se debe usar un rango muestral para estimar el valor del rango poblacional.

SU TURNO Resuelva el ejercicio 13 "Distribución muestral del rango".

¿Por qué muestrear con reemplazo? Todos los ejemplos de esta sección involucraron muestreo *con reemplazo*. El muestreo *sin reemplazo* tendría la ventaja práctica de evitar la duplicación inútil cuando se selecciona el mismo artículo más de una vez. Muchos de los procedimientos estadísticos discutidos en los siguientes capítulos se basan en el supuesto de que el muestreo se lleva a cabo con reemplazo debido a estas dos importantes razones:

1. Al seleccionar una muestra relativamente pequeña de una población grande, no hay ninguna diferencia significativa si muestreamos con reemplazo o sin reemplazo.
2. El muestreo con reemplazo genera eventos *independientes* que no se ven afectados por resultados anteriores, a la vez que los eventos independientes son más fáciles de analizar y resultan en cálculos y fórmulas más simples.

6-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Nacimientos Hay alrededor de 11,000 nacimientos cada día en Estados Unidos, y la proporción de niños nacidos en dicho país es 0.512. Suponga que cada día se seleccionan al azar 100 nacimientos y se registra la proporción de niños.

a. ¿Qué sabe usted sobre la media de las proporciones muestrales?

b. ¿Qué sabe usted sobre la forma de la distribución de las proporciones muestrales?

2. Muestreo con reemplazo El Centro de Investigación Médica Orangetown selecciona al azar 100 nacimientos en Estados Unidos cada día, y se registra la proporción de niños en cada muestra.

a. ¿Cree usted que los nacimientos se seleccionan al azar con reemplazo o sin reemplazo?

b. Indique dos razones por las cuales los métodos estadísticos tienden a basarse en el supuesto de que el muestreo se lleva a cabo *con* reemplazo, en vez de sin reemplazo.

3. Estimadores no sesgados En el apéndice B, el conjunto de datos 4 "Nacimientos" incluye pesos al nacer de 400 bebés. Si calculamos los valores de los estadísticos muestrales de esa muestra, ¿cuáles de los siguientes estadísticos son estimadores *no sesgados* de los parámetros poblacionales correspondientes: media muestral, mediana muestral, rango muestral;,varianza muestral, desviación estándar muestral y proporción muestral?

4. Distribución muestral En el apéndice B, el conjunto de datos 4 "Nacimientos" incluye una muestra de pesos al nacer. Si exploramos esta muestra de 400 pesos al nacer construyendo un histograma y determinamos la media y la desviación estándar, ¿describen esos resultados la distribución muestral de la media? ¿Por qué sí o por qué no?

5. ¿Buena muestra? Un genetista está investigando la proporción de niños nacidos en la población mundial. Debido a que el estudio tiene su base en China, obtiene los datos muestrales en ese país. ¿Es la proporción muestral resultante un buen estimador de la proporción poblacional de los niños nacidos en todo el mundo? ¿Por qué sí o por qué no?

6. Presidentes universitarios Hay aproximadamente 4200 presidentes universitarios en Estados Unidos, y tienen ingresos anuales con una distribución que está sesgada en vez de ser normal. Una gran cantidad de muestras diferentes de 40 presidentes universitarios se seleccionan al azar, y se calcula el ingreso anual medio para cada muestra.

a. ¿Cuál es la forma aproximada de la distribución de las medias muestrales (uniforme, normal, sesgada, otra)?

b. ¿A qué valor se dirigen las medias muestrales? Es decir, ¿cuál es la media de todas esas medias muestrales?

En los ejercicios 7 a 10, use la misma población de {4, 5, 9} que se usó en los ejemplos 2 y 5. Como en los ejemplos 2 y 5, suponga que se seleccionan al azar muestras de tamaño $n = 2$ con reemplazo.

7. Distribución muestral de la varianza muestral

a. Determine el valor de la varianza poblacional σ^2.

b. La tabla 6-2 describe la distribución muestral de la media muestral. Construya una tabla similar que represente la distribución muestral de la varianza muestral s^2. Después, combine los valores de s^2 que sean iguales, como en la tabla 6-3 (*Sugerencia:* Vea las tablas 6-2 y 6-3 en el ejemplo 2 de la página 258, que describen la distribución muestral de la media muestral).

c. Encuentre la media de la distribución muestral de la varianza muestral.

d. Con base en los resultados anteriores, ¿es la varianza muestral un estimador imparcial de la varianza poblacional? ¿Por qué sí o por qué no?

8. Distribución muestral de la desviación estándar muestral En los ejercicios siguientes, redondee los resultados a tres decimales.

a. Determine el valor de la desviación estándar poblacional.

b. La tabla 6-2 describe la distribución muestral de la media muestral. Construya una tabla similar que represente la distribución muestral de la desviación estándar muestral s. Después, combine los valores de s que sean iguales, como en la tabla 6-3 (*Sugerencia:* Consulte las tablas 6-2 y 6-3 en el ejemplo 2 de la página 258, que describen la distribución muestral de la media muestral).

c. Encuentre la media de la distribución muestral de la desviación estándar muestral.

d. Con base en los resultados anteriores, ¿es la desviación estándar muestral un estimador no sesgado de la desviación estándar poblacional? ¿Por qué sí o por qué no?

9. Distribución muestral de la mediana muestral

a. Encuentre el valor de la mediana poblacional.

b. La tabla 6-2 describe la distribución muestral de la media muestral. Construya una tabla similar que represente la distribución muestral de la mediana muestral. Después, combine los valores de la mediana que sean iguales, como en la tabla 6-3. (*Sugerencia:* Vea las tablas 6-2 y 6-3 en el ejemplo 2 de la página 258, que describen la distribución muestral de la media muestral).

c. Encuentre la media de la distribución muestral de la mediana muestral.

d. Con base en los resultados anteriores, ¿es la mediana muestral un estimador no sesgado de la mediana poblacional? ¿Por qué sí o por qué no?

10. Distribución muestral de la proporción muestral

a. Para la población, encuentre la proporción de números impares.

b. La tabla 6-2 describe la distribución muestral de la media muestral. Construya una tabla similar que represente la distribución muestral de la proporción muestral de números impares. Después, combine los valores de la proporción muestral que sean iguales, como en la tabla 6-3. (*Sugerencia:* Vea las tablas 6-2 y 6-3 en el ejemplo 2 de la página 258, que describen la distribución muestral de la media muestral).

c. Encuentre la media de la distribución muestral de la proporción muestral de números impares.

d. Con base en los resultados anteriores, ¿la proporción muestral es un estimador no sesgado de la proporción poblacional? ¿Por qué sí o por qué no?

En los ejercicios 11 a 14, use la población de {34, 36, 47, 51} de las cantidades de cafeína (mg/12 oz) en las bebidas Coca-Cola Zero, Diet Pepsi, Dr Pepper y Mellow Yello Zero. Suponga que se seleccionan muestras aleatorias de tamaño $n = 2$ con reemplazo.

11. Distribución muestral de la media muestral

a. Después de identificar las 16 diferentes muestras posibles, determine la media de cada muestra, luego construya una tabla que represente la distribución muestral de la media muestral. En la tabla, combine los valores de las medias muestrales que sean iguales. (*Sugerencia:* Consulte la tabla 6-3 en el ejemplo 2 de la página 258).

b. Compare la media de la población {34, 36, 41, 51} con la media de la distribución muestral de la media muestral.

c. ¿Las medias muestrales se dirigen al valor de la media poblacional? En general, ¿las medias muestrales constituyen buenos estimadores de las medias poblacionales? ¿Por qué sí o por qué no?

12. Distribución muestral de la mediana Repita el ejercicio 11 usando medianas en lugar de medias.

13. Distribución muestral del rango Repita el ejercicio 11 usando rangos en lugar de medias.

14. Distribución muestral de la varianza Repita el ejercicio 11 utilizando varianzas en lugar de medias.

15. Nacimientos: Distribución muestral de la proporción muestral Cuando se seleccionan aleatoriamente dos nacimientos, el espacio muestral para los géneros es hh, hm, mh y mm (donde h = hombre y m = mujer). Suponga que esos cuatro resultados son igualmente probables. Construya una tabla que describa la distribución muestral de la proporción muestral de las niñas en dos nacimientos. ¿La media de las proporciones muestrales es igual a la proporción de niñas en dos nacimientos? ¿El resultado sugiere que una proporción muestral es un estimador no sesgado de una proporción poblacional?

16. Nacimientos: Distribución muestral de la proporción muestral Para tres nacimientos, suponga que los géneros son igualmente probables. Construya una tabla que describa la distribución muestral de la proporción muestral de niñas en tres nacimientos. ¿La media de las proporciones muestrales es igual a la proporción de niñas en tres nacimientos? (*Sugerencia:* Vea el ejercicio 15 para dos nacimientos).

17. Exámenes SAT y ACT Debido a que las preguntas de opción múltiple permiten procedimientos eficientes para calificar las respuestas, se utilizan comúnmente en exámenes estandarizados, como SAT o ACT. Estas preguntas suelen tener cinco opciones, una de las cuales es correcta. Supongamos que usted debe hacer conjeturas aleatorias para dos de estas preguntas. Supongamos que la respuesta correcta para ambas preguntas es "a".

a. Después de listar las 25 muestras posibles, encuentre la proporción de respuestas correctas en cada muestra, luego construya una tabla que describa la distribución muestral de las proporciones muestrales de respuestas correctas.

b. Encuentre la media de la distribución muestral de la proporción muestral.

c. ¿Es la media de la distribución muestral [del inciso (b)] igual a la proporción poblacional de respuestas correctas? ¿La media de la distribución muestral de las proporciones es *siempre* igual a la proporción poblacional?

18. Hibridación Un experimento de hibridación comienza con cuatro chícharos con vainas amarillas y un chícharo con vaina verde. Dos de los chícharos se seleccionan al azar con el *reemplazo* de esta población.

a. Después de identificar las 25 muestras posibles, encuentre la proporción de chícharos con vainas amarillas en cada una de ellas, luego construya una tabla para describir la distribución muestral de las proporciones de chícharos con vainas amarillas.

b. Encuentre la media de la distribución muestral.

c. ¿La media de la distribución muestral [del inciso (b)] es igual a la proporción poblacional de chícharos con vainas amarillas? ¿La media de la distribución muestral de las proporciones es *siempre* igual a la proporción poblacional?

6-3 Más allá de lo básico

19. Uso de una fórmula para describir una distribución muestral El ejercicio 15 "Nacimientos" requiere la construcción de una tabla que describa la distribución muestral de las proporciones de niñas en dos nacimientos. Considere la fórmula que se muestra aquí, y evalúela usando las proporciones muestrales (representadas por x) de 0, 0.5 y 1. Con base en los resultados, ¿describe la fórmula la distribución muestral? ¿Por qué sí o por qué no?

$$P(x) = \frac{1}{2(2-2x)!(2x)!} \quad \text{donde } x = 0, 0.5, 1$$

20. Desviación absoluta media ¿Es la desviación media absoluta muestral un buen estadístico para estimar la desviación absoluta media poblacional? ¿Por qué sí o por qué no? (*Sugerencia:* Vea el ejemplo 5).

6-4 Teorema del límite central

Concepto clave En la sección anterior vimos que la distribución muestral de las medias muestrales tiende a ser una distribución normal a medida que aumenta el tamaño de la muestra. En esta sección introducimos y aplicamos el *teorema del límite central*, el cual permite utilizar una distribución normal para algunas aplicaciones muy importantes y significativas.

Dada una población con cualquier distribución (uniforme, asimétrica, o cualquier otra), la distribución de las medias muestrales $\overline{x}$ puede aproximarse mediante una distribución normal cuando las muestras son lo suficientemente grandes con $n > 30$. (Hay algunos casos especiales de distribuciones muy distintas a la normal para las que el requisito de $n > 30$ no es suficiente; en esos casos muy raros se requiere un número mayor que 30).

TEOREMA DEL LÍMITE CENTRAL

Para todas las muestras del mismo tamaño n con $n > 30$, la distribución muestral de $\bar{x}$ puede aproximarse mediante una distribución normal con media μ y desviación estándar $\sigma/\sqrt{n}$.

EJEMPLO 1 Profundidades de terremoto

Las figuras 6-21 y 6-22 ilustran el teorema del límite central.

- **Datos originales:** La figura 6-21 es un histograma de las profundidades (km) de 1392 terremotos, y este histograma muestra claramente que esas profundidades tienen una distribución *no* normal.
- **Medias muestrales:** La figura 6-22 es un histograma de 100 *medias muestrales*. Cada muestra incluye 50 profundidades de terremoto (km), y este histograma muestra que las medias muestrales tienen una distribución que es aproximadamente normal.

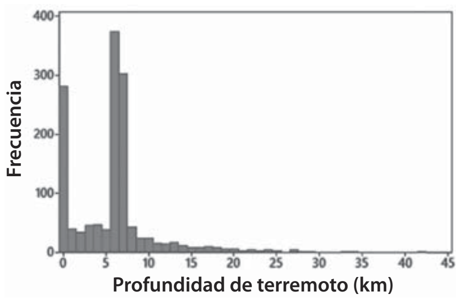


FIGURA 6-21 Distribución no normal: Profundidades (en km) de 1392 terremotos

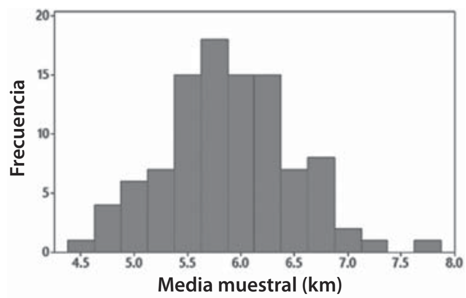


FIGURA 6-22 Distribución aproximadamente normal: Medias en muestras con tamaño $n = 50$ de profundidades de terremotos

INTERPRETACIÓN

Las profundidades originales de los terremotos que se muestran en la figura 6-21 tienen una distribución sesgada, pero cuando recolectamos muestras y calculamos sus medias, las medias muestrales tienden a tener una distribución *normal*.

Una verdad universal El ejemplo 1 y el teorema del límite central son verdaderamente notables porque describen una regla de la naturaleza que funciona a través del universo. Si pudiéramos enviar una nave espacial a un planeta distante “en una galaxia muy muy lejana”, y si recogiéramos muestras de rocas (todas las muestras del mismo tamaño grande) y las pesáramos, la muestra tendría una distribución que sería aproximadamente normal. ¡Piense en el significado de eso! ¡Increíble!

Los siguientes puntos clave constituyen la base para estimar los parámetros poblacionales y las pruebas de hipótesis-temas que se analizan a detalle en los siguientes capítulos.

ELEMENTOS CLAVE

El teorema del límite central y la distribución muestral de $\bar{x}$

Dado que

1. La población (con cualquier distribución) tiene media μ y desviación estándar σ.
2. Se seleccionan muestras aleatorias simples del mismo tamaño n de la población.

Reglas prácticas para aplicaciones reales que involucran una media muestral $\bar{x}$

Requisitos: La población tiene una distribución normal *o* $n > 30$:

Media de todos los valores de $\bar{x}$: $\mu_{\bar{x}} = \mu$

Desviación estándar de todos los valores de $\bar{x}$: $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

Conversión de z a una puntuación $\bar{x}$: $z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

La población original no se distribuye normalmente y $n \leq 30$: La distribución de $\bar{x}$ no puede aproximarse mediante una distribución normal, y los métodos de esta sección no se aplican. Utilice otros métodos, como los no paramétricos (capítulo 13) o los métodos de remuestreo (sección 7-4).

Consideraciones para la resolución de problemas prácticos

1. **Requisitos de verificación:** Cuando trabaje con la media de una muestra, verifique que la distribución normal pueda ser utilizada confirmando que la población original tiene una distribución normal o que el tamaño de la muestra es $n > 30$.
2. **¿Valor individual o media de una muestra?** Determine si está utilizando una distribución normal con un valor único x o la media $\bar{x}$ de una muestra de n valores. Vea lo siguiente.
   - Valor individual: Cuando trabaje con un valor *individual* de una población normalmente distribuida, utilice los métodos de la sección 6-2 con $z = \dfrac{x - \mu}{\sigma}$.
   - Media de una muestra de valores: Cuando trabaje con una media para alguna *muestra* de n valores, asegúrese de usar el valor de $\sigma/\sqrt{n}$ para la desviación estándar de las medias muestrales, así que utilice $z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$.

La siguiente notación nueva se utiliza para la media y la desviación estándar de la distribución de $\bar{x}$.

NOTACIÓN PARA LA DISTRIBUCIÓN MUESTRAL DE $\bar{X}$

Si se seleccionan todas las posibles muestras aleatorias simples de tamaño n de una población con media μ y desviación estándar σ, la media de todas las medias muestrales se denota $\mu_{\bar{x}}$, y la desviación estándar de todas las medias muestrales se denota $\sigma_{\bar{x}}$.

Media de todos los valores de $\bar{x}$: $\mu_{\bar{x}} = \mu$

Desviación estándar de todos los valores de $\bar{x}$: $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

Nota: $\sigma_{\bar{x}}$ se denomina el *error estándar de la media* y se denota en ocasiones como EEM.

Aplicación del teorema del límite central

Muchos problemas prácticos pueden resolverse mediante el teorema del límite central. El ejemplo 2 es una buena ilustración del teorema del límite central porque podemos ver la diferencia entre trabajar con un valor *individual* en el inciso (a) y emplear la *media* de una muestra en el inciso (b). Estudie cuidadosamente el ejemplo 2 para entender la diferencia fundamental entre los procedimientos usados en los incisos (a) y (b). En particular, tenga en cuenta que cuando se trabaja con un valor *individual*, se utiliza $z = \dfrac{x - \mu}{\sigma}$, y cuando se trabaja con la media $\bar{x}$ para una colección de valores *muestrales*, se usa $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$.

EJEMPLO 2 Carga segura en ascensores

El ascensor en el edificio de alquiler de autos en el Aeropuerto Internacional de San Francisco tiene un cartel que indica que la capacidad máxima es “4000 lb—27 pasajeros”. Debido a que 4000/27 = 148, esto se convierte en un peso medio por pasajero de 148 lb, cuando el elevador está lleno. Asumiremos el peor escenario en el que el ascensor está lleno de 21 hombres adultos. Con base en el conjunto de datos 1 “Datos corporales” del apéndice B, suponga que los hombres adultos tienen pesos que se distribuyen normalmente con una media de 189 lb y una desviación estándar de 39 lb.

a. Encuentre la probabilidad de que 1 varón adulto seleccionado al azar tenga un peso mayor que 148 lb.

b. Encuentre la probabilidad de que una muestra de 27 varones adultos seleccionados al azar tenga un peso promedio mayor de 148 lb.

SOLUCIÓN

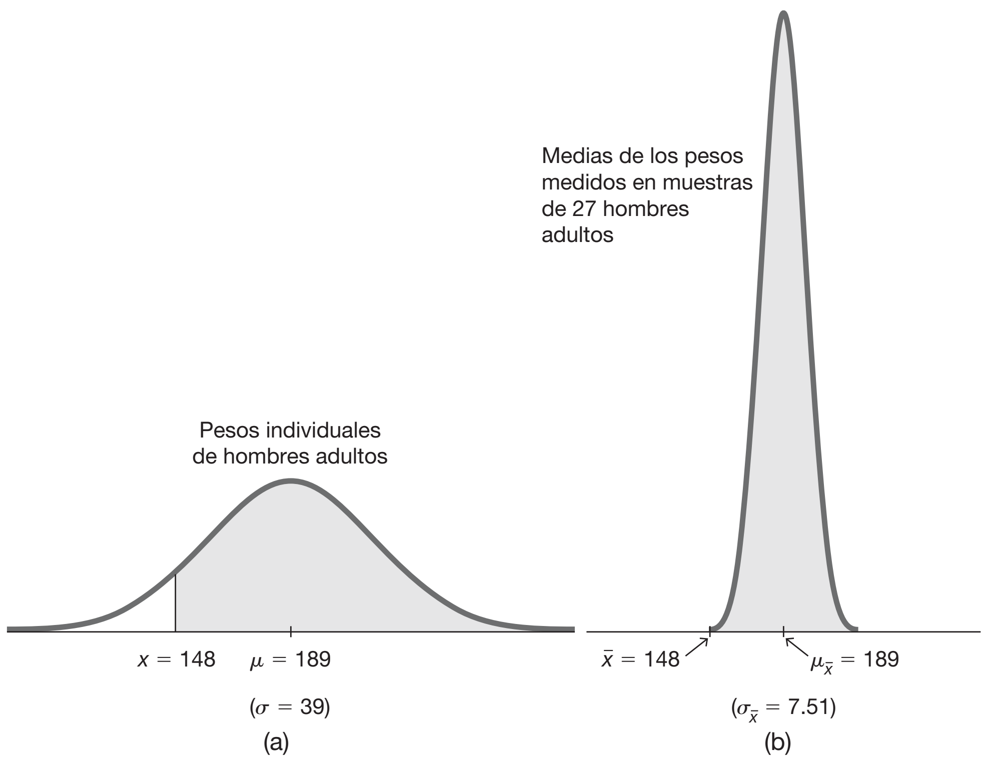


FIGURA 6-23 Pesos en ascensor

a. Enfoque utilizado para un valor individual: Use los métodos presentados en la sección 6-2 porque estamos tratando con un valor *individual* de una población distribuida normalmente. Buscamos el área de la región sombreada en la figura 6-23(a).

Tecnología: Si utilizamos tecnología (como se describe al final de la sección 6-2), encontramos que el área sombreada es 0.8534.

Tabla A-2: Si se utilizamos la tabla A-2, convertimos el peso de $x = 148$ lb en la puntuación z correspondiente de $z = -1.05$, como se muestra aquí:

$$z = \frac{x - \mu}{\sigma} = \frac{148 - 189}{39} = -1.05$$

Con referencia a la tabla A-2 encontramos que el área acumulada a la *izquierda* de $z = -1.05$ es 0.1496, por lo que el área sombreada de la figura 6-23(a) es 1 – 0.1469 = 0.8531. (El resultado de 0.8534 con la tecnología es más exacto que el resultado encontrado a partir de la tabla A-2).

b. Enfoque utilizado para la media de los valores muestrales: Use el teorema del límite central porque estamos tratando con la media de una muestra de 27 hombres, no con un hombre individual.

REQUISITOS PARA EL INCISO B Es posible utilizar la distribución normal si la población original se distribuye normalmente o $n > 30$. El tamaño de la muestra no es mayor que 30, pero la población original de pesos de los hombres tiene una distribución normal, por lo que las muestras de cualquier tamaño darán medias que se distribuyen normalmente.

Debido a que ahora estamos tratando con una distribución de las medias muestrales, debemos utilizar los parámetros $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$, que se evalúan de la siguiente manera:

$$\mu_{\bar{x}} = \mu = 189$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{39}{\sqrt{27}} = 7.51$$

Queremos encontrar el área sombreada que se muestra en la figura 6-23(b).

Tecnología: Si usamos tecnología, el área sombreada de la figura 6-23(b) es 0.99999998.

Tabla A-2: Si utilizamos la tabla A-2, convertimos el valor de $\bar{x} = 148$ a la puntuación correspondiente de $z = -5.46$, como se muestra aquí:

$$z = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{148 - 189}{\frac{39}{\sqrt{27}}} = \frac{-41}{7.51} = -5.46$$

En la tabla A-2 se observa que el área acumulada a la izquierda de $z = -5.46$ es 0.0001, por lo que el área sombreada de la figura 6-23(b) es 1 – 0.0001 = 0.9999. Estamos bastante seguros de que 27 varones adultos seleccionados al azar tienen un peso medio mayor de 148 lb.

INTERPRETACIÓN

Hay una probabilidad de 0.8534 de que un individuo masculino pese más de 148 libras y una probabilidad 0.99999998 de que 27 hombres seleccionados al azar tengan un peso medio de más de 148 libras. Dado que la capacidad de seguridad del elevador es de 4000 lb, es casi seguro que tendrá sobrepeso si se llena con 27 varones adultos seleccionados al azar.

SU TURNO Resuelva el ejercicio 5 “Uso del teorema del límite central”.

El difuso teorema del límite central

En *The Cartoon Guide to Statistics*, los autores Gonick y Smith describen el “difuso” teorema del límite central de la siguiente manera: “Los datos que se ven influidos por efectos aleatorios muy pequeños y sin relación entre sí se distribuyen aproximadamente de manera normal. Esto explica por qué la normalidad está en todos lados: en las fluctuaciones del mercado bursátil, en los pesos de estudiantes, en los promedios anuales de temperatura y en las calificaciones del SAT. Todos son el resultado de muchos efectos diferentes”. La estatura de las personas, por ejemplo, es el resultado de factores hereditarios, ambientales y nutricionales, además de otros como el cuidado de la salud, la región geográfica y ciertas influencias que, cuando se combinan, producen valores distribuidos de forma normal.

El ejemplo 2 muestra que podemos usar los mismos procedimientos básicos de la sección 6-2, pero debemos ajustar correctamente la desviación estándar cuando se trabaja con una *media muestral* en lugar de un valor individual de la muestra. Los cálculos utilizados en el ejemplo 2 son exactamente el tipo de cálculos utilizados por los ingenieros cuando diseñan teleféricos, escaleras mecánicas, aviones, embarcaciones, paseos en parques de diversiones y otros dispositivos que transportan personas.

Introducción a la prueba de hipótesis

Examine cuidadosamente las conclusiones a las que se llega en el siguiente ejemplo, el cual ilustra el tipo de pensamiento que es la base del importante procedimiento de las pruebas de hipótesis (introducido formalmente en el capítulo 8). El ejemplo 3 usa la regla del evento raro para la estadística inferencial, presentada por primera vez en la sección 4-1:

Identificación de resultados significativos con probabilidades: La regla del evento raro para la estadística inferencial

Si, bajo un supuesto dado, la probabilidad de un evento observado particular es muy pequeña y el evento observado ocurre *significativamente menor* o *significativamente mayor* de lo que típicamente esperaríamos con ese supuesto, concluimos que el supuesto probablemente no es correcto.

El siguiente ejemplo ilustra la regla del evento raro descrita anteriormente y utiliza el conjunto de datos preferido del autor.

EJEMPLO 3 Temperaturas corporales

Suponga que la población de las temperaturas corporales humanas tiene una media de 98.6 °F, como se cree comúnmente. También suponga que la desviación estándar poblacional es 0.62 °F (según datos de investigadores de la Universidad de Maryland). Si se selecciona aleatoriamente una muestra de tamaño $n = 106$, determine la probabilidad de obtener una media de 98.2 °F o menor. (El valor de 98.2 °F en realidad obtenido de los investigadores; vea las temperaturas de medianoche para el día 2 en el conjunto de datos 3 “Temperaturas corporales” en el apéndice B).

SOLUCIÓN

Trabajamos bajo el supuesto de que la población de las temperaturas corporales humanas tiene una media de 98.6 °F. No se nos proporcionó la distribución de la población, pero debido a que el tamaño de la muestra $n = 106$ excede 30, usamos el teorema del límite central y concluimos que la distribución de las medias muestrales es una distribución normal con los siguientes parámetros:

$$\mu_{\bar{x}} = \mu = 98.6 \text{ (por suposición)}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.62}{\sqrt{106}} = 0.0602197$$

La figura 6-24 muestra el área sombreada (vea la diminuta cola izquierda del gráfico) correspondiente a la probabilidad que buscamos. Una vez encontrados los parámetros que se aplican a la distribución mostrada en la figura 6-24, podemos determinar el área sombreada empleando los mismos procedimientos desarrollados en la sección 6-2.

Tecnología: Si usamos la tecnología para encontrar el área sombreada en la figura 6-24, obtenemos 0.0000000000155, que puede expresarse como 0+.

Tabla A-2: Si utilizamos la tabla A-2 para encontrar el área sombreada en la figura 6-24, debemos primero convertir el valor de $x = 98.20$ °F a la puntuación z correspondiente:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{98.20 - 98.6}{0.0602197} = -6.64$$

Con referencia a la tabla A-2, encontramos que $z = -6.64$ está fuera de la gráfica, pero para valores de z por debajo de -3.49, utilizamos un área de 0.0001 para el área izquierda acumulada hasta $z = -3.49$. Por lo tanto, concluimos que la región sombreada en la figura 6-24 es 0.0001.

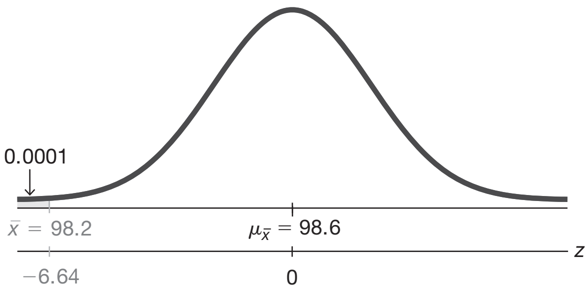


FIGURA 6.24 Medias de las temperaturas corporales en muestras de tamaño $n = 106$

INTERPRETACIÓN

El resultado muestra que si la media de nuestras temperaturas corporales es realmente de 98.6 °F, como asumimos, entonces hay una probabilidad extremadamente pequeña de obtener una muestra con media de 98.2 °F o menos, cuando 106 sujetos se seleccionan al azar. Los investigadores de la Universidad de Maryland obtuvieron tal media muestral y después de confirmar que la muestra es sólida, surgen dos explicaciones factibles: (1) La media de la población es realmente de 98.6 °F y su muestra representa un suceso casual que es extremadamente raro; (2) la media de la población es realmente menor que el valor supuesto de 98.6 °F y, por lo tanto, su muestra es típica. Debido a que la probabilidad es tan baja, es más razonable concluir que la media de la población es menor que 98.6 °F. En realidad, parece que la verdadera temperatura corporal media ¡está más cerca de 98.2 °F!

Este es el tipo de razonamiento utilizado en la *prueba de hipótesis*, que se introducirá en el capítulo 8. Por ahora, debemos centrarnos en el uso del teorema del límite central para encontrar la probabilidad de 0.0001; pero también debemos tener en cuenta que este teorema se utilizará más tarde en la aplicación de algunos conceptos muy importantes en estadística.

SU TURNO Resuelva el ejercicio 9 “Seguridad en el ascensor”.

Corrección para una población finita

Al aplicar el teorema del límite central, nuestro uso de $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ supone que la población tiene infinitamente muchos miembros. Cuando muestreamos con reemplazo, la población es efectivamente infinita. Cuando muestreamos sin reemplazo en una población finita, es posible que necesitemos ajustar $\sigma_{\bar{x}}$. La siguiente es una regla común:

> **Cuando muestree sin reemplazo y el tamaño n de la muestra sea mayor que 5% del tamaño de la población finita N (es decir, $n > 0.05N$), ajuste la desviación estándar de las medias muestrales $\sigma_{\bar{x}}$ multiplicándolas por el siguiente *factor de corrección de población finita*:**

$$\sqrt{\frac{N-n}{N-1}}$$

A excepción del ejercicio 21 “Corrección para una población finita”, los ejemplos y ejercicios de esta sección suponen la *no* aplicación del factor de corrección de población finita, porque estamos muestreando con reemplazo, o la población es infinita, o el tamaño de la muestra no supera 5% del tamaño de la población.

6-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Requisitos Un investigador recolecta una muestra aleatoria simple de las calificaciones promedio de estudiantes de estadística, y calcula la media de esa muestra. ¿En qué condiciones se puede tratar esa media muestral como un valor poblacional que posee una distribución normal?

2. Muestra pequeña Los pesos de los perros Golden Retriever se distribuyen normalmente. Se recolectan muestras aleatorias, cada una con tamaño $n = 15$, de los pesos de perros Golden Retriever y se encuentran las medias muestrales. ¿Es correcto concluir que las medias muestrales no pueden ser tratadas como una distribución normal porque el tamaño de la muestra es demasiado pequeño? Explique.

3. Notación En general, ¿qué representan los símbolos $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$? ¿Cuáles son los valores de $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$ para muestras de tamaño 64 seleccionadas al azar de la población de puntajes de IQ con media poblacional de 100 y desviación estándar de 15?

4. Ingresos anuales Se sabe que los ingresos anuales tienen una distribución que está sesgada hacia la derecha en lugar de distribuirse normalmente. Suponga que recopilamos una gran muestra aleatoria de ingresos anuales ($n > 30$). ¿Puede la distribución de los ingresos en esa muestra aproximarse mediante una distribución normal porque la muestra es grande? ¿Por qué sí o por qué no?

Uso del teorema del límite central. ***En los ejercicios 5 a 8, suponga que las mujeres tienen pulsos distribuidos normalmente con una media de 74.0 latidos por minuto y una desviación estándar de 12.5 latidos por minuto (con base en el conjunto de datos 1 "Datos corporales" del apéndice B).***

5. a. Si se selecciona al azar a una mujer adulta, determine la probabilidad de que su pulso sea inferior a 80 latidos por minuto.

b. Si se seleccionan al azar 16 mujeres adultas, encuentre la probabilidad de que tengan pulsos con una media de menos de 80 latidos por minuto.

c. ¿Por qué se puede utilizar la distribución normal en el inciso (b), aunque el tamaño de muestra no sea mayor que 30?

6. a. Si se selecciona al azar a una mujer adulta, encuentre la probabilidad de que su pulso sea mayor que 70 latidos por minuto.

b. Si se seleccionan al azar 25 mujeres adultas, encuentre la probabilidad de que tengan pulsos con una media mayor que 70 latidos por minuto.

c. ¿Por qué se puede utilizar la distribución normal en el inciso (b), aunque el tamaño de muestra no sea mayor que 30?

7. a. Si se selecciona al azar a una mujer adulta, encuentre la probabilidad de que su pulso esté entre 72 y 76 latidos por minuto.

b. Si se seleccionan al azar 4 mujeres adultas, encuentre la probabilidad de que tengan pulsos con una media entre 72 y 76 latidos por minuto.

c. ¿Por qué se puede utilizar la distribución normal en el inciso (b), aunque el tamaño de muestra no sea mayor que 30?

8. a. Si se selecciona al azar 1 mujer adulta, encuentre la probabilidad de que su pulso esté entre 78 y 90 latidos por minuto.

b. Si se seleccionan al azar 16 mujeres adultas, encuentre la probabilidad de que tengan pulsos con una media entre 78 y 90 latidos por minuto.

c. ¿Por qué se puede utilizar la distribución normal en el inciso (b), aunque el tamaño de la muestra no sea mayor que 30?

9. Seguridad en el ascensor El ejemplo 2 se refiere a un ascensor con capacidad máxima de 4000 libras. Cuando se clasifican los ascensores, es común utilizar un factor de seguridad del 25%, por lo que el ascensor debe ser capaz de transportar una carga de 25% mayor que el límite establecido. La capacidad máxima de 4000 libras se convierte en 5000 libras después de aumentarle un 25%, por lo que 27 pasajeros varones adultos pueden tener un peso medio de hasta 185 libras. Si el ascensor se carga con 27 pasajeros adultos varones, encuentre la probabilidad de que esté sobrecargado porque los pasajeros tienen un peso medio mayor que 185 libras. (Al igual que en el ejemplo 2, suponga que los pesos de los hombres se distribuyen normalmente con una media de 189 lb y una desviación estándar de 39 lb). ¿Este ascensor será seguro?

10. Seguridad en el ascensor El ejercicio 9 utiliza $\mu = 189$ lb, que se basa en el conjunto de datos 1 "Datos corporales" del apéndice B. Repita el ejercicio 9 usando $\mu = 174$ libras (en lugar de 189 libras), que es el peso promedio supuesto que se usaba comúnmente hace sólo algunos años. ¿Qué concluye usted sobre el efecto de usar una media antigua que es sustancialmente más baja de lo que debería ser?

11. Mensa La membresía en Mensa requiere una puntuación que se encuentre dentro del 2% superior en una prueba de inteligencia estándar. El test Wechsler IQ está diseñado para una media de 100 y una desviación estándar de 15, y las puntuaciones se distribuyen normalmente.

a. Encuentre el puntaje mínimo en la prueba Wechsler IQ que satisfaga el requisito de Mensa.

b. Si 4 adultos seleccionados al azar realizan la prueba Wechsler IQ, determine la probabilidad de que su puntuación media sea al menos de 131.

c. Si 4 sujetos realizan la prueba de Wechsler y tienen una media de 132, pero las puntuaciones individuales se pierden, ¿podemos concluir que los 4 son elegibles para Mensa?

12. Diseño de alcantarillas De acuerdo con el sitio web www.torchmate.com, "las tapas de alcantarilla deben tener un diámetro mínimo de 22 pulgadas, pero pueden tener un diámetro de hasta 60 pulgadas. Suponga que una alcantarilla se construye con una abertura circular de 22 pulgadas de diámetro. Los hombres tienen anchuras de hombros que se distribuyen normalmente con una media de 18.2 pulgadas y una desviación estándar de 1.0 pulgadas (según datos de la Encuesta Nacional de Salud y Nutrición).

a. ¿Qué porcentaje de hombres cabrá en la boca de la alcantarilla?

b. Suponga que la compañía Eversource de Connecticut emplea a 36 hombres que trabajan en las alcantarillas. Si se seleccionan al azar 36 hombres, ¿cuál es la probabilidad de que la anchura media de sus hombros sea menor que 18.5 pulgadas? ¿Este resultado sugiere que se puede ahorrar dinero haciendo alcantarillas más pequeñas con un diámetro de 18.5 pulgadas? ¿Por qué sí o por qué no?

13. Seguridad en taxi acuático Algunos pasajeros murieron cuando un taxi acuático se hundió en el muelle interior de Baltimore. Los hombres suelen ser más pesados que las mujeres y los niños, por lo que al cargar un taxi acuático, podemos suponer que el peor escenario es aquel en el que todos los pasajeros son hombres. Suponga que los pesos de los hombres se distribuyen normalmente con una media de 189 libras y una desviación estándar de 39 libras (según el conjunto de datos 1 "Datos corporales" en el apéndice B). El taxi acuático que se hundió tenía una capacidad declarada de 25 pasajeros, y el barco estaba clasificado para un límite de carga de 3500 libras.

a. Dado que el taxi acuático que se hundió fue clasificado para un límite de carga de 3500 libras, ¿cuál es el peso medio máximo de los pasajeros si el barco está lleno hasta la capacidad declarada de 25 pasajeros?

b. Si el taxi acuático está lleno de 25 hombres seleccionados al azar, ¿cuál es la probabilidad de que su peso medio exceda el valor del inciso (a)?

c. Después de que el taxi acuático se hundiera, los supuestos de peso fueron modificados, para que la nueva capacidad fuera de 20 pasajeros. Si el taxi acuático está lleno de 20 hombres seleccionados al azar, ¿cuál es la probabilidad de que su peso promedio exceda 175 libras, que es el peso medio máximo antes de que la carga total sea mayor que 3500 libras?

d. ¿Es segura la nueva capacidad de 20 pasajeros?

14. Máquinas expendedoras Las monedas de ¢25 se fabrican ahora de manera que tengan un peso medio de 5.670 g, una desviación estándar de 0.062 g y sus pesos se distribuyan normalmente. Una máquina expendedora está configurada para aceptar solamente aquellas monedas de ¢25 que pesan entre 5.550 y 5.790 g.

a. Si una moneda de ¢25 seleccionada aleatoriamente se inserta en la máquina expendedora, ¿cuál es la probabilidad de que sea aceptada?

continúa

b. Si 4 monedas de ¢25 seleccionadas al azar se insertan en la máquina expendedora, ¿cuál es la probabilidad de que su peso medio esté entre 5.550 y 5.790 g?

c. Si usted es dueño de la máquina expendedora, ¿cuál resultado es más importante: el resultado del inciso (a) o el del inciso (b)?

15. Asientos de Southwest Airlines Southwest Airlines tiene actualmente una anchura de asiento de 17 pulgadas. Los hombres tienen anchuras de cadera que se distribuyen normalmente con una media de 14.4 pulgada y una desviación estándar de 1.0 pulgada (según datos antropométricos de la encuesta de Gordon, Churchill y otros).

a. Encuentre la probabilidad de que si un hombre individual se selecciona al azar, su anchura de cadera sea mayor que 17 pulgadas.

b. Southwest Airlines utiliza un Boeing 737 para algunos de sus vuelos, y ese avión tiene capacidad para 122 pasajeros. Si el avión está lleno con 122 hombres seleccionados al azar, encuentre la probabilidad de que estos hombres tengan una anchura media de cadera mayor que 17 pulgadas.

c. ¿Qué resultado debe considerarse para cualquier cambio en el diseño de los asientos: el resultado del inciso (a) o el del inciso (b)?

16. Latas de Coca-Cola Suponga que las latas de Coca-Cola se llenan de modo que las cantidades reales tengan una media de 12.00 onzas y una desviación estándar de 0.11 onzas.

a. Encuentre la probabilidad de que una sola lata de Coca Cola contenga por lo menos 12.19 onzas.

b. Las 36 latas de Coca-Cola en el conjunto de datos 26 “Pesos y volúmenes de bebidas de Cola” en el apéndice B tienen una media de 12.19 onzas. Encuentre la probabilidad de que 36 latas de Coca Cola seleccionadas al azar tengan una media de por lo menos 12.19 onzas.

c. Dado el resultado del inciso (b), ¿es razonable creer que las latas están realmente llenas con una media igual a 12.00 onzas? Si la media no es igual a 12.00 onzas, ¿se engaña a los consumidores?

17. Rediseño de asientos de eyección Cuando finalmente se permitió a las mujeres ser pilotos de aviones de combate, los ingenieros necesitaban rediseñar los asientos de eyección porque habían sido diseñados originalmente sólo para hombres. Los asientos de eyección ACES-II fueron diseñados para hombres que pesan entre 140 y 211 libras. Los pesos de las mujeres ahora se distribuyen normalmente con una media de 171 lb y una desviación estándar de 46 lb (según el conjunto de datos 1 “Datos corporales” del apéndice B).

a. Si una mujer se selecciona al azar, encuentre la probabilidad de que su peso esté entre 140 y 211 lb.

b. Si se seleccionan al azar 25 mujeres diferentes, encuentre la probabilidad de que su peso promedio esté entre 140 y 211 lb.

c. Al rediseñar los asientos de eyección más ajustados para acomodarse mejor a las mujeres, ¿qué probabilidad es más relevante: el resultado del inciso (a) o el del inciso (b)? ¿Por qué?

18. Carga de un barco turístico El barco Ethan Allen se hundió en Lake George, Nueva York, y 20 de los 47 pasajeros murieron ahogados. Con base en un supuesto de 1960 de un peso medio de 140 libras para los pasajeros, el barco estaba clasificado para llevar a 50 pasajeros. Después de que el barco se hundiera, el estado de Nueva York cambió el peso promedio asumido de 140 lb a 174 lb.

a. Dado que la embarcación estaba clasificada para 50 pasajeros con un peso promedio supuesto de 140 libras, su límite de carga de pasajeros era de 7000 libras. Suponga que el barco estaba cargado con 50 pasajeros masculinos, y asuma que los pesos de los hombres se distribuyen normalmente con un peso medio de 189 libras y una desviación estándar de 39 libras (según el conjunto de datos 1 “Datos corporales” en el apéndice B). Determine la probabilidad de que el barco esté sobrecargado porque los 50 pasajeros masculinos tienen un peso medio superior a 140 libras.

b. El barco fue clasificado después para llevar solamente a 14 pasajeros, y el límite de carga fue cambiado a 2436 libras. Si los 14 pasajeros son todos varones, encuentre la probabilidad que el barco esté sobrecargado porque su peso medio es mayor que 174 libras (de modo que su peso total es mayor que la capacidad máxima de 2436 lb). ¿Las nuevas clasificaciones parecen seguras cuando el barco está cargado con 14 pasajeros masculinos?

19. Altura de puerta El avión Boeing 157-200 ER transporta 200 pasajeros y tiene puertas con una altura de 72 pulgadas. Las alturas de los hombres se distribuyen normalmente con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas (según el conjunto de datos 1 "Datos corporales" en el apéndice B).

a. Si un pasajero masculino es seleccionado al azar, encuentre la probabilidad de que pueda pasar a través de la puerta sin inclinarse.

b. Si la mitad de los 200 pasajeros son hombres, encuentre la probabilidad de que la altura media de los 100 hombres sea inferior a 72 pulgadas.

c. Al considerar la comodidad y seguridad de los pasajeros, ¿cuál es el resultado más relevante: la probabilidad del inciso (a) o la probabilidad del inciso (b)? ¿Por qué?

d. Al considerar la comodidad y seguridad de los pasajeros, ¿por qué en este caso las mujeres son ignoradas?

20. Carga de aviones Antes de cada vuelo, un piloto de avión debe verificar que el peso total de la carga sea menor que la carga máxima permitida para la aeronave. El avión Bombardier Dash 8 puede transportar 37 pasajeros y un vuelo tiene combustible y equipaje que permite una carga total de pasajeros de 6200 libras. El piloto ve que el avión está lleno y todos los pasajeros son hombres. La aeronave estará sobrecargada si el peso medio de los pasajeros es mayor que 6200 lb/37 = 167.6 lb. ¿Cuál es la probabilidad de que la aeronave esté sobrecargada? ¿Debería el piloto tomar alguna acción para corregir una aeronave sobrecargada? Suponga que los pesos de los hombres se distribuyen normalmente con una media de 189 libras y una desviación estándar de 39 libras (según el conjunto de datos 1 "Datos corporales" en el apéndice B).

6-4 Más allá de lo básico

21. Corrección para una población finita En un estudio de bebés con muy bajo peso al nacer, 275 niños realizaron pruebas de IQ a los 8 años, y sus puntuaciones se aproximaron a una distribución normal con $\mu = 95.5$ y $\sigma = 16.0$ (de acuerdo con datos de "Resultados del comportamiento neurológico en niños en edad escolar, nacidos con bajo peso o de manera prematura", de Anderson *et al*., *Journal of the American Medical Association*, vol. 289, núm. 24). Para realizar un estudio de seguimiento, 50 de esos niños deben seleccionarse al azar y sin reemplazo.

a. Al considerar la distribución de las puntuaciones medias de IQ para las muestras de 50 niños, ¿debe $\sigma_{\bar{x}}$ ser corregida mediante el factor de corrección para una población finita? ¿Por qué sí o por qué no? ¿Cuál es el valor de $\sigma_{\bar{x}}$?

b. Determine la probabilidad de que el puntaje medio de IQ en la muestra de seguimiento esté entre 95 y 105.

6-5 Evaluación de la normalidad

Concepto clave Los siguientes capítulos incluyen métodos estadísticos importantes que requieren que los datos muestrales provengan de una población con distribución *normal*. En esta sección presentamos los criterios para determinar si se satisface el requisito de una distribución normal. Los criterios implican (1) inspección visual de un histograma para ver si tiene aproximadamente una forma de campana; (2) identificación de cualesquiera datos atípicos; y (3) construcción de una *gráfica cuantilar normal*.

PARTE 1 Conceptos básicos de la evaluación de la normalidad

Al intentar determinar si una colección de datos tiene una distribución que es aproximadamente normal, podemos inspeccionar visualmente un histograma para ver si tiene aproximadamente una forma de campana (como se indicó en la sección 2-2), también es posible identificar valores atípicos, además de usar una *gráfica cuantilar normal* (analizada brevemente en la sección 2-2).

DEFINICIÓN

Una **gráfica cuantilar normal** (o **gráfica de probabilidad normal**) es una gráfica de puntos (x, y) donde cada valor x proviene del conjunto original de datos muestrales y cada valor y es la puntuación z correspondiente que se espera de la distribución normal estándar.

Procedimiento para determinar si es razonable suponer que los datos muestrales provienen de una población con distribución normal

1. *Histograma:* Construya un histograma. Si el histograma se aparta dramáticamente de una forma de campana, concluya que los datos no tienen una distribución normal.
2. *Valores atípicos:* Identifique los valores atípicos. Si hay más de uno de estos valores, concluya que los datos podrían no tener una distribución normal. (Un solo valor atípico podría ser un error o resultado de una variación aleatoria, pero tenga cuidado, porque incluso un solo valor atípico puede tener un efecto dramático en los resultados).
3. *Gráfica cuantilar normal:* Si el histograma es básicamente simétrico y el número de valores atípicos es 0 o 1, utilice la tecnología para generar una *gráfica cuantilar normal*. Aplique los siguientes criterios para determinar si la distribución es normal. (Estos criterios se pueden utilizar de forma poco flexible para muestras pequeñas, pero deben usarse más estrictamente para muestras grandes).

   Distribución normal: La distribución de la población es normal si el patrón de los puntos está razonablemente cerca de una línea recta y los puntos no muestran algún patrón sistemático que no sea un patrón lineal.

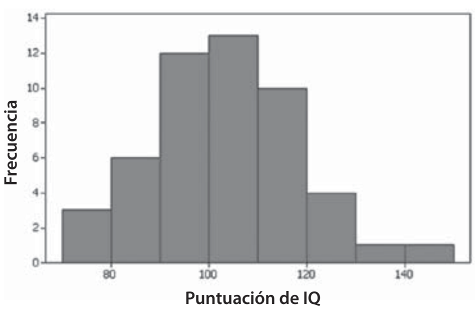
   No es una distribución normal: La distribución de la población *no* es normal si se aplica una o dos de las siguientes condiciones:

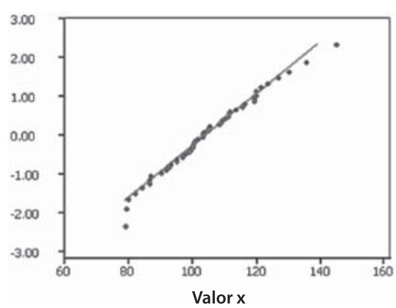
   - Los puntos no se encuentran razonablemente cerca de una línea recta.
   - Los puntos muestran algún *patrón sistemático* que no es un patrón en línea recta.

Histogramas y gráficas cuantilares normales

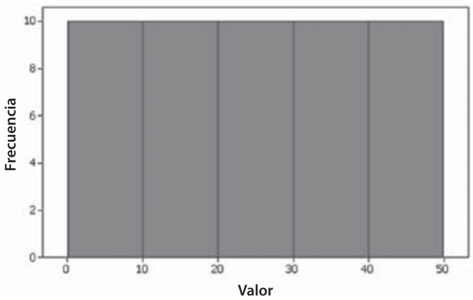
En la parte 2 de esta sección se describe el proceso de construcción de una gráfica cuantilar normal, pero por ahora nos centramos en la interpretación de una gráfica de este tipo. Las siguientes pantallas muestran histogramas de datos junto con las gráficas cuantilares normales correspondientes.

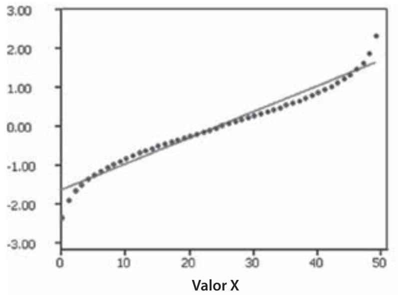
Normal: El primer caso muestra un histograma de puntuaciones de IQ que están cerca de tener una forma de campana, por lo que el histograma sugiere que las puntuaciones de IQ tienen una distribución normal. La gráfica cuantilar normal correspondiente muestra puntos que están razonablemente cerca de un patrón en línea recta, y los puntos no muestran ningún otro patrón sistemático diferente de una línea recta. Es seguro asumir que estas puntuaciones de IQ provienen de una población con distribución normal.



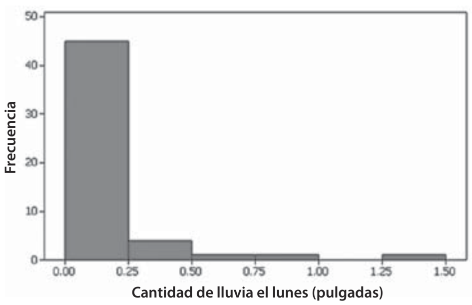


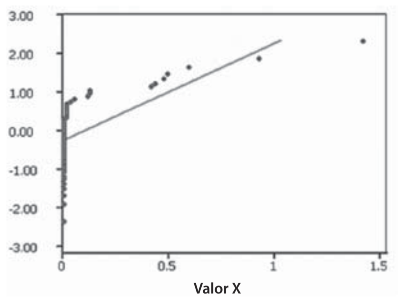
Uniforme: El segundo caso muestra un histograma de datos que tiene una distribución uniforme (plana). La gráfica cuantilar normal correspondiente sugiere que los puntos no se distribuyen normalmente. Aunque el patrón de puntos está razonablemente cerca de un patrón lineal, *hay otro patrón sistemático diferente del lineal*. Concluimos que estos valores muestrales son de una población que tiene una distribución que no es normal.





Asimétrica: El tercer caso muestra un histograma de la cantidad de lluvia (en pulgadas) en Boston para cada lunes durante un año. La forma del histograma es asimétrica hacia la derecha, no en forma de campana. La gráfica cuantilar normal correspondiente muestra puntos que no están en absoluto cerca de un patrón lineal. Estas cantidades de lluvia provienen de una población que tiene una distribución que no es normal.





Herramientas para determinar la normalidad

- **Histograma/valores atípicos:** Si el requisito de una distribución normal no es demasiado estricto, basta con mirar un histograma y encontrar el número de valores atípicos. Si el histograma tiene forma de campana y el número de valores atípicos es 0 o 1, considere que la población tiene una distribución normal.
- **Gráfica cuantilar normal:** Las gráficas cuantilares normales pueden ser difíciles de construir por su cuenta, pero pueden generarse con una calculadora TI-83/84 Plus o un software adecuado, como Statdisk, Minitab, Excel o StatCrunch.
- **Métodos avanzados:** Además de los procedimientos analizados en esta sección, existen otros procedimientos más avanzados para evaluar la normalidad, como las pruebas de bondad de ajuste de la chi-cuadrada, de Kolmogorov-Smirnov, de Lilliefors, de Anderson-Darling, de Jarque-Bera y de Ryan-Joiner (que se estudia brevemente en la parte 2).

PARTE 2 Construcción manual de gráficas cuantilares normales

El siguiente es un procedimiento relativamente sencillo para construir manualmente una gráfica cuantilar normal, y es el mismo procedimiento utilizado por Statdisk y la calculadora TI-83/84 Plus. Algunos paquetes estadísticos utilizan varios otros enfoques, pero la interpretación de la gráfica es esencialmente la misma.

Construcción manual de una gráfica cuantilar normal

Paso 1: Primero ordene los datos, en un arreglo de menor a mayor valor.

Paso 2: Con una muestra de tamaño n, cada valor representa una proporción de $1/n$ de la muestra. Utilizando el tamaño de muestra conocido n, encuentre los valores de $\frac{1}{2n}, \frac{3}{2n}, \frac{5}{2n}$, y así sucesivamente, hasta obtener n valores. Estos valores son las áreas acumuladas a la izquierda de los valores muestrales correspondientes.

Paso 3: Utilice la distribución normal estándar (software, una calculadora o la tabla A-2) para encontrar las puntuaciones z correspondientes a las áreas acumuladas desde la izquierda encontradas en el paso 2. (Estas son las puntuaciones z que se esperan de una muestra distribuida normalmente).

Paso 4: Haga coincidir los valores de los datos ordenados originales con las correspondientes puntuaciones z encontradas en el paso 3, luego grafique los puntos (x, y), donde cada x es un valor original de la muestra y y es la puntuación correspondiente.

Paso 5: Examine la gráfica cuantilar normal y utilice los criterios dados en la parte 1. Concluya que la población tiene una distribución normal si el patrón de puntos está razonablemente cerca de una línea recta y los puntos no muestran algún patrón sistemático diferente a una línea recta.

EJEMPLO 1 Tiempos de erupción del Old Faithful

El conjunto de datos 23 “Old Faithful” en el apéndice B incluye los tiempos de duración (segundos) de las erupciones del géiser Old Faithful. Considere la siguiente muestra de cinco tiempos de erupción: 125, 229, 236, 257, 234. Con sólo cinco valores, un histograma no será muy útil para revelar la distribución de los datos. En vez de eso, construya una gráfica cuantilar normal para estos cinco valores y determine si parecen proceder de una población que se distribuye normalmente.

SOLUCIÓN

Los siguientes pasos corresponden a los listados en el procedimiento anterior para construir una gráfica cuantilar normal.

Paso 1: Primero, ordene los datos de menor a mayor valor. Obtenemos 125, 229, 234, 236, 257.

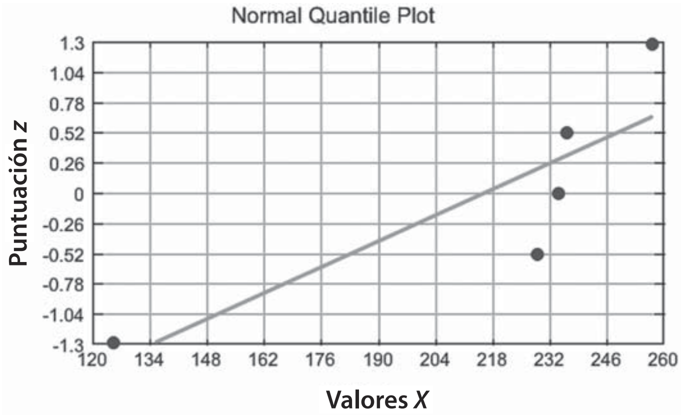
Paso 2: Con una muestra de tamaño $n = 5$, cada valor representa una proporción de 1/5 de la muestra, así que procedemos a identificar las áreas acumuladas a la izquierda de los valores muestrales correspondientes. Las áreas acumuladas desde la izquierda, que se expresan en general como $\frac{1}{2n}, \frac{3}{2n}, \frac{5}{2n}$, etcétera, se convierten en las siguientes áreas específicas para este ejemplo con $n = 5$: $\frac{1}{10}, \frac{3}{10}, \frac{5}{10}, \frac{7}{10}, \frac{9}{10}$. Estas áreas acumuladas a la izquierda expresadas en forma decimal son 0.1, 0.3, 0.5, 0.7 y 0.9.

Paso 3: Ahora utilizamos la tecnología (o la tabla A-2) con las áreas acumuladas a la izquierda de 0.1000. 0.3000, 0.5000, 0.7000 y 0.9000 para encontrar las puntuaciones z correspondientes: -1.28, -0.52, 0, 0.52 y 1.28. (Por ejemplo, la puntuación z de -1.28 tiene un área de 0.1000 a su izquierda).

Paso 4: Ahora emparejamos los tiempos originales de duración de las erupciones, en forma ordenada, con sus correspondientes puntuaciones z. Obtenemos las siguientes coordenadas (x, y), que se trazan en la pantalla de Statdisk adjunta:

$$(125, -1.28), (229, -0.52), (234, 0), (236, 0.52), (257, 1.28)$$

Statdisk



INTERPRETACIÓN

Examinamos la gráfica cuantilar normal en la pantalla de Statdisk. Los puntos no parecen estar razonablemente cerca de la línea recta, por lo que concluimos que la muestra de cinco tiempos de erupción no parece ser de una población normalmente distribuida.

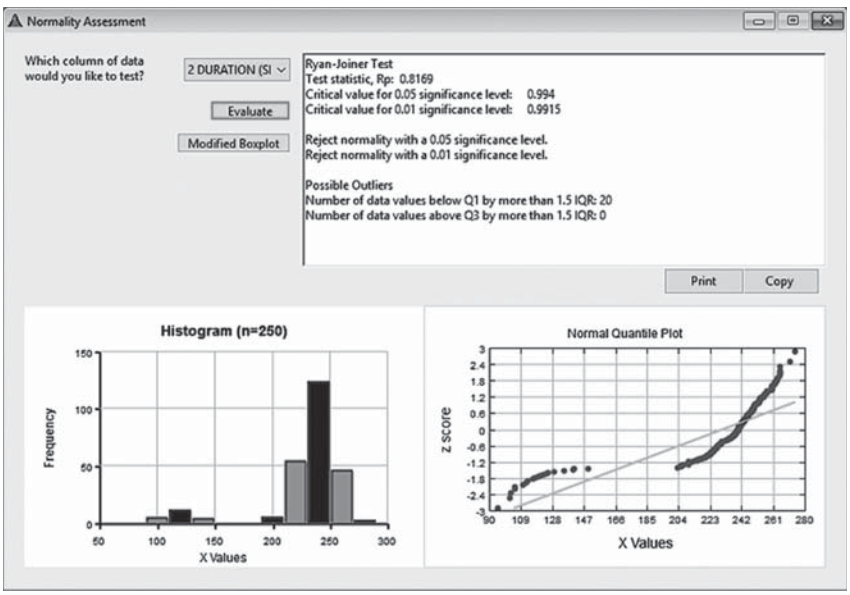
SU TURNO Resuelva el ejercicio 17 "Circunferencia del brazo femenino".

Prueba de Ryan-Joiner La prueba de Ryan-Joiner es una de varias pruebas formales de normalidad, cada una con sus propias ventajas y desventajas. Statdisk tiene una característica denominada **Normality Assessment** que muestra un histograma, una gráfica cuantilar normal, el número de posibles valores atípicos y los resultados de la prueba de Ryan-Joiner.

EJEMPLO 2 Tiempos de erupción del Old Faithful

En el ejemplo 1 se utilizaron sólo cinco de los tiempos de erupción listados en el conjunto de datos 23 "Old Faithful" del apéndice B. Podemos utilizar la función de **evaluación de la normalidad** de Statdisk con los 250 tiempos de erupción para obtener la pantalla adjunta.

Statdisk



continúa

Usaremos la pantalla con los tres criterios para evaluar la normalidad.

1. *Histograma:* Podemos ver que el histograma es *asimétrico* a la izquierda y está lejos de tener una forma de campana.
2. *Valores atípicos:* La pantalla muestra que hay 20 valores atípicos posibles. Si examinamos una lista ordenada de los 250 tiempos de erupción, los 20 tiempos más bajos parecen ser valores atípicos.
3. *Gráfica cuantilar normal*: ¡Increíble! Los puntos en la gráfica cuantilar normal están muy alejados de un patrón lineal. Concluimos que los 250 tiempos de erupción no parecen pertenecer a una población con distribución normal.

SU TURNO Resuelva el ejercicio 19 “Volúmenes cerebrales”.

Transformaciones de datos Muchos conjuntos de datos tienen una distribución que no es normal, pero podemos *transformar* los datos para que los valores modificados sí se distribuyan normalmente. Una transformación común consiste en convertir cada valor de x tomando su logaritmo. (Puede utilizar logaritmos naturales o logaritmos con base 10. Si cualquier valor original es 0, tome logaritmos de valores de $x + 1$). Si la distribución de los logaritmos de los valores es una distribución normal, la distribución de los valores originales se denomina **distribución lognormal**. Aparte de las transformaciones con logaritmos, hay otras transformaciones, como reemplazar cada valor de x por $\sqrt{x}$, o $1/x$, o x^2. Además de obtener una distribución normal requerida cuando los valores de datos originales no se distribuyen normalmente, tales transformaciones pueden usarse para corregir deficiencias, como un requisito (que se encuentra en capítulos posteriores) de que diferentes conjuntos de datos tengan la misma varianza.

CENTRO DE TECNOLOGÍA

Gráficas cuantilares normales

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Data** en el menú superior.
2. Seleccione **Normal Quantile Plot** en el menú desplegable.
3. Seleccione la columna de datos deseada y haga clic en **Plot**.

SUGERENCIA: Seleccione **Normality Assesment** en el menú desplegable bajo **Data** para obtener la gráfica cuantilar normal junto con otros resultados útiles para evaluar la normalidad.

Minitab

Minitab genera una gráfica de probabilidad que es similar a la gráfica cuantilar normal y puede ser interpretada usando los mismos criterios dados en esta sección.

Gráfica de probabilidad

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **Normality Test** en el submenú.
3. Seleccione la columna deseada en el cuadro *Variable* y haga clic en **OK**.

Gráfica de probabilidad con fronteras

1. Haga clic en **Graph** en el menú superior.
2. Seleccione **Probability Plot** en el menú desplegable, seleccione **Single** y haga clic en **OK**.
3. Seleccione la columna deseada en el cuadro *Graph Variables* y haga clic en **OK**.
4. Si todos los puntos están dentro de las fronteras, concluya que los datos se distribuyen normalmente. Si los puntos están fuera de las fronteras, concluya que los datos no se distribuyen normalmente.

StatCrunch

StatCrunch genera una gráfica QQ que es similar a la gráfica cuantilar normal y puede ser interpretada usando los mismos criterios dados en esta sección.

1. Haga clic en **Graph** en el menú superior.
2. Seleccione **QQ Plot** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Compute!**

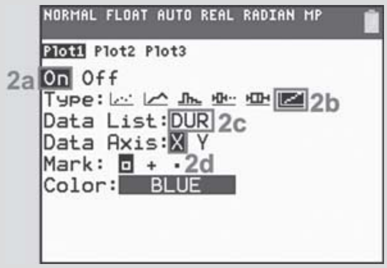
CENTRO DE TECNOLOGÍA *continuación*

Gráficas cuantilares normales

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

1. Abra el menú **STAT PLOTS** pulsando 2ND, Y=.
2. Presione ENTER para acceder a la pantalla de configuración de la Gráfica 1 como se muestra:
   a. Seleccione **ON** y pulse ENTER.
   b. Seleccione el último tipo de gráfico, pulse ENTER.
   c. Introduzca el nombre de la lista que contiene los datos.
   d. Para *Data Axis*, seleccione **X**.
3. Presione ZOOM y luego 9 (ZoomStat) para generar la gráfica cuantilar normal.
4. Presione WINDOW para personalizar la gráfica y, a continuación, presione GRAPH para ver la gráfica cuantilar normal.



Excel

Complemento XLSTAT (Requerido)

1. Haga clic en la ficha **XLSTAT** en la cinta de opciones y, a continuación, haga clic en **Describing Data**.
2. Seleccione **Normality Tests** en el menú desplegable.
3. Ingrese el rango de datos deseado en el cuadro de datos. Si la primera fila de datos contiene una etiqueta, marque la casilla **Sample Labels**.
4. Haga clic en la ficha **Charts** y confirme que está marcada la casilla **Normal Q-Q plots**.
5. Haga clic en **OK** y desplace hacia abajo los resultados para ver la gráfica Q-Q normal.

6-5 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

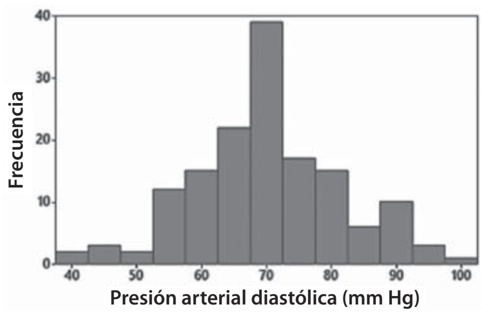
1. Gráfica cuantilar normal El conjunto de datos 1 "Datos corporales" en el apéndice B incluye las estaturas de 147 mujeres seleccionadas al azar, y las estaturas de las mujeres se distribuyen normalmente. Si se construye un histograma de las 147 estaturas del conjunto de datos 1, ¿qué forma espera que tenga el histograma? Si usted fuera a construir una gráfica cuantilar normal de esas mismas estaturas, ¿qué patrón esperaría ver en la gráfica?

2. Gráfica cuantilar normal Después de construir un histograma de las edades de las 147 mujeres incluidas en el conjunto de datos 1 "Datos corporales" del apéndice B, verá que el histograma está lejos de tener una forma de campana. ¿Qué sabe ahora sobre el patrón de puntos en la gráfica cuantilar normal?

3. Muestra pequeña El conjunto de datos 29 "Pesos de monedas" en el apéndice B incluye pesos de 20 monedas de un dólar. Dado que el tamaño de la muestra es menor de 30, ¿qué requisito se debe cumplir para tratar la media muestral como un valor de una población normalmente distribuida? Identifique tres herramientas para verificar ese requisito.

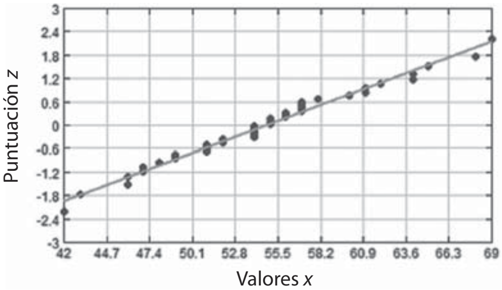
4. Evaluación de la normalidad El histograma adjunto se construye a partir de las mediciones de la presión arterial diastólica de las 147 mujeres incluidas en el conjunto de datos 1 "Datos corporales" del apéndice B. Si planea realizar más pruebas estadísticas y hay un requisito holgado de una población normalmente distribuida, ¿qué concluye usted sobre la distribución de la población basada en este histograma?

Minitab

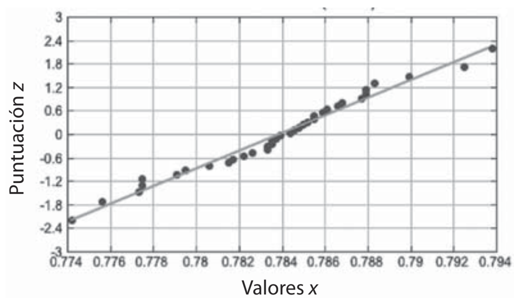


Interpretación de gráficas cuantilares normales. ***En los ejercicios 5 a 8, examine la gráfica cuantilar normal y determine si los datos de la muestra parecen ser de una población con distribución normal.***

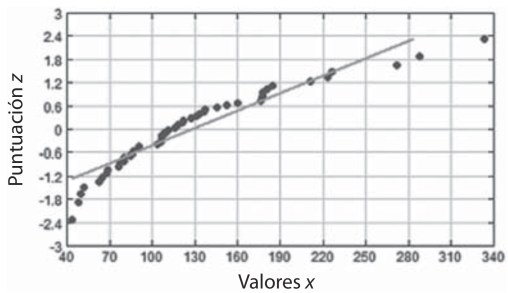
5. Edades de los presidentes La gráfica cuantilar normal representa las edades de los presidentes de Estados Unidos al momento de tomar sus cargos. Los datos provienen del conjunto de datos 15 “Presidentes” en el apéndice B.



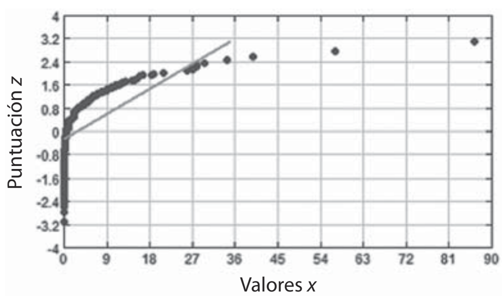
6. Diet Pepsi La gráfica cuantilar normal representa los pesos (libras) del contenido de la latas de Diet Pepsi en el conjunto de datos 26 “Pesos y volúmenes de bebidas de cola” en el apéndice B.



7. Tiempos de servicio en Dunkin’ Donuts La gráfica cuantilar normal representa los tiempos de servicio durante la cena en Dunkin’ Donuts (del conjunto de datos 25 “Comida rápida” en el apéndice B).



8. Tornados La gráfica cuantilar normal representa las distancias (en millas) que viajan los tornados (del conjunto de datos 22 “Tornados” en el apéndice B).



Determinación de la normalidad. ***En los ejercicios 9 a 12, refiérase a los datos muestrales indicados y determine si parecen ser de una población con una distribución normal. Suponga que este requisito es holgado en el sentido de que la distribución de la población no tiene que ser exactamente normal, sino que debe ser una distribución que parezca aproximadamente de campana.***

9. Galletas El número de chispas de chocolate en las galletas Chips Ahoy (reducidas en grasa), que se indica en el conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B.

10. Edades de las mejores actrices Las edades de actrices al momento de ganar el Oscar, como se indica en el conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B.

11. Basura Los pesos (en libras) de desechos de jardinería, según se listan en el conjunto de datos 31 "Peso de la basura" en el apéndice B.

12. Diet Coke Los pesos (en libras) de los contenidos en las latas de Diet Coke, como se listan en el conjunto de datos 26 "Pesos y volúmenes de bebidas de cola" en el apéndice B.

Uso de la tecnología para generar gráficas cuantilares normales. ***En los ejercicios 13 a 16, utilice los datos del ejercicio indicado en esta sección. Use software (como Statdisk, Minitab, Excel o StatCrunch) o una calculadora TI-83/84 Plus para generar una gráfica cuantilar normal. Después, determine si los datos provienen de una población normalmente distribuida.***

13. Ejercicio 9 "Galletas"

14. Ejercicio 10 "Edades de las mejores actrices"

15. Ejercicio 11 "Basura"

16. Ejercicio 12 "Diet Coke"

Construcción de gráficas cuantilares normales. ***En los ejercicios 17 a 20, use los valores de datos dados para identificar las puntuaciones z correspondientes que se usan para una gráfica cuantilar normal, luego identifique las coordenadas de cada punto en la gráfica, construya la gráfica, y después determine si los datos parecen ser de una población con distribución normal.***

17. Circunferencias del brazo femenino Una muestra de las circunferencias de los brazos (en cm) de las mujeres del conjunto de datos 1 "Datos corporales" en el apéndice B: 40.7, 44.3, 34.2, 32.5, 38.5.

18. Profundidad de terremotos Se obtiene una muestra de profundidades (en km) de los terremotos del conjunto de datos 21 "Terremotos" en el apéndice B: 17.3, 7.0, 7.0, 7.0, 8.1, 6.8.

19. Volúmenes cerebrales Se obtiene una muestra de volúmenes del cerebro humano (cm^3) de los listados en el conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B: 1027, 1029, 1034, 1070, 1079, 1079, 963, 1439.

20. Tiempos de servicio de McDonald's en la cena Una muestra de los tiempos de servicio en auto (segundos) en McDonald's durante la cena, como se indica en el conjunto de datos 25 "Comida rápida" del apéndice B: 84, 121, 119, 146, 266, 181, 123, 152, 162.

6-5 Más allá de lo básico

21. Transformaciones Las estaturas (en pulgadas) de los hombres que se listan en el conjunto de datos 1 "Datos corporales" del apéndice B tienen una distribución que es aproximadamente normal, por lo que parece que esas alturas provienen de una población normalmente distribuida.

a. Si se añaden 2 pulgadas a cada altura, ¿las nuevas estaturas también se distribuyen normalmente?

b. Si cada estatura se convierte de pulgadas a centímetros, ¿se distribuyen normalmente las estaturas en centímetros?

c. ¿Están normalmente distribuidos los logaritmos de las estaturas normalmente distribuidas?

22. Distribución Lognormal Los siguientes son valores del patrimonio neto (en miles de dólares) de los integrantes del gobierno de Estados Unidos. Pruebe la normalidad de estos valores, luego tome el logaritmo de cada valor y pruebe su normalidad. ¿Qué se puede concluir?

237,592 16,068 15,350 11,712 7304 6037 4483 4367 2658 1361 311

6-6 Distribución normal como una aproximación a la binomial

Concepto clave En la sección 5-2 se introdujeron las distribuciones de probabilidad binomiales, y en esta sección se presenta un método para usar una distribución normal como una aproximación a una distribución de probabilidad binomial, de modo que algunos problemas que involucran proporciones pueden resolverse usando una distribución normal. A continuación se describen los dos puntos principales de esta sección:

- Dadas las probabilidades p y q (donde $q = 1 - p$) y el tamaño de muestra n, si se satisfacen las condiciones $np \geq 5$ y $nq \geq 5$, entonces las probabilidades de una distribución de probabilidad binomial pueden aproximarse razonablemente bien usando una distribución normal con los siguientes parámetros:

$$\mu = np$$
$$\sigma = \sqrt{npq}$$

- La distribución de probabilidad binomial es *discreta* (con números enteros para la variable aleatoria x), pero la aproximación normal es *continua*. Para compensar, utilizamos una "corrección de continuidad" con un número entero x representado por el intervalo desde $x - 0.5$ hasta $x + 0.5$.

Breve repaso de la distribución de probabilidad binomial En la sección 5-2 vimos que una *distribución de probabilidad binomial* tiene (1) un número fijo de ensayos; (2) ensayos que son independientes; (3) ensayos que se clasifican en dos categorías comúnmente denominadas *éxito* y *fracaso*; y (4) ensayos con la propiedad de que la probabilidad de éxito permanece constante. La sección 5-2 también introdujo la siguiente notación.

Notación

n = el número fijo de ensayos

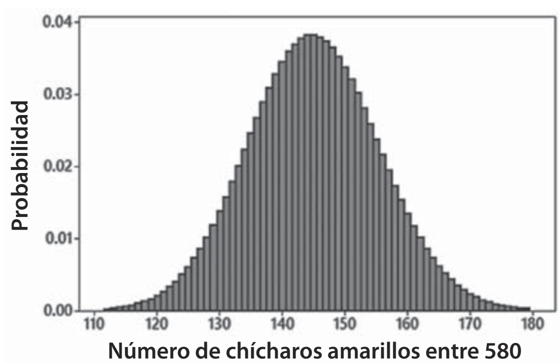
x = el número específico de éxitos en n ensayos

p = probabilidad de *éxito* en *uno* de los n ensayos

q = probabilidad de *fracaso* en *uno* de los n ensayos (por lo que $q = 1 - p$)

Justificación para el uso de una aproximación normal Vimos en la sección 6-3 que la distribución muestral de una proporción muestral tiende a aproximarse a una distribución normal. Además, vea el siguiente histograma de probabilidad para una distribución binomial con $n = 580$ y $p = 0.25$. (En uno de los famosos experimentos de hibridación de Mendel, esperaba que 25% de sus 580 chícharos fueran amarillos, pero obtuvo 152 de ellos, para una proporción de 26.2%). La forma de campana de esta gráfica sugiere que podemos usar un patrón de distribución normal para aproximar la distribución binomial.

Minitab



ELEMENTOS CLAVE

Distribución normal como aproximación a la distribución binomial

Requisitos

1. La muestra es una muestra aleatoria simple de tamaño n obtenida de una población en la que la proporción de éxitos es p, o la muestra es el resultado de la realización de n ensayos independientes de un experimento binomial en el que la probabilidad de éxito es p.
2. $np \geq 5$ y $nq \geq 5$.

(Los requisitos de $np \geq 5$ y $nq \geq 5$ son comunes, pero hay quienes recomiendan usar 10 en lugar de 5).

Aproximación normal

Si se cumplen los requisitos anteriores, entonces la distribución de probabilidad de la variable aleatoria x puede aproximarse mediante una distribución normal con los siguientes parámetros:

- $\mu = np$
- $\sigma = \sqrt{npq}$

Corrección de continuidad

Al usar la aproximación normal, ajuste el número entero discreto x usando una *corrección de continuidad* para que cualquier valor individual x esté representado en la distribución normal por el intervalo desde $x - 0.5$ hasta $x + 0.5$.

Procedimiento para utilizar una distribución normal para aproximar una distribución binomial

1. Compruebe los requisitos que $np \geq 5$ y $nq \geq 5$.
2. Encuentre $\mu = np$ y $\sigma = \sqrt{npq}$ para usarlos en la distribución normal.
3. Identifique el número entero discreto x que es relevante para el problema de probabilidad binomial que se está considerando, y represente ese valor por la región delimitada por $x - 0.5$ y $x + 0.5$.
4. Represente gráficamente la distribución normal y sombree el área deseada limitada por $x - 0.5$ o $x + 0.5$ según corresponda.

EJEMPLO 1 ¿Mendel se equivocó?

En uno de los famosos experimentos de hibridación de Mendel, esperaba que entre 580 chícharos, 145 de ellos (o 25%) sería amarillo, pero en realidad obtuvo 152 chícharos amarillos. Suponiendo que la proporción de Mendel de 25% es correcta, encuentre la probabilidad de obtener 152 o más chícharos amarillos por azar. Es decir, dadas $n = 580$ y $p = 0.25$, determine P(al menos 152 chícharos amarillos). ¿Son 152 chícharos amarillos significativamente altos?

SOLUCIÓN

Paso 1: Verificación de requisitos: Con $n = 580$ y $p = 0.25$, obtenemos $np = (580)(0.25) = 145$ y $nq = (580)(0.75) = 435$, entonces se satisfacen los requisitos de que $np \geq 5$ y $nq \geq 5$.

Paso 2: Ahora encontramos los valores de μ y σ necesarios para la distribución normal:

$$\mu = np = 580 \cdot 0.25 = 145$$

$$\sigma = \sqrt{npq} = \sqrt{580 \cdot 0.25 \cdot 0.75} = 10.4283$$

continúa

Paso 3: Queremos la probabilidad de al menos 152 chícharos amarillos, por lo que el número entero discreto relevante para este ejemplo es $x = 152$. Utilizamos la corrección de continuidad mientras representamos el valor discreto de 152 en la gráfica de la distribución normal por el intervalo entre 151.5 y 152.5 (como se muestra en la parte superior de la figura 6-25).

Paso 4: Vea la parte inferior de la figura 6-25, que muestra la distribución normal y el área a la derecha de 151.5 (que representa "152 o más" chícharos amarillos).

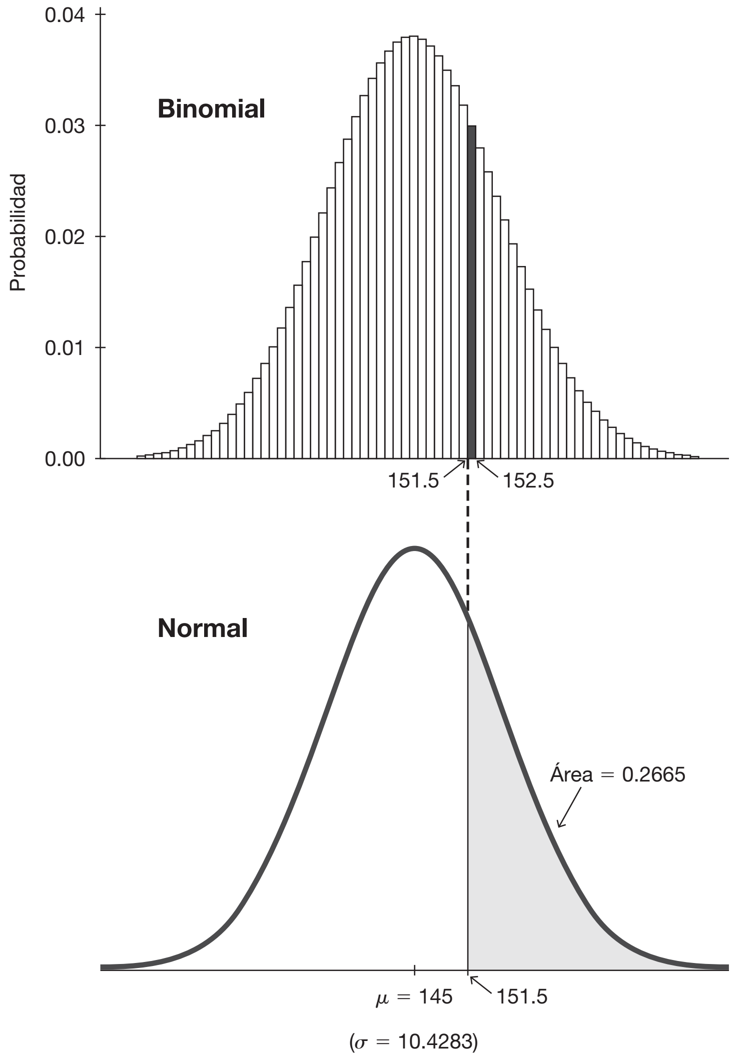


FIGURA 6-25 Número de chícharos amarillos entre 580

Queremos conocer el área a la derecha de 151.5 en la parte inferior de la figura 6-25.

Tecnología: Si se utiliza tecnología, encontramos que el área sombreada es 0.2665.

Tabla A-2: Si se usa la tabla A-2, primero debemos encontrar la puntuación z usando $x = 151.5$, $\mu = 145$ y $\sigma = 10.4283$ como sigue:

$$z = \frac{x - \mu}{\sigma} = \frac{151.5 - 145}{10.4283} = 0.62$$

Con base en la tabla A-2, encontramos que $z = 0.62$ corresponde a un área acumulada a la izquierda de 0.7324, por lo que la región sombreada en la porción inferior de la figura 6-25 es $1 - 0.7324 = 0.2676$. (El resultado de 0.2665 de la tecnología es más exacto).

INTERPRETACIÓN

El resultado de Mendel de 152 chícharos amarillos es mayor que los 145 que esperaba con su teoría de los híbridos, pero con P(152 o más chícharos amarillos) = 0.2665, vemos que 152 chícharos amarillos *no son significativamente altos*. Ése es un resultado que podría ocurrir fácilmente con una proporción verdadera de 25% para los chícharos amarillos. Este experimento no contradice la teoría de Mendel.

SU TURNO Realice el ejercicio 9 "Autos blancos".

Corrección de continuidad

DEFINICIÓN

Cuando usamos la distribución normal (que es una distribución de probabilidad *continua*) como una aproximación a la distribución binomial (que es *discreta*), se hace una **corrección de continuidad** a un número entero discreto x en la distribución binomial, representando x por el *intervalo* de $x - 0.5$ a $x + 0.5$ (es decir, sumando y restando 0.5).

El ejemplo 1 utilizó una corrección de continuidad cuando el valor discreto de 152 se representó en la distribución normal por el área entre 151.5 y 152.5. Debido a que queríamos la probabilidad de "152 o más" chícharos amarillos, usamos el área a la derecha de 151.5. Los siguientes son otros usos de la corrección de continuidad:

Enunciado sobre el valor *discreto*	Área de la distribución normal *continua*
Al menos 152 (incluye 152 y mayores)	A la *derecha* de 151.5
Más de 152 (no incluye 152)	A la *derecha* de 152.5
A lo más 152 (incluye 152 y menores)	A la *izquierda* de 152.5
Menos de 152 (no incluye 152)	A la *izquierda* de 151.5
Exactamente 152	Entre 151.5 y 152.5

EJEMPLO 2 Exactamente 252 chícharos amarillos

Usando la misma información del ejemplo 1, encuentre la probabilidad de *exactamente* 152 chícharos amarillos entre los 580 chícharos descendientes. Es decir, dada $n = 580$ y suponiendo que $p = 0.25$, encuentre P(exactamente 152 chícharos amarillos). ¿Es útil este resultado para determinar si 152 chícharos amarillos son *significativamente altos*?

SOLUCIÓN

Vea la figura 6-26 en la página siguiente, que muestra la distribución normal con $\mu = 145$ y $\sigma = 10.4283$. El área sombreada aproxima la probabilidad de *exactamente* 752 chícharos amarillos. Esa región es la franja vertical entre 151.5 y 152.5, como se muestra. Podemos encontrar esa área usando los mismos métodos introducidos en la sección 6-2.

Tecnología: Si se usa tecnología, el área sombreada es 0.0305.

Tabla A-2: Usando la tabla A-2, convertimos 151.5 y 152.5 en $z = 0.62$ y $z = 0.72$, que producen áreas acumuladas a la izquierda de 0.7324 y 0.7642. Debido a que ambos son áreas acumuladas a la izquierda, la región sombreada en la figura 6-26 es $0.7642 - 0.7324 = 0.0318$. La probabilidad de exactamente 152 chícharos amarillos es 0.0318.

continúa

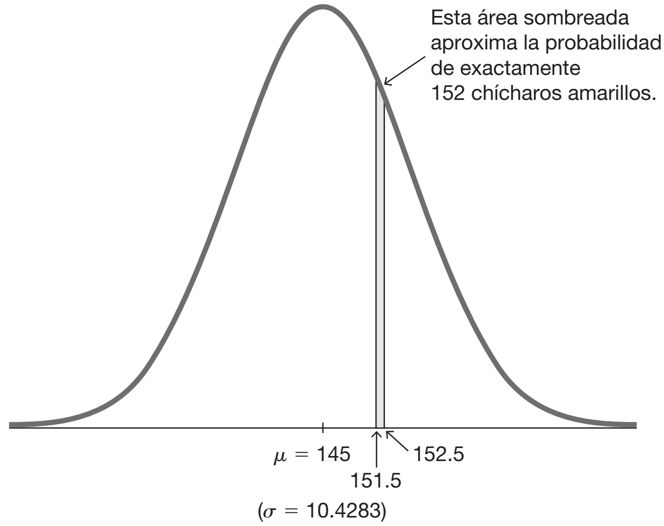
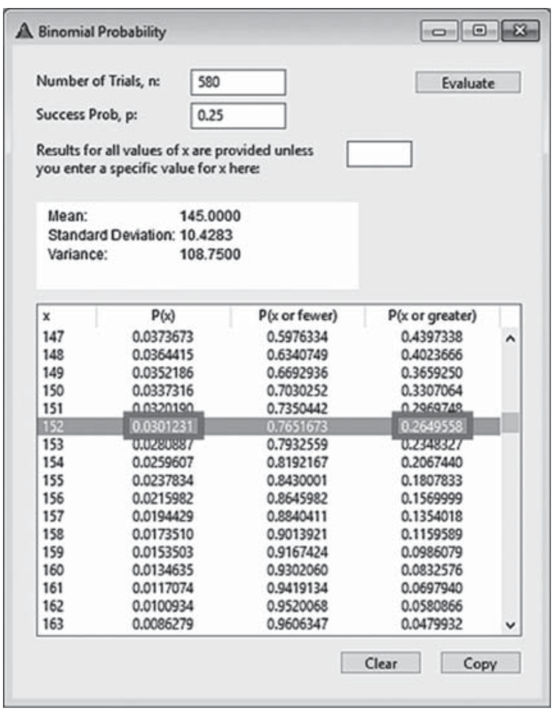


FIGURA 6-26 Probabilidad de exactamente 152 chícharos amarillos

INTERPRETACIÓN

En la sección 4-1 vimos que x éxitos entre n ensayos son significativamente altos si la probabilidad de *x o más* éxitos es improbable con una probabilidad de 0.05 o menos. Para determinar si el resultado de Mendel de 152 chícharos amarillos contradice su teoría de que el 25% de la descendencia deben ser chícharos amarillos, debemos considerar la probabilidad de 152 *o más* chícharos amarillos, no la probabilidad de exactamente 152 chícharos. El resultado de 0.0305 no es la probabilidad relevante; ésta es 0.2665, que se encontró en el ejemplo 1. En general, el resultado relevante es la probabilidad de obtener un resultado *al menos tan extremo* como el obtenido.

SU TURNO Resuelva el ejercicio 11 “Autos rojos”.



Tecnología para probabilidades binomiales

Este tema de usar una distribución normal para aproximar una distribución binomial fue bastante importante en su momento, pero ahora podemos usar la tecnología para encontrar probabilidades binomiales que antiguamente estaban fuera de nuestras capacidades. Por ejemplo, vea la pantalla de Statdisk adjunta que muestra que para el ejemplo 1 la probabilidad de 152 o más chícharos amarillos es de 0.2650, y para el ejemplo 2 la probabilidad de exactamente 152 chícharos amarillos es 0.0301, por lo que no existe una necesidad real de usar una aproximación. Sin embargo. hay casos donde es necesario utilizar una aproximación normal, y la sección 8-2 usa una aproximación normal a una distribución binomial para un método estadístico importante introducido en esa sección.

6-6 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Corrección de continuidad Al probar la hipótesis de que la probabilidad de tener un niño es de 0.512, un genetista obtiene una muestra aleatoria de 1000 nacimientos y descubre que 502 de ellos son varones. Utilizando la corrección de continuidad, describa el área bajo la gráfica de una distribución normal correspondiente a lo siguiente. (Por ejemplo, el área correspondiente a “la probabilidad de al menos 502 niños” es el área a la derecha de 501.5)

a. La probabilidad de 502 o menos niños

b. La probabilidad de exactamente 502 niños

c. La probabilidad de más de 502 niños

2. Comprobación de requisitos Las pruebas comunes como SAT, ACT, la prueba de Admisión a la Facultad de Derecho (LSAT, por sus siglas en inglés) y la Prueba de Admisión a la Facultad de Medicina (MCAT, por sus siglas en inglés) usan preguntas de opción múltiple, cada una con posibles respuestas de a, b, c, d, e; y cada pregunta tiene sólo una respuesta correcta. Queremos encontrar la probabilidad de obtener al menos 25 respuestas correctas para alguien que hace conjeturas aleatorias en sus respuestas a un bloque de 100 preguntas. Si planeamos utilizar los métodos de esta sección usando una distribución normal para aproximar una distribución binomial, ¿se satisfacen los requisitos necesarios? Explique.

3. Notación Las pruebas comunes como SAT, ACT, LSAT y MCAT usan preguntas de opción múltiple, cada una con posibles respuestas de a, b, c, d, e; y cada pregunta tiene sólo una respuesta correcta. Para las personas que hacen conjeturas aleatorias al dar sus respuestas a un bloque de 100 preguntas, identifique los valores de p, q, μ y σ. ¿Cuánto miden μ y σ?

4. Distribución de proporciones Cada semana, Nielsen Media Research realiza una encuesta a 5000 hogares y registra la proporción de hogares que sintonizan el programa *60 minutos*. Si obtenemos una gran colección de esas proporciones y construimos un histograma con ellas, ¿cuál es la forma aproximada del histograma?

Uso de la aproximación normal. ***En los ejercicios 5 a 8, haga lo siguiente: Si se satisfacen los requisitos $np \geq 5$ y $nq \geq 5$, estime la probabilidad indicada usando la distribución normal como una aproximación a la distribución binomial; si $np < 5$ o $nq < 5$, indique que la aproximación normal no debe ser usada.***

5. Nacimientos de niños Con $n = 20$ nacimientos y $p = 0.512$ de tener un niño, encuentre P(menos de 8 niños).

6. Nacimientos de niños Con $n = 8$ nacimientos y $p = 0.512$ de tener un niño, encuentre P(exactamente 5 niños).

7. Conjeturas en pruebas estándar Con $n = 20$ conjeturas y $p = 0.2$ de dar una respuesta correcta, busque P(al menos 6 respuestas correctas).

8. Conjeturas en pruebas estándar Con $n = 50$ conjeturas y $p = 0.2$ de dar una respuesta correcta, encuentre P(exactamente 12 respuestas correctas).

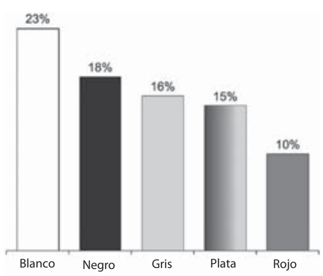
Colores de auto. ***En los ejercicios 9 a 12, suponga que se seleccionan 100 autos al azar. Consulte la gráfica adjunta, la cual muestra los colores más populares de automóvil y los porcentajes de autos con esos colores (de acuerdo con PPG Industries).***

9. Autos blancos Encuentre la probabilidad de que menos de 20 autos sean blancos. ¿Es 20 un número significativamente bajo de autos blancos?

10. Autos negros Encuentre la probabilidad de que al menos 25 autos sean negros. ¿Es 25 un número significativamente alto de autos negros?

11. Autos rojos Encuentre la probabilidad de que exactamente 14 autos sea rojos. ¿Por qué no se puede utilizar el resultado para determinar si 14 es un número significativamente alto de autos rojos?

12. Autos grises Encuentre la probabilidad de que exactamente 10 autos sean grises. ¿Por qué no se puede utilizar el resultado para determinar si 10 es un número significativamente bajo de autos grises?



Colores populares de autos

13. Repetición en el tenis En el año en que se escribió este ejercicio, se hicieron 879 desafíos a las decisiones del árbitro en los partidos individuales de tenis. Entre esos desafíos, 231 resultaron en una decisión revocada. Supongamos que en general, 25% de los desafíos concluyen exitosamente con una decisión revocada.

a. Si la proporción de 25% es correcta, encuentre la probabilidad de que entre los 879 desafíos, el número de decisiones revocadas sea exactamente 231.

b. Si la proporción de 25% es correcta, encuentre la probabilidad de que entre los 879 desafíos, el número de decisiones revocadas sea 231 o más. ¿Son 231 decisiones revocadas entre 879 desafíos un resultado que es significativamente alto?

14. Repetición en el tenis Repita el ejercicio anterior después de cambiar la proporción supuesta de decisiones revocadas de 25% a 22%.

15. Teléfonos inteligentes Con base en una encuesta de los teléfonos inteligentes LG, suponga que 51% de los adultos con teléfonos inteligentes los utilizan en cines. En un estudio por separado de 250 adultos con teléfonos inteligentes, se encuentra que 109 los utilizan en cines.

a. Si la proporción de 51% es correcta, encuentre la probabilidad de obtener 109 o menos propietarios de teléfonos inteligentes que los usen en cines.

b. ¿El resultado de 109 es significativamente bajo?

16. Color de los ojos Con base en un estudio realizado por el Dr. P Sorita de la Universidad de Indiana, suponemos que 12% de los estadounidenses tiene ojos verdes. En un estudio de 650 personas, se encuentra que 86 de ellos tienen ojos verdes.

a. Encuentre la probabilidad de obtener al menos 86 personas con ojos verdes entre 650 personas seleccionadas al azar.

b. ¿Son 86 personas con ojos verdes significativamente altas?

17. Genética mendeliana Cuando Mendel condujo sus famosos experimentos genéticos con chícharos, una muestra de descendientes constaba de 929 chícharos, de los cuales 705 tenían flores rojas. Si suponemos, como lo hizo Mendel, que bajo estas circunstancias hay una probabilidad de 3/4 de que un chícharo tenga una flor roja, esperaríamos que 696.75 (o aproximadamente 697) de los chícharos tuvieran flores rojas, por lo que el resultado de 705 chícharos con flores rojas es más de lo esperado.

a. Si la probabilidad supuesta por Mendel es correcta, encuentre la probabilidad de obtener 705 o más chícharos con flores rojas.

b. ¿Son 705 chícharos con flores rojas significativamente altos?

c. ¿Qué sugieren estos resultados acerca del supuesto de Mendel de que 3/4 de los chícharos tendrán flores rojas?

18. Sonambulismo Supongamos que 29.2% de las personas han sufrido sonambulismo (según "Prevalencia y comorbilidad del sonambulismo en la población general adulta de Estados Unidos", de Ohayon *et al*., en *Neurology*, vol. 78, núm. 20). Suponga que en una muestra aleatoria de 1480 adultos, 455 han sufrido sonambulismo.

a. Suponiendo que la proporción de 29.2% es correcta, determine la probabilidad de que 455 o más de los 1480 adultos hayan experimentado sonambulismo.

b. ¿Es ese resultado de 455 o más significativamente alto?

c. ¿Qué sugiere el resultado sobre la proporción de 29.2%?

19. ¿Los votantes mienten? En una encuesta a 1002 personas, 701 dijeron que votaron en una reciente elección presidencial (según datos del ICR Research Group). Los registros electorales mostraron que, en realidad, votaron 61% de los votantes elegibles.

a. Dado que 61% de los votantes elegibles realmente votaron, encuentre la probabilidad de que entre 1002 votantes elegibles al azar, al menos 701 realmente votaran.

b. ¿Qué sugiere el resultado?

20. Teléfonos celulares y cáncer cerebral En un estudio de 420,095 usuarios de teléfonos celulares en Dinamarca, se encontró que 135 desarrollaron cáncer en el cerebro o en el sistema nervioso. Para aquellos que no usan teléfonos celulares, existe una probabilidad de 0.000340 de que una persona desarrolle cáncer cerebral o del sistema nervioso. Por lo tanto, esperamos unos 143 casos de estos tipos de cáncer en un grupo de 420,095 personas seleccionadas al azar.

a. Encuentre la probabilidad de 135 o menos casos de estos tipos de cáncer en un grupo de 420,095 personas.

b. ¿Qué sugieren estos resultados sobre los informes de los medios que sugieren que los teléfonos celulares causan cáncer cerebral o del sistema nervioso?

6-6 Más allá de lo básico

21. Nacimientos La probabilidad de que un bebé nazca siendo niño es 0.512. Considere el problema de encontrar la probabilidad de exactamente 7 niños en 11 nacimientos. Resuelva el problema usando (1) la aproximación normal a la binomial empleando la tabla A-2; (2) la aproximación normal a la binomial empleando tecnología, en lugar de la tabla A-2; (3) usando la tecnología con la distribución binomial en vez de emplear una aproximación normal. Compare los resultados. Dado que los requisitos para el uso de la aproximación normal se cumplen con dificultad, ¿son las aproximaciones muy alejadas?

22. Sobreventa de un Boeing 767-300 Un avión Boeing 767-300 tiene 213 asientos. Cuando alguien compra un boleto para un vuelo, hay una probabilidad de 0.0995 de que la persona no se presente al vuelo (según datos de un trabajo de investigación de IBM de Lawrence, Hong y Cherrier). ¿Cuántas reservaciones podrían aceptarse para un Boeing 767-300 de modo que haya al menos una probabilidad de 0.95 de que todos los poseedores de reservaciones que lleguen al vuelo encuentren lugar?

Examen rápido del capítulo

Prueba de densidad ósea. ***En los ejercicios 1 a 4, suponga que las puntuaciones en un test de densidad mineral ósea se distribuyen normalmente con una media de 0 y una desviación estándar de 1.***

1. Densidad ósea Trace una gráfica que muestre la forma de la distribución de los resultados de la prueba de densidad ósea.

2. Densidad ósea Encuentre la puntuación que separa el 9% más bajo de las puntuaciones del 91% más alto.

3. Densidad ósea Para un sujeto seleccionado al azar, encuentre la probabilidad de una puntuación mayor que -2.93.

4. Densidad ósea Para un sujeto seleccionado al azar, encuentre la probabilidad de una puntuación entre 0.87 y 1.78.

5. Notación

a. Identifique los valores de μ y σ para la distribución normal estándar.

b. ¿Qué representan los símbolos $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$?

En los ejercicios 6 a 10, suponga que las mujeres tienen medidas de presión arterial diastólica que se distribuyen normalmente con una media de 70.2 mm Hg y una desviación estándar de 11.2 mm Hg (de acuerdo con el conjunto de datos 1 “Datos corporales” en el apéndice B).

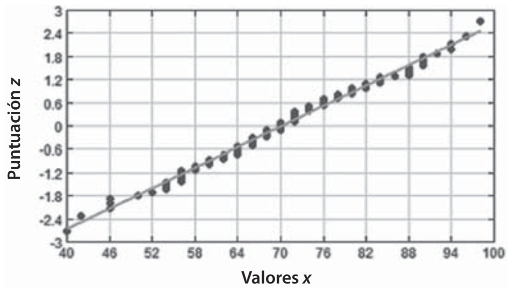
6. Presión arterial diastólica Encuentre la probabilidad de que una mujer seleccionada al azar tenga un nivel normal de presión arterial diastólica, que es inferior a 80 mm Hg.

7. Presión arterial diastólica Encuentre la probabilidad de que una mujer seleccionada al azar tenga un nivel de presión arterial diastólica entre 60 mm Hg y 80 mm Hg.

8. Presión arterial diastólica Encuentre P_{90}, el percentil 90 para los niveles de presión arterial diastólica de las mujeres.

9. Presión arterial diastólica Si 16 mujeres se seleccionan al azar, encuentre la probabilidad de que la media de sus niveles de presión arterial diastólica sea inferior a 75 mm Hg.

10. Presión arterial diastólica La gráfica cuantilar normal adjunta se construyó a partir de los niveles de presión arterial diastólica de una muestra de mujeres. ¿Qué sugiere esta gráfica sobre los niveles de presión arterial diastólica en las mujeres?



Ejercicios de repaso

1. Prueba de densidad ósea Se utiliza una prueba de densidad mineral ósea para identificar una enfermedad en los huesos. El resultado de una prueba de densidad ósea se mide comúnmente como una puntuación z, y la población de puntuaciones z se distribuye normalmente con una media de 0 y una desviación estándar de 1.

a. Para un sujeto seleccionado al azar, determine la probabilidad de una puntuación en la prueba de densidad ósea menor de 1.54.

b. Para un sujeto seleccionado al azar, encuentre la probabilidad de una puntuación en la prueba de densidad ósea mayor de -1.54.

c. Para un sujeto seleccionado al azar, determine la probabilidad de una puntuación en la prueba de densidad ósea entre -1.33 y 2.33.

d. Determine Q_1, la puntuación de la prueba de densidad ósea que separa el 25% inferior del 75% superior.

e. Si se encuentra la puntuación media de la prueba de densidad ósea para 9 sujetos seleccionados al azar, determine la probabilidad de que la media sea mayor que 0.50.

2. Seguridad biométrica Al diseñar un sistema de seguridad basado en el reconocimiento de ojos (iris), debemos considerar la altura de los ojos de las mujeres, que normalmente se distribuyen con una media de 59.7 pulgadas y una desviación estándar de 2.5 pulgadas (según datos antropométricos de Gordon Churchill y otros).

a. Si un sistema de seguridad de reconocimiento de ojos está situado a una altura que resulta incómoda para aquellas mujeres que tienen una altura de sus ojos menor a 54 pulgadas, ¿qué porcentaje de las mujeres encuentra la altura incómoda?

b. Al posicionar el sistema de seguridad de reconocimiento de ojos, queremos que sea adecuado para el 95% más bajo en cuanto a las alturas de los ojos de las mujeres. ¿Qué altura de los ojos de las mujeres separa el 95% más bajo del 5% más alto?

3. Seguridad biométrica La altura de los ojos de los hombres normalmente se distribuye con una media de 64.3 pulgadas y una desviación estándar de 2.6 pulgadas (según datos antropométricos de la encuesta de Gordon, Churchill y otros).

a. Si un sistema de seguridad de reconocimiento ocular está situado a una altura que es incómoda para aquellos hombres que tienen una altura de los ojos de más de 70 pulgadas, ¿qué porcentaje de los hombres encuentra esa altura incómoda?

b. Al posicionar el sistema de seguridad de reconocimiento ocular, queremos que sea adecuado para el 98% más alto en cuanto a la altura de los ojos de los hombres. ¿Qué altura de los ojos de los hombres separa al 98% más alto del 2% más bajo?

4. Distribuciones de muestreo Las puntuaciones en la Escala de Puntuación de Autismo de Gilliam (GARS) se distribuyen normalmente con una media de 100 y una desviación estándar de 15. Se selecciona al azar una muestra de 64 puntuaciones de GARS y se calcula la media de la muestra.

a. Describa la distribución de dichas medias muestrales.

b. ¿Cuál es la media de todas estas medias muestrales?

c. ¿Cuál es la desviación estándar de todas las medias muestrales?

5. Estimadores no sesgados

a. ¿Qué es un estimador no sesgado?

b. Para los siguientes estadísticos, identifique aquellos que son estimadores no sesgados: media, mediana, rango, varianza, proporción.

c. Determine si la siguiente afirmación es verdadera o falsa: "La desviación estándar muestral es un estimador sesgado, pero el sesgo es relativamente pequeño en muestras grandes, por lo que s se utiliza a menudo para estimar σ".

6. Monorriel de Disney El monorriel Mark VI utilizado en Disney World tiene puertas con una altura de 72 pulgadas. Las alturas de los hombres se distribuyen normalmente con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas (de acuerdo con el conjunto de datos I "Datos corporales" en el apéndice B).

a. ¿Qué porcentaje de hombres adultos puede pasar a través de las puertas sin inclinarse? ¿El diseño de las puertas con una altura de 72 pulgadas parece ser adecuado? Explique.

b. ¿Qué altura de la puerta permitiría que 99% de los hombres adultos entraran sin inclinarse?

7. Monorriel de Disney Considere el mismo monorriel Mark VI descrito en el ejercicio anterior. Una vez más, suponga que las alturas de los hombres se distribuyen normalmente con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas.

a. Al determinar la idoneidad de la altura de la puerta del monorriel, ¿por qué tiene sentido considerar a los hombres mientras que las mujeres son ignoradas?

b. Los monorrieles Mark VI tienen una capacidad de 60 pasajeros. Si un vagón está cargado con 60 hombres seleccionados al azar, ¿cuál es la probabilidad de que su estatura media sea menor de 72 pulgadas?

c. ¿Por qué no se puede utilizar el resultado del inciso (b) para determinar qué tan bien la altura de la puerta se ajusta a los hombres?

8. Evaluación de la normalidad A continuación se listan los salarios recientes (en millones de dólares) de los jugadores el equipo de baloncesto profesional San Antonio Spurs.

a. ¿Parece que estos salarios provienen de una población que tiene una distribución normal? ¿Por qué sí o por qué no?

b. ¿Se puede tratar la media de esta muestra como un valor de una población que tiene una distribución normal? ¿Por qué sí o por qué no?

12.5	10.4	9.3	7.0	4.0	3.1	2.9	2.1	1.8	1.4	1.1	1.1	0.9

9. Experimento de hibridación En uno de los experimentos de Mendel con plantas, 1064 descendientes consistieron en 787 plantas con tallos largos. De acuerdo con la teoría de Mendel, 3/4 de las plantas descendientes debían tener tallos largos. Suponiendo que la proporción de Mendel de 3/4 es correcta, encuentre la probabilidad de obtener 787 o menos plantas con tallos largos entre 1064 plantas descendientes. Con base en el resultado, ¿son 787 plantas descendientes con tallos largos significativamente bajas? ¿Qué implica el resultado sobre la proporción de Mendel de 3/4?

10. Clubes de altos La organización social internacional Clubes de Altos tiene el requisito de que las mujeres deben tener al menos 70 pulgadas de altura. Suponga que las mujeres tienen estaturas normalmente distribuidas con una media de 63.7 pulgadas y una desviación estándar de 2.9 pulgadas (de acuerdo con el conjunto de datos 1 "Datos corporales" en el apéndice B).

a. Encuentre el porcentaje de mujeres que satisfacen el requisito de estatura.

b. Si se cambia el requisito de estatura para que el 2.5% de las mujeres más altas sea elegible, ¿cuál es el nuevo requisito de altura?

Ejercicios de repaso acumulado

En los ejercicios 1 a 3, utilice los siguientes salarios anuales recientes (en millones de dólares) para los jugadores del equipo de baloncesto profesional NY Knicks.

23.4 22.5 11.5 7.1 6.0 4.1 3.3 2.8 2.6 1.7 1.6 1.3 0.9 0.9 0.6

1. Salarios de los NY Knicks

a. Determine la media $\bar{x}$.

b. Encuentre la mediana.

c. Determine la desviación estándar s.

d. Encuentre la varianza.

e. Convierta el salario más alto en una puntuación z.

f. ¿Qué nivel de medición (nominal, ordinal, de intervalo, de razón) describe este conjunto de datos?

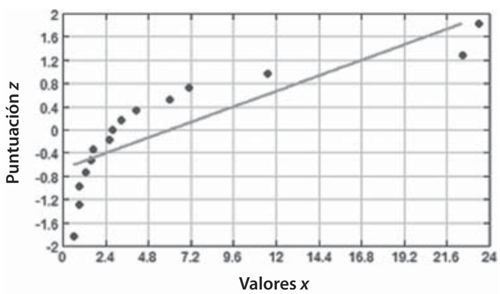
g. ¿Son los salarios datos discretos o datos continuos?

2. Salarios de los NY Knicks

a. Encuentre Q_1, Q_2 y Q_3.

b. Construya una gráfica de caja y bigotes.

c. Con base en la gráfica cuantilar normal de los salarios, ¿qué conclusión obtiene sobre la muestra de salarios?



3. Salarios de los NY Knicks Redondee cada uno de los salarios al millón de dólares más cercano, luego construya una gráfica de puntos. ¿Los valores parecen ser de una población que tiene una distribución normal?

4. Ojos azules Suponga que 35% de los estadounidenses tienen ojos azules (basado en un estudio del Dr. P. Soria en la Universidad de Indiana).

a. Sea B el evento de seleccionar a alguien con ojos azules. ¿Qué denota el evento $\overline{B}$?

b. Determine el valor de $P(\overline{B})$.

c. Encuentre la probabilidad de seleccionar al azar a tres personas diferentes y descubrir que todas tienen ojos azules.

d. Encuentre la probabilidad de que entre 100 personas seleccionadas al azar, al menos 40 tengan ojos azules.

e. Si 35% de los estadounidenses realmente tiene ojos azules, ¿el resultado de 40 personas con ojos azules entre 100 personas seleccionadas al azar es un resultado significativamente alto?

5. Longitudes de los pies de mujeres Suponga que las longitudes de los pies de las mujeres se distribuyen normalmente con una media de 9.6 pulgadas y una desviación estándar de 0.5 pulgadas, con base en datos de la Encuesta Antropométrica del Ejército de Estados Unidos (ANSUR, por sus siglas en inglés).

a. Encuentre la probabilidad de que una mujer seleccionada al azar tenga una longitud de pie inferior a 10.0 pulgadas.

b. Determine la probabilidad de que una mujer seleccionada al azar tenga una longitud de pie entre 8.0 y 11.0 pulgadas.

c. Encuentre P_{95}.

d. Determine la probabilidad de que 25 mujeres tengan longitudes de los pies con una media mayor que 9.8 pulgadas.

Proyectos de tecnología

Algunos métodos en este capítulo son fáciles de ejecutar con tecnología pero muy difíciles de aplicar sin ella. Los dos proyectos que siguen ilustran lo fácil que es usar la tecnología para evaluar la normalidad y encontrar probabilidades binomiales.

1. Evaluación de la normalidad A menudo es necesario determinar si los datos muestrales parecen ser de una población normalmente distribuida, y esa determinación se apoya en la construcción de un histograma y una gráfica cuantilar normal. Consulte el conjunto de datos 1 "Datos corporales" en el apéndice B. Para cada una de las 13 columnas de datos (sin incluir la edad o el sexo), determine si los datos parecen provenir de una población normalmente distribuida. Utilice Statdisk o cualquier otra tecnología. (Descargue una copia gratuita de Statdisk desde www.statdisk.org)

2. Probabilidades binomiales La sección 6-6 describe un método para usar una distribución normal con el objeto de aproximar una distribución binomial. Muchas tecnologías son capaces de generar probabilidades para una distribución binomial. Las instrucciones para estas diferentes tecnologías se encuentran en el Centro de tecnología al final de la sección 5-2 en la página 208. En lugar de usar una aproximación normal a una distribución binomial, utilice la tecnología para encontrar las probabilidades binomiales exactas en los ejercicios 9 a 12 de la sección 6-6.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: Diseño de un ascensor en el dormitorio del campus

Un estudiante universitario de Ohio murió cuando trató de escapar de un ascensor del dormitorio que estaba sobrecargado con 24 pasajeros. El ascensor estaba clasificado para un peso máximo de 2500 libras.

Consideremos este ascensor con un peso permisible de 2500 libras. Tengamos en cuenta también los parámetros para los pesos de los adultos, como se muestra en la tabla adjunta (con base en el conjunto de datos 1 "Datos corporales" del apéndice B).

continúa

Pesos de adultos

	Hombres	Mujeres
μ	189 lb	171 lb
σ	39 lb	46 lb
Distribución	Normal	Normal

Podríamos considerar características de diseño como el tipo de música que se podía reproducir en el ascensor. Podríamos seleccionar canciones como “Imagine” o “Daydream Believer”. En vez de eso, nos centraremos en la característica crítica del diseño: el peso.

a. En primer lugar, los elevadores suelen tener un margen de error del 25%, por lo que pueden transportar con seguridad una carga que es un 25% mayor que la carga indicada. ¿Qué cantidad es 25% mayor que 2500 libras? Hagamos referencia a esta cantidad como “la carga máxima segura”, mientras que el límite de 2,500 libras es la “carga máxima declarada”.

b. Ahora necesitamos determinar el número máximo de pasajeros que se debe permitir. ¿Debemos basar nuestros cálculos en la carga máxima segura o en la carga máxima de 2500 libras?

c. Los pesos dados en la tabla adjunta son pesos de adultos que no incluyen ropa o libros de texto. Añada otras 10 libras para la ropa de cada estudiante y los libros de texto. ¿Cuál es el número máximo de pasajeros que se deben permitir en el ascensor?

d. ¿Cree usted que los pesos de los estudiantes universitarios son diferentes de los pesos de los adultos de la población en general? Si es así, ¿de qué forma? ¿Cómo afectaría esto el diseño del ascensor?

Actividades de cooperación en grupo

1. Actividades fuera de la clase Utilice Internet para encontrar los resultados de la lotería “Pick 4” para 50 selecciones diferentes. Encuentre las 50 medias diferentes. Grafique un histograma de los 200 dígitos originales que se seleccionaron y represente gráficamente un histograma de las 50 medias muestrales. ¿Qué principio importante observa?

2. Actividad en clase Divídanse en grupos de tres o cuatro estudiantes y aborde los siguientes aspectos que afectan el diseño de tapas de alcantarilla.

• ¿Cuál de los siguientes temas es más relevante para determinar si el diámetro de una boca de inspección de 24 pulgadas es suficientemente grande: pesos de hombres, pesos de mujeres, estaturas de hombres, estaturas de mujeres, anchuras de cadera de hombres, anchuras de cadera de mujeres, anchuras de hombros de hombres, anchura de hombros de mujeres?

• ¿Por qué las tapas de alcantarilla son generalmente redondas? (Alguna vez, esta fue una pregunta común en las entrevistas a los solicitantes de empleo en IBM, y hay al menos tres buenas respuestas. Aquí basta con una sola respuesta buena).

3. Actividades fuera de clase Divídanse en grupos de tres o cuatro estudiantes. En cada grupo, desarrolle un procedimiento original para ilustrar el teorema del límite central. El objetivo principal es mostrar que cuando se seleccionan aleatoriamente muestras de una población, las medias de esas muestras tienden a distribuirse *normalmente*, sin importar la naturaleza de la distribución de la población. Para esta ilustración, comience con una población de valores que no tenga una distribución normal.

4. Actividad en clase Divídanse en grupos de tres o cuatro estudiantes. Utilizando una moneda para simular nacimientos, cada miembro del grupo debe simular 25 nacimientos y registrar el número de niñas simuladas. Combinen todos los resultados del grupo y registren n = número total de nacimientos y x = número de niñas. Dados lotes de n nacimientos, calcule la media y la desviación estándar para el número de niñas. ¿Es inusual el resultado simulado? ¿Por qué sí o por qué no?

5. Actividad en clase Divídanse en grupos de tres o cuatro estudiantes. Seleccione un conjunto de datos de uno de estos conjuntos de datos del apéndice B: 2, 3, 4, 5, 6, 7, 13, 16, 17, 18, 19, 20, 24, 27, 30, 32. (Estos son conjuntos de datos que no se utilizaron en ejemplos o ejercicios de la sección 6-5). Utilice los métodos de la sección 6-5 para construir un histograma y una gráfica cuantilar normal, y luego determine si el conjunto de datos parece provenir de una población normalmente distribuida.

6. Actividades fuera de la clase Divídanse en grupos de tres o cuatro estudiantes. Cada estudiante debe obtener una muestra de autos y registrar el número de automóviles blancos. Combine los resultados y use los métodos de este capítulo para comparar los resultados con los esperados en la gráfica de los colores más populares de auto que acompaña a los ejercicios 9 a 12 de la sección 6-6.

7

ESTIMACIÓN DE PARÁMETROS Y DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

PROBLEMA DEL CAPÍTULO

Encuestas: la ventana a la evolución de las tecnologías

Importancia de las encuestas En este mundo moderno regido por datos, es esencial tener la capacidad de analizar y comprender las encuestas y sondeos, que tienen un papel central y creciente en la orientación del entretenimiento, la política, el desarrollo de productos y casi todas las demás facetas de nuestras vidas. Este capítulo presenta las herramientas para desarrollar esa capacidad. A continuación se describen cuatro encuestas recientes que se centran en la evolución de las tecnologías:

- **Vehículos autónomos:** En una encuesta de TE Connectivity aplicada a 1000 adultos, 29% dijo que se sentiría cómodo en un vehículo autoconducido (como también se le conoce).
- **Propiedad de teléfono celular:** En una encuesta del Pew Research Center aplicada a 2076 adultos, 91% dijo que poseía un teléfono celular.
- **Facebook:** En una encuesta de Gallup aplicada a 1487 adultos, 43% dijo que tenían una página de Facebook.

continúa

- **Seguridad biométrica:** En una encuesta de *USA Today* aplicada a 510 personas, 53% dijo que deberíamos reemplazar las contraseñas por seguridad biométrica, como las huellas dactilares.

Debido a que las encuestas están ahora tan generalizadas y son tan extensas, y puesto que frecuentemente se aceptan sin dudar, debemos analizarlas considerando aspectos como los siguientes.

- ¿Qué método se utilizó para seleccionar los sujetos de la encuesta?
- ¿Cómo usamos los resultados de la muestra para estimar los valores de los parámetros poblacionales?
- ¿Cuál es la exactitud de los resultados de las encuestas?
- Los informes típicos de los medios sobre las encuestas carecen de un elemento muy importante de la información relevante. ¿De qué carecen por lo general?
- ¿Cómo interpretamos correctamente los resultados de la encuesta?

Por ejemplo, la encuesta "seguridad biométrica" ya citada, se basa en una ***muestra de respuesta voluntaria*** (descrita en la sección 1-1), por lo que su validez fundamental es muy cuestionable. Las otras tres encuestas presentan métodos de muestreo adecuados, por lo que las consideraremos para revisar los restantes aspectos ya listados, lo cual es precisamente el enfoque de este capítulo.

OBJETIVOS DEL CAPÍTULO

En este capítulo comenzamos el estudio de los métodos de ***estadística inferencial***. Las siguientes son las principales actividades de la estadística inferencial, y en este capítulo se presentan métodos para realizar la primera actividad, en la que se utilizan datos muestrales para estimar parámetros poblacionales. El capítulo 8 presentará los métodos básicos para probar hipótesis sobre los parámetros poblacionales.

Actividades principales de la estadística inferencial

1. Utilizar datos muestrales para *estimar valores de los parámetros poblacionales* (como una proporción o una media poblacionales).
2. Utilizar datos muestrales para *probar hipótesis* sobre los parámetros poblacionales.

Los objetivos del capítulo son:

7-1 Estimación de una proporción poblacional

- Construir una estimación del intervalo de confianza de una proporción poblacional e interpretar tal estimación del intervalo de confianza.
- Identificar los requisitos necesarios para el procedimiento que se utiliza y determinar si tales requisitos se satisfacen.
- Desarrollar la capacidad de determinar el tamaño de muestra necesario para estimar una proporción poblacional.

7-2 Estimación de una media poblacional

- Construir una estimación del intervalo de confianza de una media poblacional y ser capaz de interpretar tal estimación del intervalo de confianza.
- Determinar el tamaño de muestra necesario para estimar una media poblacional.

7-3 Estimación de una desviación estándar o varianza poblacionales

- Desarrollar la capacidad de construir una estimación del intervalo de confianza de una desviación estándar o varianza poblacionales, y ser capaz de interpretar tal estimación del intervalo de confianza.

7-4 Bootstrap: Uso de software para realizar estimaciones

- Desarrollar la capacidad de utilizar la tecnología junto con el método de bootstrap para construir una estimación del intervalo de confianza de una proporción, media, desviación estándar y varianza poblacionales.

7-1 Estimación de una proporción poblacional

Concepto clave En esta sección se presentan los métodos para usar una proporción muestral con el fin de hacer una inferencia sobre el valor de la proporción poblacional correspondiente. Esta sección se centra en la proporción poblacional p; también es posible trabajar con probabilidades o porcentajes. Al utilizar porcentajes, realizaremos los cálculos con el valor de proporción equivalente. Los tres conceptos principales que se incluyen en esta sección son:

- **Estimación puntual:** La proporción muestral (expresada por $\hat{p}$) es la mejor *estimación puntual* (o estimación de un solo valor) de la proporción poblacional p.
- **Intervalo de confianza:** Podemos usar una proporción muestral para construir una estimación del *intervalo de confianza* del verdadero valor de una proporción poblacional, y debemos saber cómo construir e interpretar tal intervalo de confianza.
- **Tamaño de la muestra:** Debemos saber cómo encontrar el tamaño de la muestra necesario para estimar una proporción poblacional.

Los conceptos presentados en esta sección se utilizan en las siguientes secciones y capítulos, por lo que es importante entender esta sección a la perfección.

PARTE 1 Estimación puntual, intervalo de confianza y tamaño de la muestra

Estimación puntual

Si deseamos estimar una proporción poblacional con un solo valor, la mejor estimación es la proporción muestral $\hat{p}$. Debido a que $\hat{p}$ consiste en un valor único que es equivalente a un punto en una línea, se denomina *estimación puntual*.

DEFINICIÓN

Una **estimación puntual** es un valor único utilizado para estimar un parámetro poblacional.

La proporción muestral $\hat{p}$ es la mejor *estimación puntual* de la proporción poblacional p.

Estimador no sesgado Utilizamos $\hat{p}$ como la estimación puntual de p porque no es sesgado y es el más consistente de los estimadores que podrían ser utilizados. (Un estimador no sesgado es un estadístico que se dirige al valor del parámetro poblacional correspondiente, en el sentido de que la distribución muestral del estadístico tiene una media que es igual al parámetro poblacional correspondiente. El estadístico $\hat{p}$ se dirige a la proporción poblacional p). La proporción muestral $\hat{p}$ es el estimador más consistente de p en el sentido de que la desviación estándar de las proporciones muestrales tiende a ser menor que la desviación estándar de otros estimadores no sesgados de p.

Sondeo de empuje

El *sondeo de empuje* es la práctica de efectuar campañas políticas fingiendo realizar un sondeo de opinión. Su nombre se deriva de su objetivo de empujar a los votantes para alejarlos de los candidatos de la oposición haciendo preguntas tendenciosas diseñadas para desacreditar a esos candidatos. He aquí un ejemplo de una pregunta de este tipo: “Dígame, por favor, si sería más o menos probable que usted votara por Roy Romer, si supiera que el gobernador Romer, desde que entró en funciones, nombró una junta de libertad bajo palabra que otorga la libertad, antes de cumplir la totalidad de su condena, a un promedio de cuatro delincuentes convictos al día”. El National Council on Public Polls considera que los sondeos de empuje son poco éticos. Los encuestadores con mayor reputación no aprueban el uso de esta práctica.

EJEMPLO 1 Facebook

El problema del capítulo incluyó la referencia a una encuesta de Gallup aplicada a 1487 adultos, en ella, 43% de los encuestados dijeron que tienen una página de Facebook. Con base en ese resultado, encuentre la mejor estimación puntual de la proporción de *todos* los adultos que tienen una página de Facebook.

SOLUCIÓN

Debido a que la proporción muestral es la mejor estimación puntual de la proporción poblacional, concluimos que la mejor estimación puntual de p es 0.43. (Si usa los resultados de la muestra para estimar el *porcentaje* de todos los adultos que tienen una página de Facebook, la mejor estimación puntual es 43%).

SU TURNO Encuentre la estimación puntual en el ejercicio 13 “Mickey D’s”.

Intervalo de confianza

¿Por qué necesitamos intervalos de confianza? En el ejemplo 1 vimos que 0.43 es nuestra *mejor* estimación puntual de la proporción poblacional p, pero no tenemos indicación de qué tan *buena* es esa mejor estimación. Un intervalo de confianza nos da una mejor idea de lo buena que es una estimación.

DEFINICIÓN

Un **intervalo de confianza** (o **estimación de intervalo**) es un rango (o un intervalo) de valores utilizados para estimar el valor real de un parámetro poblacional. En ocasiones, un intervalo de confianza se abrevia como IC.

DEFINICIÓN

El **nivel de confianza** es la probabilidad $1 - \alpha$ (por ejemplo 0.95, o 95%) de que el intervalo de confianza realmente contenga el parámetro poblacional asumiendo que el proceso de estimación se repite un gran número de veces. (El nivel de confianza también se denomina **grado de confianza** o **coeficiente de confianza**).

La siguiente tabla muestra la relación entre el nivel de confianza y el valor correspondiente de α. El nivel de confianza de 95% es el valor utilizado con mayor frecuencia.

Niveles de confianza más comunes	Valores correspondientes de α
Nivel de confianza del 90% (o 0.90):	$\alpha = 0.10$
Nivel de confianza del 95% (o 0.95):	$\alpha = 0.05$
Nivel de confianza del 99% (o 0.99):	$\alpha = 0.01$

A continuación se presenta un ejemplo de un intervalo de confianza que se encontrará más adelante en el ejemplo 3:

La estimación del intervalo de confianza de 0.95 (o 95%) para la proporción de poblacional p es $0.405 < p < 0.455$.

Interpretación de un intervalo de confianza

Debemos tener cuidado de interpretar los intervalos de confianza correctamente. Existe una interpretación correcta y muchas interpretaciones incorrectas y creativas del intervalo de confianza $0.405 < p < 0.455$.

Correcta: **“Tenemos una confianza del 95% de que el intervalo de 0.405 a 0.455 realmente contiene el valor verdadero de la proporción poblacional *p*”**.
Esta es una manera reducida y aceptable de decir que si seleccionáramos diferentes muestras aleatorias de tamaño 1487 (del ejemplo 3) y construyéramos los correspondientes intervalos de confianza, el 95% contendría la proporción poblacional p. En esta interpretación correcta, el nivel de confianza del 95% se refiere a la *tasa de éxito del proceso* utilizado para estimar la proporción poblacional.

Incorrecta: **“Existe un 95% de posibilidades de que el valor verdadero de *p* se encuentre entre 0.405 y 0.455”**.
Esto es incorrecto porque p es un parámetro poblacional con un valor fijo; no es una variable aleatoria con valores que cambian.

Incorrecta: **“95% de las proporciones muestrales estarán entre 0.405 y 0.455”.** Esto es incorrecto porque los valores de 0.405 y 0.455 resultan de una muestra; no son parámetros que describen el comportamiento de todas ellas.

Nivel de confianza: la tasa de éxito del proceso Un nivel de confianza del 95% nos dice que el proceso que estamos usando debería, a largo plazo, resultar en límites de intervalo de confianza que contienen la verdadera proporción poblacional el 95% de las veces. Supongamos que la proporción verdadera de adultos con páginas de Facebook es $p = 0.50$. Vea la figura 7-1, que muestra que 19 de cada 20 (o 95%) intervalos de confianza diferentes contienen el valor asumido de $p = 0.50$. La figura 7-1 trata de contar esta historia: con un nivel de confianza del 95%, esperamos que aproximadamente 19 de cada 20 intervalos de confianza (o 95%) contengan el valor verdadero de p.

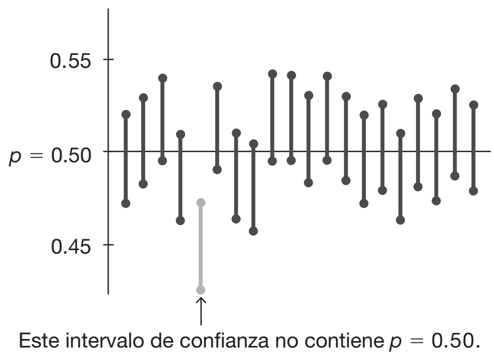


FIGURA 7-1 Intervalos de confianza de 20 muestras

Valores críticos

Los valores críticos se definen formalmente en la página siguiente y están basados en las siguientes observaciones:

1. Cuando se cumplen ciertos requisitos, la distribución muestral de las proporciones muestrales puede aproximarse mediante una distribución normal, como se muestra en la figura 7-2.
2. Una puntuación z asociada a una proporción muestral tiene una probabilidad de $\alpha/2$ de caer en la cola derecha de la figura 7-2.
3. La puntuación z en el límite de la cola derecha se denomina comúnmente $z_{\alpha/2}$ y se denomina *valor crítico* porque está en la frontera que separa las puntuaciones z que son significativamente altos.

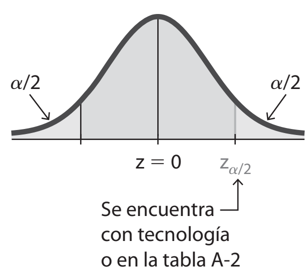


FIGURA 7-2 Valor crítico $z_{\alpha/2}$ en la distribución normal estándar

Vocabulario de Shakespeare

Según Bradley Efron y Ronald Thisted, los escritos de Shakespeare incluyeron 31,534 palabras distintas. Utilizaron la teoría de la probabilidad para concluir que Shakespeare probablemente conocía por lo menos otras 35,000 palabras que no utilizó en sus escritos. El estimar el tamaño de una población es un problema importante que a menudo se encuentra en los estudios de ecología, pero el resultado dado aquí es otra aplicación interesante. (Vea "Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?"; en *Biometrika*, vol. 63, núm. 3).

DEFINICIÓN

Un **valor crítico** es el número en la frontera que separa los estadísticos muestrales que son significativamente altos o bajos de aquellas que no son significativos. El número $z_{\alpha/2}$ es un valor crítico que es una puntuación z con la propiedad de que está en el límite que separa un área de $\alpha/2$ en la cola derecha de la distribución normal estándar (como en la figura 7-2).

EJEMPLO 2 Determinación de un valor crítico

Encuentre el valor crítico $z_{\alpha/2}$ correspondiente a un nivel de confianza del 95%.

SOLUCIÓN

Un nivel de confianza del 95% corresponde a $\alpha = 0.05$, por lo que $\alpha/2 = 0.025$. La figura 7-3 muestra que el área en cada una de las colas con sombreado es $\alpha/2 = 0.025$. Encontramos $z_{\alpha/2} = 1.96$ al observar que el área acumulada a su izquierda debe ser $1 - 0.025$, o 0.975. Podemos utilizar la tecnología o consultar la tabla A-2 para encontrar que el área acumulada a la izquierda de 0.9750 corresponde a $z = 1.96$. Por lo tanto, para un nivel de confianza del 95%, el valor crítico es $z_{\alpha/2} = 1.96$.

Tenga en cuenta que cuando se encuentra la puntuación crítica z para un nivel de confianza del 95%, se utiliza un área acumulada a la izquierda de 0.9750 (*no* de 0.95). Piénselo de la siguiente manera:

Este es nuestro nivel de confianza:		El área en *ambas* colas es:		El área en la cola *derecha* es:		El área acumulada desde la izquierda, excluyendo la cola derecha es:
95%	→	$\alpha = 0.05$	→	$\alpha/2 = 0.025$	→	$1 - 0.025 = 0.975$

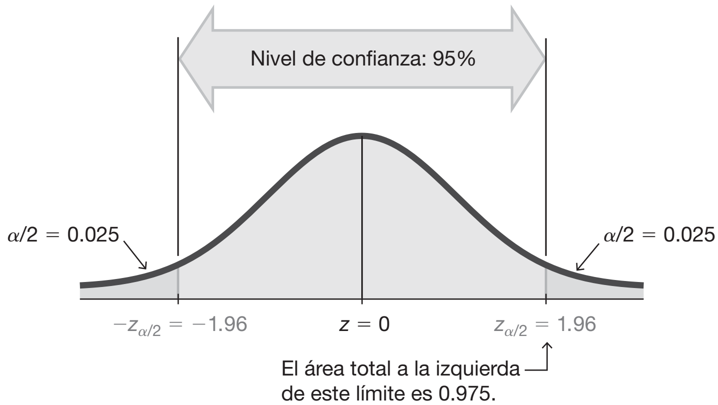


FIGURA 7-3 **Determinación del valor crítico $z_{\alpha/2}$ para un nivel de confianza del 95%**

SU TURNO Resuelva el ejercicio 5 "Determinación de valores críticos".

El ejemplo 2 mostró que un nivel de confianza del 95% da como resultado un valor crítico de $z_{\alpha/2} = 1.96$. Éste es el valor crítico más común y se lista junto con otros dos valores comunes en la siguiente tabla.

Nivel de confianza	α	Valor crítico, $z_{\alpha/2}$
90%	0.10	1.645
95%	0.05	1.96
99%	0.01	2.575

Margen de error

Ahora definimos formalmente el *margen de error E* del que todos hemos oído hablar tan a menudo en los informes de los medios de comunicación.

DEFINICIÓN

Cuando se utilizan datos de una muestra aleatoria simple para estimar una proporción poblacional p, la diferencia entre la proporción muestral $\hat{p}$ y la proporción poblacional p es un error. La cantidad máxima probable de ese error es el **margen de error**, expresado por ***E***. Existe una probabilidad de $1 - \alpha$ (por ejemplo 0.95) de que la diferencia entre $\hat{p}$ y p sea E o menos. El margen de error E también se denomina *error máximo de la estimación* y se puede encontrar multiplicando el valor crítico por la desviación estándar estimada de las proporciones muestrales, como lo indica la fórmula 7-1.

En cifras

$1,000,000: Diferencia estimada de las ganancias de por vida entre una persona que tiene un título universitario y alguien sin un título universitario.

FÓRMULA 7-1

$$E = \underbrace{z_{\alpha/2}}_{\text{Valor crítico}} \underbrace{\sqrt{\frac{\hat{p}\hat{q}}{n}}}_{\text{Desviación estándar estimada de las proporciones muestrales}} \quad \text{margen de error para proporciones}$$

ELEMENTOS CLAVE

Intervalo de confianza para estimar una proporción poblacional *p*

Objetivo

Construir un intervalo de confianza utilizado para estimar una proporción poblacional p.

Notación

p = proporción *poblacional*
$\hat{p}$ = proporción *muestral*
n = número de valores muestrales
E = margen de error
$z_{\alpha/2}$ = valor crítico: la puntuación z que separa un área de $\alpha/2$ en la cola derecha de la distribución normal estándar

Requisitos

1. La muestra es una muestra aleatoria simple.
2. Se cumplen las condiciones para la distribución binomial: hay un número fijo de ensayos, los ensayos son independientes, hay dos categorías de resultados y las probabilidades permanecen constantes para cada ensayo (como en la sección 5-2).
3. Hay al menos 5 éxitos y al menos 5 fracasos. (Este requisito es una forma de verificar que $np \geq 5$ y $nq \geq 5$, por lo que la distribución normal es una aproximación adecuada a la distribución binomial).

Estimación del intervalo de confianza de *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 se expresa a menudo en los siguientes dos formatos equivalentes:

$$\hat{p} \pm E \quad \text{o} \quad (\hat{p} - E, \hat{p} + E)$$

Regla de redondeo para estimaciones del intervalo de confianza de *p*

Redondee los límites del intervalo de confianza para p a tres dígitos significativos.

¿Sesgo en las encuestas por Internet?

Aprovechando el uso generalizado de la tecnología y las redes sociales, existe una tendencia creciente a realizar encuestas utilizando sólo Internet en vez de realizar entrevistas en persona o llamadas telefónicas a sujetos seleccionados al azar. Las encuestas por Internet son más rápidas y mucho menos costosas, y proporcionan importantes ventajas en el diseño y la administración de la información. Pero ¿las encuestas por Internet están sesgadas porque usan solamente sujetos seleccionados al azar entre 90% de la población estadounidense que usa Internet? El Pew Research Center estudió este cuestionamiento comparando los resultados de encuestas en línea con encuestas que incluían a la población fuera de línea. Se encontró que las diferencias eran generalmente muy pequeñas, pero los temas relacionados con Internet y la tecnología produjeron diferencias mucho mayores. Debemos tener cuidado de considerar las consecuencias del sesgo en las encuestas por Internet.

Procedimiento para construir un intervalo de confianza para *p*

1. Verifique que se cumplen los requisitos enunciados en el recuadro anterior de Elementos clave.
2. Utilice la tecnología o la tabla A-2 para encontrar el valor crítico $z_{\alpha/2}$ que corresponde al nivel de confianza deseado.
3. Evalúe el margen de error $E = z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$.
4. Con base en el valor del margen de error E calculado y el valor de la proporción muestral $\hat{p}$, encuentre los valores de los *límites del intervalo de confianza* $\hat{p} - E$ y $\hat{p} + E$. Sustituya esos valores en el formato general del intervalo de confianza.
5. Redondee los límites del intervalo de confianza resultante a tres dígitos significativos.

EJEMPLO 3 Construcción de un intervalo de confianza: resultados de una encuesta

En el problema del capítulo observamos que una encuesta de Gallup aplicada a 1487 adultos demostró que 43% de los encuestados tiene páginas de Facebook. Los resultados de la muestra son $n = 1487$ y $\hat{p} = 0.43$.

a. Encuentre el margen de error E que corresponde a un nivel de confianza del 95%.

b. Determine la estimación del intervalo de confianza del 95% para la proporción poblacional p.

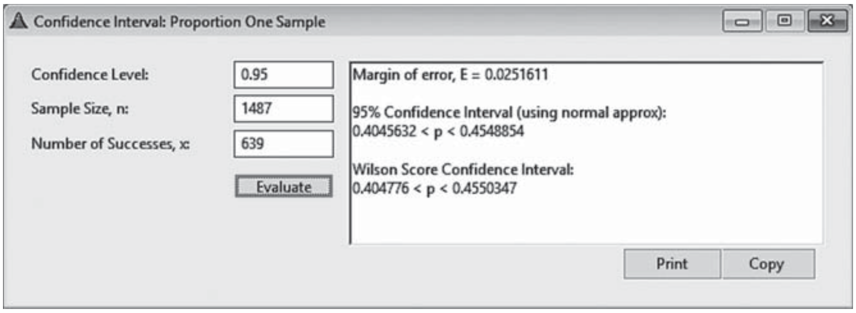
c. Con base en los resultados, ¿podemos concluir con seguridad que menos de 50% de los adultos tienen páginas en Facebook? Asumiendo que usted es un periodista, escriba un breve artículo que describa con precisión los resultados e incluya toda la información relevante.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los métodos de sondeo utilizados por la organización Gallup generan muestras que pueden considerarse muestras aleatorias simples. (2) Las condiciones para un experimento binomial se satisfacen porque hay un número fijo de ensayos (1487), los ensayos son independientes (porque la respuesta de una persona no afecta la probabilidad de la respuesta de otra persona), hay dos categorías de resultados (el sujeto tiene una página de Facebook o no), y la probabilidad permanece constante, porque P (tener una página de Facebook) es fija para un punto dado en el tiempo. (3) Dado que 43% de los encuestados tienen páginas de Facebook, la cantidad con páginas de Facebook es 639 (o 43% de 1487). Si 639 de los 1487 sujetos tienen páginas de Facebook, los otros 848 no, por lo que el número de éxitos (639) y el número de fracasos (848) son al menos de 5. La verificación de requisitos es completada con éxito.

Tecnología: El intervalo de confianza y el margen de error se pueden encontrar fácilmente utilizando tecnología. En la pantalla de Statdisk de la página siguiente podemos ver las entradas requeridas a la izquierda y los resultados mostrados a la derecha. Como la mayoría de las tecnologías, Statdisk requiere un valor para el número de éxitos, por lo que simplemente encuentre el 43% de 1487 y redondee el resultado de 639.41 al número entero 639. Los resultados muestran que el margen de error es $E = 0.025$ (redondeado) y que el intervalo de confianza es $0.405 < p < 0.455$ (redondeado). (El intervalo de confianza Wilson Score incluido en la pantalla se discutirá más adelante en la parte 2 de esta sección).

Statdisk



Cálculo manual A continuación se muestra cómo encontrar el intervalo de confianza mediante cálculos manuales:

a. El margen de error se encuentra usando la fórmula 7-1 con $z_{\alpha/2} = 1.96$ (como en el ejemplo 2), $\hat{p} = 0.43$, $\hat{q} = 0.57$ y $n = 1487$.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{(0.43)(0.57)}{1487}} = 0.0251636$$

b. La construcción del intervalo de confianza es realmente fácil ahora que sabemos que $\hat{p} = 0.43$ y $E = 0.0251636$. Simplemente sustituya esos valores para obtener el siguiente resultado:

$$\hat{p} - E < p < \hat{p} + E$$
$$0.43 - 0.0251636 < p < 0.43 + 0.0251636$$
$$0.405 < p < 0.455 \quad \text{(redondeado a tres dígitos significativos)}$$

Este mismo resultado podría expresarse en el formato 0.43 ± 0.025 o (0.405, 0.455). Si queremos el intervalo de confianza del 95% para el *porcentaje* de la población real, podríamos expresar el resultado como $40.5\% < p < 45.5\%$.

c. Con base en el intervalo de confianza obtenido en el inciso (b), parece que menos del 50% de los adultos tienen una página de Facebook porque el intervalo de valores de 0.405 a 0.455 es un intervalo que está completamente por debajo de 0.50.

Aquí hay una declaración que resume los resultados: 43% de los adultos tienen páginas de Facebook. Ese porcentaje se basa en una encuesta Gallup de 1487 adultos seleccionados al azar en EUA. En teoría, en el 95% de esas encuestas, el porcentaje no debería diferir en más de 2.5 puntos porcentuales en cualquier dirección, a partir del porcentaje que se obtendría al entrevistar a todos los adultos.

SU TURNO Encuentre el intervalo de confianza en el ejercicio 13 "Mickey D's",

Análisis de encuestas El ejemplo 3 trata sobre una encuesta típica. Al analizar los resultados de las encuestas, considere lo siguiente:

1. La muestra debe ser una muestra aleatoria simple, no una muestra inadecuada (como una muestra de respuesta voluntaria).
2. Se debe proporcionar el nivel de confianza. (A menudo es el 95%, pero los informes de los medios de comunicación generalmente omiten identificar el nivel de confianza).
3. Se debe proporcionar el tamaño de la muestra. (Con frecuencia, los reportan los medios de comunicación, pero no siempre).
4. Con excepción de casos relativamente raros, la calidad de los resultados de la encuesta depende del método de muestreo y del tamaño de la muestra, pero el tamaño de la población no suele ser un factor.

Cómo se realizó la encuesta

El *New York Times* es bastante bueno reportando resultados de encuestas. A menudo, este periódico informa sobre los resultados de las encuestas con un artículo anexo bajo el título "Cómo se realizó la encuesta". La descripción típicamente incluye el tamaño de muestra, el margen de error y el siguiente enunciado que revela que el nivel de confianza es del 95%: "En teoría, en 19 casos de 20, los resultados globales basados en dichas muestras no serán mayores que . . . ". Un informe reciente también proporcionó información de que la encuesta incluyó adultos registrados para votar, los números de teléfono de línea fija fueron seleccionados aleatoriamente por computadora, los números de teléfono celular también fueron generados al azar y los resultados se ponderaron de acuerdo con la región geográfica, el sexo, la raza, el estado civil, la edad y el grado de educación. En la opinión no tan humilde de este autor, las descripciones de "Cómo se realizó la encuesta" son un modelo para todos los medios que informan sobre este tipo de resultados.

PRECAUCIÓN Nunca piense que los resultados de la encuesta no son confiables si el *tamaño de la muestra* es un pequeño porcentaje del *tamaño de la población*. El tamaño de la población no suele ser un factor para determinar la confiabilidad de una encuesta.

Determinación de la estimación puntual y de *E* a partir de un intervalo de confianza

En ocasiones queremos entender mejor un intervalo de confianza que pudo haber sido tomado de un artículo de revista o de la aplicación de una tecnología. Si ya conocemos los límites del intervalo de confianza, la proporción muestral (o la mejor estimación puntual) y el margen de error E se pueden encontrar de la siguiente manera:

Estimación puntual de p:

$$\hat{p} = \frac{(\text{límite superior del intervalo de confianza}) + (\text{límite inferior del intervalo de confianza})}{2}$$

Margen de error:

$$E = \frac{(\text{límite superior del intervalo de confianza}) - (\text{límite inferior del intervalo de confianza})}{2}$$

EJEMPLO 4 Determinación de la proporción muestral y el margen de error

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, 70% se abstuvieron de fumar a las 8 semanas (intervalo de confianza [IC] del 95%, 58% a 81%)". Use esta afirmación para encontrar la estimación puntual y el margen de error E.

SOLUCIÓN

Obtenemos el intervalo de confianza del 95% de $0.58 < p < 0.81$ del rango dado "58% a 81%". La estimación puntual $\hat{p}$ es el valor que está a la mitad entre los límites superior e inferior del intervalo de confianza, por lo que obtenemos

$$\hat{p} = \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2}$$
$$= \frac{0.81 + 0.58}{2} = 0.695$$

El margen de error se puede encontrar como sigue:

$$E = \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2}$$
$$= \frac{0.81 - 0.58}{2} = 0.115$$

Uso de intervalos de confianza para pruebas de hipótesis

Un intervalo de confianza puede usarse para abordar de manera *informal* alguna afirmación hecha sobre una proporción poblacional. Por ejemplo, si los resultados de la muestra consisten en 70 caras en 100 lanzamientos de una moneda, el intervalo de confianza del 95% resultante de $0.610 < p < 0.790$ puede usarse para apoyar *informalmente* una afirmación de que la proporción de caras es *diferente al* 50% (porque 0.50 no está contenido dentro del intervalo de confianza).

Determinación del tamaño de muestra

Si planeamos recolectar datos muestrales para estimar alguna proporción poblacional, ¿cómo sabemos *cuántas* unidades de muestra debemos recolectar? Si despejamos el tamaño de muestra n de la fórmula para el margen de error E (fórmula 7-1), obtenemos la fórmula 7-2 que se presenta a continuación. La fórmula 7-2 requiere de $\hat{p}$ como una estimación de la proporción poblacional p, pero si no se conoce tal estimación (como suele ser el caso), reemplazamos $\hat{p}$ por 0.5 y $\hat{q}$ por 0.5, con el resultado dado en la fórmula 7-3 . Si se reemplazan $\hat{p}$ y $\hat{q}$ por 0.5 se obtiene el tamaño de muestra más grande posible, por lo que estamos seguros de que el tamaño de muestra es adecuado para estimar p.

ELEMENTOS CLAVE

Determinación del tamaño de muestra requerido para estimar una proporción poblacional

Objetivo

Determinar cuán grande debe ser el tamaño de muestra n para estimar la proporción poblacional p.

Notación:

p = proporción poblacional

$\hat{p}$ = proporción muestral

n = número de valores muestrales

E = margen de error deseado

$z_{\alpha/2}$ = puntuación z que separa un área de $\alpha/2$ en la cola derecha de la distribución normal estándar

Requisitos

La muestra debe ser una muestra aleatoria simple de unidades muestrales independientes.

Cuando se conoce una estimación $\hat{p}$: **Fórmula 7-2** $$n = \frac{[z_{\alpha/2}]^2\hat{p}\hat{q}}{E^2}$$

Cuando no se conoce una estimación $\hat{p}$: **Fórmula 7-3** $$n = \frac{[z_{\alpha/2}]^2\,0.25}{E^2}$$

Si se puede hacer una estimación razonable de $\hat{p}$ mediante el uso de muestras previas, un estudio piloto o el conocimiento experto de alguien, utilice la fórmula 7-2. Si no se conoce nada sobre el valor de $\hat{p}$, utilice la fórmula 7-3.

Regla de redondeo para determinar el tamaño de la muestra

Si el tamaño de muestra calculado n no es un número entero, redondee el valor de n al siguiente número entero más *grande*, de manera que el tamaño de la muestra sea suficiente en lugar de ligeramente insuficiente. Por ejemplo, redondee 1067.11 a 1068.

EJEMPLO 5 **¿Qué porcentaje de adultos hace compras en línea?**

Cuando el autor estaba realizando investigaciones para este capítulo, no pudo encontrar información sobre el porcentaje de adultos que realizan compras en línea, pero esa información es extremadamente importante para las tiendas en línea, como lo es para las tiendas con una instalación física. Si el autor tiene que realizar su propia encuesta, ¿cuántos adultos deberían ser encuestados para estar 95% seguros de que el porcentaje muestral tiene un error no mayor de tres puntos porcentuales?

a. Suponga que una encuesta reciente mostró que 80% de los adultos realizan compras en línea.

b. Suponga que no tenemos información previa que sugiera un posible valor de la proporción poblacional.

continúa

Falsificación de datos

El glosario del censo define la *falsificación de datos* (o *curbstoning*) como "la práctica por medio de la cual un censor fabrica un cuestionario para una vivienda, sin visitarla". La falsificación de datos ocurre cuando un censor se sienta en la acera (o en cualquier otro lado) y llena las formas inventando las respuestas. Puesto que estos datos no son reales, afectan la validez del censo. En varios estudios se ha investigado la magnitud de la falsificación; uno de ellos reveló que aproximadamente el 4% de los censores realizan esta práctica al menos en una ocasión. Los métodos de la sección 7-1 suponen que los datos muestrales se reunieron de una forma adecuada, así que si gran parte de los datos se obtuvieron a través de falsificaciones, entonces las estimaciones de los intervalos de confianza resultantes podrían estar muy errados.

SOLUCIÓN

a. Con un nivel de confianza del 95%, tenemos $\alpha = 0.05$, así que $z_{\alpha/2} = 1.96$. Además, el margen de error es $E = 0.03$, que es el equivalente decimal de "tres puntos porcentuales". La encuesta anterior sugiere que $\hat{p} = 0.80$, por lo que $\hat{q} = 0.20$ (determinada a partir de $\hat{q} = 1 - 0.80$). Debido a que tenemos un valor estimado de $\hat{p}$, usamos la fórmula 7-2 de la manera siguiente:

$$n = \frac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2} = \frac{[1.96]^2 (0.80)(0.20)}{0.03^2}$$
$$= 682.951 = 683 \text{ (redondeada hacia arriba)}$$

Debemos obtener una muestra aleatoria simple que incluya al menos 683 adultos.

b. Sin conocimiento previo de $\hat{p}$ (o $\hat{q}$), usamos la fórmula 7-3 como sigue:

$$n = \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2} = \frac{[1.96]^2 \cdot 0.25}{0.03^2}$$
$$= 1067.11 = 1068 \text{ (redondeada hacia arriba)}$$

Debemos obtener una muestra aleatoria simple que incluya al menos 1068 adultos.

INTERPRETACIÓN

Para tener un 95% de confianza en que nuestro porcentaje muestral está dentro de los tres puntos porcentuales del verdadero porcentaje para todos los adultos, debemos obtener una muestra aleatoria simple de 1068 adultos, asumiendo que no hay conocimiento previo. Al comparar este resultado con el tamaño de muestra de 683 que se encontró en el inciso (a), podemos ver que si no tenemos conocimiento de un estudio previo, se requiere una muestra más grande para obtener los mismos resultados que cuando es posible estimar el valor de $\hat{p}$.

SU TURNO Resuelva el ejercicio 31 "Zurdos".

DEFINICIÓN

Trate de evitar los siguientes tres errores comunes al calcular el tamaño de la muestra:

1. No cometa el error de usar $E = 3$ como el margen de error correspondiente a "tres puntos porcentuales". Si el margen de error es de tres puntos porcentuales, use $E = 0.03$.
2. Asegúrese de sustituir la puntuación crítica z por $z_{\alpha/2}$. Por ejemplo, al trabajar con una confianza del 95%, asegúrese de reemplazar $z_{\alpha/2}$ por 1.96. No cometa el error de reemplazar $z_{\alpha/2}$ por 0.95 o 0.05.
3. Asegúrese de redondear hasta el siguiente entero superior; no redondee usando las reglas habituales. Redondee 1067.11 a 1068.

Papel del tamaño *N* de la población Las fórmulas 7-2 y 7-3 son notables porque muestran que el tamaño de la muestra no depende del tamaño (N) de la población; el tamaño de la muestra depende del nivel de confianza deseado, el margen de error deseado y, en ocasiones, de la estimación conocida $\hat{p}$. (Vea el ejercicio 39 "Factor de corrección de población finita" para tratar con casos en los que se selecciona una muestra relativamente grande sin reemplazo de una población finita, de modo que el tamaño n de la muestra depende del tamaño N de la población).

PARTE 2 Intervalos de confianza de mejor desempeño

Desventaja del intervalo de confianza de Wald

Probabilidad de cobertura La **probabilidad de cobertura** de un intervalo de confianza es la proporción real de tales intervalos de confianza que contienen la proporción poblacional real. Si seleccionamos un nivel de confianza específico, como 0.95 (o 95%), nos gustaría obtener la probabilidad de cobertura *real* igual a nuestro nivel de confianza *deseado*. Sin embargo, para el intervalo de confianza descrito en la parte 1 (denominado "intervalo de confianza de Wald"), la probabilidad de cobertura real suele ser menor o igual al nivel de confianza que seleccionamos y podría ser sustancialmente menor. Por ejemplo, si seleccionamos un nivel de confianza del 95%, usualmente obtenemos un 95% o *menos* de intervalos de confianza que contienen la proporción de población p. (Esto a veces se considera "demasiado liberal.") Por esta razón, el intervalo de confianza de Wald se usa raramente en aplicaciones y revistas profesionales.

Intervalos de confianza de mejor desempeño

Nota importante sobre los ejercicios: Con la excepción de algunos ejercicios del tipo Más allá de lo básico, los ejercicios para esta sección 7-1 se basan en el método para construir un intervalo de confianza de Wald como se describe en la parte 1, no los intervalos de confianza descritos aquí. Se recomienda que los estudiantes aprendan los métodos presentados anteriormente, pero que reconozcan que hay mejores métodos disponibles, y que pueden ser utilizados con el software adecuado.

Método más cuatro El *intervalo de confianza más cuatro* se desempeña mejor que el intervalo de confianza de Wald en el sentido de que su probabilidad de cobertura está más cerca del nivel de confianza que se utiliza. El intervalo de confianza más cuatro utiliza el siguiente procedimiento muy simple: añada 2 al número de éxitos x, añada 2 al número de fracasos (de modo que el número de ensayos n aumente en 4) y luego busque el intervalo de confianza de Wald como se describe en la parte 1 de esta sección. El intervalo de confianza más cuatro es muy fácil de calcular y tiene probabilidades de cobertura similares a las del intervalo de confianza de la puntuación de Wilson que se presenta a continuación.

Puntuación de Wilson Otro intervalo de confianza que se comporta mejor que el IC de Wald es el *intervalo de confianza de la puntuación de Wilson*:

$$\frac{\hat{p} + \dfrac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q} + \frac{z_{\alpha/2}^2}{4n}}{n}}}{1 + \dfrac{z_{\alpha/2}^2}{n}}$$

El intervalo de confianza de la puntuación de Wilson se comporta mejor que el IC de Wald en el sentido de que la probabilidad de cobertura es más cercana al nivel de confianza. Con un nivel de confianza del 95%, el intervalo de confianza de la puntuación de Wilson nos acercaría a una probabilidad de 0.95 de contener el parámetro p. Sin embargo, dada su complejidad, es fácil ver por qué este intervalo de confianza no se utiliza mucho en los cursos de estadística introductoria. La complejidad de la expresión anterior puede ser eludida usando algunas tecnologías, como Statdisk, que proporcionan resultados del intervalo de confianza de la puntuación de Wilson.

Método Clopper-Pearson El método de Clopper-Pearson es un método "exacto" en el sentido de que se basa en la distribución binomial exacta en lugar de en la aproximación de una distribución, y se le critica por ser *demasiado conservador* en el siguiente sentido: cuando seleccionamos un nivel de confianza específico, la probabilidad de cobertura suele ser mayor o igual al nivel de confianza seleccionado. Seleccione un nivel de confianza de 0.95 y la probabilidad de cobertura real sea normalmente de 0.95 o mayor, de modo que 95% o más de dichos intervalos de confianza contendrán p. Los cálculos de este método son demasiado complicados para ser considerados aquí.

¿Cuál es el mejor método? Existen otros métodos para construir intervalos de confianza que no se analizan aquí. No hay un acuerdo universal sobre cuál es el mejor método para construir una estimación del intervalo de confianza para p.

- El intervalo de confianza de Wald funciona mejor como una herramienta de enseñanza para introducir a los estudiantes a los intervalos de confianza.
- El intervalo de confianza más cuatro es casi tan fácil como el de Wald y se comporta mejor que éste por tener una probabilidad de cobertura más cercana al nivel de confianza seleccionado.

Una vez más, tenga en cuenta que a excepción de algunos ejercicios del tipo Más allá de lo básico, los ejercicios que siguen se basan en el intervalo de confianza de Wald explicado con anterioridad, y no en los intervalos de confianza de mejor desempeño aquí analizados.

CENTRO DE TECNOLOGÍA

Proporciones: Intervalos de confianza y determinación del tamaño de muestra

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Intervalo de confianza

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Confidence Intervals** en el menú desplegable y seleccione **Proportion One Sample** del submenú.
3. Introduzca el nivel de confianza, el tamaño de la muestra y el número de éxitos.
4. Haga clic en **Evaluate**.

Determinación del tamaño de muestra

1. Haga clic en **Analysis** del menú superior.
2. Seleccione **Sample Size Determination** en el menú desplegable y seleccione **Estimate Proportion** del submenú.
3. Introduzca el nivel de confianza, el margen de error E, la estimación de p si la conoce, y el tamaño de la población si es conocido.
4. Haga clic en **Evaluate**.

StatCrunch

Intervalo de confianza

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Proportion Stats** en el menú desplegable y seleccione **One Sample-With Summary** en el submenú.
3. Introduzca el número de éxitos y el número de observaciones.
4. Seleccione **Confidence Interval for p** e ingrese el nivel de confianza. Seleccione el método **Standard-Wald**.
5. Haga clic en **Compute!**

Determinación del tamaño de la muestra

No está disponible.

Minitab

Intervalo de confianza

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **1 Proportion** en el submenú.
3. Seleccione **Summarized data** en la ventana desplegable e ingrese el número de eventos (éxitos) y el número de ensayos. Revise que la opción *Perform hypothesis test* no está marcada.
4. Haga clic en el botón **Options** e ingrese el nivel de confianza deseado. Para *Alternative Hypothesis*, seleccione $\neq$ y, para *Method*, seleccione **Normal aproximation** para acceder a los métodos de esta sección.
5. Haga clic dos veces en **OK**.

Determinación del tamaño de la muestra

Minitab determina el tamaño de la muestra usando la distribución binomial (no la distribución normal), por lo que los resultados serán diferentes de los encontrados usando los métodos de esta sección.

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Power y Sample Size** en el menú desplegable y seleccione **Sample Size for Estimation** en el submenú.
3. Para *Parameter*, seleccione **Proportion (Binomial)** e ingrese una estimación de la proporción si la conoce o ingrese 0.5 si no la conoce.
4. Seleccione **Estimate simple sizes** en el menú desplegable e ingrese el margen de error para los intervalos de confianza.
5. Haga clic en el botón **Options** para ingresar el nivel de confianza y seleccione un intervalo de confianza del tipo **two sided**.
6. Haga clic dos veces en **OK**.

CENTRO DE TECNOLOGÍA *continuación*

Proporciones: Intervalos de confianza y determinación del tamaño de muestra

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

Intervalo de confianza

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **1-PropZInt** en el menú y presione ENTER.
3. Introduzca el número de éxitos x, el número de observaciones n, y el nivel de confianza (*C-Level*).
4. Seleccione **Calculate** y pulse ENTER.

Determinación del tamaño de la muestra

No está disponible.

Excel

Complemento XLSTAT (requerido)

1. Haga clic en la ficha **XLSTAT** de la cinta de opciones y después en **Parametric Tests**.
2. Seleccione **Test for one proportion** del menú desplegable.
3. En *Data Format* seleccione **Frequency** si conoce el número de éxitos x o seleccione **Proportion** si conoce la proporción muestral $\hat{p}$.
4. Introduzca la frecuencia o proporción muestral, el tamaño de la muestra y **0.5** para *Test proportion*.
5. Marque **z test** y desmarque **Continuity correction**.
6. Haga clic en la ficha **Options**.
7. Bajo *Alternative hypothesis*, seleccione $\neq$ **D**. Introduzca **0** para *Hypothesized difference* e ingrese el nivel de significancia deseado (ingrese **5** para el intervalo de confianza del 95%). Bajo *Variance (confidence interval)* seleccione **Sample** y bajo *Confidence interval* seleccione **Wald**.
8. Haga clic en **OK** para mostrar el resultado del intervalo de confianza para la proporción (Wald).

Determinación del tamaño de la muestra

No está disponible.

7-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Resultados de encuestas en los medios de comunicación *USA Today* proporcionó los resultados de una encuesta a 1000 adultos, a quienes se les pidió que identificaran su pastel favorito. Entre los 1000 encuestados, 14% eligió pastel de chocolate, y el margen de error fue dado como ± 4 puntos porcentuales. ¿Qué característica importante de la encuesta se omitió?

2. Margen de error Para la encuesta descrita en el ejercicio 1, describa qué significa la afirmación "el margen de error fue dado como ± 4 puntos porcentuales".

3. Notación Para la encuesta descrita en el ejercicio 1, ¿qué valores representan $\hat{p}$, $\hat{q}$, n, E y p? Si el nivel de confianza es del 95%, ¿cuál es el valor de α?

4. Niveles de confianza Dados datos muestrales específicos, como los datos del ejercicio 1, ¿qué intervalo de confianza es más amplio: el intervalo de confianza del 95% o el intervalo de confianza del 80%? ¿Por qué es más amplio?

Determinación de valores críticos. ***En los ejercicios 5 a 8, determine el valor crítico $z_{\alpha/2}$ que corresponde al nivel de confianza dado.***

5. 90% **6.** 99% **7.** 99.5% **8.** 98%

Formatos de intervalos de confianza. ***En los ejercicios 9 a 12, exprese el intervalo de confianza usando el formato indicado. (Los intervalos de confianza se basan en las proporciones de M&Ms de color rojo, naranja, amarillo y azul en el conjunto de datos 27 "Pesos de M&Ms" en el apéndice B).***

9. M&Ms rojos Exprese $0.0434 < p < 0.217$ en la forma de $\hat{p} \pm E$.

10. M&Ms naranjas Exprese $0.179 < p < 0.321$ en la forma de $\hat{p} \pm E$.

11. M&Ms amarillos Exprese el intervalo de confianza (0.0169, 0.143) en la forma de $\hat{p} - E < p < \hat{p} + E$.

12. M&Ms azules Exprese el intervalo de confianza $0.270 + 0.073$ en la forma de $\hat{p} - E < p < \hat{p} + E$.

Construcción e interpretación de intervalos de confianza. ***En los ejercicios 13 a 16, utilice los datos muestrales y el nivel de confianza dados. En cada caso, (a) encuentre la mejor estimación puntual de la proporción poblacional p; (b) identifique el valor del margen de error E; (c) construya el intervalo de confianza; (d) escriba un enunciado que interprete correctamente el intervalo de confianza.***

13. Mickey D's En un estudio sobre la exactitud de las órdenes de comida rápida en auto, McDonald's tuvo 33 órdenes que no fueron exactas entre 362 órdenes observadas (según datos de la revista *QSR*). Construya un intervalo de confianza del 95% para la proporción de pedidos que no son exactos.

14. Eliquis El fármaco Eliquis (apixaban) se utiliza para ayudar a prevenir coágulos de la sangre en ciertos pacientes. En ensayos clínicos, entre 5924 pacientes tratados con Eliquis, 153 desarrollaron la reacción adversa de náuseas (según datos de Bristol-Myers Squibb Co). Construya un intervalo de confianza del 99% para la proporción de reacciones adversas.

15. Tasa de retorno de la encuesta En un estudio sobre el uso del teléfono celular y el predominio hemisférico del cerebro, se envió una encuesta por Internet a 5000 sujetos seleccionados aleatoriamente de un grupo en línea relacionado con el estudio de los oídos. Se devolvieron 717 encuestas. Construya un intervalo de confianza del 90% para la proporción de encuestas devueltas.

16. Negligencia médica En un estudio de 1228 demandas legales por negligencia médica seleccionadas al azar, se encontró que 856 de ellas fueron descartadas o desechadas (con base en datos de la Asociación de Aseguradores de Médicos de EUA). Construya un intervalo de confianza del 95% para la proporción de demandas por negligencia médica que se descartan o desechan.

Pensamiento crítico. ***En los ejercicios 17 a 28, use los datos y el nivel de confianza para construir una estimación de intervalo de confianza de p, después responda las preguntas.***

17. Nacimientos Una muestra aleatoria de 860 nacimientos en el Estado de Nueva York incluyó a 426 niños. Construya una estimación del intervalo de confianza del 95% para la proporción de varones en todos los nacimientos. Se cree que la proporción de niños entre todos los nacimientos es 0.512. ¿Estos resultados muestrales proporcionan evidencia fuerte contra esa creencia?

18. Genética mendeliana Uno de los famosos experimentos genéticos de Mendel arrojó 580 chícharos, de los cuales 428 eran verdes y 152 amarillos.

a. Encuentre una estimación del intervalo de confianza del 99% para el *porcentaje* de chícharos verdes.

b. Con base en su teoría de la genética, Mendel esperaba que 75% de los chícharos descendientes fueran verdes. Dado que el porcentaje de chícharos de la descendencia no es de 75%, ¿los resultados contradicen la teoría de Mendel? ¿Por qué sí o por qué no?

19. Exactitud en la comida rápida. En un estudio sobre la exactitud de las órdenes de comida rápida. Burger King tuvo 264 órdenes exactas y 54 que no lo fueron (con base en datos de la revista *QSR*).

a. Construya una estimación del intervalo de confianza del 99% para el porcentaje de órdenes que no son exactas.

b. Compare el resultado del inciso (a) con el intervalo de confianza del 99% para el porcentaje de órdenes que no son exactas en Wendy's: $6.2\% < p < 15.9\%$. ¿Qué se puede concluir?

20. OxyContin El medicamento OxyContin (oxycodone) se usa para tratar el dolor, pero es peligroso porque es adictivo y puede ser letal. En ensayos clínicos, 227 sujetos fueron tratados con OxyContin y 52 de ellos desarrollaron náuseas (de acuerdo con datos de Purdue Pharma L.P.).

a. Construya una estimación del 95% del intervalo de confianza para el porcentaje de usuarios de OxyContin que desarrollan náuseas.

b. Compare el resultado del inciso (a) con este intervalo de confianza del 95% para 5 sujetos que desarrollaron náuseas entre los 45 sujetos que recibieron placebo en lugar de OxyContin: $1.93\% < p < 20.3\%$. ¿Qué se puede concluir?

21. Terapia de contacto Cuando tenía 9 años de edad, Emily Rosa hizo un experimento en la feria de ciencias en el que probó a terapeutas de contacto profesionales para ver si podían detectar su campo de energía. Lanzaba una moneda para seleccionar su mano derecha o su mano izquierda y luego pedía a los terapeutas que identificaran la mano seleccionada colocando su mano justo debajo de la mano de Emily sin verla y sin tocarla. En 280 ensayos, los terapeutas de contacto acertaron 123 veces (con base en datos de "Un vistazo cercano a la terapia de contacto", *Journal of the American Medical Association*, vol. 279. núm. 13).

a. Dado que Emily utilizó un lanzamiento de monedas para seleccionar su mano derecha o su mano izquierda, ¿qué proporción de respuestas correctas se esperaría si los terapeutas de contacto hicieran conjeturas al azar?

b. Con base en los resultados muestrales de Emily, ¿cuál es la mejor estimación puntual de la tasa de éxito de los terapeutas?

c. Con base en los resultados muestrales de Emily, construya una estimación del intervalo de confianza del 99% de la proporción de respuestas correctas hechas por los terapeutas de contacto.

d. ¿Qué sugieren los resultados acerca de la capacidad de los terapeutas de contacto para seleccionar la mano correcta mediante la detección de un campo de energía?

22. Uso de medicamentos En una encuesta a 3005 adultos de 57 a 85 años, se encontró que el 81.7% de ellos usó al menos un medicamento recetado (con base en datos de "Use of Prescription and Over-the-Counter Medications and Dietary Supplements Among Older Adults in the United State", de Qato *et al*., *Journal of the American Medical Association*, vol. 300, núm. 24).

a. ¿Cuántos de los 3005 sujetos usaron al menos un medicamento recetado?

b. Construya una estimación del intervalo de confianza del 90% para el *porcentaje* de adultos de 57 a 85 años que usan al menos un medicamento recetado.

c. ¿Qué indican los resultados sobre la proporción de estudiantes universitarios que usan al menos un medicamento recetado?

23. Teléfonos celulares y cáncer Un estudio realizado a 420,095 usuarios daneses de teléfonos celulares encontró que 0.0321% de ellos desarrollaron cáncer cerebral o del sistema nervioso. Antes del estudio sobre el uso de teléfonos celulares, se encontró que la tasa de este tipo de cáncer era 0.0340% para aquellos que no utilizan teléfonos celulares. Los datos provienen de la *Journal of the National Cancer Institute*.

a. Utilice los datos muestrales para construir una estimación del intervalo de confianza del 90% para el porcentaje de usuarios de teléfonos celulares que desarrollan cáncer cerebral o del sistema nervioso.

b. ¿Los usuarios de teléfonos celulares parecen tener una proporción de cáncer cerebral o del sistema nervioso diferente de la proporción de cáncer entre los que no utilizan teléfonos celulares? ¿Por qué sí o por qué no?

24. No votantes que dicen haber votado En una encuesta de 1002 personas, 70% dijo que votaron en una reciente elección presidencial (según datos del *ICR Research Group*). Los registros electorales muestran que 61% de los votantes elegibles votaron.

a. Encuentre una estimación del intervalo de confianza del 98% para la proporción de personas que dicen haber votado.

b. ¿Los resultados de la encuesta son consistentes con la participación electoral efectiva del 61%? ¿Por qué sí o por qué no?

25. Lipitor En ensayos clínicos del fármaco Lipitor (atorvastatina), 270 sujetos recibieron placebo y 7 de ellos tuvieron reacciones alérgicas. Entre los 863 sujetos tratados con 10 mg del fármaco, 8 experimentaron reacciones alérgicas. Construya las dos estimaciones del intervalo de confianza del 95% para los porcentajes de reacciones alérgicas. Compare los resultados. ¿Qué se puede concluir?

26. Selección de género Antes de que sus ensayos clínicos fueran interrumpidos, el Instituto Genetics & IVF llevó a cabo un ensayo clínico del método XSORT diseñado para aumentar la probabilidad de concebir a una niña y, entre los 945 bebés nacidos de padres que usaban el método XSORT, hubo

continúa

879 niñas. El método YSORT fue diseñado para aumentar la probabilidad de concebir a un niño y, entre los 291 bebés nacidos de padres que usaban el método YSORT, hubo 239 niños. Construya las dos estimaciones del intervalo de confianza del 95% para los porcentajes de éxito. Compare los resultados. ¿Qué se puede concluir?

27. Dejar de fumar En un programa diseñado para ayudar a pacientes que quieren dejar de fumar, 198 pacientes recibieron atención *sostenida*, y el 82.8% de ellos ya no fumaban después de un mes. Entre los 199 pacientes que recibieron atención estándar, el 62.8% ya no fumaban después de un mes (según datos de "Sustained Care Intervention and Postdischarge Smoking Cessation Among Hospitalized Adults", Rigotti *et al*., *Journal of the American Medical Association*, vol. 312. núm. 7). Construya las dos estimaciones del intervalo de confianza del 95% para los porcentajes de éxito. Compare los resultados. ¿Qué se puede concluir?

28. Resultados medidos contra resultados reportados El mismo estudio citado en el ejercicio anterior produjo los siguientes resultados después de seis meses para los 198 pacientes que recibieron atención sostenida: el 25.8% ya no fumaba y los resultados fueron confirmados bioquímicamente, pero 40.9% de los pacientes informaron que ya no fumaban. Construya los dos intervalos de confianza del 95%. Compare los resultados. ¿Qué se puede concluir?

Uso de los conjuntos de datos del apéndice B. ***En los ejercicios 29 y 30, utilice el conjunto de datos indicado en el apéndice B.***

29. Estaturas de presidentes Consulte el conjunto de datos 15 "Presidentes" en el apéndice B. Trate los datos como una muestra y encuentre la proporción de presidentes que eran más altos que sus oponentes. Utilice ese resultado para construir una estimación del intervalo de confianza del 95% para el porcentaje de la población. Con base en el resultado, ¿parece que una mayor estatura es una ventaja para los candidatos presidenciales? ¿Por qué sí o por qué no?

30. M&M verdes El conjunto de datos 27 "Pesos de M&Ms" en el apéndice B incluye datos de 100 caramelos M&Ms, y 19 de ellos son verdes, la compañía de caramelos Mars afirma que 16% de sus caramelos M&M son verdes. Utilice los datos muestrales para construir una estimación del intervalo de confianza del 95% para el porcentaje de M&Ms verdes. ¿Qué concluye usted acerca de la afirmación del 16%?

Determinación del tamaño de la muestra. ***En los ejercicios 31 a 38, use los datos dados para encontrar el tamaño de muestra mínimo requerido para estimar una proporción o porcentaje poblacional.***

31. Zurdos Encuentre el tamaño de muestra necesario para estimar el porcentaje de residentes de California que son zurdos. Utilice un margen de error de tres puntos porcentuales y utilice un nivel de confianza del 99%.

a. Suponga que $\hat{p}$ y $\hat{q}$ son desconocidos.

b. Suponga que con base en estudios anteriores, alrededor del 10% de los californianos son zurdos.

c. ¿Cómo cambian los resultados de los incisos (a) y (b) si se usa todo EUA en lugar de California?

32. Varicela Usted planea realizar una encuesta para estimar el porcentaje de adultos que han tenido varicela. Encuentre el número de personas a encuestar si desea estar 90% seguro de que el porcentaje muestral está dentro de dos puntos porcentuales del porcentaje verdadero para toda la población de adultos.

a. Suponga que no se sabe nada sobre la prevalencia de la varicela.

b. Suponga que alrededor de 95% de los adultos han tenido varicela.

c. ¿El conocimiento adicional del inciso (b) tiene mucho efecto sobre el tamaño de la muestra?

33. Licenciatura en cuatro años En un estudio sobre la ayuda financiera del gobierno para estudiantes universitarios, se hace necesario estimar el porcentaje de estudiantes universitarios de tiempo completo que obtienen una licenciatura en cuatro años o menos. Encuentre el tamaño de muestra necesario para estimar ese porcentaje. Utilice un margen de error de 0.05 y utilice un nivel de confianza del 95%.

a. Suponga que no se conoce nada sobre el porcentaje a estimar.

b. Suponga que estudios anteriores han demostrado que alrededor de 40% de los estudiantes a tiempo completo obtienen títulos de licenciatura en cuatro años o menos.

c. ¿El conocimiento adicional del inciso (b) tiene mucho efecto sobre el tamaño de la muestra?

34. Astrología Un sociólogo planea realizar una encuesta para estimar el porcentaje de adultos que creen en la astrología. ¿Cuántas personas deben ser encuestadas si queremos un nivel de confianza del 99% y un margen de error de cuatro puntos porcentuales?

a. Suponga que no se conoce nada sobre el porcentaje a estimar.

b. Utilice la información de una encuesta anterior de Harris en la que 26% de los encuestados dijeron que creían en la astrología.

35. Asientos de aerolínea Usted es el gerente de operaciones de American Airlines y está considerando aplicar un nivel tarifario más alto para pasajeros en los asientos del pasillo. Desea estimar el porcentaje de pasajeros que ahora prefieren los asientos de pasillo. ¿Cuántos pasajeros aéreos seleccionados aleatoriamente deben ser encuestados? Suponga que desea estar 95% seguro de que el porcentaje muestral está dentro de 2.5 puntos porcentuales del porcentaje de poblacional real.

a. Suponga que no se sabe nada sobre el porcentaje de pasajeros que prefieren asientos de pasillo.

b. Suponga que una encuesta anterior sugiere que aproximadamente 38% de los pasajeros aéreos prefieren un asiento de pasillo (según una encuesta de *3M Privacy Filters*).

36. iOS Marketshare Usted planea desarrollar una nueva aplicación de juego social iOS que, cree, superará el éxito de Angry Birds y Facebook combinados. En la prospección de ingresos, debe calcular el porcentaje de todos los dispositivos de teléfonos inteligentes y tablets que utilizan el sistema operativo iOS en comparación con Android y otros sistemas operativos. ¿Cuántos teléfonos inteligentes y tablets deben ser encuestados para estar seguros de que su estimación tiene un error no mayor que dos puntos porcentuales?

a. Suponga que no se sabe nada sobre el porcentaje de dispositivos portátiles que utilizan el sistema operativo iOS.

b. Suponga que una encuesta reciente sugiere que aproximadamente 43% de los teléfonos inteligentes y las tabletas utilizan el sistema operativo iOS (según datos de NetMarketShare).

c. ¿La información adicional de la encuesta en el inciso (b) tiene mucho efecto sobre el tamaño de la muestra requerido?

37. Videojuegos Una inversionista está considerando la posibilidad de financiar un nuevo videojuego y desea conocer el porcentaje mundial de personas que juegan videojuegos, por lo que está planeando una encuesta. ¿Cuántas personas deben ser encuestadas para tener una confianza del 90% de que el porcentaje estimado está dentro de tres puntos porcentuales del porcentaje de poblacional real?

a. Suponga que no se sabe nada sobre el porcentaje mundial de personas que juegan videojuegos.

b. Suponga que alrededor de 16% de las personas juegan videojuegos (según un informe de Spil Games).

c. Dado que el tamaño de muestra requerido es relativamente pequeño, ¿podría la inversionista simplemente encuestar a las personas que conoce?

38. Mujeres que dan a luz Un epidemiólogo planea realizar una encuesta para estimar el porcentaje de mujeres que dan a luz. ¿Cuántas mujeres deben ser encuestadas para estar 99% confiados en que el porcentaje estimado tiene un error no mayor que dos puntos porcentuales?

a. Suponga que no se conoce nada sobre el porcentaje a estimar.

b. Suponga que un estudio previo realizado por la Oficina del Censo de EUA mostró que 82% de las mujeres dan a luz.

c. ¿Qué hay de erróneo en encuestar a mujeres adultas seleccionadas al azar?

7-1 Más allá de lo básico

39. Factor de corrección de población finita Para las fórmulas 7-2 y 7-3, suponemos que la población es infinita o muy grande y que estamos muestreando con reemplazo. Cuando muestreamos sin reemplazo de una población relativamente pequeña con tamaño N, modificamos E para incluir el factor de *corrección de población finita* mostrado aquí, y podemos despejar n para obtener el resultado aquí dado. Utilice este resultado para repetir el inciso (b) del ejercicio 38, suponiendo que limitamos nuestra población a un condado con 2500 mujeres que han completado el tiempo durante el cual pueden dar a luz.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\sqrt{\frac{N-n}{N-1}} \qquad n = \frac{N\hat{p}\hat{q}[z_{\alpha/2}]^2}{\hat{p}\hat{q}[z_{\alpha/2}]^2 + (N-1)E^2}$$

40. Intervalo de confianza unilateral Una afirmación unilateral sobre una proporción de la población es que la proporción es menor que (o mayor que) un valor específico. Dicha afirmación puede abordarse formalmente mediante el uso de un *intervalo de confianza unilateral* para p, que puede expresarse como $p < \hat{p} + E$ o $p > \hat{p} - E$, donde el margen de error E se modifica reemplazando $z_{\alpha/2}$ por z_{α}. (En lugar de dividir α entre dos colas de la distribución normal estándar, se coloca todo en una cola). El problema del capítulo se refiere a una encuesta Gallup a 1487 adultos que muestra que 43% de los encuestados tienen páginas de Facebook. Utilice esos datos para construir un intervalo de confianza unilateral del 95% que sería adecuado para ayudar a determinar si la proporción de todos los adultos que tienen páginas de Facebook es inferior al 50%.

41. Manejo de la falta de éxito De acuerdo con la *Regla de Tres*, cuando tenemos un tamaño de muestra n con $x = 0$ éxitos, tenemos 95% de confianza de que la proporción poblacional verdadera tiene un límite superior de $3/n$. (Vea “A Look at the Rule of Three”, de Jovanovic y Levy, en *American Statistician*, vol. 51, núm. 2).

a. Si n ensayos independientes resultan en ningún éxito, ¿por qué no podemos encontrar límites del intervalo de confianza usando los métodos descritos en esta sección?

b. Si 40 parejas usan un método de selección de género y cada pareja tiene una niña, ¿cuál es el límite superior del 95% para p, la proporción de todos los bebés que son niños?

7-2 Estimación de un promedio poblacional

Concepto clave El objetivo principal de esta sección es presentar métodos para usar una media muestral $\bar{x}$ con el fin de hacer una inferencia sobre el valor de la media poblacional correspondiente μ. Hay tres conceptos principales incluidos en esta sección:

- **Estimación puntual:** La media muestral $\bar{x}$ es la mejor *estimación puntual* (o estimación de valor único) de la media poblacional μ.
- **Intervalo de confianza:** Utilice datos muestrales para construir e interpretar una estimación del *intervalo de confianza* para el valor real de una media poblacional μ.
- **Tamaño de muestra:** Encuentre el tamaño de muestra necesario para estimar una media poblacional.

La parte 1 de esta sección trata del caso real que se usa comúnmente para estimar μ, donde se desconoce la desviación estándar σ de la población. La parte 2 incluye una breve descripción del procedimiento utilizado cuando se conoce σ, lo cual es muy raro.

PARTE 1 Estimación de un promedio poblacional cuando no se conoce σ

Es poco frecuente que queramos estimar el valor desconocido de una media poblacional μ pero que de alguna manera conozcamos el valor de la desviación estándar de la población σ, por lo que la parte 1 se centra en la situación real en la que σ no se conoce.

Estimación puntual Como se analizó en la sección 6-3, la media muestral $\bar{x}$ es un *estimador no sesgado* de la media poblacional μ. Además, para muchas poblaciones, las medias muestrales tienden a variar menos que otras medidas de tendencia central. Por estas razones, la media muestral $\bar{x}$ suele ser la mejor estimación puntual de la media poblacional μ.

La media muestral $\bar{x}$ es la mejor estimación puntual de la media poblacional μ.

Debido a que incluso la mejor estimación puntual no proporciona ninguna indicación de lo exacta que es, utilizamos un *intervalo de confianza* (o *estimación del intervalo*), que consiste en un intervalo (o rango) de valores en vez de un solo valor.

Intervalo de confianza El cuadro adjunto incluye los elementos clave para construir una estimación del intervalo de confianza para una media poblacional μ en la situación común donde σ no se conoce.

ELEMENTOS CLAVE

Intervalo de confianza para estimar una media de población con σ no conocida

Objetivo

Construir un intervalo de confianza utilizado para estimar una media poblacional.

Notación

μ = media poblacional
$\bar{x}$ = media muestral
s = desviación estándar muestral
n = número de valores muestrales
E = margen de error

Requisitos

1. La muestra es una muestra aleatoria simple.
2. Se cumple cualquiera de estas dos condiciones, o ambas: la población se distribuye normalmente o $n > 30$.

Intervalo de confianza

Formatos: $\bar{x} - E < \mu < \bar{x} + E$ o $\bar{x} \pm E$ o $(\bar{x} - E, \bar{x} + E)$

- **Margen de error:** $E = t_{\alpha/2} \cdot \dfrac{s}{\sqrt{n}}$ (Utilice gl $= n - 1$).
- **Nivel de confianza:** El intervalo de confianza se asocia con un nivel de confianza, como 0.95 (o 95%), y α es el complemento del nivel de confianza. Para un nivel de confianza de 0.95 (o 95%), $\alpha = 0.05$.
- **Valor crítico:** $t_{\alpha/2}$ es el valor crítico de t que separa un área de $\alpha/2$ en la cola derecha de la distribución t de Student.
- **Grados de libertad:** gl $= n - 1$ es el número de grados de libertad utilizados para encontrar el valor crítico.

Regla de redondeo

1. *Datos originales:* Al utilizar un *conjunto original de valores de datos*, redondee los límites del intervalo de confianza a un decimal más que el utilizado para ese conjunto original de datos.
2. *Estadísticos resumidos*: Cuando se usan los *datos estadísticos resumidos* n, $\bar{x}$ y s, redondee los límites del intervalo de confianza al mismo número de decimales utilizados para la media muestral.

Requisito de "Normalidad o $n > 30$"

Normalidad El método para encontrar una estimación del intervalo de confianza de μ es *robusto* contra un alejamiento de la normalidad, lo que significa que el requisito de normalidad es *holgado*. No es necesario que la distribución tenga una forma de campana perfecta, pero debe parecer algo simétrica, con una moda y sin valores atípicos.

Tamaño de muestra $n > 30$ Esta es una directriz común, pero los tamaños de muestra de 15 a 30 son adecuados si la población parece tener una distribución que no está lejos de ser normal y no hay valores atípicos. Para algunas distribuciones de población que están extremadamente lejos de lo normal, el tamaño de muestra podría necesitar ser mayor de 30. Este texto utiliza el criterio simplificado de $n > 30$ como justificación para tratar la distribución de las medias muestrales como una distribución normal.

Distribución *t* de Student

En esta sección usamos una *distribución t de Student*, que comúnmente se conoce como una "distribución *t*". Fue desarrollada por William Gosset (1816-1937), un empleado de la fábrica de cerveza Guinness que necesitaba una distribución que pudiera usarse con pequeñas muestras. La cervecería prohibió la publicación de los resultados de la investigación, pero Gosset consiguió evadir la prohibición publicándolos bajo el seudónimo "Student". (Estrictamente en el interés de servir mejor a sus lectores, el autor visitó la cervecería Guinness y se sintió obligado a probar algunos de los productos). A continuación se presentan algunos puntos clave sobre la distribución *t* de Student:

- **Distribución *t* de Student** Si una población tiene una distribución normal, entonces la distribución de

$$t = \frac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}}$$

  es una **distribución *t* de Student** para todas las muestras de tamaño *n*. Una distribución *t* de Student se conoce comúnmente como una **distribución *t***.

- **Grados de libertad** La determinación de un valor crítico $t_{\alpha/2}$ requiere un valor para los **grados de libertad** (o **gl**). En general, el número de grados de libertad para una colección de datos muestrales es la cantidad de valores muestrales que pueden variar después de que se hayan impuesto ciertas restricciones a todos los valores de datos. (*Ejemplo*: si 10 calificaciones de examen tienen la restricción de que su media sea 80, entonces su suma debe ser 800, y podemos asignar libremente valores a las primeras 9 calificaciones, pero después se determinará la décima puntuación; entonces, en este caso hay 9 grados de libertad). Para los métodos de esta sección, el número de grados de libertad es el tamaño de muestra menos 1.

$$\textbf{Grados de libertad} = \boldsymbol{n - 1}$$

- **Determinación del valor crítico $t_{\alpha/2}$** Un valor crítico $t_{\alpha/2}$ se puede encontrar usando la tecnología o la tabla A-3. La tecnología se puede usar con cualquier cantidad de grados de libertad, pero la tabla A-3 puede emplearse únicamente para ciertos números seleccionados de grados de libertad. Si usa la tabla A-3 para encontrar un valor crítico $t_{\alpha/2}$, pero la tabla no incluye el número exacto de grados de libertad, usted podría usar el valor más cercano, también podría ser conservador y usar el siguiente número menor de grados de libertad encontrado en la tabla, o bien podría interpolar.

- La distribución *t* de Student es diferente para los distintos tamaños de muestra. (Vea la figura 7-4 para los casos $n = 3$ y $n = 12$).

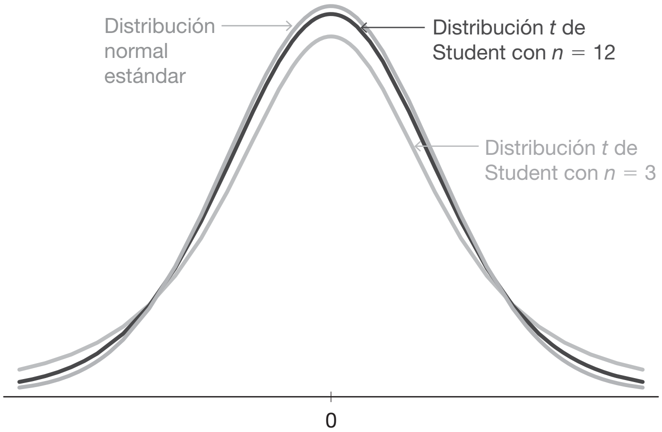


FIGURA 7.4 Distribuciones *t* de Student para *n* = 3 y *n* = 12

La distribución *t* de Student tiene la misma forma general y simetría que la distribución normal estándar, pero tiene la mayor variabilidad esperada con muestras pequeñas.

- La distribución t de Student tiene la misma forma general de campana simétrica que la distribución normal estándar, pero posee más variabilidad (con distribuciones más amplias), como se espera con muestras pequeñas.
- La distribución t de Student tiene una media de $t = 0$ (al igual que la distribución normal estándar tiene una media de $z = 0$).
- La desviación estándar de la distribución t de Student varía según el tamaño de la muestra, pero es mayor que 1 (a diferencia de la distribución normal estándar, que tiene $\sigma = 1$).
- Conforme el tamaño de muestra n se hace más grande, la distribución t de Student se acerca a la distribución normal estándar.

Procedimiento para construir un intervalo de confianza para μ

Los intervalos de confianza pueden construirse fácilmente con la tecnología o de manera manual mediante el siguiente procedimiento.

1. Verifique que se satisfagan los dos requisitos: la muestra es una muestra aleatoria simple y la población se distribuye normalmente o $n > 30$.
2. Con σ desconocida (como suele ser el caso), use $n - 1$ grados de libertad y utilice tecnología o una tabla de la distribución t (como la tabla A-3) para encontrar el valor crítico $t_{\alpha/2}$ que corresponde al nivel de confianza deseado.
3. Evalúe el margen de error usando $E = t_{\alpha/2} \cdot s/\sqrt{n}$.
4. Usando el valor del margen de error calculado E y el valor de la media muestral $\bar{x}$, sustituya esos valores en uno de los formatos para el intervalo de confianza: $\bar{x} - E < \mu < \bar{x} + E$ o $\bar{x} \pm E$ o $(\bar{x} - E, \bar{x} + E)$.

Redondee los límites del intervalo de confianza resultante de la siguiente manera: con un *conjunto original de valores de datos*, redondee los límites del intervalo de confianza a un lugar decimal más de lo que se usa para ese conjunto original de datos, pero cuando use los *estadísticos de resumen* n, $\bar{x}$ y s, redondee los límites del intervalo de confianza al mismo número de decimales utilizados para la media muestral.

Estimación de tamaños de población en la vida silvestre

La Ley Nacional de Administración de Bosques de EUA protege a especies en peligro de extinción, entre las que destaca el búho moteado del norte; la ley impidió que la industria silvícola talara vastas regiones de árboles en la costa noroeste del Pacífico. Se pidió a biólogos y a especialistas en estadística que analizaran el problema, y ellos concluyeron que estaban disminuyendo las tasas de supervivencia y los tamaños de las poblaciones de los búhos hembra, quienes desempeñan un papel importante en la supervivencia de la especie. Los biólogos y especialistas en estadística también estudiaron el salmón en los ríos Snake y Columbia del estado de Washington, así como a los pingüinos en Nueva Zelanda. En el artículo “Sampling Wildlife Populations” (*Chance*, vol. 9, núm. 2), los autores Brian Manly y Lyman McDonald comentan que, en estudios de esta clase, “los biólogos ganan habilidades de modelado, que son una característica distintiva de la buena estadística. Por su parte, los especialistas en estadística aprenden a compenetrarse en la realidad de los problemas, ya que los biólogos los introducen en asuntos cruciales”.

Estimación del azúcar en las naranjas

En Florida, los miembros de la industria de los cítricos usan profusamente métodos estadísticos. Una aplicación específica tiene que ver con la forma en que se paga a los agricultores por las naranjas que se usan para elaborar el jugo de naranja. Cuando llega un camión cargado con naranjas, primero se pesa la carga en la planta receptora, luego se elige al azar una muestra de una docena de naranjas. La muestra se pesa, se exprime y se mide la cantidad de azúcar que contiene el jugo. Con base en los resultados de la muestra, se estima la cantidad total de azúcar contenida en toda la carga del camión. El pago por la carga de naranjas se basa en la estimación de la cantidad de azúcar, ya que las naranjas más dulces son más valiosas que las menos dulces, aun cuando las cantidades de jugo sean iguales.

EJEMPLO 1 Determinación de un valor crítico $t_{\alpha/2}$

Encuentre el valor crítico $t_{\alpha/2}$ correspondiente a un nivel de confianza del 95%, dado que la muestra tiene el tamaño $n = 15$.

SOLUCIÓN

Debido a que $n = 15$, el número de grados de libertad es $n - 1 = 14$. El nivel de confianza del 95% corresponde a $\alpha = 0.05$, por lo que hay un área de 0.025 en cada una de las dos colas de la distribución t, como se muestra en la figura 7-5.

Uso de la tecnología La tecnología se puede usar para encontrar que para 14 grados de libertad y un área de 0.025 en cada cola, el valor crítico es $t_{\alpha/2} = t_{0.025} = 2.145$.

Uso de la tabla A-3 Para encontrar el valor crítico usando la tabla A-3, considere la columna con 0.05 para el "Área en dos colas" (o la misma columna con 0.025 para el "Área en una cola"). El número de grados de libertad es gl $= n - 1 = 14$. Obtenemos $t_{\alpha/2} = t_{0.025} = 2.145$.

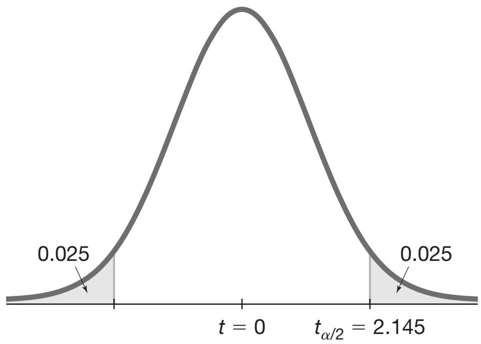


FIGURA 7-5 Valor crítico $t_{\alpha/2}$

SU TURNO Encuentre el valor crítico para el ejercicio 2 "Grados de libertad".

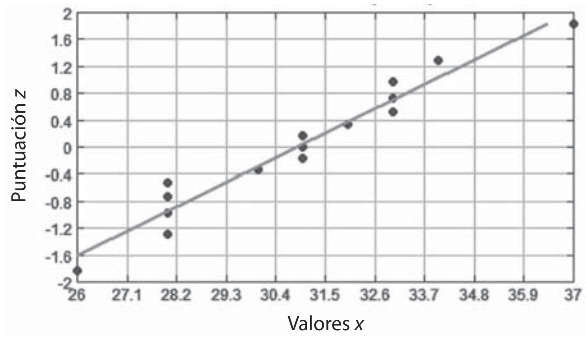
EJEMPLO 2 Intervalo de confianza usando pesos al nacer

A continuación se listan los pesos (hectogramas o hg) de niñas seleccionadas al azar en el momento del nacimiento, según los datos del Centro Nacional de Estadísticas de la Salud. Los estadísticos de resumen son: $n = 15$, $\bar{x} = 30.9$ hg, $s = 2.9$ hg. Use los datos muestrales para construir un intervalo de confianza del 95% para el peso medio de las niñas.

33 28 33 37 31 32 31 28 34 28 33 26 30 31 28

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Primero debemos verificar que los requisitos sean satisfechos. (1) La muestra es una muestra aleatoria simple. (2) Dado que el tamaño de la muestra es $n = 15$, el requisito de que "la población se distribuya normalmente o el tamaño de muestra sea mayor que 30" puede satisfacerse sólo si los datos muestrales parecen provenir de una población normalmente distribuida, por ello necesitamos investigar la normalidad. La gráfica cuantilar normal adjunta muestra que los datos muestrales parecen provenir de una población normalmente distribuida, por lo que se satisface este segundo requisito.



Uso de software Se puede usar software para construir automáticamente el intervalo de confianza. (Vea las instrucciones al final de esta sección). Aquí se muestra un ejemplo de la pantalla de StatCrunch resultante de los 15 pesos al nacer. La pantalla muestra el límite inferior del intervalo de confianza (29.247163) y el límite superior del intervalo de confianza (32.486171). Después de redondear a un decimal (como en la media de la muestra), podemos expresar el intervalo de confianza del 95% como 29.2 hg $< \mu <$ 32.5 hg.

StatCrunch

95% confidence interval results:
μ : Mean of variable

Variable	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
BirthWt	30.866667	0.75508856	14	29.247163	32.486171

Uso de la tabla de la distribución *t* Con base en la tabla A-3, el valor crítico es $t_{0.025} = 2.145$ como se muestra en el ejemplo 1. Ahora encontramos el margen de error E como se muestra a continuación:

$$E = t_{\alpha/2}\frac{s}{\sqrt{n}} = 2.145 \cdot \frac{2.9}{\sqrt{15}} = 1.606126$$

Con $\bar{x} = 30.9$ hg y $E = 1.606126$ hg, construimos el intervalo de confianza de la siguiente manera:

$$\bar{x} - E < \mu < \bar{x} + E$$

$$30.9 - 1.606126 < \mu < 30.9 + 1.606126$$

$$29.3 \text{ hg} < \mu < 32.5 \text{ hg} \quad \text{(redondeado a un decimal)}$$

El límite inferior del intervalo de confianza (29.3 hg) es en realidad de 29.2 hg si usamos tecnología y también si usamos los estadísticos de resumen con más decimales que el lugar decimal utilizado en el cálculo anterior.

INTERPRETACIÓN

Estamos 95% seguros de que los límites de 29.2 hg y 32.5 hg en realidad contienen el valor de la media poblacional μ. Si tuviéramos que recolectar muchas muestras aleatorias diferentes de 15 niñas recién nacidas y encontrar el peso promedio en cada muestra, aproximadamente el 95% de los intervalos de confianza resultantes deberían contener el valor del peso promedio de todas las niñas recién nacidas.

SU TURNO Resuelva el ejercicio 15 “Genes”.

Interpretación del intervalo de confianza

El intervalo de confianza está asociado con un **nivel de confianza**, como 0.95 (o 95%). Al interpretar una estimación del intervalo de confianza de μ, tenga en cuenta que el nivel de confianza proporciona la *tasa de éxito del procedimiento* utilizado para construir el intervalo de confianza.

Los números de serie de tanques capturados revelan el tamaño de la población

Durante la Segunda Guerra Mundial, especialistas en espionaje del bando de los Aliados querían determinar el número de tanques que Alemania estaba produciendo. Las técnicas de espionaje tradicionales produjeron resultados poco confiables, pero los especialistas en estadística obtuvieron estimaciones exactas al analizar los números de serie de los tanques capturados. Por ejemplo, los registros indican que, en junio de 1941, Alemania realmente produjo 271 tanques. La estimación basada en los números de serie fue de 244, en tanto que los métodos de espionaje tradicionales dieron como resultado una estimación extrema de 1550. (Vea “An Empirical Approach to Economic Intelligence in World War II”, de Ruggles y Brodie, en *Journal of the American Statistical Association*, vol. 42).

Por ejemplo, la estimación del intervalo de confianza del 95% de 29.2 hg $< \mu <$ 32.5 hg se puede interpretar de la siguiente manera:

> **"Estamos 95% seguros de que el intervalo de 29.2 hg a 32.5 hg en realidad contiene el verdadero valor de μ".**

Con "95% seguros" queremos decir que si tuviéramos que seleccionar muchas muestras diferentes del mismo tamaño y construir los intervalos de confianza correspondientes, en el largo plazo, 95% de los intervalos de confianza debería contener el valor de μ.

Estimación puntual y margen de error *E* a partir de un intervalo de confianza

Los artículos de tecnología y las revistas suelen expresar un intervalo de confianza en un formato como (10.0, 30.0). La media muestral $\bar{x}$ es el valor intermedio entre esos límites, y el margen de error E es la mitad de la diferencia entre esos límites (porque el límite superior es $\bar{x} + E$ y el límite inferior es $\bar{x} - E$, la distancia que los separa es $2E$).

$$\text{Estimación puntual de } \mu\text{: } \bar{x} = \frac{(\text{límite de confianza superior}) + (\text{límite de confianza inferior})}{2}$$

$$\text{Margen de error: } E = \frac{(\text{límite de confianza superior}) - (\text{límite de confianza inferior})}{2}$$

Por ejemplo, el intervalo de confianza (10.0, 30.0) produce $\bar{x} = 20.0$ y $E = 10.0$.

Uso de intervalos de confianza para describir, explorar o comparar datos

En algunos casos, los intervalos de confianza pueden estar entre las diferentes herramientas utilizadas para describir, explorar o comparar conjuntos de datos, como en el siguiente ejemplo.

EJEMPLO 3 Fumadores pasivos

La figura 7-6 muestra gráficas de estimaciones del intervalo de confianza para el nivel medio de cotinina en cada una de tres muestras: (1) personas que fuman, (2) personas que no fuman pero que están expuestas al humo de tabaco en el hogar o en el trabajo, (3) personas que no fuman y no están expuestas al humo. (Los datos de la muestra se listan en el conjunto de datos 12 "Fumadores pasivos y activos" del apéndice B). Debido a que la cotinina es producida por el cuerpo cuando absorbe nicotina, es una buena indicación de la ingesta de nicotina. La figura 7-6 nos ayuda a ver los efectos del humo en los fumadores pasivos. En la figura 7-6, vemos que el intervalo de confianza para fumadores no se superpone a los otros intervalos de confianza, por lo que parece que el nivel medio de cotinina de los fumadores es diferente al de los otros dos grupos. Los dos grupos de no fumadores tienen intervalos de confianza que se superponen, por lo que es posible que tengan el mismo nivel de cotinina. Resulta útil comparar los intervalos de confianza o sus gráficas, pero tales comparaciones no deben usarse para formular conclusiones formales y finales sobre la igualdad de medias. En los capítulos 9 y 12 se introducen métodos mejores para realizar comparaciones formales entre medias.

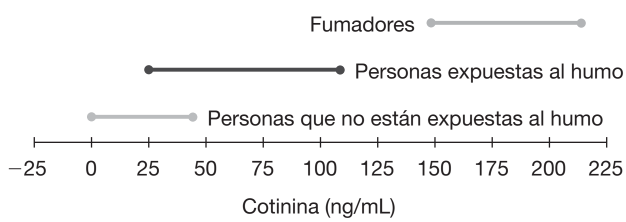


FIGURA 7-6 Comparación de intervalos de confianza

SU TURNO Resuelva el ejercicio 25 "Pulsos".

Estimación del tamaño de multitudes

Existen métodos complejos para analizar el tamaño de una multitud. Se pueden emplear fotografías aéreas y medidas de densidad demográfica con una exactitud bastante razonable. Sin embargo, los reportes de estimaciones del tamaño de multitudes a menudo son simples conjeturas. Después de que los Medias Rojas de Boston ganaron la Serie Mundial por primera vez en 86 años, las autoridades de la ciudad de Boston estimaron que a la celebración callejera acudieron 3.2 millones de aficionados. La policía de Boston hizo una estimación de alrededor de un millón de personas, pero aceptó que este cálculo se basaba en conjeturas de los comandantes de la policía. Un análisis fotográfico produjo una estimación de alrededor de 150,000. El profesor Farouk El-Baz de la Universidad de Boston utilizó imágenes del U.S. Geological Survey para llegar a una estimación de casi 400,000. El físico Bill Donnelly del MIT dijo que "es un problema serio que la gente indique un número cualquiera. Esto significa que otros tantos asuntos no se investigan de manera cuidadosa".

PRECAUCIÓN Los intervalos de confianza pueden usarse de manera *informal* para comparar diferentes conjuntos de datos, pero la *superposición de los intervalos de confianza no debe utilizarse para formular conclusiones formales y finales sobre la igualdad de medias*.

Determinación del tamaño de muestra

Si queremos recolectar una muestra que se usará para estimar una población promedio μ, ¿*cuántos* valores muestrales necesitamos? Al determinar el tamaño de muestra necesario para estimar una media poblacional, debemos tener un valor estimado o conocido de la desviación estándar poblacional σ, de modo que podamos usar la fórmula 7-4 que se muestra en el cuadro adjunto Elementos clave.

ELEMENTOS CLAVE

Determinación del tamaño de muestra requerido para estimar una media poblacional

Objetivo

Determine el tamaño de muestra n necesario para estimar el valor de una población promedio μ.

Notación

μ = media poblacional

σ = desviación estándar poblacional

$\bar{x}$ = media muestral

E = margen de error deseado

$z_{\alpha/2}$ = puntuación z que separa un área de $\alpha/2$ en la cola derecha de la distribución normal estándar

Requisito

La muestra debe ser una muestra aleatoria simple.

Tamaño de la muestra

El tamaño de muestra requerido se encuentra mediante el uso de la fórmula 7-4.

Fórmula 7-4
$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2$$

Regla de redondeo

Si el tamaño de muestra calculado n no es un número entero, redondee el valor de n al siguiente número entero mayor.

Tamaño de la población La fórmula 7-4 no depende del tamaño (N) de la población (excepto en los casos en que se selecciona una muestra relativamente grande sin reemplazo de una población finita).

Redondeo El tamaño muestral debe ser un número entero porque es el número de valores muestrales que se debe encontrar, pero la fórmula 7-4 generalmente da un resultado que no es un número entero. La regla de redondeo se basa en el principio de que, cuando es necesario, el tamaño de muestra requerido debe redondearse hacia *arriba* para que sea al menos suficientemente grande en lugar de ser demasiado pequeño.

Tratamiento de una σ desconocida al encontrar el tamaño de muestra La fórmula 7-4 requiere que sustituyamos un valor conocido por la desviación estándar de la población σ, aunque sea más bien desconocida. Cuando se determina el tamaño de muestra requerido (sin construir un intervalo de confianza), aquí hay algunas maneras en que podemos resolver el problema de no conocer el valor de σ:

1. Utilice la regla práctica del rango (vea la sección 3-2) para estimar la desviación estándar de la siguiente manera: $\sigma \approx \text{rango}/4$, donde el rango se determina a partir de los datos muestrales. (Con una muestra de 87 o más valores seleccionados al azar de una población normalmente distribuida, el rango/4 dará como resultado un valor mayor o igual que 0 al menos el 95% del tiempo).
2. Inicie el proceso de recolección de muestras sin conocer σ y, usando los primeros valores, calcule la desviación estándar muestral s y úsela en lugar de σ. El valor estimado de σ puede mejorarse a medida que se obtienen más datos muestrales, y el tamaño de muestra requerido se puede ajustar a medida que recolecta más datos muestrales.
3. Estime el valor de σ utilizando los resultados de algún otro estudio anterior. En ocasiones podemos ser creativos en nuestro uso de otros resultados conocidos. Por ejemplo. Las pruebas de IQ de Wechsler están diseñadas para que la desviación estándar sea de 15. Los estudiantes de estadística tienen puntuaciones de IQ con una desviación estándar menor que 15 porque son un grupo más homogéneo que las personas seleccionadas al azar de la población general. No sabemos el valor específico de σ para estudiantes de estadística, pero podemos estar seguros al usar $\sigma = 15$. La utilización de un valor para σ que sea mayor que el valor verdadero hará que el tamaño de muestra sea más grande de lo necesario, pero usando un valor para σ que sea demasiado pequeño daría como resultado que el tamaño de muestra sea inadecuado. *Al determinar el tamaño de muestra* n*, cualquier error siempre debe ser conservador, es decir, que tienda a que el tamaño de muestra sea muy grande en lugar de muy pequeño.*

EJEMPLO 4 Puntuaciones de IQ de estudiantes de estadística

Suponga que queremos estimar la puntuación promedio de IQ para la población de estudiantes de estadística. ¿Cuántos estudiantes deben seleccionarse al azar para las pruebas de IQ si queremos un 95% de confianza de que la media de la muestra esté dentro de 3 puntos de IQ de la media de la población?

SOLUCIÓN

Para un intervalo de confianza del 95%, tenemos $\alpha = 0.05$, por lo que $z_{\alpha/2} = 1.96$. Debido a que queremos que la muestra esté dentro de 3 puntos de IQ desde μ, el margen de error es $E = 3$. También podemos suponer que $\sigma = 15$ (vea el análisis anterior a este ejemplo). Si utilizamos la fórmula 7-4, resulta

$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2 = \left[\frac{1.96 \cdot 15}{3}\right]^2 = 96.04 = 97 \quad \text{(redondeado \textit{hacia arriba})}$$

INTERPRETACIÓN

Entre los miles de estudiantes de estadística, necesitamos obtener una muestra aleatoria simple de al menos 97 de sus puntuaciones de IQ. Con una muestra aleatoria simple de sólo 97 estudiantes de estadística, estaremos 95% seguros de que la media de la muestra $\overline{x}$ se encuentra dentro de 3 puntos de IQ de la media poblacional real μ.

SU TURNO Resuelva el ejercicio 29 “IQ promedio de profesores universitarios”.

PARTE 2 Estimación de una media poblacional cuando se conoce σ

En el mundo real de los estadísticos profesionales y las revistas e informes profesionales, es extremadamente raro que queramos estimar el valor desconocido de una media poblacional μ, pero que de alguna manera conozcamos el valor de la desviación estándar de la población σ. Si esto es así, el intervalo de confianza se construye utilizando la distribución normal estándar en lugar de la distribución t de Student, por lo que se puede usar el mismo procedimiento de la parte 1 con el siguiente margen de error:

$$\text{Margen de error: } E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad \text{(utilizado con } \sigma \text{ conocida)}$$

EJEMPLO 5 Estimación del intervalo de confianza de μ con σ conocida

Use los mismos 15 pesos al nacer de niñas que se dieron en el ejemplo 2, para los cuales $n = 15$ y $\bar{x} = 30.9$ hg. Construya una estimación del intervalo de confianza del 95% para el peso promedio al nacer de todas las niñas suponiendo que se sabe que σ es 2.9 hg.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Los requisitos se revisaron en el ejemplo 2. Por lo tanto, los requisitos se cumplen. ✓

Con un nivel de confianza del 95%, tenemos $\alpha = 0.05$, y obtenemos $z_{\alpha/2} = 1.96$ (como en el ejemplo 2 de la sección 7-1). Usando $z_{\alpha/2} = 1.96$, $\sigma = 2.9$ hg, y $n = 15$, encontramos el valor del margen de error E:

$$\begin{aligned} E &= z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 1.96 \cdot \frac{2.9}{\sqrt{15}} = 1.46760 \end{aligned}$$

Con $\bar{x} = 30.9$ y $E = 1.46760$, encontramos el intervalo de confianza del 95% de la siguiente manera:

$$\begin{aligned} \bar{x} - E < \mu &< \bar{x} + E \\ 30.9 - 1.46760 < \mu &< 30.9 + 1.46760 \\ 29.4 \text{ hg} < \mu &< 32.4 \text{ hg} \quad \text{(redondeado a un decimal)} \end{aligned}$$

El intervalo de confianza encontrado aquí usando la distribución normal es ligeramente más estrecho que el intervalo de confianza encontrado usando la distribución t en el ejemplo 2. Debido a que $z_{\alpha/2} = 1.96$ es menor que $t_{\alpha/2} = 2.145$, el margen de error E es menor y el intervalo de confianza es más estrecho. El valor crítico $t_{\alpha/2}$ es mayor porque la distribución t incorpora la cantidad de variación mayor que se obtiene al utilizar muestras más pequeñas.

Recuerde, este ejemplo ilustra la situación en que se conoce la desviación estándar de la población σ, lo cual es raro. La situación más realista con una σ desconocida se analizó en la parte 1 de esta sección.

SU TURNO Resuelva el ejercicio 37 "Pesos al nacer de niñas".

Elección de la distribución adecuada

Cuando se construye una estimación del intervalo de confianza de la media poblacional μ, es importante usar la distribución correcta. La tabla 7-1 en la página siguiente resume los puntos clave a considerar.

TABLA 7-1 Elección entre las distribuciones *t* de Student y *z* (normal)

Condiciones	Método
σ desconocida y población normalmente distribuida o σ desconocida y $n > 30$	Use la distribución ***t*** de Student
σ conocida y población normalmente distribuida o σ conocida y $n > 30$ (en realidad, σ se conoce pocas veces)	Use la distribución ***normal (z)***
La población no se distribuye normalmente y $n \leq 30$	Use el método de bootstrap (sección 7-4) o un método no paramétrico

CENTRO DE TECNOLOGÍA

Medias: Intervalos de confianza y determinación del tamaño de muestra

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Intervalo de confianza

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Confidence Intervals** en el menú desplegable y seleccione **Mean One-Sample** en el submenú.
3. *Si utiliza estadísticos de resumen*, seleccione la pestaña **Use Summary Statistics** e ingrese el nivel de confianza deseado, el tamaño de muestra, la media muestral y la desviación estándar muestral.

   Si utiliza datos muestrales, seleccione la pestaña **Use Data** y seleccione la columna de datos deseada.
4. Haga clic en **Evalute**.

SUGERENCIA: Statdisk elegirá automáticamente entre las distribuciones normal y *t*, lo cual dependerá de si se ingresa o no un valor para la desviación estándar poblacional.

Determinación del tamaño de muestra

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Sample Size Determination** en el menú desplegable y seleccione **Estimate Mean** en el submenú.
3. Ingrese el nivel de confianza, el margen de error *E* y la desviación estándar poblacional. Ingrese también el tamaño de la población conocida si está muestreando sin reemplazo de una población finita.
4. Haga clic en **Evaluate**.

Minitab

Intervalo de confianza

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **1-Sample t** en el submenú.
3. *Si utiliza estadísticos de resumen*, seleccione **Summarized data** del menú desplegable e ingrese el tamaño de muestra, la media muestral y la desviación estándar.

   Si utiliza datos muestrales, seleccione **One or more samples, each in a column** del menú desplegable y seleccione la(s) columna(s) de datos deseada(s).
4. Confirme que la opción *Perform hypothesis test* no esté marcada.
5. Haga clic en el botón **Options** e ingrese el nivel de confianza. Para la *Alternative Hypothesis*, seleccione $\neq$.
6. Haga clic en **OK** dos veces.

Determinación del tamaño de muestra

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Power and Sample Size** en el menú desplegable y seleccione **Sample Size for Estimation** en el submenú.
3. Para *Parameter*, seleccione **Mean (Normal)** e ingrese la desviación estándar.
4. Seleccione **Estimate sample sizes** del menú desplegable e ingrese el margen de error deseado para los intervalos de confianza.
5. Haga clic en el botón **Options** para ingresar el nivel de confianza y seleccione un intervalo de confianza del tipo **two-sided**.
6. Haga clic en **OK** dos veces.

StatCrunch

Intervalo de confianza

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **T Stats** en el menú desplegable, luego seleccione **One Sample** en el submenú.
3. *Si utiliza estadísticos de resumen*, seleccione *With Summary* en el submenú e ingrese la media muestral, la desviación estándar muestral y el tamaño de muestra.

   Si utiliza datos muestrales, seleccione *With Data* en el submenú y seleccione la(s) columna(s) de datos deseado(s).
4. Seleccione el **Confidence interval for** μ e ingrese el nivel de confianza deseado.
5. ¡Haga clic en **Compute!**

Determinación del tamaño de muestra

No está disponible.

CENTRO DE TECNOLOGÍA *continuación*

Medias: Intervalos de confianza y determinación del tamaño de muestra

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

Intervalo de confianza

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **TInterval** del menú si σ no se conoce. (Elija **ZInterval** si σ es conocida).
3. *Si utiliza estadísticos de resumen*, seleccione **Stats**, presione ENTER e ingrese los valores para la media muestral $\bar{x}$, la desviación estándar muestral *Sx* y el tamaño de muestra *n*.

   Si utiliza datos muestrales, seleccione **Data**, presione ENTER e ingrese el nombre de la lista que contiene los datos muestrales. La *frecuencia* debe establecerse en **1**.
4. Introduzca el nivel de confianza deseado, *C-Level*.
5. Seleccione **Calculate** y presione ENTER.

Determinación del tamaño de muestra

No está disponible.

Excel

Intervalo de confianza

Complemento XLSTAT (requerido)

Requiere datos muestrales originales, no funciona con datos resumidos.

1. Haga clic en la pestaña **XLSTAT** en la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **One-Sample t-test and z-test** del menú desplegable.
3. En *Data*, ingrese el rango de celdas que contiene los datos muestrales. Para *Data format* seleccione **One Sample**. Si la primera fila de datos contiene una etiqueta, también marque el cuadro **Column labels**.
4. Seleccione **Student's t test**.
5. Haga clic en la pestaña **Options**.
6. En *Alternative hypothesis* seleccione $\neq$ **Theoretical Mean**. Ingrese **0** para la media teórica e ingrese el nivel de significancia deseado (ingrese **5** para el intervalo con 95% de confianza).
7. Haga clic en **OK** para mostrar el resultado bajo "intervalo de confianza para la media".

Determinación del tamaño de muestra

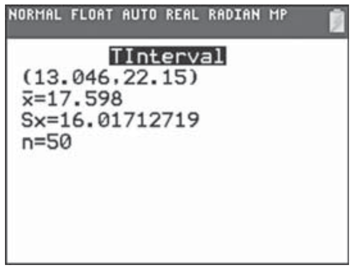
No está disponible.

7-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

En los ejercicios 1 a 3, consulte la pantalla adjunta que resulta de las velocidades de transmisión de datos en el aeropuerto para Verizon (Mbps) a partir del conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B. Se utilizó un nivel de confianza del 95%.

TI-83/84 Plus



1. Velocidades de datos en aeropuertos Consulte la pantalla adjunta.

a. Exprese el intervalo de confianza en el formato que usa el símbolo "menor que". Dado que los datos de la lista original utilizan un lugar decimal, redondee los límites del intervalo de confianza en consecuencia.

b. Identifique la mejor estimación puntual de μ y el margen de error.

c. Al construir la estimación del intervalo de confianza para μ, ¿por qué no es necesario confirmar que los datos muestrales parecen provenir de una población con distribución normal?

2. Grados de libertad

a. ¿Cuál es la cantidad de grados de libertad que se deben usar para encontrar el valor crítico $t_{\alpha/2}$?

b. Encuentre el valor crítico $t_{\alpha/2}$ correspondiente a un nivel de confianza del 95%.

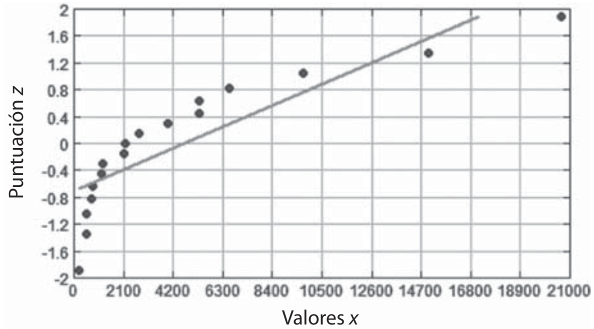
c. Dé una descripción general breve del número de grados de libertad.

3. Interpretación de un intervalo de confianza Los resultados en la pantalla se basan en un nivel de confianza del 95%. Escriba un enunciado que interprete correctamente el intervalo de confianza.

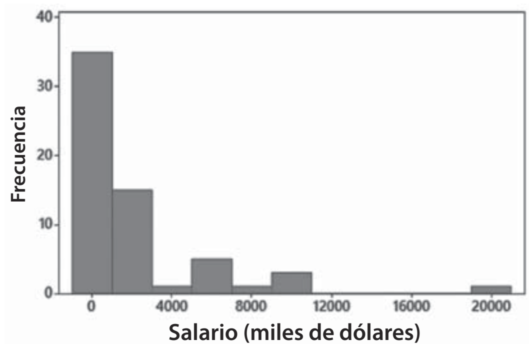
4. Requisito de normalidad ¿Qué significa cuando decimos que los métodos del intervalo de confianza presentados en esta sección son sólidos contra las desviaciones de la normalidad?

Uso de distribución correcta. ***En los ejercicios 5 a 8, suponga que queremos construir un intervalo de confianza. Realice una de las siguientes acciones, según corresponda: (a) Encuentre el valor crítico superior, (b) encuentre el valor crítico $z_{\alpha/2}$, o (c) indique que no se aplica ni la distribución normal ni la distribución t.***

5. Salarios del Miami Heat El nivel de confianza es del 95%, no se conoce σ, y la gráfica cuantilar normal de los 17 salarios (en miles de dólares) de los jugadores de baloncesto del Miami Heat es como se muestra.



6. Salarios de los Broncos de Denver El nivel de confianza es del 90%, σ no se conoce, y el histograma de los salarios de 61 jugadores (miles de dólares) es el siguiente.



7. Salarios de los Broncos de Denver El nivel de confianza es del 99%, $\sigma = 3342$ miles de dólares, y el histograma de 61 salarios de los jugadores (miles de dólares) se muestra en el ejercicio 6.

8. Pesos al nacer A continuación se presentan los datos estadísticos de resumen para pesos seleccionados al azar de niñas recién nacidas: $n = 205$, $\bar{x} = 30.4$ hg, $s = 7.1$ hg (según el conjunto de datos 4 "Nacimientos" en el apéndice B). El nivel de confianza es del 95%.

Intervalos de confianza. ***En los ejercicios 9 a 24, construya la estimación del intervalo de confianza de la media.***

9. Pesos al nacer de niñas Use los siguientes datos estadísticos de resumen que se dan en el ejercicio 8: $n = 205$, $\bar{x} = 30.4$ hg, $s = 7.1$ hg. Use un nivel de confianza del 95%. ¿Los resultados son muy diferentes de los encontrados en el ejemplo 2 con sólo 15 valores muestrales?

10. Pesos al nacer de niños Use los siguientes estadísticos de resumen para el peso al nacer de 195 niños: $\bar{x} = 32.7$ hg, $s = 6.6$ hg (según el conjunto de datos 4 "Nacimientos" en el apéndice B). Use un nivel de confianza del 95%. ¿Los resultados son muy diferentes de los encontrados en el ejercicio 9? ¿Parece que los niños y las niñas tienen pesos muy diferentes al nacer?

11. Temperatura corporal media El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye una muestra de 106 temperaturas corporales con una media de 98.20 °F y una desviación estándar de 0.62 °F. Construya una estimación del intervalo de confianza del 95% para la temperatura corporal media de toda la población. ¿Qué sugiere el resultado sobre la creencia común de que la temperatura media del cuerpo es de 98.6 °F?

12. Programa de pérdida de peso Atkins En una prueba a programas de pérdida de peso, 40 adultos utilizaron el programa de pérdida de peso Atkins. Después de 12 meses, se descubrió que su pérdida de peso promedio era de 2.1 lb, con una desviación estándar de 4.8 lb. Construya una estimación del intervalo de confianza del 90% para la pérdida de peso media de todos estos sujetos. ¿Parece que el programa Atkins es efectivo? ¿Parece ser práctico?

13. Tratamiento del insomnio Se realizó un ensayo clínico para probar la efectividad del fármaco zopiclona en el tratamiento del insomnio en sujetos mayores. Antes del tratamiento con zopiclona, 16 sujetos tuvieron un tiempo de vigilia medio de 102.8 minutos. Después del tratamiento con zopiclona, los 16 sujetos tuvieron un tiempo de vigilia medio de 98.9 minutos y una desviación estándar de 42.3 minutos (con base en los datos de "Cognitive Behavioral Therapy *vs* Zoplicone for Treatment of Chronic Primary Insomnia in Older Adults", de Sivertsen *et al*., en *Journal of the American Medical Association*, vol. 295, núm. 24). Suponga que los 16 valores muestrales parecen ser de una población normalmente distribuida y construya una estimación del intervalo de confianza del 98% para el tiempo de vigilia medio de una población con tratamiento de zopiclona. ¿Qué sugiere el resultado sobre el tiempo de vigilia medio de 102.8 minutos antes del tratamiento? ¿La zopiclona parece ser efectiva?

14. Ajo para reducir el colesterol En una prueba de la efectividad del ajo para bajar el colesterol, 49 sujetos fueron tratados con ajo crudo. Los niveles de colesterol se midieron antes y después del tratamiento. Los cambios (antes *menos* después) en sus niveles de colesterol LDL (en mg/dL) tuvieron una media de 0.4 y una desviación estándar de 21.0 (según datos de "Effect of Raw Garlic *vs* Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults with Moderate Hypercholesterolemia", de Gardner *et al*., en *Archives of Internal Medicine*, vol. 167). Construya una estimación del intervalo de confianza del 98% para el cambio neto medio en el colesterol LDL después del tratamiento con ajo. ¿Qué sugiere el intervalo de confianza sobre la efectividad del ajo en la reducción del colesterol LDL?

15. Genes Se recolectan muestras de ADN y las cuatro bases de ADN de A, G, C y T se codifican como 1, 2, 3 y 4, respectivamente. Los resultados se listan a continuación. Construya una estimación del intervalo de confianza del 95% para la media. ¿Cuál es el uso práctico del intervalo de confianza?

2 2 1 4 3 3 3 3 4 1

16. Arsénico en el arroz A continuación se listan las cantidades de arsénico (μg, o microgramos, por porción) en muestras de arroz integral de California (según datos de la FDA, Administración de Drogas y Alimentos). Use un nivel de confianza del 90%. La Administración de Drogas y Alimentos también midió cantidades de arsénico en muestras de arroz integral de Arkansas. ¿Se puede usar el intervalo de confianza para describir los niveles de arsénico en Arkansas?

5.4 5.6 8.4 7.3 4.5 7.5 1.5 5.5 9.1 8.7

17. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a los sujetos masculinos que calificaran el atractivo de sus citas femeninas, y una muestra de los resultados se detalla a continuación (1 = no atractiva; 10 = extremadamente atractiva). Use un nivel de confianza del 99%. ¿Qué nos dicen los resultados sobre la media de las calificaciones del atractivo hechas por la población de todas las mujeres adultas?

7 8 2 10 6 5 7 8 8 9 5 9

18. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a los sujetos femeninos que calificaran el atractivo de sus citas masculinas, y una muestra de los resultados se detalla a continuación (1 = no atractivo; 10 = extremadamente atractivo). Use un nivel de confianza del 99%. ¿Puede usarse el resultado para estimar el atractivo medio de la población de todos los adultos masculinos?

5 8 3 8 6 10 3 7 9 8 5 5 6 8 8 7 3 5 5 6 8 7 8 8 8 7

19. Mercurio en el sushi Una directriz de la FDA es que el mercurio en el pescado debe ser inferior a 1 parte por millón (ppm). A continuación se listan las cantidades de mercurio (ppm) encontrados en el sushi de atún que se muestreó en diferentes establecimientos de la ciudad de Nueva York. El estudio fue patrocinado por el *New York Times*, y las establecimientos (en orden) son D'Agostino, Eli's Manhattan, Fairway, Food Emporium, Gourmet Garage, Grace's Marketplace y Whole Foods. Construya una estimación del intervalo de confianza del 98% para la cantidad media de mercurio en la población. ¿Se puede inferir que hay demasiado mercurio en el sushi de atún?

0.56 0.75 0.10 0.95 1.25 0.54 0.88

20. Años en la universidad A continuación se listan las cantidades de años que requirió una muestra aleatoria de estudiantes universitarios para obtener títulos de licenciatura (según datos del Centro Nacional de Estadísticas de Educación). Elabore una estimación del intervalo de confianza del 95% para el tiempo medio requerido por todos los estudiantes universitarios para obtener una licenciatura. ¿Se puede inferir que los estudiantes universitarios generalmente obtienen títulos de licenciatura en cuatro años? ¿Hay algo acerca de los datos que sugiera que el intervalo de confianza podría no ser un buen resultado?

4 4 4 4 4 4 4.5 4.5 4.5 4.5 4.5 4.5

6 6 8 9 9 13 13 15

21. Patrimonio de celebridades A continuación se listan los montos del patrimonio neto (en millones de dólares) de las diez celebridades más ricas: Tom Cruise, Will Smith, Robert De Niro, Drew Carey, George Clooney, John Travolta, Samuel L. Jackson, Larry King, Demi Moore y Bruce Willis. Construya un intervalo de confianza del 98%. ¿Qué nos dice el resultado sobre la población de todas las celebridades? ¿Los datos parecen ser de una población normalmente distribuida como es requerido?

250 200 185 165 160 160 150 150 150 150

22. Cafeína en bebidas refrescantes A continuación se listan las cantidades medidas de cafeína (mg por 12 onzas de bebida) obtenidas en una lata de cada una de 20 marcas (7UP, A&W Root Beer, Cherry Coke, . . . , TaB). Use un nivel de confianza del 99%. ¿El intervalo de confianza nos da información buena sobre la población de todas las latas de las mismas 20 marcas que se consumen? ¿La muestra parece provenir de una población normalmente distribuida? De lo contrario, ¿cómo se ven afectados los resultados?

0 0 34 34 34 45 41 51 55 36 47 41 0 0 53 54 38 0 41 47

23. Evaluaciones de estudiantes A continuación se listan las calificaciones que algunos estudiantes dan a sus cursos, donde una calificación de 5 es para "excelente". Las calificaciones se obtuvieron en la Universidad de Texas en Austin (consulte el conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B). Use un nivel de confianza del 90%. ¿Qué nos dice el intervalo de confianza sobre la población de estudiantes universitarios en Texas?

3.8 3.0 4.0 4.8 3.0 4.2 3.5 4.7 4.4 4.2 4.3 3.8 3.3 4.0 3.8

24. Llegadas de vuelo A continuación se listan los retrasos de llegada (en minutos) de vuelos de American Airlines seleccionados, desde Nueva York (JFK) hasta Los Ángeles (LAX). Los números negativos corresponden a vuelos que llegaron antes de la hora de llegada programada. Use un intervalo de confianza del 95%. ¿Qué tan bueno es el desempeño en cuanto al tiempo?

−5 −32 −13 −9 −19 49 −30 −23 14 −21 −32 11

Conjuntos de datos del apéndice B. ***En los ejercicios 25 a 28, use los conjuntos de datos del apéndice B para construir estimaciones del intervalo de confianza para la media.***

25. Pulsos Consulte el conjunto de datos 1 "Datos corporales" y construya una estimación del intervalo de confianza del 95% para el pulso medio de las mujeres adultas: luego haga lo mismo para los hombres adultos. Compare los resultados.

26. Velocidades de datos en aeropuertos Consulte el conjunto de datos 32 "Velocidades de datos en aeropuertos" y construya una estimación del intervalo de confianza del 95% para la velocidad media de Sprint, luego haga lo mismo para T-Mobile. Compare los resultados.

27. Tiempos de servicio de comida rápida Consulte el conjunto de datos 25 "Comida rápida" y construya una estimación del intervalo de confianza del 95% para el tiempo medio de servicio en auto en las cenas de McDonald's; luego haga lo mismo para las cenas en Burger King. Compare los resultados.

28. Galletas con chispas de chocolate Consulte el conjunto de datos 28 "Galletas con chispas de chocolate" y construya una estimación del intervalo de confianza del 95% para el número medio de chispas de chocolate en las galletas Chips Ahoy regulares; luego haga lo mismo para las galletas Keebler. Compare los resultados.

Tamaño de muestra. ***En los ejercicios 29 a 36, encuentre el tamaño de muestra requerido para estimar la media poblacional.***

29. IQ medio de profesores universitarios La prueba Wechsler de IQ está diseñada para que la media sea 100 y la desviación estándar sea 15 para la población de adultos normales. Encuentre el tamaño de muestra necesario para estimar la puntuación promedio de IQ de los profesores universitarios. Queremos tener el 99% de confianza de que nuestra media muestral está dentro de 4 puntos del coeficiente intelectual de la media verdadera. La media para esta población es claramente mayor que 100, la desviación estándar es menor que 15 porque es un grupo con menos variación que un grupo seleccionado al azar de la población general; por lo tanto, si usamos $\sigma = 15$, estamos siendo conservadores al usar un valor que hará que el tamaño de muestra sea tan grande como se necesite. Suponga entonces que $\sigma = 15$ y determine el tamaño de muestra requerido. ¿Se puede inferir que el tamaño de muestra es práctico?

30. IQ medio de abogados Vea el ejercicio anterior, en el que podemos suponer que $\sigma = 15$ para los puntajes de IQ. Los abogados son un grupo con puntajes de IQ que varían menos que las puntuaciones de la población general. Encuentre el tamaño de muestra necesario para estimar el coeficiente intelectual medio de los abogados, dado que queremos un 98% de confianza de que la media muestral está dentro de 3 puntos del IQ medio poblacional. ¿Se puede inferir que el tamaño de muestra es práctico?

31. Media de calificaciones promedio Suponga que todas las calificaciones promedio deben estandarizarse en una escala entre 0 y 4. ¿Cuántas calificaciones promedio se deben obtener para que la media muestral esté dentro de 0.01 de la media poblacional? Suponga que se desea un nivel de confianza del 95%. Si usamos la regla práctica, podemos estimar σ como rango/4 = $(4 - 0)/4 = 1$. ¿Se puede inferir que el tamaño de la muestra es práctico?

32. Peso medio de estudiantes varones de estadística El conjunto de datos 1 “Datos corporales” en el apéndice B incluye pesos de 153 hombres adultos seleccionados al azar, y esos pesos tienen una desviación estándar de 17.65 kg. Debido a que es razonable suponer que los pesos de los estudiantes varones de estadística tienen menos variación que los pesos de la población de varones adultos, sea $\sigma = 17.65$ kg. ¿Cuántos estudiantes varones de estadística deben pesarse para estimar el peso promedio de todos los estudiantes de estadística del sexo masculino? Suponga que deseamos tener un 90% de confianza en que la media muestral está dentro de 1.5 kg desde la media poblacional. ¿Parece razonable suponer que los pesos de los estudiantes varones de estadística tienen menos variación que los pesos de la población de hombres adultos?

33. Edad media de las estudiantes de estadística El conjunto de datos 1 “Datos corporales” en el apéndice B incluye las edades de 147 mujeres adultas seleccionadas al azar, y esas edades tienen una desviación estándar de 17.7 años. Suponga que las edades de las estudiantes de estadística tienen menos variación que las edades de las mujeres en la población general, por lo que $\sigma = 17.7$ años para el cálculo del tamaño de la muestra. ¿De cuántas estudiantes femeninas de estadística se debe obtener la edad para estimar la edad media de todas las estudiantes de estadísticas? Suponga que deseamos tener un 95% de confianza en que la media muestral está dentro del medio año desde la media poblacional. ¿Se puede inferir que las edades de las estudiantes de estadística tienen menos variación que las edades de las mujeres en la población general?

34. Pulso medio de mujeres El conjunto de datos 1 “Datos corporales” en el apéndice B incluye los pulsos de 147 mujeres adultas seleccionadas al azar, y esos pulsos varían desde un mínimo de 36 lpm hasta un máximo de 104 lpm. Encuentre el tamaño de muestra mínimo requerido para estimar el pulso promedio de las mujeres adultas. Suponga que deseamos tener un 99% de confianza en que la media de la muestra está dentro de 2 lpm desde la media poblacional.

a. Encuentre el tamaño de muestra usando la regla práctica del rango para estimar σ.

b. Suponga que $\sigma = 12.5$ lpm, con base en el valor de $s = 12.5$ lpm para la muestra de 147 pulsos femeninos.

c. Compare los resultados de los incisos (a) y (b). ¿Cuál resultado parece ser mejor?

35. Pulso medio de hombres El conjunto de datos 1 “Datos corporales” en el apéndice B incluye los pulsos de 153 hombres adultos seleccionados al azar, y esos pulsos varían desde un mínimo de 40 lpm hasta un máximo de 104 lpm. Encuentre el tamaño de muestra mínimo requerido para estimar el pulso medio de los hombres adultos. Suponga que deseamos tener un 99% de confianza en que la media muestral está dentro de 2 lpm desde la media poblacional.

continúa

a. Encuentre el tamaño de muestra usando la regla práctica del rango para estimar σ.

b. Suponga que $\sigma = 11.3$ lpm, con base en el valor de $x = 11.3$ lpm para la muestra de 153 pulsos masculinos.

c. Compare los resultados de los incisos (a) y (b). ¿Cuál resultado parece ser mejor?

36. Temperatura corporal media El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye 106 temperaturas corporales de adultos para el Día 2 a las 12 am, y varían desde un mínimo de 96.5 °F hasta un máximo de 99.6 °F. Encuentre el tamaño de muestra mínimo requerido para estimar la temperatura corporal media de todos los adultos. Suponga que deseamos tener un 98% de confianza en que la media muestral está dentro de 0.1 °F desde la media poblacional.

a. Encuentre el tamaño de muestra usando la regla práctica del rango para estimar σ.

b. Suponga que $\sigma = 0.62$ °F, con base en el valor de $s = 0.62$ °F para la muestra de 106 temperaturas corporales.

c. Compare los resultados de los incisos (a) y (b). ¿Qué resultado parece ser mejor?

7.2 Más allá de lo básico

Intervalo de confianza con σ conocida. ***En los ejercicios 37 y 38, encuentre el intervalo de confianza usando el valor conocido de σ.***

37. Pesos al nacer de niñas Construya el intervalo de confianza para el ejercicio 9 "Pesos al nacer de niñas", suponiendo que se sabe que σ es 7.1 hg.

38. Pesos al nacer de niños Construya el intervalo de confianza para el ejercicio 10 "Pesos al nacer de niños", suponiendo que se sabe que σ es 6.6 hg.

39. Factor de corrección de población finita Si se selecciona una muestra aleatoria simple de tamaño n y sin reemplazo de una población finita de tamaño N, y el tamaño de la muestra supera el 5% del tamaño de la población (n / 0.05N), pueden obtenerse mejores resultados utilizando el factor de corrección de población finita, lo que implica multiplicar el margen de error E por $\sqrt{(N - n)/(N - 1)}$. Para la muestra de un peso de 100 para los caramelos M&M en el conjunto de datos 27 "Peso de M&Ms" del apéndice B, obtenemos $\overline{x} = 0.8565$ g y $s = 0.0518$ g. Primero construya una estimación del intervalo de confianza del 95% para μ, suponiendo que la población es grande; luego construya una estimación del intervalo de confianza del 95% para el peso medio de los M&Ms en la bolsa completa de la que se tomó la muestra. La bolsa completa tiene 465 M&Ms. Compare los resultados.

7-3 Estimación de una desviación estándar o varianza poblacionales

Concepto clave Esta sección presenta métodos para usar una desviación estándar muestral s (o una varianza muestral s^2) para estimar el valor de la desviación estándar poblacional σ (o la varianza poblacional σ^2) correspondiente. A continuación se enuncian los conceptos principales incluidos en esta sección:

- **Estimación puntual:** La varianza muestral s^2 es la mejor *estimación puntual* (o estimación de valor único) de la varianza poblacional σ^2. La desviación estándar muestral s se usa comúnmente como una estimación puntual de σ, aunque es un estimador sesgado, como se describió en la sección 6-3.
- **Intervalo de confianza:** Al construir una estimación del *intervalo de confianza* de una desviación estándar poblacional (o varianza poblacional), construimos el intervalo de confianza utilizando la *distribución* χ^2. (La letra griega χ se pronuncia "ji").

Distribución ji cuadrada

Los puntos clave sobre la distribución χ^2 (ji cuadrada o chi cuadrada) son:

- En una población normalmente distribuida con varianza σ^2, si seleccionamos aleatoriamente muestras independientes de tamaño n y, para cada muestra, calculamos la varianza muestral s^2, el estadístico muestral $\chi^2 = (n-1)s^2/\sigma^2$ tiene una distribución muestral llamada distribución ji cuadrada, como se muestra en la fórmula 7-5.

FÓRMULA 7.5

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

- **Valores críticos de χ^2** Expresamos un valor crítico de la cola derecha por χ^2_R y un valor crítico de la cola izquierda por χ^2_L. Esos valores críticos se pueden encontrar mediante el uso de tecnología o de la tabla A-4, y requieren que primero determinemos un valor para la cantidad de *grados de libertad*.
- **Grados de libertad** Para los métodos de esta sección, el número de grados de libertad es el tamaño de la muestra menos 1.

Grados de libertad: gl = $n - 1$

- La distribución ji cuadrada es asimétrica a la derecha, a diferencia de las distribuciones normal y t de Student (vea la figura 7-7).
- Los valores de la ji cuadrada pueden ser cero o positivos, pero no pueden ser negativos, como se muestra en la figura 7-7.
- La distribución ji cuadrada es diferente para cada cantidad de grados de libertad, como se ilustra en la figura 7-8. A medida que aumenta la cantidad de grados de libertad, la distribución ji cuadrada se acerca a una distribución normal.

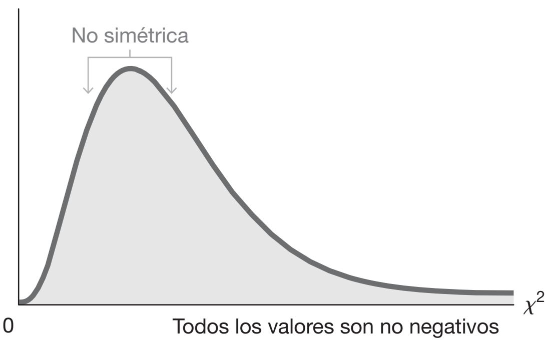


FIGURA 7-7 Distribución ji cuadrada

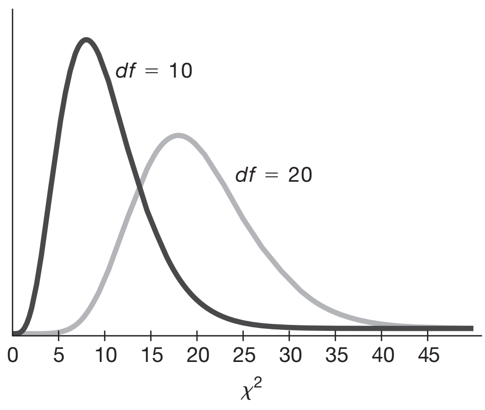


FIGURA 7-8 Distribución ji cuadrada para gl = 10 y gl = 20

Debido a que la distribución ji cuadrada no es simétrica, una estimación del intervalo de confianza para σ^2 no se ajusta a un formato del tipo $s^2 - E < \sigma^2 < s^2 + E$, por lo que debemos hacer cálculos separados para los límites superior e inferior del intervalo de confianza. Si usted usa la tabla A-4 para encontrar los valores críticos, tenga en cuenta la siguiente característica de diseño de dicha tabla:

En la tabla A-4, cada valor crítico de χ^2 en el cuerpo de la tabla corresponde a un área dada en la fila superior de la tabla, y cada área en esa fila superior es un *área acumulada a la derecha* del valor crítico.

PRECAUCIÓN La tabla A-2 para la distribución normal estándar proporciona áreas acumuladas desde la *izquierda*, pero la tabla A-4 para la distribución ji cuadrada usa áreas acumuladas desde la *derecha*.

EJEMPLO 1 Determinación de los valores críticos de χ^2

Se obtiene una muestra aleatoria simple de 22 puntuaciones de IQ (como en el ejemplo 2, que sigue). La construcción de un intervalo de confianza para la desviación estándar poblacional σ requiere los valores críticos a la izquierda y a la derecha de χ^2 correspondientes a un nivel de confianza del 95% y un tamaño de muestra de $n = 22$. Encuentre χ^2_L (el valor crítico de χ^2 que separa un área de 0.025 en la cola izquierda), y encuentre χ^2_R (el valor crítico de χ^2 que separa un área de 0.025 en la cola derecha).

SOLUCIÓN

Con un tamaño de muestra de $n = 22$, el número de grados de libertad es gl $= n - 1 = 21$. Vea la figura 7-9.

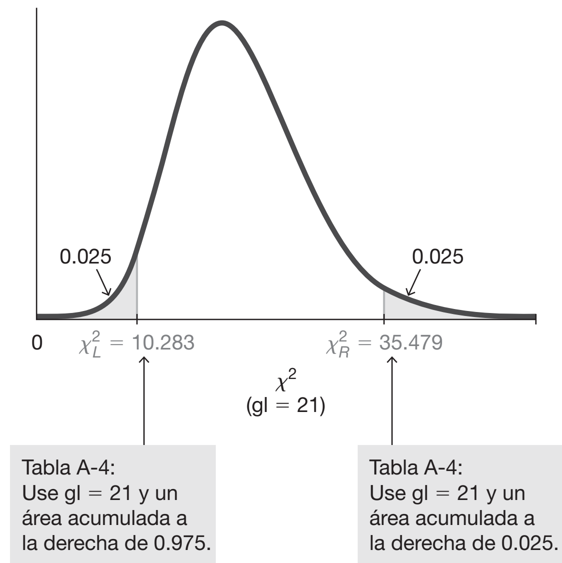


FIGURA 7-9 Determinación de valores críticos de χ^2

El valor crítico a la derecha ($\chi^2_R = 35.479$) se obtiene de la tabla A-4 de manera directa al ubicar 21 en la columna de grados de libertad a la izquierda y 0.025 a lo largo de la fila superior. El valor crítico más a la izquierda de $\chi^2_L = 10.283$ también corresponde a 21 en la columna de grados de libertad, pero debemos ubicar 0.975 (o $1 - 0.025$) en la fila superior porque los valores en la fila superior siempre son *áreas a la derecha* del valor crítico. Consulte la figura 7-9 y vea que el área total a la derecha de $\chi^2_L = 10.283$ es 0.975.

SU TURNO Encuentre los valores críticos en el ejercicio 5 "Nicotina en cigarrillos mentolados".

Al obtener valores críticos de χ^2 con base en la tabla A-4, si no se encuentra un número de grados de libertad específico, usted puede ser conservador usando el siguiente número menor de grados de libertad, también puede usar el valor crítico más cercano en la tabla, o bien puede obtener un resultado aproximado mediante interpolación. Para cantidades de grados de libertad superiores a 100, use la ecuación dada en el ejercicio 23 "Determinación de valores críticos" en la página 342, también puede usar una tabla más extensa, o bien use tecnología.

Aunque s^2 es la mejor estimación puntual de σ^2, no hay una indicación de cuán buena es, por lo que usamos un intervalo de confianza que nos proporciona un rango de valores asociados con un nivel de confianza.

ELEMENTOS CLAVE

Intervalo de confianza para estimar una desviación estándar o varianza poblacionales

Objetivo

Construir una estimación del intervalo de confianza para una desviación estándar o varianza poblacionales.

Notación

σ = desviación estándar poblacional

s = desviación estándar muestral

n = número de valores muestrales

χ^2_L = valor crítico de la cola izquierda de χ^2

σ^2 = varianza poblacional

s^2 = varianza muestral

E = margen de error

χ^2_R = valor crítico de la cola derecha de χ_2

Requisitos

1. La muestra debe ser una muestra aleatoria simple.
2. La población debe tener valores normalmente distribuidos (incluso si la muestra es grande). El requisito de una distribución normal es mucho más estricto aquí que en las secciones anteriores, por lo que las grandes desviaciones de las distribuciones normales pueden generar grandes errores. (Si el requisito de normalidad no se cumple, use el método de bootstrap descrito en la sección 7-4).

Intervalo de confianza para la varianza poblacional σ^2

$$\frac{(n-1)\,s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

Intervalo de confianza para la desviación estándar poblacional σ

$$\sqrt{\frac{(n-1)\,s^2}{\chi^2_R}} < \sigma < \sqrt{\frac{(n-1)\,s^2}{\chi^2_L}}$$

Regla de redondeo

1. *Datos originales:* Cuando utilice el *conjunto original de valores de datos*, redondee los límites del intervalo de confianza a un lugar decimal más de lo usado en los datos originales.
2. *Estadísticos de resumen:* Cuando utilice los *estadísticos de resumen* (n, s), redondee los límites del intervalo de confianza al mismo número de decimales utilizados para la desviación estándar muestral.

PRECAUCIÓN Un intervalo de confianza se puede expresar en un formato del tipo $11.0 < \sigma < 20.4$ o como (11.0, 20.4), *pero no en un formato del tipo* $s \pm E$.

Procedimiento para construir un intervalo de confianza para σ o σ^2

Los intervalos de confianza pueden construirse fácilmente con tecnología o bien utilizando la tabla A-4 con el siguiente procedimiento.

1. Verifique que se satisfagan los dos requisitos: La muestra es una muestra aleatoria de una población normalmente distribuida.

continúa

2. Mediante el uso de $n - 1$ grados de libertad, encuentre los valores críticos χ^2_R y χ^2_L que corresponden al nivel de confianza deseado (como en el ejemplo 1).
3. Para obtener una estimación del intervalo de confianza de σ^2, use la siguiente expresión:

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

4. Para obtener una estimación del intervalo de confianza de σ, calcule la raíz cuadrada de cada componente del intervalo de confianza anterior.
5. Redondee los límites del intervalo de confianza usando la regla de redondeo dada en el recuadro anterior de Elementos clave.

Uso de intervalos de confianza para realizar comparaciones o pruebas de hipótesis

Comparaciones Los intervalos de confianza pueden usarse de manera *informal* para comparar la variación en diferentes conjuntos de datos, pero la *superposición de los intervalos de confianza no debe usarse para formular conclusiones formales y finales sobre la igualdad de varianzas o desviaciones estándar*.

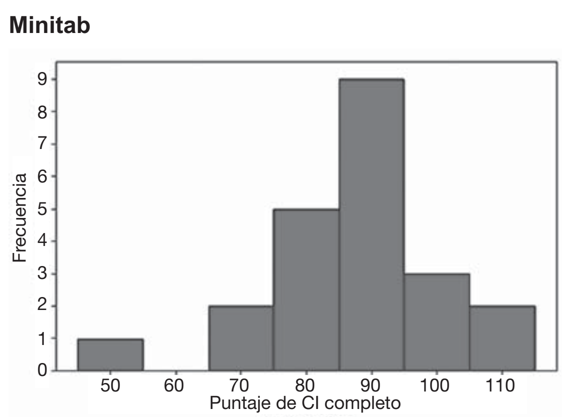
EJEMPLO 2 Intervalo de confianza para estimar σ de puntuaciones de IQ

El conjunto de datos 7 "IQ y plomo" en el apéndice B lista las puntuaciones de IQ para sujetos en tres diferentes grupos de exposición al plomo. Las 22 puntuaciones totales de IQ para el grupo con exposición media al plomo (grupo 2) tienen una desviación estándar de 14.29263. Considere que la muestra es una muestra aleatoria simple y construya una estimación del intervalo de confianza del 95% para σ, la desviación estándar de la población a partir de la cual se obtuvo la muestra.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS

Paso 1: Verifique los requisitos. (1) La muestra puede tratarse como una muestra aleatoria simple. (2) El histograma adjunto tiene una forma muy cercana a la forma de campana de una distribución normal, por lo que se cumple el requisito de normalidad.



Paso 2: Uso de tecnología El intervalo de confianza se puede encontrar usando algún software de estadística. La pantalla de StatCrunch muestra los límites inferior y superior del intervalo

de confianza para la estimación del intervalo de confianza del 95% para σ^2, por lo que obtenemos $120.9 < \sigma^2 < 417.2$. Al obtener las raíces cuadradas, resulta $11.0 < \sigma < 20.4$

StatCrunch

95% confidence interval results:
σ^2 : Variance of variable

Variable	Sample Var.	DF	L. Limit	U. Limit
IQ	204.27922	21	120.91318	417.18431

Uso de la tabla A-4: Si usamos la tabla A-4, primero consideramos el tamaño de muestra $n = 22$ para encontrar los grados de libertad: gl $= n - 1 = 21$. En la tabla A-4, consulte la fila correspondiente a 21 grados de libertad, y las columnas con áreas de 0.975 y 0.025. (Para un nivel de confianza de 95%, dividimos $\alpha = 0.05$ equitativamente entre las dos colas de la distribución ji cuadrada (χ^2), y consultamos los valores de 0.975 y 0.025 en la fila superior de la tabla A-4). Los valores críticos son $\chi^2_L = 10.283$ y $\chi^2_R = 35.479$ (como se muestra en el ejemplo 1).

Paso 3: Mediante el uso de los valores críticos de 10.283 y 35.479, la desviación estándar muestral $s = 14.29263$ y el tamaño de la muestra de $n = 22$, construimos el intervalo de confianza del 95% evaluando lo siguiente:

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

$$\frac{(22-1)(14.29263)^2}{35.479} < \sigma^2 < \frac{(22-1)(14.29263)^2}{10.283}$$

Paso 4: La evaluación de la expresión anterior resulta en $120.9 < \sigma^2 < 417.2$. Al determinar la raíz cuadrada de cada parte (antes del redondeo), para después redondear a un decimal, se obtiene la siguiente estimación del intervalo de confianza del 95% para la desviación estándar poblacional: $11.0 < \sigma < 20.4$.

INTERPRETACIÓN

Con base en este resultado, tenemos un 95% de confianza de que los límites de 11.0 y 20.4 contienen el verdadero valor de σ. El intervalo de confianza también se puede expresar como (11.0, 20.4), *pero no se puede expresar en un formato del tipo* $s \pm E$.

SU TURNO Encuentre el intervalo de confianza en el ejercicio 5 "Nicotina en cigarrillos mentolados".

En cifras

1 minuto y 41 segundos: Tiempo medio que dura una llamada de servicio al cliente personalizado, al llamar al número 800 de uno de los 100 principales proveedores de Internet.

Justificación del intervalo de confianza Vea la figura 7-9 en la página 334 para dar sentido a esta afirmación: si seleccionamos muestras aleatorias de tamaño n de una población distribuida normalmente con varianza σ^2, existe una probabilidad de $1 - \alpha$ de que el estadístico $(n-1)s^2/\sigma^2$ se encuentre entre los valores críticos de χ^2_L y χ^2_R. Se deduce que hay una probabilidad de $1 - \alpha$ de que las dos expresiones siguientes sean verdaderas:

$$\frac{(n-1)s^2}{\sigma^2} < \chi^2_R \quad \text{y} \quad \frac{(n-1)s^2}{\sigma^2} > \chi^2_L$$

Multiplique las dos desigualdades anteriores por σ^2, luego divida cada desigualdad entre el valor crítico adecuado de χ^2, de modo que ambas desigualdades puedan expresarse en estas formas equivalentes:

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 \quad \text{y} \quad \frac{(n-1)s^2}{\chi^2_L} > \sigma^2$$

continúa

TABLA 7-2 Determinación del tamaño de muestra

σ	
Para estar 95% seguros de que *s* está dentro de . . .	**del valor de σ, el tamaño de muestra *n* debe ser al menos de . . .**
1%	19,205
5%	768
10%	192
20%	48
30%	21
40%	12
50%	8
Para estar 99% seguros de que *s* está dentro de . . .	**del valor de σ, el tamaño de muestra *n* debe ser al menos de . . .**
1%	33,218
5%	1,336
10%	336
20%	85
30%	38
40%	22
50%	14

Las dos desigualdades anteriores se pueden combinar en una desigualdad para obtener el formato del intervalo de confianza utilizado en esta sección:

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

Determinación del tamaño de muestra

Los procedimientos para encontrar el tamaño de muestra necesario para estimar σ son mucho más complejos que los procedimientos proporcionados anteriormente para las medias y proporciones. Con poblaciones normalmente distribuidas, se puede usar la tabla 7-2 o la fórmula dada en el ejercicio 24 "Determinación del tamaño de muestra" en la página 342.

Statdisk también proporciona tamaños de muestra. Con Statdisk, seleccione **Analysis, Sample Size Determination** y luego **Estimate Standard Deviation**. Minitab, Excel, StatCrunch y la calculadora TI-83/84 Plus no proporcionan tamaños de muestra.

EJEMPLO 3 Determinación del tamaño de muestra para estimar σ

Queremos estimar la desviación estándar σ de todas las puntuaciones de IQ de las personas con exposición al plomo. Deseamos tener el 99% de confianza de que nuestra estimación está dentro de 5% del valor real de σ. ¿Qué tan grande debería ser la muestra? Suponga que la población se distribuye normalmente.

SOLUCIÓN

Podemos ver en la tabla 7-2 que el 99% de confianza y un error del 5% para σ corresponden a una muestra de tamaño 1336. Deberíamos obtener una muestra aleatoria simple de 1336 IQ de la población de sujetos expuestos al plomo.

SU TURNO Resuelva el ejercicio 19 "IQ de profesores de estadística".

CENTRO DE TECNOLOGÍA

Estimación del intervalo de confianza para la desviación estándar o la varianza

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Intervalo de confianza

1. Verifique que la distribución sea normal usando **Data—Normality Assessment**.
2. Haga clic en **Analysis** en el menú superior.
3. Seleccione **Confidence Intervals** en el menú desplegable y seleccione **Standard Deviation One Sample** en el submenú.
4. Introduzca el nivel de confianza deseado.
5. *Si utiliza estadísticos de resumen:* seleccione la pestaña **Use Summary Statistics** e ingrese el tamaño de muestra y la desviación estándar muestral.

   Si utiliza datos muestrales: Seleccione la pestaña **Use Data** y seleccione la columna de datos deseada.
6. Haga clic en **Evaluate**.

Minitab

Intervalo de confianza

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **1 Variance** en el submenú.
3. *Si utiliza estadísticos de resumen:* seleccione **Sample standard deviation** o **Sample Variance** del menú desplegable e ingrese el tamaño de muestra y la desviación estándar o la varianza muestrales.

   Si utiliza datos muestrales: seleccione **One or more samples, each in a column** del menú desplegable y seleccione la(s) columna(s) de datos deseada(s).
4. Confirme que *Perform hypothesis test* no esté marcada.
5. Haga clic en el botón **Options**, ingrese el nivel de confianza deseado y, para *Alternative Hypothesis*, seleccione $\neq$.
6. Haga clic en **OK** dos veces. Los resultados incluyen un intervalo de confianza para la desviación estándar y otro para la varianza.

Excel

Ni Excel ni XLSTAT tienen una función para generar una estimación del intervalo de confianza para la desviación estándar o la varianza.

CENTRO DE TECNOLOGÍA *continuación*

Estimación del intervalo de confianza para la desviación estándar o la varianza

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

StatCrunch

Intervalo de confianza

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Variance Stats** en el menú desplegable, luego seleccione **One Sample** en el submenú.
3. *Si utiliza estadísticos de resumen:* seleccione *With Summary* en el submenú e ingrese la varianza muestral y el tamaño de muestra.

   Si utiliza datos muestrales: seleccione *With Data* en el submenú y seleccione la(s) columna(s) de datos deseada(s).
4. Seleccione **Confidence Interval for σ^2** e ingrese el nivel de confianza deseado.
5. ¡Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Descargue e instale los programas Michael Lloyd *S2INT* y *ZZINEWT* (disponible en www.pearsonenespañol.com/triola) en su calculadora TI-83/84 Plus.
2. Presione PRGM, seleccione **S2INT** en el menú y presione ENTER dos veces.
3. Ingrese la varianza muestral Sx^2, el tamaño de muestra n y el nivel de confianza *C-Level*. Presione ENTER después de cada entrada.
4. Espere a que se muestre el intervalo de confianza σx^2. (Sea paciente, ¡puede llevar cierto tiempo!).
5. Presione ENTER para ver el intervalo de confianza de σx.

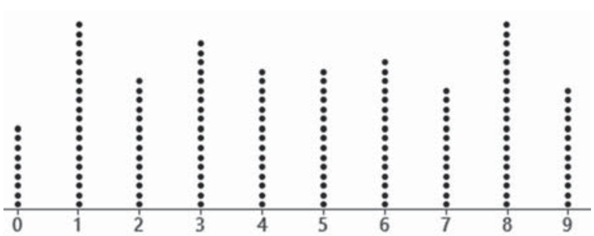
7-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Volumen del cerebro Mediante el uso de todos los volúmenes cerebrales listados en el conjunto de datos 8 "IQ y tamaño del cerebro", obtenemos la siguiente estimación del intervalo de confianza del 95%: $9027.8 < \sigma^2 < 33{,}299.8$, y las unidades de medida son $(\text{cm}^3)^2$. Identifique la estimación del intervalo de confianza correspondiente para σ e incluya las unidades apropiadas. Dado que los valores originales son números enteros, redondee los límites usando la regla de redondeo dada en esta sección. Escriba un enunciado que interprete correctamente la estimación del intervalo de confianza para σ.

2. Expresión de intervalos de confianza El ejemplo 2 mostró cómo los datos estadísticos de $n = 22$ y $s = 14.3$ dan como resultado una estimación del intervalo de confianza del 95% para σ: $11.0 < \sigma < 20.4$. Ese intervalo de confianza también se puede expresar como (11.0, 20.4), pero no se puede expresar como 15.7 ± 4.7. Dado que 15.7 ± 4.7 resulta en valores de 11.0 y 20.4, ¿por qué es incorrecto expresar el intervalo de confianza como 15.7 ± 4.7?

3. Análisis del último dígito La gráfica de puntos que se presenta a continuación representa los últimos dígitos de los pesos de 153 hombres en el conjunto de datos 1 "Datos corporales". ¿Será posible que esos dígitos procedan de una población normalmente distribuida? De no ser así, ¿el tamaño de muestra grande de $n = 153$ justifica tratar los valores como si fueran de una distribución normal? ¿Se puede usar la muestra para construir una estimación del intervalo de confianza del 95% para la σ poblacional de todos esos dígitos?



4. Requisito de normalidad ¿Qué diferencia hay entre el requisito de normalidad para una estimación del intervalo de confianza de σ y el requisito de normalidad para una estimación del intervalo de confianza de μ?

Determinación de valores críticos e intervalos de confianza. ***En los ejercicios 5 a 8, use la información dada para encontrar la cantidad de grados de libertad, los valores críticos χ^2_L y χ^2_R, y la estimación del intervalo de confianza para σ. Las muestras son del apéndice B y es razonable suponer que se ha seleccionado una muestra aleatoria simple de una población con una distribución normal.***

5. Nicotina en cigarrillos mentolados 95% de confianza; $n = 25$, $s = 0.24$ mg.

6. Peso de los centavos 95% de confianza; $n = 37$, $s = 0.01648$ g.

7. Conteo de plaquetas de mujeres 99% de confianza; $n = 147$, $s = 65.4$.

8. Estaturas de hombres 99% de confianza; $n = 153$, $s = 7.10$ cm.

Determinación de intervalos de confianza. ***En los ejercicios 9 a 16, suponga que cada muestra es una muestra aleatoria simple obtenida de una población con una distribución normal.***

9. Temperatura corporal El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye una muestra de 106 temperaturas corporales con una media de 98.20 °F y una desviación estándar de 0.62 °F (para el día 2 a las 12 AM). Construya una estimación del intervalo de confianza del 95% para la desviación estándar de las temperaturas corporales de toda la población.

10. Programa de pérdida de peso Atkins En una prueba a programas de pérdida de peso, 40 adultos utilizaron el programa de pérdida de peso Atkins. Después de 12 meses, su pérdida de peso promedio fue de 2.1 lb, con una desviación estándar de 4.8 lb. Construya una estimación del intervalo de confianza del 90% para la desviación estándar de la pérdida de peso de todos los sujetos. ¿El intervalo de confianza proporciona información sobre la efectividad de la dieta?

11. Tratamiento del insomnio Se realizó un ensayo clínico para probar la efectividad del fármaco zopiclona para tratar el insomnio en sujetos mayores. Después del tratamiento con zopiclona, 16 sujetos tuvieron un tiempo de vigilia medio de 98.9 minutos y una desviación estándar de 42.3 minutos (de acuerdo con los datos de "Cognitive Behavioral Therapy *vs* Zopiclone for Treatment of Chronic Primary Insomnia in Older Adults", de Sivertsen *et al*, en *Journal of the American Medical Association*, vol. 295, núm. 24). Suponga que los 16 valores muestrales parecen ser de una población distribuida normalmente y construya una estimación del intervalo de confianza del 98% para la desviación estándar de los tiempos de vigilia para una población con tratamientos de zopiclona. ¿El resultado indica si el tratamiento es efectivo?

12. Ajo para reducir el colesterol En una prueba de la efectividad del ajo para reducir el colesterol, 49 sujetos fueron tratados con ajo crudo. Los niveles de colesterol se midieron antes y después del tratamiento. Los cambios (antes menos después) en sus niveles de colesterol LDL (en mg/dL) tuvieron una media de 0.4 y una desviación estándar de 21.0 (según datos de "Effect of Raw Garlic *vs* Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults with Moderate Hypercholesterolemia", de Gardner *et al*., en *Archives of Internal Medicine*, vol. 167). Construya una estimación del intervalo de confianza del 98% para la desviación estándar de los cambios en el colesterol LDL después del tratamiento con ajo. ¿El resultado indica si el tratamiento es efectivo?

13. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a los sujetos masculinos que calificaran el atractivo de sus citas femeninas, y una muestra de los resultados se detalla a continuación (1 = no atractiva; 10 = extremadamente atractiva). Construya una estimación del intervalo de confianza del 95% para la desviación estándar de la población a partir de la cual se obtuvo la muestra.

7 8 2 10 6 5 7 8 8 9 5 9

14. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a las mujeres que calificaran el atractivo de sus citas masculinas, y una muestra de los resultados se detalla a continuación (1 = no atractivo; 10 = extremadamente atractivo). Construya una estimación del intervalo de confianza del 95% para la desviación estándar de la población a partir de la cual se obtuvo la muestra.

5 8 3 8 6 10 3 7 9 8 5 5 6 8 8 7 3 5 5 6 8 7 8 8 8 7

15. Velocidades en carretera A continuación se listan las velocidades (mi/h) medidas en el tráfico hacia el sur de la carretera I-280 cerca de Cupertino, California (según datos de SigAlert). Esta muestra simple aleatoria se obtuvo a las 3:30 PM en un día laborable. Use los datos muestrales para construir una estimación del intervalo de confianza del 95% para la desviación estándar poblacional. ¿El intervalo de confianza describe la desviación estándar para todas las horas de la semana?

62 61 61 57 61 54 59 58 59 69 60 67

16. Comparación de líneas de espera

a. Los valores que se listan a continuación son los tiempos de espera (en minutos) de los clientes en el banco Jefferson Valley, donde los clientes ingresan a una sola fila de espera que alimenta tres ventanillas. Construya un intervalo de confianza del 95% para la desviación estándar poblacional σ.

6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7

b. Los valores que se listan a continuación son los tiempos de espera (en minutos) de los clientes en el banco Providence, donde los clientes pueden ingresar a una de tres diferentes filas que se han formado en tres ventanillas. Construya un intervalo de confianza del 95% para la desviación estándar poblacional σ.

4.2 5.1 5.8 6.2 6.7 7.7 7.7 8.5 9.3 10.0

c. Interprete los resultados encontrados en los incisos (a) y (b). ¿Los intervalos de confianza sugieren una diferencia en la variación entre los tiempos de espera? ¿Qué arreglo parece mejor: el sistema de una sola fila o el sistema de varias filas?

Grandes conjuntos de datos del apéndice B. ***En los ejercicios 17 y 18, use el conjunto de datos del apéndice B. Suponga que cada muestra es una muestra aleatoria simple obtenida de una población con una distribución normal.***

17. Evaluaciones de estudiantes Consulte el conjunto de datos 17 “Evaluaciones de cursos” en el apéndice B.

a. Use las 93 evaluaciones de curso para construir una estimación del intervalo de confianza del 95% para la desviación estándar de la población de la que se obtuvo la muestra.

b. Repita el inciso (a) usando las 93 evaluaciones a los profesores.

c. Compare los resultados del inciso (a) y del inciso (b).

18. Pesos al nacer Consulte el conjunto de datos 4 “Nacimientos” en el apéndice B.

a. Use los 205 pesos al nacer de niñas para construir una estimación del intervalo de confianza del 95% para la desviación estándar de la población de la que se obtuvo la muestra.

b. Repita el inciso (a) usando los 195 pesos al nacer de niños.

c. Compare los resultados del inciso (a) y del inciso (b).

Determinación del tamaño de muestra. ***En los ejercicios 19 a 22, suponga que cada muestra es una muestra aleatoria simple obtenida de una población normalmente distribuida. Utilice la tabla 7-2 de la página 338 para encontrar el tamaño de muestra indicado.***

19. IQ de profesores de estadística Usted desea estimar σ para la población de puntuaciones de IQ de los profesores de estadística. Encuentre el tamaño mínimo de muestra necesario para estar 95% seguro de que la desviación estándar muestral s está dentro de 1% de σ. ¿Es práctico este tamaño de la muestra?

20. *Space Mountain* Desea estimar σ para la población de tiempos de espera para entrar a la atracción Space Mountain en Walt Disney World. Desea estar 99% seguro de que la desviación estándar muestral está dentro de 1% de σ. Encuentre el tamaño de muestra mínimo. ¿Es práctico este tamaño de la muestra?

21. Ingresos de estudiantes de estadística Usted desea estimar la desviación estándar de los ingresos anuales de todos los estudiantes actuales de estadística. Encuentre el tamaño mínimo de muestra necesario para estar 95% seguro de que la desviación estándar de la muestra está dentro de 20% de la desviación estándar poblacional. ¿Es probable que esos ingresos satisfagan el requisito de tener una distribución normal?

22. Monedas de ¢25 Al establecer las especificaciones de las monedas de ¢25 que se aceptarán en una máquina expendedora, usted debe estimar la desviación estándar de la población de monedas de esa denominación en uso. Encuentre el tamaño de muestra mínimo necesario para estar seguro de que la desviación estándar muestral está dentro de 10% de la desviación estándar poblacional.

7-3 Más allá de lo básico

23. Determinación de valores críticos Al construir intervalos de confianza para σ o σ^2, es posible utilizar la tabla A-4 para encontrar los valores críticos χ_L^2 y χ_R^2 sólo para valores seleccionados de n hasta 101, por lo que el número de grados de libertad es 100 o menos. Para números más grandes de grados de libertad, podemos aproximar χ_L^2 y χ_R^2 usando

$$\chi^2 = \frac{1}{2}\left[\pm z_{\alpha/2} + \sqrt{2k - 1}\right]^2$$

donde k es el número de grados de libertad y $z_{\alpha/2}$ es la puntuación z crítica que se describió en la sección 7-1. Use esta aproximación para encontrar los valores críticos χ_L^2 y χ_R^2 en el ejercicio 8 "Estaturas de hombres", donde el tamaño de muestra es 153 y el nivel de confianza es del 99%. ¿Cómo se comparan los resultados con los valores críticos reales de $\chi_L^2 = 110.846$ y $\chi_R^2 = 200.657$?

24. Determinación del tamaño de muestra En lugar de utilizar la tabla 7-2 con el fin de determinar el tamaño de muestra requerido para estimar una desviación estándar poblacional σ, es posible emplear la siguiente fórmula

$$n = \frac{1}{2}\left(\frac{z_{\alpha/2}}{d}\right)^2$$

donde $z_{\alpha/2}$ corresponde al nivel de confianza y d es la forma decimal del error porcentual. Por ejemplo, para estar 95% seguro de que s está dentro de 15% del valor de σ, use $z_{\alpha/2} = 1.96$ y $d = 0.15$ con el fin de obtener un tamaño de muestra $n = 86$. Encuentre el tamaño de muestra requerido para estimar σ, suponiendo que queremos tener un 98% de confianza de que s está dentro de 15% de σ.

7-4 Bootstrap: Uso de la tecnología para realizar estimaciones

Concepto clave Las secciones anteriores presentaron métodos para estimar proporciones, medias y desviaciones estándar (o varianzas) poblacionales. Todos estos métodos tienen ciertos requisitos que limitan las situaciones en las que se pueden usar. Cuando algunos de los requisitos no se cumplen, a menudo podemos usar el método bootstrap para estimar un parámetro con un intervalo de confianza. El método bootstrap normalmente requiere el uso de software.

Requisito de muestreo Los métodos anteriores de este capítulo tienen todos un requisito de que la muestra debe ser una muestra aleatoria simple. Si la muestra no se recolecta de forma adecuada, existe una buena probabilidad de que no se pueda hacer *nada* para obtener una estimación útil del intervalo de confianza para un parámetro. *Los métodos bootstrap no se corrigen para los malos métodos de muestreo*.

Requisitos A continuación se listan los importantes requisitos de las secciones anteriores de este capítulo:

- **IC para proporción (sección 7-1):** Hay al menos 5 éxitos y al menos 5 fracasos, o $np \geq 5$ y $nq \geq 5$.
- **IC para media (sección 7-2):** La población se distribuye normalmente o $n > 30$.
- **IC para σ o σ^2 (sección 7-3):** La población debe tener valores normalmente distribuidos, incluso si la muestra es grande.

Cuando los requisitos anteriores no se cumplen, no deberíamos utilizar los métodos presentados en las secciones previas de este capítulo, pero podemos emplear en su lugar el método bootstrap. Este método no requiere muestras grandes, ni que la muestra se recolecte a partir de una distribución normal o de cualquier otra distribución particular, por lo que se denomina **método no paramétrico** o **libre de distribución**; en el capítulo 13, se estudian otros métodos no paramétricos.

DEFINICIÓN

Dada una muestra aleatoria simple de tamaño *n*, una **muestra bootstrap** es otra muestra aleatoria de *n* valores obtenidos *con reemplazo* de la muestra original.

Sin reemplazo, cada muestra sería idéntica a la muestra original, por lo que las proporciones, medias, desviaciones estándar o varianzas serían iguales, y no habría "intervalo" de confianza.

PRECAUCIÓN Tenga en cuenta que una muestra bootstrap implica muestrear con *reemplazo*, de modo que cuando se selecciona un valor muestral, se repite antes de realizar la siguiente selección.

EJEMPLO 1 Muestra bootstrap de ingresos

Cuando el autor recolectó los ingresos anuales de estudiantes de estadística actuales, obtuvo los siguientes resultados (en miles de dólares): 0, 2, 3, 7.

Muestra original	Muestra bootstrap
0	7
2	2
3	2
7	3

La muestra de {7, 2, 2, 3} es una muestra bootstrap obtenida de la muestra original. Otras muestras bootstrap pueden ser diferentes.

Los ingresos tienden a tener distribuciones asimétricas en lugar de ser normales, por lo que no deberíamos usar los métodos de la sección 7-2 con una pequeña muestra de ingresos. Esta es una situación en la que el método bootstrap viene al rescate.

SU TURNO Resuelva el ejercicio 3 "Muestra bootstrap".

¿Por qué se llama "bootstrap"? El término "bootstrap" (que significa literalmente lengüeta de la bota) se utiliza porque los datos "se estiran a sí mismos por sus propias lengüetas" para generar nuevos conjuntos de datos. En los días de antaño, "jalarse a uno mismo por la lengüeta" significaba que de alguna manera se lograba una tarea imposible, y el método bootstrap descrito en esta sección puede parecer imposible, ¡pero funciona!

¿Cuántos? Con el fin de proporcionar ejemplos manejables que no ocupen varias páginas cada uno, los ejemplos en esta sección incluyen conjuntos de datos muy pequeños y no más de 20 muestras bootstrap, pero deberíamos usar al menos 1000 muestras cuando usamos este método en aplicaciones reales. Los estadísticos profesionales comúnmente usan 10,000 o más muestras bootstrap.

Procedimiento bootstrap para estimar el intervalo de confianza de un parámetro

1. Dada una muestra aleatoria simple de tamaño *n*, obtenga muchas (como 1000 o más) muestras bootstrap del mismo tamaño *n*.
2. Para el parámetro que se estima, busque el estadístico correspondiente para cada una de las muestras bootstrap. (Ejemplo: para una estimación de confianza de μ, encuentre la *media muestral* $\bar{x}$ de cada muestra bootstrap).
3. Clasifique la lista de estadísticos muestrales de menor a mayor.

continúa

¿A cuántas personas conoce?

Es difícil para cualquiera contar el número de personas que conoce, pero es posible utilizar métodos estadísticos para estimar el número medio de personas que todos conocemos. El enfoque simple de preguntar a alguien a cuántas personas conoce no ha funcionado bien con anterioridad. Un enfoque mucho más útil es seleccionar una muestra representativa de personas y preguntarle a cada una a cuántas personas conoce que se llamen Mario, Ginny, Rachel o Todd. (Los nombres poco comunes son más efectivos porque las personas con nombres más típicos son difíciles de recordar con precisión). Las respuestas se usan para proyectar el número total de personas que se conocen. (Si los sujetos de la muestra conocen un promedio de 1.76 personas con esos nombres, y sabemos que 0.288% de la población tiene tales nombres, entonces el número medio de personas conocidas es 1.76/0.00288 = 611). Según una estimación, el número medio de personas conocidas es 611 y la mediana es 472. (Vea "How Much People Do You Know?, Efficiently Estimating Personal Network Size", de McCormick, Salganik y Zheng, en *Journal of the American Statistical Association*, vol. 105, núm. 4).

4. Con base en la lista ordenada de los estadísticos, cree el intervalo de confianza encontrando los valores de percentil correspondientes. Los procedimientos para encontrar los percentiles se dan en la sección 3-3. (Ejemplo: Con base en una lista de medias muestrales ordenadas, los límites del intervalo de confianza del 90% son P_5 y P_{95}. La estimación del intervalo de confianza del 90% para μ es $P_5 < \mu < P_{95}$).

Utilidad de los resultados Con el fin de ilustrar el procedimiento bootstrap, los ejemplos 2, 3 y 4 implican muestras muy pequeñas con sólo 20 muestras bootstrap. En consecuencia, los intervalos de confianza resultantes incluyen casi todo el rango de valores muestrales, y esos intervalos de confianza no son muy útiles. Las muestras más grandes con 1000 o más muestras bootstrap proporcionarán resultados mucho mejores que las de los ejemplos 2, 3 y 4.

Proporciones

Cuando se trabaja con proporciones, es muy útil representar los datos de las dos categorías mediante el uso de 0 y 1, como en el siguiente ejemplo.

EJEMPLO 2 Encuesta del color de ojos: IC bootstrap para la proporción

En una encuesta, se preguntó a cuatro sujetos seleccionados al azar si tenían ojos marrones, y los resultados fueron: 0, 0, 1, 0 (donde 0 = no y 1 = sí). Use el procedimiento de remuestreo bootstrap para construir una estimación del intervalo de confianza del 90% para la proporción poblacional p, la proporción de personas con ojos marrones en la población.

TABLA 7-3 Muestras bootstrap para p

Muestra bootstrap				$\hat{p}$	$\hat{p}$ ordenadas	
1	0	0	1	0.50	0.00	→ $P_5 = 0.00$
1	0	1	0	0.50	0.00	
0	1	1	1	0.75	0.00	
0	0	0	0	0.00	0.00	
0	1	0	0	0.25	0.25	
1	0	0	0	0.25	0.25	
0	1	0	1	0.50	0.25	
1	0	0	0	0.25	0.25	
0	0	0	0	0.00	0.25	
0	0	1	1	0.50	0.25	**90% de intervalo de confianza:**
0	0	0	1	0.25	0.25	$0.00 < p < 0.75$
0	0	1	0	0.25	0.25	
1	1	1	0	0.75	0.50	
0	0	0	0	0.00	0.50	
0	0	0	0	0.00	0.50	
0	1	1	0	0.50	0.50	
0	0	1	0	0.25	0.50	
1	0	0	0	0.25	0.75	
1	1	1	0	0.75	0.75	→ $P_{95} = 0.75$
0	0	0	1	0.25	0.75	

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS La muestra es una muestra aleatoria simple. (No hay requisito de al menos 5 éxitos y al menos 5 fallas, o $np \geq 5$ y $nq \geq 5$. No hay requisito de que la muestra deba provenir de una población distribuida normalmente).

Paso 1: En la tabla 7-3, creamos 20 muestras bootstrap de la muestra original 0, 0, 1, 0.

Paso 2: Como queremos una estimación del intervalo de confianza para la proporción poblacional p, necesitamos la proporción muestral $\hat{p}$ para cada una de las 20 muestras bootstrap, y esas proporciones muestrales se presentan en la columna más a la derecha de las muestras bootstrap.

Paso 3: La columna de datos que se muestra más a la derecha es una lista de las 20 proporciones muestrales ordenadas de menor a mayor.

Paso 4: Puesto que queremos un nivel de confianza del 90%, tenemos que encontrar los percentiles P_5 y P_{95}. Recuerde que P_5 separa el 5% más bajo de los valores, y P_{95} separa el 5% más alto de los valores. Usando los métodos de la sección 3-3 para encontrar percentiles, empleamos la lista ordenada de proporciones muestrales bootstrap para encontrar que $P_5 = 0.00$ y $P_{95} = 0.75$. La estimación del intervalo de confianza del 90% para la proporción poblacional es $0.00 < p < 0.75$.

INTERPRETACIÓN

El intervalo de confianza de $0.00 < p < 0.75$ es bastante amplio. Después de todo, cada intervalo de confianza para cada proporción debe caer entre 0 y 1, por lo que el intervalo de confianza resultante no parece ser útil, pero se basa sólo en cuatro valores muestrales.

SU TURNO Resuelva el ejercicio 5 "Compra en línea".

SUGERENCIA: El ejemplo 2 utiliza sólo 20 muestras bootstrap, pero la aplicación efectiva del método requiere el uso de software para generar 1000 o más muestras bootstrap.

Medias

En la sección 7-2 observamos que cuando se construye una estimación del intervalo de confianza para una media poblacional, existe un requisito de que la muestra provenga de una población normalmente distribuida o que el tamaño de la muestra sea mayor que 30. El método bootstrap se puede usar aun cuando tal requisito no esté satisfecho.

EJEMPLO 3 Ingresos: IC bootstrap para la media

Cuando el autor recolectó una muestra aleatoria simple de los ingresos anuales de sus estudiantes de estadística, obtuvo los siguientes resultados (en miles de dólares): 0, 2, 3, 7. Utilice el procedimiento de remuestreo bootstrap para construir una estimación del intervalo de confianza del 90% para el ingreso anual medio de la población de todos los estudiantes de estadística del autor.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS La muestra es una muestra aleatoria simple y no existe el requisito de que la muestra deba provenir de una población distribuida normalmente. Debido a que las distribuciones de ingresos suelen ser asimétricas en lugar de normales, no deberíamos usar los métodos de la sección 7-2 para encontrar el intervalo de confianza, pero se puede usar el método bootstrap. ✓

Paso 1: En la tabla 7-4, creamos 20 muestras bootstrap (¡con reemplazo!) de la muestra original 0, 2, 3, 7. (Aquí usamos sólo 20 muestras bootstrap, por lo que tenemos un ejemplo manejable que no ocupará muchas páginas de texto, pero generalmente se requieren al menos 1000 muestras bootstrap).

Paso 2: Como queremos una estimación del intervalo de confianza para la media poblacional μ, necesitamos la media muestral $\bar{x}$ para cada una de las 20 muestras bootstrap, y esas medias muestrales se presentan en la columna más a la derecha de la tabla.

Paso 3: La columna de datos que se muestra más a la derecha es una lista de las 20 medias muestrales ordenadas de menor a mayor.

Paso 4: Puesto que queremos un nivel de confianza del 90%, tenemos que encontrar los percentiles P_5 y P_{95}. Recuerde que P_5 separa el 5% más bajo de los valores, y P_{95} separa el 5% más alto de los valores. Con base en los métodos de la sección 3-3 para encontrar percentiles, usamos la lista *ordenada* de medias muestrales bootstrap para encontrar que $P_5 = 1.75$ y $P_{95} = 4.875$. La estimación del intervalo de confianza del 90% para la media poblacional es $1.75 < \mu < 4.875$, donde los valores se dan en miles de dólares.

TABLA 7-4 Muestras bootstrap para μ

Muestra bootstrap				$\bar{x}$	$\bar{x}$ ordenadas	
3	3	0	2	2.00	1.75	
0	3	2	2	1.75	1.75	→ $P_5 = 1.75$
7	0	2	7	4.00	1.75	
3	2	7	3	3.75	2.00	
0	0	7	2	2.25	2.00	
7	0	0	3	2.50	2.25	
3	0	3	2	2.00	2.50	
3	7	3	7	5.00	2.50	
0	3	2	2	1.75	2.50	Intervalo de confianza del 90%:
0	3	7	0	2.50	2.75	
0	7	2	2	2.75	3.00	$1.75 < \mu < 4.875$
7	2	2	3	3.50	3.25	
7	2	3	7	4.75	3.25	
2	7	2	7	4.50	3.50	
0	7	2	3	3.00	3.75	
7	3	7	2	4.75	4.00	
3	7	0	3	3.25	4.50	
0	0	3	7	2.50	4.75	
3	3	7	0	3.25	4.75	
2	0	2	3	1.75	5.00	→ $P_{95} = 4.875$

SU TURNO Resuelva el ejercicio 7 "Novato 15".

En cifras

61.68 °F es la temperatura global media en el mes de julio, al momento de escribir estas palabras. Este mes fue el *mes más caluroso de la historia*, según la Administración Nacional Oceánica y Atmosférica.

Desviaciones estándar

En la sección 7-3 observamos que cuando se construyen estimaciones del intervalo de confianza para las desviaciones estándar o varianzas poblacionales, existe el requisito de que la muestra debe provenir de una población con valores distribuidos normalmente. Incluso si la muestra es grande, este requisito de normalidad es mucho más estricto que el utilizado para estimar las medias poblacionales. En consecuencia, el método bootstrap se vuelve más importante para las estimaciones del intervalo de confianza de σ o σ^2.

EJEMPLO 4 Ingresos: IC bootstrap para la desviación estándar

Considere los mismos ingresos (en miles de dólares) del ejemplo 3: 0, 2, 3, 7. Utilice el procedimiento de remuestreo bootstrap a fin de construir una estimación del intervalo de confianza del 90% para la desviación estándar poblacional σ; es decir, la desviación estándar de los ingresos anuales de los estudiantes de estadística del autor.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Aquí se aplica la misma verificación de requisitos utilizada en el ejemplo 3.

También se usa el mismo procedimiento básico. El ejemplo 3 ya incluye 20 muestras bootstrap, así que se encontrará la *desviación estándar* de cada muestra, y luego se clasificarán para obtener la lista ordenada de desviaciones estándar muestrales:

1.26	1.26	1.26	1.41	1.41	2.22	2.31	2.38	2.63	2.63
2.87	2.87	2.89	2.94	2.99	3.30	3.32	3.32	3.32	3.56

Los límites del intervalo de confianza del 90% se encuentran en esta lista ordenada de desviaciones estándar mediante la determinación de P_5 y P_{95}. Con base en los métodos de la sección 3-3, obtenemos $P_5 = 1.26$ y $P_{95} = 3.44$. La estimación del intervalo de confianza del 90% para la desviación estándar poblacional σ es $1.26 < \sigma < 3.44$, donde los valores se dan en miles de dólares.

SU TURNO Resuelva el inciso (b) del ejercicio 8 “Radiación del teléfono celular”.

Una vez más, tenga en cuenta que, por razones prácticas, los ejemplos de esta sección incluyeron conjuntos de datos muy pequeños y no más de 20 muestras bootstrap, pero se recomienda utilizar al menos 1000 muestras de este tipo. El uso de 10,000 o más muestras bootstrap es común.

CENTRO DE TECNOLOGÍA

Remuestreo bootstrap

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Ingrese los valores muestrales listados en la columna 1 del editor de muestras.
2. Haga clic en **Analysis** en el menú superior y seleccione **Bootstrap Resampling** en el menú desplegable.
3. Ingrese la cantidad deseada de muestras bootstrap, luego haga clic en **Resample**.
4. Las medias muestrales ordenadas se listan en la columna 2; las desviaciones estándar muestrales ordenadas se listan en la columna 3.

Minitab

Minitab aún no tiene una función de bootstrap, pero las muestras se pueden generar mediante una serie de comandos existentes en Minitab.

Visite www.pearsonenespañol.com/triola para obtener más información.

StatCrunch

1. Ingrese los valores muestrales listados en la columna *var1*.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Resample-Statistic** en el menú desplegable.
4. En *Columns to resample*, seleccione **var1**.
5. En *Statistic* haga clic en **Build** y cree la expresión deseada, como **mean(var1)** o **std(var1)**. Haga clic en **Okay** cuando haya terminado.
6. Complete el cuadro de diálogo y haga clic en **Compute**!

CENTRO DE TECNOLOGÍA *continuación*

Remuestreo bootstrap

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus	Excel
No disponible.	**Complemento XLSTAT (requerido)** 1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego en **Describing data**. 2. Seleccione **Resampled statistics** en el menú desplegable. 3. En *Quantitative Data* introduzca el rango de celdas que contiene los valores muestrales. Si la primera fila contiene una etiqueta, marque la casilla **Sample Labels**. 4. En *Method*, seleccione **Bootstrap**. 5. Introduzca el número deseado de muestras. 6. Haga clic en la pestaña **Outputs** e ingrese el nivel de confianza (%) de **95**; confirme que **Standard bootstrap interval** esté seleccionado. También seleccione **Mean** y **Standard deviation ($n - 1$)**. 7. Haga clic en **OK.**

7-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Sustitución ¿Por qué el método bootstrap requiere muestreo con reemplazo? ¿Qué sucedería si usáramos los métodos de esta sección pero muestreamos sin reemplazo?

2. Muestra bootstrap La siguiente es una muestra aleatoria de los tiempos de salida (min) para los vuelos de American Airlines en el aeropuerto JFK: 12, 19, 13, 43, 15. Para esta muestra, ¿cuál es una muestra bootstrap?

3. Muestra bootstrap Considere los datos muestrales del ejercicio 2, ¿cuál(es) de las siguientes opciones no son posibles muestras bootstrap?

a. 12, 19, 13, 43, 15 **b.** 12, 19, 15 **c.** 12, 12, 12, 43, 43

d. 14, 20, 12, 19, 15 **e.** 12, 13, 13, 12, 43, 15, 19

4. ¿Cuántos? Los ejemplos en esta sección implican no más de 20 muestras bootstrap, ¿cuántos se deben usar en aplicaciones reales?

En los ejercicios 5 a 8, use el número relativamente pequeño de muestras bootstrap proporcionadas para construir el intervalo de confianza.

5. Compras en línea En una encuesta del Centro de Investigación de Informes del Consumidor, a las mujeres se les preguntó si compraban libros en línea, y las respuestas incluían lo siguiente: no, sí, no, no. Si "sí" = 1 y "no" = 0, aquí hay diez muestras bootstrap para tales respuestas: {0, 0, 0, 0}, {1, 0, 1, 0}, {1, 0, 1, 0}, {0, 0, 0, 0}, {0, 0, 0, 0}, {0, 1, 0, 0}, {0, 0, 0, 0}, {0, 0, 0, 0}, {0, 1, 0, 0}, {1, 1, 0, 0}. Utilizando sólo las diez muestras bootstrap proporcionadas, construya una estimación del intervalo de confianza del 90% para la proporción de mujeres que afirman comprar libros en línea.

6. Elección del asiento En una encuesta de 3M Privacy Filters, se pidió a los encuestados que identificaran su asiento favorito cuando vuelan, y el resultado incluye las siguientes respuestas: ventanilla, ventanilla, otro y otro. Sí "ventanilla" = 1 y "otro" = 0, las siguientes son diez muestras bootstrap para tales respuestas: {0, 0, 0, 0}, {0, 1, 0, 0}, {0, 1, 0, 1}, {0, 0, 1, 0}, {1, 1, 1, 0}, {0, 1, 1, 0}, {1, 0, 0, 1}, {0, 1, 1, 1}, {1, 0, 1, 0}, {1, 0, 0, 1}. Usando sólo las diez muestras bootstrap proporcionadas, construya una estimación del intervalo de confianza del 80% para la proporción de encuestados que indicaron que su asiento favorito es "ventanilla".

continúa

7. Novato 15 La siguiente es una muestra de los cambios de peso (kg) de estudiantes universitarios en su primer año (del conjunto de datos 6 "Novato 15" en el apéndice B): 11, 3, 0, −2, donde −2 representa una *pérdida* de 2 kg y los valores positivos representan el peso ganado. A continuación se presentan diez muestras bootstrap: $\{11, 11, 11, 0\}$, $\{11, -2, 0, 11\}$, $\{11, -2, 3, 0\}$, $\{3, -2, 0, 11\}$, $\{0, 0, 0, 3\}$, $\{3, -2, 3, -2\}$, $\{11, 3, -2, 0\}$, $\{-2, 3, -2, 3\}$, $\{-2, 0, -2, 3\}$, $\{3, 11, 11, 11\}$.

a. Usando sólo las diez muestras bootstrap proporcionadas, construya una estimación del intervalo de confianza del 80% para el cambio de peso medio en la población.

b. Usando sólo las diez muestras bootstrap proporcionadas, construya una estimación del intervalo de confianza del 80% para la desviación estándar de los cambios de peso en la población.

8. Radiación del teléfono celular La siguiente es una muestra de las emisiones de radiación medidas (cW/kg) para teléfonos celulares (según datos del Grupo de Trabajo Ambiental): 38, 55, 86, 145. A continuación se presentan diez muestras bootstrap: $\{38, 145, 55, 86\}$, $\{86, 38, 145, 145\}$, $\{145, 86, 55, 55\}$, $\{55, 55, 55, 145\}$, $\{86, 86, 55, 55\}$, $\{38, 38, 86, 86\}$, $\{145, 38, 86, 55\}$, $\{55, 86, 86, 86\}$, $\{145, 86, 55, 86\}$, $\{38, 145, 86, 55\}$.

a. Usando sólo las diez muestras bootstrap proporcionadas, construya una estimación del intervalo de confianza del 80% para la media poblacional.

b. Usando sólo las diez muestras bootstrap proporcionadas, construya una estimación del intervalo de confianza del 80% para la desviación estándar poblacional.

En los ejercicios 9 a 22 use la tecnología para crear la gran cantidad de muestras bootstrap.

9. Novato 15 Repita el ejercicio 7 "Novato 15" usando un nivel de confianza del 90% para los incisos (a) y (b) y empleando 1000 muestras bootstrap en lugar de las 10 que se dieron en ese ejercicio.

10. Radiación del teléfono celular Repita el ejercicio 8 "Radiación del teléfono celular" usando un nivel de confianza del 90% para los incisos (a) y (b) y empleando 1000 muestras bootstrap en lugar de las 10 que se dieron en ese ejercicio.

11. Citas rápidas Use las siguientes medidas masculinas del atractivo femenino dadas en el ejercicio 17 "Citas rápidas" de la sección 7-2 en la página 329: 7, 8, 2, 10, 6, 5, 7, 8, 8, 9, 5, 9. Use el método bootstrap con 1000 muestras.

a. Construya una estimación del intervalo de confianza del 99% para la media de poblacional. ¿Es dramáticamente diferente el resultado del intervalo de confianza al 99% encontrado en el ejercicio 17 de la sección 7-2?

b. Construya una estimación del intervalo de confianza del 95% para la desviación estándar poblacional. ¿Es dramáticamente diferente el resultado del intervalo de confianza al 95% que se encontró en el ejercicio 13 "Citas rápidas" de la sección 7-3 en la página 340?

12. Citas rápidas Use las siguientes medidas femeninas del atractivo masculino dadas en el ejercicio 18 "Citas rápidas" de la sección 7-2 de la página 329: 5, 8, 3, 8, 6, 10, 3, 7, 9, 8, 5, 5, 6, 8, 8, 7, 3, 5, 5, 6, 8, 7, 8, 8, 8, 7. Use el método bootstrap con 1000 muestras.

a. Construya una estimación del intervalo de confianza del 99% para la media poblacional. ¿Es dramáticamente diferente el resultado al intervalo de confianza del 99% encontrado en el ejercicio 18 de la sección 7-2?

b. Construya una estimación del intervalo de confianza del 95% para la desviación estándar poblacional. ¿Es dramáticamente diferente el resultado al intervalo de confianza del 95% que se encontró en el ejercicio 14 "Citas rápidas" de la sección 7-3 en la página 340?

13. Mickey D's En un estudio sobre la precisión de los pedidos de comida rápida, McDonald's tuvo 33 pedidos que no fueron precisos entre los 362 pedidos observados (según datos de la revista *QSR*). Utilice el método bootstrap para construir una estimación del intervalo de confianza del 95% para la proporción de pedidos que no son precisos. Use 1000 muestras bootstrap. ¿Cómo se compara el resultado con el intervalo de confianza encontrado en el ejercicio 13 "Mickey D's" de la sección 7-1?

14. Eliquis El medicamento Eliquis (apixaban) se usa para ayudar a prevenir los coágulos en la sangre de ciertos pacientes. En los ensayos clínicos, entre 5924 pacientes tratados con Eliquis, 153 desarrollaron la reacción adversa de náuseas (según datos de Bristol-Myers Squibb Co.). Utilice el método bootstrap para construir una estimación del intervalo de confianza del 99% para la proporción de pacientes que experimentan náuseas. Use 1000 muestras bootstrap. ¿Cómo se compara este resultado con el intervalo de confianza encontrado en el ejercicio 14 "Eliquis" de la sección 7-1 en la página 312?

15. Tasa de retorno de la encuesta En un estudio sobre el uso de teléfonos celulares y el dominio hemisférico del cerebro, se envió un cuestionario por Internet a 5000 sujetos seleccionados aleatoriamente de un grupo otológico en línea (centrado en los oídos) y 717 encuestas fueron devueltas. Use el método bootstrap para construir una estimación del intervalo de confianza del 90% para la proporción de encuestas devueltas. Use 1000 muestras bootstrap. ¿Cómo se compara este resultado con el del intervalo de confianza encontrado en el ejercicio 15 "Tasa de retorno de la encuesta" de la sección 7-1 en la página 312?

16. Negligencia médica En un estudio de 1228 demandas legales por negligencia médica seleccionadas aleatoriamente, se descubrió que 856 de ellas fueron descartadas o desechadas (con base en datos de la Asociación de Aseguradores de Médicos de EUA). Utilice el método bootstrap para construir una estimación del intervalo de confianza del 95% para la proporción de demandas que se descartan o desechan. Use 1000 muestras bootstrap. ¿Cómo se compara este resultado con el del intervalo de confianza encontrado en el ejercicio 16 "Negligencia médica" de la sección 7-1 en la página 312?

17. Evaluaciones de estudiantes A continuación se listan las calificaciones dadas por estudiantes a sus cursos, donde una calificación de 5 equivale a "excelente". Las calificaciones se obtuvieron en la Universidad de Texas en Austin. (Consulte el conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B). Utilizando el método bootstrap con 1000 muestras, construya una estimación del intervalo de confianza del 90% para μ. ¿Cómo se compara este resultado con el del intervalo de confianza encontrado en el ejercicio 23 "Evaluaciones de estudiantes" de la sección 7-2 en la página 330?

3.8 3.0 4.0 4.8 3.0 4.2 3.5 4.7 4.4 4.2 4.3 3.8 3.3 4.0 3.8

18. Cafeína en bebidas refrescantes A continuación se listan las cantidades medidas de cafeína (mg por 12 onzas de bebida) obtenidas en una lata de cada una de 20 marcas. Utilizando el método bootstrap con 1000 muestras, construya una estimación del intervalo de confianza del 99% para μ. ¿Cómo se compara este resultado con el del intervalo de confianza encontrado en el ejercicio 22 "Cafeína en bebidas refrescantes" de la sección 7-2 en la página 330?

0 0 34 34 34 45 41 51 55 36 47 41 0 0 53 54 38 0 41 47

19. Old Faithful Use los siguientes tiempos de duración de las erupciones del Old Faithful (en segundos):

125 203 205 221 225 229 233 233 235 236 236 237 238 238 239 240 240 240 240 241

241 242 242 242 243 243 244 245 245 245 245 246 246 248 248 248 249 249 250 251

252 253 253 255 255 256 257 258 262 264

a. Utilice el método bootstrap con 1000 muestras para encontrar una estimación del intervalo de confianza del 95% para μ.

b. Encuentre la estimación del intervalo de confianza del 95% para μ determinada mediante los métodos de la sección 7-2.

c. Compare los resultados.

20. Old Faithful Repita el ejercicio 19 "Old Faithful" usando la desviación estándar en lugar de la media. Compare el intervalo de confianza que se encontraría usando los métodos de la sección 7-3. Si los dos intervalos de confianza son muy diferentes, ¿cuál es mejor? ¿Por qué?

21. Análisis de los últimos dígitos El peso de varios encuestados se registró como parte de la Encuesta Entrevista de Salud en California. A continuación se listan los últimos dígitos de los pesos de 50 encuestados seleccionados al azar.

5 0 1 0 2 0 5 0 5 0 3 8 5 0 5 0 5 6 0 0 0 0 0 0 8

5 5 0 4 5 0 0 4 0 0 0 0 0 8 0 9 5 3 0 5 0 0 0 5 8

a. Utilice el método bootstrap con 1000 muestras a fin de encontrar una estimación del intervalo de confianza del 95% para σ.

b. Encuentre la estimación del intervalo de confianza del 95% para σ determinada mediante los métodos de la sección 7-3.

c. Compare los resultados. Si los dos intervalos de confianza son diferentes, ¿cuál es mejor? ¿Por qué?

22. Análisis de los últimos dígitos Repita el ejercicio 21 "Análisis de los últimos dígitos" utilizando la media en vez de la desviación estándar. Compare el intervalo de confianza con el que se encontraría usando los métodos de la sección 7-2.

7-4 Más allá de lo básico

23. Efecto del número de muestras bootstrap Repita el ejercicio 21 "Análisis de los últimos dígitos" utilizando 10,000 muestras bootstrap en lugar de 1000. ¿Qué sucede?

24. Formas de distribución Use los datos muestrales que se proporcionan en el ejercicio 21 "Análisis de los últimos dígitos".

a. ¿Los valores muestrales originales parecen ser de una población distribuida normalmente? Explique.

b. ¿Las 1000 muestras bootstrap parecen tener medias que provienen de una población distribuida normalmente? Explique.

c. ¿Las 1000 muestras bootstrap parecen tener desviaciones estándar que provienen de una población distribuida normalmente? Explique.

Examen rápido del capítulo

1. Celebridades y la ley A continuación se presenta una estimación del intervalo de confianza del 95% para la proporción de adultos que dicen que la ley es permisiva para las celebridades: $0.692 < p < 0.748$ (según datos de una encuesta de Rasmussen Reports). ¿Cuál es la mejor estimación puntual de la proporción de adultos en la población que dice que la ley es permisiva para las celebridades?

2. Interpretación del IC Escriba un enunciado breve que interprete correctamente el intervalo de confianza dado en el ejercicio 1 "Celebridades y la ley".

3. Valor crítico Para la encuesta descrita en el ejercicio 1 "Celebridades y la ley", encuentre el valor crítico que se usaría para construir una estimación del intervalo de confianza del 99% para la proporción poblacional.

4. Cambio suelto *USA Today* informó que 40% de las personas encuestadas planeaba usar el cambio suelto acumulado para pagar algunas cuentas. El margen de error se dio como ± 3.1 puntos porcentuales. Identifique el intervalo de confianza que corresponde a esa información.

5. Tamaño de muestra para la proporción Encuentre el tamaño de muestra requerido para estimar el porcentaje de estudiantes universitarios que toman un curso de estadística. Suponga que queremos un 95% de confianza de que la proporción muestral se encuentre dentro de cuatro puntos porcentuales del porcentaje real poblacional.

6. Tamaño de muestra para la media Encuentre el tamaño de muestra requerido para estimar el IQ promedio de músicos profesionales. Suponga que queremos un 98% de confianza de que la media muestral se encuentre dentro de tres puntos de IQ de la media real poblacional. También suponga que $\sigma = 15$.

7. Requisitos Un analista de control de calidad ha recolectado una muestra aleatoria de 12 baterías de teléfonos inteligentes y planea probar su nivel de voltaje y construir una estimación del intervalo de confianza del 95% para el nivel de voltaje medio en la población de baterías. ¿Qué requisitos se deben cumplir para construir el intervalo de confianza usando el método de la distribución t?

8. Grados de libertad En general, ¿a qué se refieren los "grados de libertad"? Para los datos muestrales descritos en el ejercicio 7 "Requisitos", encuentre la cantidad de grados de libertad, suponiendo que se desea construir una estimación del intervalo de confianza de μ usando la distribución t.

9. Valor crítico Consulte el ejercicio 7 "Requisitos" y suponga que se cumplen los requisitos. Encuentre el valor crítico que se usaría para construir una estimación del intervalo de confianza del 95% para μ usando la distribución t.

10. ¿Qué método? Consulte el ejercicio 7 "Requisitos" y suponga que la muestra de 12 niveles de voltaje parece ser de una población con una distribución que está sustancialmente lejos de ser normal. ¿Debería construirse una estimación del intervalo de confianza del 95% para σ usando la distribución χ^2? De no ser así, ¿qué otro método se podría emplear para encontrar una estimación del intervalo de confianza del 95% para σ?

Ejercicios de repaso

1. Noticias en línea En una encuesta de Harris aplicada a 2036 adultos, el 40% dijo que prefieren recibir sus noticias en línea. Construya una estimación del intervalo de confianza del 95% para el *porcentaje* de todos los adultos que dicen que prefieren recibir sus noticias en línea. ¿Podemos decir con seguridad que menos del 50% de los adultos prefieren recibir sus noticias en línea?

2. Computadoras Con el fin de planificar mejor los recursos de los estudiantes, el presidente del departamento de matemáticas del Broward College quiere calcular el porcentaje de estudiantes que poseen una computadora. Si deseamos estimar ese porcentaje según los resultados de la encuesta, ¿a cuántos estudiantes debemos aplicar la encuesta para tener el 90% de confianza de que estamos dentro de cuatro puntos porcentuales desde el porcentaje poblacional? Suponga que el número de estudiantes es muy grande.

3. Magnitudes de sismos A continuación se listan las magnitudes en la escala Richter de sismos seleccionados al azar.

a. Identifique la mejor estimación puntual de la media poblacional μ.

b. Construya una estimación del intervalo de confianza del 95% para la magnitud media de la población de sismos.

c. Escriba un enunciado que interprete el intervalo de confianza.

3.09 2.76 2.65 3.44 3.01 2.94 3.45 2.72 2.69 2.89 2.71 2.76

4. Zurdos Se han llevado a cabo varios estudios en un intento por identificar las formas en que las personas zurdas son diferentes de los diestros. Suponga que desea calcular el IQ medio de todos los adultos zurdos. ¿Cuántos adultos zurdos al azar deben ser evaluados para estar 99% seguros de que el IQ medio del grupo muestral está dentro de cuatro puntos desde el IQ medio de todos los adultos zurdos? Suponga que se sabe que σ es 15.

5. Distribuciones Identifique la distribución (normal, *t* de Student, ji cuadrada) que debe usarse en cada una de las siguientes situaciones. Si no se puede usar ninguna de las tres distribuciones, ¿qué otro método podría emplearse?

a. Al construir un intervalo de confianza para μ, usted tiene 75 valores muestrales y parecen provenir de una población con una distribución asimétrica. Se desconoce la desviación estándar de la población.

b. Al construir una estimación del intervalo de confianza para μ, usted tiene 75 valores muestrales y parecen provenir de una población con una distribución asimétrica. Se sabe que la desviación estándar poblacional es de 18.2 cm.

c. Al construir una estimación del intervalo de confianza para σ, usted tiene 75 valores muestrales y parecen provenir de una población con una distribución asimétrica.

d. Al construir una estimación de intervalo de confianza para σ, usted tiene 75 valores muestrales y parecen pertenecer a una población con distribución normal.

e. Al construir una estimación del intervalo de confianza para *p*, usted tiene 1200 encuestados y el 5% respondió "sí" a la primera pregunta.

6. Tamaño de muestra Su nuevo empleador lo ha contratado para encuestar adultos sobre sus suscripciones a periódicos impresos.

a. Si desea calcular el porcentaje de adultos que tienen una suscripción pagada a un periódico impreso, ¿a cuántos adultos debe encuestar si desea un 95% de confianza de que su porcentaje tenga un margen de error de tres puntos porcentuales?

b. Si desea estimar la cantidad media que gastaron los adultos en periódicos impresos durante el año pasado, ¿a cuántos adultos debe encuestar si desea un 95% de confianza de que su promedio muestral tenga un error no mayor a \$5? (Según los resultados de un estudio piloto, suponga que la desviación estándar de los montos gastados en periódicos impresos en el último año es de \$47).

c. Si planea obtener las estimaciones descritas en los incisos (a) y (b) con una sola encuesta que tenga varias preguntas, ¿cuántos adultos deben ser encuestados?

7. Precisión de relojes de pulsera Los estudiantes del autor recopilaron datos que miden la precisión de los relojes de pulsera. Los siguientes tiempos (en segundos) muestran la discrepancia entre la hora real y la hora indicada en el reloj de pulsera. Los valores negativos corresponden a los relojes que están atrasados. Los datos parecen ser de una población normalmente distribuida. Construya una estimación del intervalo de confianza del 95% para la discrepancia media de la población de relojes de pulsera.

$-85 \quad 325 \quad 20 \quad 305 \quad -93 \quad 15 \quad 282 \quad 27 \quad 555 \quad 570 \quad -241 \quad 36$

8. Precisión de relojes de pulsera Use los datos muestrales del ejercicio 7 "Precisión de relojes de pulsera" y construya una estimación del intervalo de confianza del 95% para σ.

9. Bootstrap para la precisión de relojes de pulsera Repita el ejercicio 7 "Precisión de relojes de pulsera" utilizando 1000 muestras bootstrap.

10. IC para la proporción En una encuesta de TE Connectivity aplicada a 1000 adultos seleccionados al azar, el 2% dijo que "no sabían" si se sentirían cómodos estando en un vehículo autónomo. Existe la necesidad de construir una estimación del intervalo de confianza del 95% para la proporción de todos los adultos en la población que no saben si se sentirían cómodos.

a. Encuentre el intervalo de confianza usando la distribución normal como una aproximación a la distribución binomial.

b. Encuentre el intervalo de confianza usando 1000 muestras bootstrap.

c. Compare los resultados.

Ejercicios de repaso acumulado

Llegadas de vuelos. ***A continuación se listan los tiempos de retraso en la llegada (minutos) de algunos vuelos de American Airlines seleccionados al azar que volaron desde JFK en Nueva York hasta LAX en Los Ángeles. Los valores negativos corresponden a los vuelos que llegaron antes de la hora de llegada programada. Use estos valores para los ejercicios 1 a 4.***

$-30 \quad -23 \quad 14 \quad -21 \quad -32 \quad 11 \quad -23 \quad 28 \quad 103 \quad -19 \quad -5 \quad -46$

1. Estadísticos Encuentre la media, la mediana, la desviación estándar y el rango. ¿Los resultados son estadísticos o parámetros?

2. Regla práctica del rango Utilice los resultados del ejercicio 1 "Estadísticos" con la regla práctica del rango para encontrar los tiempos de llegada que separan aquellos que son significativamente bajos y significativamente altos. ¿El tiempo de retraso en la llegada de 103 min es significativamente alto?

3. Nivel de medición ¿Cuál es el nivel de medición (nominal, ordinal, de intervalo, de razón) de estos datos? ¿Los tiempos de llegada originales no redondeados son datos continuos o discretos?

4. Intervalo de confianza Construya una estimación del intervalo de confianza del 95% para el tiempo medio de retraso en la llegada de la población de todos los vuelos de American Airlines desde JFK a LAX.

5. Distribución normal Mediante el uso de un conjunto de datos mayor que el dado para los ejercicios 1 a 4, suponga que los retrasos en la llegada de las líneas aéreas se distribuyen normalmente con una media de -5.0 minutos y una desviación estándar de 30.4 minutos.

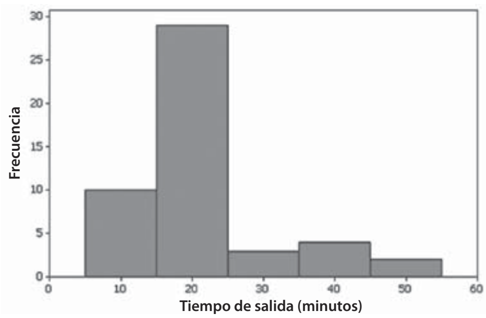
a. Encuentre la probabilidad de que un vuelo seleccionado aleatoriamente tenga un tiempo de retraso en la llegada de más de 15 minutos.

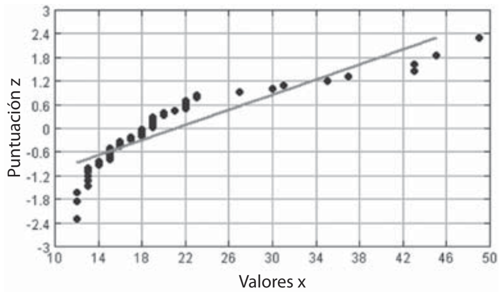
b. Encuentre el valor de Q_3, el tiempo de retraso en la llegada que es el tercer cuartil.

6. Tamaño de muestra Encuentre el tamaño de muestra necesario para calcular el tiempo medio de retraso en la llegada de todos los vuelos de American Airlines desde JFK a LAX. Suponga que queremos un 95% de confianza de que la media muestral tenga un error no mayor de 5 minutos. Con base en una muestra más grande que la dada para los ejercicios 1 a 4, suponga que todos los tiempos de retraso en la llegada tienen una desviación estándar de 30.4 minutos.

7. Llegadas a tiempo De acuerdo con la Oficina de Transporte, American Airlines tuvo una tasa de llegada a tiempo del 80.3% para un año determinado. Suponga que esta estadística se basa en una muestra de 1000 vuelos de American Airlines seleccionados al azar. Encuentre una estimación del intervalo de confianza del 99% para el *porcentaje* de todos los vuelos de American Airlines que llegan a tiempo.

8. Evaluación de la normalidad Una muestra aleatoria consiste en 48 tiempos (en minutos) requeridos para la salida (despegue) de los vuelos de American Airlines. Todos los vuelos son de American Airlines de Nueva York (JFK) a Los Ángeles y todos ocurrieron en enero de un año reciente. A continuación se representan los 48 tiempos de salida en un histograma y en una gráfica cuantilar normal. Con base en esas gráficas, ¿se puede inferir que los tiempos de salida son de una población con distribución normal? Dé una explicación de la distribución que se muestra. ¿Se puede inferir que los tiempos de salida satisfacen los requisitos necesarios para la construcción de una estimación del intervalo de confianza para la desviación estándar poblacional de todos esos tiempos?





Proyecto de tecnología

Profundidades de sismos El conjunto de datos 21 "Terremotos" en el apéndice B incluye las profundidades (en km) de las fuentes de 600 terremotos. Use tecnología para lo siguiente.

a. Encuentre la media y la desviación estándar de las 600 profundidades.

b. Genere un histograma y una gráfica cuantilar normal de las 600 profundidades. ¿Parece que las profundidades son de una población que tiene una distribución normal? Explique.

c. Si se desea obtener una estimación del intervalo de confianza del 95% para la profundidad de todos los terremotos, ¿se cumplen los requisitos para usar una distribución *t*? Explique.

d. Encuentre una estimación del intervalo de confianza del 95% para la profundidad de todos los terremotos utilizando la distribución *t*.

e. Encuentre una estimación del intervalo de confianza del 95% para la profundidad de todos los sismos utilizando 1000 muestras bootstrap.

f. Compare los resultados de los incisos (d) y (e). ¿Qué intervalo de confianza parece ser mejor? ¿Por qué?

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿Qué nos dice la encuesta?

Las encuestas se han convertido en una parte integral de nuestras vidas. Debido a que es tan importante que cada ciudadano tenga la capacidad de interpretar los resultados de este tipo de estudios, las encuestas son el foco del presente proyecto.

El Pew Research Center llevó a cabo recientemente una encuesta aplicada a 1007 adultos de EUA y descubrió que 85% de los encuestados conocen Twitter.

Análisis de los datos

1. Use los resultados de la encuesta para construir una estimación del intervalo de confianza del 95% para el porcentaje de todos los adultos que conocen Twitter.

2. Identifique el margen de error para esta encuesta.

3. Explique por qué estaría bien, o no, que un periódico hiciera esta afirmación: "Con base en los resultados de una encuesta reciente, más de 3 de cada 4 adultos conocen Twitter".

4. Suponga que usted es un reportero. Escriba una descripción de los resultados de la encuesta para su periódico.

5. Una crítica común a las encuestas es que sondean sólo un porcentaje muy pequeño de la población y, por lo tanto, no pueden ser precisas. ¿Tiene una muestra de sólo 1007 adultos tomada de una población de todos los adultos un tamaño demasiado pequeño? Escriba una breve explicación de por qué el tamaño de muestra de 1007 es demasiado pequeño, o no.

6. Con referencia a otro estudio, el presidente de una compañía escribió a Associated Press sobre una encuesta nacional de 1223 sujetos; él escribió:

> Cuando usted u otra persona intenta decirme a mí y a mis socios que 1223 personas explican nuestras opiniones y gustos aquí en EUA, ¡me vuelvo loco! ¡Cómo se atreve! Cuando usted o alguien más me dice que 1223 personas representan a EUA, es asombroso e injusto y debería estar prohibido.

El autor de esa carta insiste más adelante que, dado que el tamaño de muestra de 1223 personas representa a 120 millones de personas, su carta representa a 98,000 (120 millones divididos por 1223) que comparten las mismas opiniones. ¿Está usted de acuerdo o en desacuerdo con esta afirmación? Escriba una respuesta que respalde o refute el enunciado.

Actividades de cooperación en grupo

1. Actividad fuera de clase Recopile datos muestrales y use los métodos de este capítulo para construir estimaciones del intervalo de confianza para los parámetros poblaciones. A continuación se dan algunas sugerencias de parámetros:

- Proporción de estudiantes en su universidad que pueden levantar una ceja sin levantar la otra.
- Edad media de los autos manejados por estudiantes de estadística y/o la edad media de los autos manejados por el profesorado.
- Longitud media de las palabras en los editoriales del *New York Times* y longitud media de las palabras en los editoriales del periódico local.
- Proporción de estudiantes en su universidad que pueden identificar correctamente al presidente, vicepresidente y secretario de estado.
- Proporción de estudiantes en su universidad que tienen más de 18 años y están registrados para votar.
- Edad media de los estudiantes de tiempo completo en su universidad.
- Tiempo medio en horas que los estudiantes de su universidad estudian cada semana.
- Proporción de autos de estudiantes pintados de blanco.
- Proporción de autos que son rojos.

2. Actividad en clase Sin utilizar ningún dispositivo de medición, cada alumno debe dibujar una línea de 3 pulgadas de longitud y otra de 3 cm de largo. Luego use reglas para medir y registrar las longitudes de las líneas dibujadas. Encuentre las medias y las desviaciones estándar de los dos conjuntos de longitudes. Use los datos muestrales con el fin de construir un intervalo de confianza para la longitud de la línea estimada en 3 pulgadas, y luego haga lo mismo para la longitud de la línea estimada en 3 cm. ¿Los límites del intervalo de confianza realmente contienen la longitud correcta? Compare los resultados. ¿Las líneas estimadas en 3 pulgadas parecen ser más precisas que las líneas de 3 cm?

3. Actividad en clase Suponga que un método de selección de género puede afectar la probabilidad de que un bebé sea niña, de modo que la probabilidad sea de 1/4. Cada estudiante debe simular 20 nacimientos escogiendo 20 cartas de un paquete de naipes barajado. Reemplace cada carta después de escogerla, luego vuelva a barajarla. Considere que los corazones son niñas y que todas las demás cartas son niños. Después de hacer 20 selecciones y registrar los "géneros" de los bebés, construya una estimación del intervalo de confianza para la proporción de niñas. ¿El resultado parece ser eficaz para identificar el verdadero valor de la proporción poblacional? (Si no hay paquetes de naipes disponibles, use alguna otra forma para simular los nacimientos, por ejemplo use el generador de números aleatorios en una calculadora o use los dígitos de los números telefónicos o los números de Seguridad Social).

4. Actividad fuera de clase En grupos de tres o cuatro estudiantes deben asistir a la biblioteca y recolectar una muestra que consista en las edades de los libros (según las fechas de copyright). Piense y describa el procedimiento de muestreo, ejecute ese procedimiento y luego use los resultados para construir una estimación del intervalo de confianza de la edad media de todos los libros en la biblioteca.

5. Actividad en clase Cada estudiante debe escribir una estimación de la edad del actual presidente de EUA. Todas las estimaciones deben recolectarse, para después calcular la media y la desviación estándar muestrales. Luego use los resultados para construir un intervalo de confianza. ¿Los límites del intervalo de confianza contienen la edad correcta del presidente?

6. Actividad en clase Debe diseñarse un proyecto en clase para un ensayo en el que cada estudiante pruebe el sabor de Coca-Cola y Pepsi. Luego, se le pedirá al estudiante que identifique cuál muestra es Coca-Cola. Después de recopilar todos los resultados, analice la afirmación de que la tasa de éxito es mejor que la esperada con estimaciones aleatorias.

7. Actividad en clase Cada estudiante debe estimar la longitud del aula. Los valores deben basarse en estimaciones visuales, sin tomar medidas reales. Una vez recolectadas las estimaciones, construya un intervalo de confianza y luego mida la longitud real. ¿El intervalo de confianza contiene la longitud verdadera del aula? ¿Hay una "sabiduría colectiva", debido a la cual la media de la clase es aproximadamente igual a la longitud real del aula?

8. Actividad en clase Divídanse en grupos de tres o cuatro. Examinen una muestra de diferentes temas de una revista actual y encuentren la proporción de páginas que incluyen publicidad. Con base en los resultados, construyan una estimación del intervalo de confianza del 95% para el porcentaje de todas las páginas que tienen publicidad. Comparen sus resultados con los de otros grupos.

9. Actividad en clase Divídanse en grupos de dos. Primero encuentren el tamaño de muestra requerido para estimar la proporción de veces que aparece cara al lanzar una moneda, suponiendo que desea un 80% de confianza de que la proporción muestral esté dentro de 0.08 desde la proporción poblacional verdadera. Luego, lancen una moneda el número requerido de veces y registren sus resultados. ¿Qué porcentaje de los intervalos de confianza debería contener realmente el valor de la proporción poblacional, que sabemos es $p = 0.5$? Verifiquen este último resultado comparando su intervalo de confianza con los intervalos de confianza encontrados en otros grupos.

10. Actividad fuera de clase Identifique un tema de interés general y coordínese con todos los miembros de la clase para realizar una encuesta. En lugar de realizar una encuesta "científica" utilizando principios sólidos de selección aleatoria, utilice una muestra de conveniencia que incluya a los sujetos que estén disponibles, como amigos, parientes y otros estudiantes. Analice e interprete los resultados. Identifique la población. Describa las deficiencias de usar una muestra de conveniencia e intente identificar cómo una muestra de sujetos seleccionados al azar de la población podría ser diferente.

11. Actividad fuera de clase Cada estudiante debe encontrar un artículo en una revista profesional que incluya un intervalo de confianza del tipo discutido en este capítulo. Redacte un breve informe que describa el intervalo de confianza y su papel en el contexto del artículo.

8

PRUEBAS DE HIPÓTESIS

PROBLEMA DEL CAPÍTULO

Drones: ¿La mayoría de los consumidores se sentirían incómodos si sus compras les fueran entregadas por drones?

Las décadas recientes han traído consigo avances notables en tecnología, incluyendo drones, dispositivos GPS, teléfonos inteligentes, televisión de alta definición, dispositivos de seguridad biométricos y el interruptor de luz activado por aplausos. Para aprovechar las tecnologías más avanzadas, Amazon ha estado desarrollando Amazon Prime Air, que utiliza drones para entregar paquetes a sus clientes en 30 minutos o menos. El uso de los drones para esta actividad ha generado mucha discusión y debate.

Podemos usar métodos estadísticos para analizar encuestas y así evaluar e implementar de mejor manera las nuevas tecnologías. En una encuesta de Pitney Bowes, se preguntó a 1009 consumidores si se sentirían cómodos si sus compras les fueran entregadas por drones, y 54% (o 545) de ellos respondieron con un “no”. Con base en el resultado de esta encuesta, es razonable afirmar que la *mayoría* de los consumidores no se sienten cómodos con que sus compras les sean entregadas por drones, pero ¿los resultados de la encuesta realmente justifican esa afirmación? Después de todo, los encuestadores obtuvieron respuestas de sólo 1009 consumidores entre los 247,696,327 consumidores de Estados Unidos.

La afirmación de que la mayoría de los consumidores no se sienten cómodos con las entregas realizadas por drones se puede abordar utilizando el método de las *pruebas de hipótesis* que se presenta en este capítulo. Tenemos la afirmación de que $p > 0.5$, que es la forma simbólica de la afirmación verbal de que la mayoría (más de la mitad o 0.5) de los consumidores no se sienten cómodos con las entregas realizadas por drones. Este capítulo presentará los métodos estándar para probar dichas hipótesis. Con este conocimiento estaremos mejor preparados para responder a la pregunta de si estamos listos para aceptar drones como vehículos de entrega.

OBJETIVOS DEL CAPÍTULO

Los objetivos del capítulo son:

8-1 Fundamentos de las pruebas de hipótesis

- Desarrollar la capacidad de identificar las hipótesis nula y alternativa cuando se da alguna afirmación sobre un parámetro poblacional (como una proporción, media, desviación estándar o varianza).
- Desarrollar la capacidad de calcular un dato estadístico de prueba, encontrar valores críticos, calcular valores *P* y establecer una conclusión final sobre la hipótesis original. A continuación se listan los componentes que debe incluir una prueba de hipótesis:
  - Enunciados de las hipótesis nula y alternativa expresados en forma simbólica
  - Valor del dato estadístico de prueba
  - Selección de la distribución muestral que se utilizará para la prueba de hipótesis
  - Identificación de un valor *P* y/o el(los) valor(es) crítico(s)
  - Declaración de una conclusión que rechace la hipótesis nula o no la rechace
  - Declaración de una conclusión final que utilice términos simples y no técnicos sobre la hipótesis original

8-2 Prueba de una hipótesis respecto a una proporción

- Desarrollar la capacidad de utilizar datos muestrales para realizar una prueba de hipótesis formal de una afirmación acerca de una proporción poblacional. El procedimiento debe incluir los componentes listados anteriormente en los objetivos de la sección 8-1.

8-3 Prueba de una hipótesis respecto a una media

- Desarrollar la capacidad de utilizar datos muestrales para realizar una prueba de hipótesis formal de una afirmación acerca de una media poblacional. El procedimiento debe incluir los mismos componentes listados anteriormente en los objetivos de la sección 8-1.

8-4 Prueba de una hipótesis respecto a una desviación estándar o varianza

- Desarrollar la capacidad de utilizar datos muestrales para realizar una prueba de hipótesis formal de una afirmación de una desviación estándar o varianza poblacionales. El procedimiento debe incluir los mismos componentes listados anteriormente en los objetivos de la sección 8-1.

8-1 Fundamentos de las pruebas de hipótesis

Concepto clave En esta sección presentamos los componentes clave de una prueba de hipótesis formal. Los conceptos en esta sección son generales y se aplican a las pruebas de hipótesis que involucran proporciones, medias, desviaciones estándar o varianzas. En la parte 1 comenzamos con el "panorama general" para comprender el enfoque básico subyacente a las pruebas de hipótesis. Luego describimos las hipótesis nulas y alternativas, el nivel de significancia, los tipos de pruebas (de dos colas, de cola izquierda, de cola derecha), el estadístico de prueba, el valor *P*, los valores críticos y las conclusiones. En la parte 2, describimos los tipos de errores (tipo I y tipo II). En la parte 3, puntualizamos la *potencia* de una prueba de hipótesis.

PARTE 1 Conceptos básicos de las pruebas de hipótesis

Comenzamos con dos definiciones muy básicas.

DEFINICIONES

En estadística, una **hipótesis** es una afirmación o declaración sobre una propiedad de una población.

Una **prueba de hipótesis** (o **prueba de significancia**) es un procedimiento para probar una hipótesis sobre una propiedad de una población.

La "propiedad de una población" a la que se hace referencia en las definiciones anteriores suele ser el valor de un parámetro poblacional, por lo que a continuación se dan algunos ejemplos de hipótesis típicas:

- $\mu < 98.6$ °F "La temperatura corporal media de los humanos es menor que 98.6 °F".
- $p > 0.5$ "La proporción de consumidores que no se sienten cómodos con las entregas realizadas por drones es mayor que 0.5".
- $\sigma = 15$ "La población de estudiantes universitarios tiene puntuaciones de IQ con una desviación estándar de 15".

EJEMPLO 1 La mayoría de los consumidores no se sienten cómodos con las entregas realizadas por drones

Considere la afirmación del problema del capítulo de que "la mayoría de los consumidores no se sienten cómodos con las entregas realizadas por drones". Si se usa p para expresar la proporción de consumidores que no se sienten cómodos con las entregas mediante drones, la declaración de "mayoría" es equivalente a la afirmación de que la proporción es mayor que la mitad, o $p > 0.5$. La expresión $p > 0.5$ es la forma simbólica de la afirmación original. (En el problema del capítulo, una encuesta a 1009 consumidores incluyó un 54% que afirmaron no sentirse cómodos con las entregas mediante drones).

Panorama general En el ejemplo 1, tenemos la afirmación de que la proporción poblacional p es tal que $p > 0.5$. Entre 1009 consumidores, ¿cuántos necesitamos para obtener una cantidad significativamente alta de consumidores que no se sienten cómodos con la entrega mediante drones? Un resultado de 506 (o 50.1%) es apenas más de la mitad, por lo que resulta claro que 506 *no es significativamente alto*. Es evidente que un resultado de 1006 (o 99.1%) es *significativamente alto*. Pero ¿qué pasa con el resultado de 545 (o 54.0%) que se obtuvo en realidad en la encuesta de Pitney Bowes? ¿Es 545 (o 54.0%) significativamente alto? El método de las pruebas de hipótesis nos permite responder a esa pregunta clave.

Uso de la tecnología Es fácil obtener resultados de las pruebas de hipótesis usando tecnología. Las pantallas siguientes muestran los resultados de cuatro tecnologías diferentes, por lo que *podemos usar computadoras o calculadoras para hacer todo el esfuerzo de cálculo*. Al examinar las cuatro pantallas, vemos algunos elementos comunes. Todos muestran un "estadístico de prueba" $z = 2.55$ (redondeado), y todos incluyen un "valor P" de 0.005 (redondeado). Estos dos resultados son importantes, pero la comprensión del procedimiento de la prueba de hipótesis tiene una importancia crucial. Concéntrese en comprender cómo funciona el procedimiento de la prueba de hipótesis y aprenda la terminología asociada. Sólo entonces los resultados de la tecnología tendrán sentido.

Statdisk

```
Alternative Hypothesis:
 p > p(hyp)

Sample proportion: 0.5401388
Test Statistic, z: 2.5500
Critical z: 1.6449
P-Value: 0.0054

90% Confidence interval:
0.5143312 < p < 0.5659463
```

Minitab

```
Test of p = 0.5 vs p > 0.5

Sample    X     N  Sample p  95% Lower Bound  Z-Value  P-Value
1       545  1009  0.540139         0.514331     2.55    0.005

Using the normal approximation.
```

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
        1-PropZTest
prop>.5
z=2.549995628
p=.005386238
p̂=.5401387512
n=1009
```

StatCrunch

Hypothesis test results:
p : Proportion of successes
$H_0 : p = 0.5$
$H_A : p > 0.5$

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	545	1009	0.54013875	0.015740714	2.5499956	0.0054

Significancia Las pruebas de hipótesis también se llaman *pruebas de significancia*. En la sección 4-1 usamos las probabilidades para determinar cuándo los resultados muestrales son *significativamente bajos o significativamente altos*. Este capítulo formaliza esos conceptos en un procedimiento unificado que se aplica con frecuencia en diferentes campos. La figura 8-1 de la página siguiente resume los procedimientos utilizados en dos métodos ligeramente diferentes para realizar una prueba de hipótesis formal. Procederemos a realizar una prueba formal de la hipótesis del ejemplo 1: $p > 0.5$. Para probar tal hipótesis n, utilizaremos los datos muestrales de la encuesta citada en el problema del capítulo, con $x = 545$ consumidores que no se sienten cómodos con el envío mediante drones, entre $n = 1009$ consumidores encuestados.

Pasos 1, 2 y 3: uso de la afirmación original para crear una hipótesis nula H_0 y una hipótesis alternativa H_1

El objetivo de los pasos 1, 2 y 3 es identificar la hipótesis nula y la hipótesis alternativa de modo que la prueba de hipótesis formal incluya estos componentes estándar que se usan con frecuencia en muchas disciplinas diferentes. La hipótesis nula incluye el supuesto de trabajo para los propósitos de realización de la prueba de hipótesis.

DEFINICIONES

La **hipótesis nula** (expresada mediante H_0) es una afirmación de que el valor de un parámetro poblacional (por ejemplo la proporción, la media o la desviación estándar) es igual a algún valor declarado.

La **hipótesis alternativa** (expresada por H_1 o H_a o H_A) es una afirmación de que el parámetro tiene un valor que difiere en alguna forma de la hipótesis nula. Para los métodos de este capítulo, la forma simbólica de la hipótesis alternativa debe usar uno de los siguientes símbolos: $<$, $>$, $\neq$.

FIGURA 8-1
Procedimiento para las pruebas de hipótesis

1. Identifique la afirmación
Identifique la afirmación que será probada y exprésela en forma simbólica.

↓

2. Proporcione la forma simbólica
Proporcione la forma simbólica que debe ser verdadera cuando la afirmación original sea falsa.

↓

3. Identifique las hipótesis nula y alternativa
Considere las dos expresiones simbólicas obtenidas hasta ahora:

- La **hipótesis alternativa H_1** es la que *NO* contiene igualdad, por lo que H_1 usa el símbolo $>$, $<$ o $\neq$.
- La **hipótesis nula H_0** es la expresión simbólica de que el parámetro es igual al valor fijo bajo consideración.

↓

4. Seleccione el nivel de significancia
Seleccione el **nivel de significancia α** en función de la gravedad de un error de tipo I. Haga α pequeño si las consecuencias de rechazar una H_0 verdadera son severas.

- Son muy comunes los valores de 0.05 y 0.01.

↓

5. Identifique el estadístico de prueba
Identifique el estadístico de prueba que sea relevante para la prueba y determine su distribución muestral (por ejemplo normal, *t*, ji cuadrada).

Método del valor *P*

6. Encuentre los valores
Encuentre el valor del **estadístico de prueba** y el **valor *P*** (vea la figura 8-3). Trace una gráfica y muestre el estadístico de prueba y el valor *P*.

↓

7. Tome una decisión

- **Rechace H_0** si el valor $P \leq \alpha$.
- **No rechace H_0** si el valor $P > \alpha$.

Método de valor crítico

6. Determine los valores
Encuentre el valor del **estadístico de prueba** y de los **valores críticos**. Trace una gráfica que muestre el estadístico de prueba, los valores críticos y la región crítica.

↓

7. Tome una decisión

- **Rechace H_0** si el estadístico de prueba está en la región crítica.
- **No rechace H_0** si el estadístico de prueba no está en la región crítica.

↓

8. Exprese de nuevo la decisión en términos no técnicos
Exprese de nuevo la decisión previa en términos simples, no técnicos y en relación con la afirmación original.

Método del intervalo de confianza

Elabore un intervalo de confianza con un nivel de confianza seleccionado en la tabla 8-1.

Debido a que una estimación del intervalo de confianza para un parámetro poblacional contiene los valores probables de ese parámetro, rechace una afirmación de que el parámetro poblacional tiene un valor que no está incluido en el intervalo de confianza.

Tabla 8-1 Nivel de confianza para el intervalo de confianza

		Prueba de dos colas	Prueba de una cola
Nivel de significancia para una prueba de hipótesis	0.01	99%	98%
	0.05	95%	90%
	0.10	90%	80%

El término *nulo* se usa para indicar que *no* hay cambio, o que no hay ningún efecto o ninguna diferencia. La prueba de hipótesis se realiza suponiendo que el parámetro es *igual* a un valor específico para que podamos trabajar con una sola distribución que tenga un valor específico.

Ejemplo: La siguiente es una hipótesis nula que involucra una proporción:

$$H_0: p = 0.5$$

Ejemplo: Las siguientes son tres hipótesis alternativas que involucran proporciones:

$$H_1: p > 0.5 \quad H_1: p < 0.5 \quad H_1: p \neq 0.5$$

Dada la afirmación en el ejemplo 1 de que “la mayoría de los consumidores no se sienten cómodos con la entrega mediante drones”, podemos aplicar los pasos 1, 2 y 3 en la figura 8-1 de la siguiente manera.

Paso 1: Identifique la afirmación que se va a probar y exprésela en forma simbólica. Si p expresa la probabilidad de seleccionar un consumidor incómodo con la entrega mediante drones, la afirmación de que “la mayoría están incómodos con la entrega mediante drones” puede expresarse simbólicamente como $p > 0.5$.

Paso 2: Proporcione la forma simbólica que debe ser verdadera cuando la afirmación original sea falsa. Si la afirmación original de $p > 0.5$ es falsa, entonces $p \leq 0.5$ debe ser verdadera.

Paso 3: Esta etapa consta de dos partes: la identificación de la hipótesis alternativa H_1 y la identificación de la hipótesis nula H_0.

- Identificación de H_1: Mediante el uso de las dos expresiones simbólicas $p > 0.5$ y $p \leq 0.5$, la hipótesis alternativa H_1 es la que no contiene igualdad. De las dos expresiones, $p > 0.5$ no contiene igualdad, entonces obtenemos

$$H_1: p > 0.5$$

- Identificación de H_0: La hipótesis nula H_0 es la expresión simbólica de que el parámetro es igual al valor fijo bajo consideración, por lo que obtenemos

$$H_0: p = 0.5$$

Los primeros tres pasos conducen a las hipótesis nula y alternativa:

$$H_0: p = 0.5 \text{ (hipótesis nula)}$$
$$H_1: p > 0.5 \text{ (hipótesis alternativa)}$$

Nota sobre la formación de sus propias afirmaciones (Hipótesis) Si usted está realizando un estudio y desea utilizar una prueba de hipótesis para *respaldar* su afirmación, ésta debe redactarse de modo que se convierta en la hipótesis alternativa (y se pueda expresar utilizando sólo los símbolos $<$, $>$ o $\neq$). No puede sostener una aformación de un parámetro es *igual a* un valor específico.

Paso 4: Seleccione el nivel de significancia α

DEFINICIONES

El **nivel de significancia α** para una prueba de hipótesis es el valor de probabilidad utilizado como punto de corte para determinar cuándo la evidencia muestral es suficientemente *significativa* contra la hipótesis nula. Por su naturaleza, el nivel de significancia α es la probabilidad de rechazar erróneamente la hipótesis nula cuando es verdadera:

Nivel de significancia α = P (rechazar H_0 cuando es verdadera)

El nivel de significancia α es igual al que se presentó en la sección 7-1, donde definimos el “valor crítico”. Las opciones típicas para α son 0.05, 0.01 y 0.10; siendo 0.05 la más común.

En cifras

140,000,000,000,000,000,000 millas: La distancia hasta la que un humano puede ver a simple vista. Lo que usted ve es luz de hace 2.4 millones de años.

Paso 5: Identifique el estadístico relevante para la prueba y determine su distribución muestral (por ejemplo normal, *t* o χ^2)

En la tabla 8-2 se listan los parámetros junto con las distribuciones muestrales correspondientes.

Ejemplo: La afirmación $p > 0.5$ trata sobre la proporción poblacional p; por lo tanto, use la distribución normal, siempre que se cumplan los requisitos. (Con $n = 1009$, $p = 0.5$ y $q = 0.5$ del ejemplo 1, $np \geq 5$ y $nq \geq 5$ son verdaderas).

TABLA 8-2

Parámetro	Distribución muestral	Requisitos	Estadístico de prueba
Proporción p	Normal (z)	$np \geq 5$ y $nq \geq 5$	$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$
Media μ	t	σ desconocida y población distribuida normalmente o σ desconocida y $n > 30$	$t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}}$
Media μ	Normal (z)	σ conocida y población distribuida normalmente o σ conocida y $n > 30$	$z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$
Des. est. σ o varianza σ^2	χ^2	Requisito estricto: población normalmente distribuida	$\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

Paso 6: Encuentre el valor del dato estadístico de prueba, después determine el valor *P* o el(los) valor(es) crítico(s)

DEFINICIÓN

El **dato estadístico de prueba** es un valor utilizado al tomar una decisión sobre la hipótesis nula. Se encuentra al convertir el estadístico muestral (como $\hat{p}$, $\bar{x}$ o s) en una puntuación (como z, t o χ^2) con el supuesto de que la hipótesis nula es verdadera.

En este capítulo usamos los datos estadísticos de prueba listados en la última columna de la tabla 8-2.

Ejemplo: En el ejemplo 1 hay una afirmación sobre la proporción poblacional p, y tenemos $n = 1009$ y $x = 545$, entonces $\hat{p} = x/n = 0.540$. Con la hipótesis nula H_0: $p = 0.5$, estamos trabajando con el supuesto de que $p = 0.5$, y se sigue que $q = 1 - p = 0.5$. Podemos evaluar el dato estadístico de prueba como se muestra a continuación (o bien usando software). El dato estadístico de prueba $z = 2.55$ en cada una de las pantallas presentadas anteriormente es más preciso que el resultado de $z = 2.54$ mostrado a continuación. (Si reemplazamos 0.540 por $545/1009 = 0.54013875$, obtenemos $z = 2.55$).

$$z = \frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} = \frac{0.540 - 0.5}{\sqrt{\dfrac{(0.5)(0.5)}{1009}}} = 2.54$$

Para determinar el valor P y/o el(los) valor(es) crítico(s) se requiere que primero consideremos si la prueba de hipótesis es de dos colas, de cola izquierda o de cola derecha, según se describe a continuación.

Dos colas, cola izquierda, cola derecha

> **DEFINICIÓN**
>
> La **región crítica** (o **región de rechazo**) es el área correspondiente a todos los valores del estadístico de prueba que conducen al rechazo de la hipótesis nula.

Dependiendo de la afirmación que esté siendo probada, la región crítica podría estar en las dos colas extremas, en la cola izquierda o en la cola derecha.

- **Prueba de dos colas:** La región crítica se encuentra en las dos regiones extremas (colas) debajo de la curva (como en la gráfica superior de la figura 8-2).
- **Prueba de cola izquierda:** La región crítica se encuentra en la región extrema (cola) izquierda debajo de la curva (como en la gráfica de en medio de la figura 8-2).
- **Prueba de cola derecha:** La región crítica se encuentra en la región extrema (cola) derecha debajo de la curva (como en la gráfica inferior de la figura 8-2).

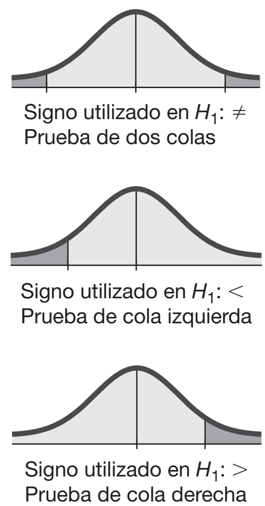


FIGURA 8-2
Región crítica en pruebas de dos colas, de cola izquierda y de cola derecha

> **SUGERENCIA** Observe el símbolo usado en la hipótesis alternativa H_1.
>
> • El símbolo $>$ apunta hacia la derecha y la prueba es de cola derecha.
>
> • El símbolo $<$ apunta hacia la izquierda y la prueba es de cola izquierda.
>
> • El símbolo $\neq$ se usa para una prueba de dos colas.
>
> *Ejemplo:* Con H_0: $p = 0.5$ y H_1: $p > 0.5$, rechazamos la hipótesis nula y aceptamos la hipótesis alternativa sólo si la proporción muestral es mayor que 0.5 en una cantidad significativa, por lo que la prueba de hipótesis en este caso es de *cola derecha*.

Método del valor *P*

Con el **método del valor *P*** para pruebas de hipótesis, tomamos una decisión comparando el valor *P* con el nivel de significancia.

> **DEFINICIÓN**
>
> En una prueba de hipótesis, el **valor *P*** es la probabilidad de obtener un valor del estadístico de prueba que sea *al menos tan extremo* como el estadístico de prueba obtenido a partir de los datos muestrales, suponiendo que la hipótesis nula es verdadera.

Para encontrar el valor *P*, primero determine el área más allá del estadístico de prueba, luego use el procedimiento dado en la figura 8-3 de la página siguiente. Ese procedimiento se puede resumir de la siguiente manera:

- Región crítica en la cola izquierda: Valor P = área a la *izquierda* del estadístico de prueba
- Región crítica en la cola derecha: Valor P = área a la *derecha* del estadístico de prueba
- Región crítica en dos colas: Valor P = *el doble* del área en la cola más allá del estadístico de prueba

Ejemplo: Con base en los datos del problema del capítulo, el estadístico de prueba es $z = 2.55$, y tiene un área en la distribución normal de 0.0054 a su derecha, por lo que una prueba de cola derecha con estadístico de prueba $z = 2.55$ tiene un valor P de 0.0054. Vea las diferentes pantallas de las tecnologías dadas previamente, y observe que cada una de ellas proporciona el mismo valor P de 0.005 después del redondeo.

¡Revista prohíbe los valores P!

El método del valor P para pruebas de hipótesis ha recibido una amplia aceptación en la comunidad de investigadores, pero los editores de la revista *Basic and Applied Social Psychology* adoptaron una postura enfática cuando dijeron que ya no publicarían artículos que incluyeran valores P. En un editorial, David Trafimow y Michael Marks expresaron su creencia de que "la barrera del valor P es demasiado fácil de pasar y, en ocasiones, sirve como excusa para realizar investigaciones de menor calidad". David Trafimow declaró que no sabía qué método estadístico debería reemplazar el uso de los valores P. Muchas reacciones a la prohibición de los valores P reconocieron que aunque pueden ser mal utilizados y malinterpretados, su uso sigue siendo una herramienta de investigación valiosa.

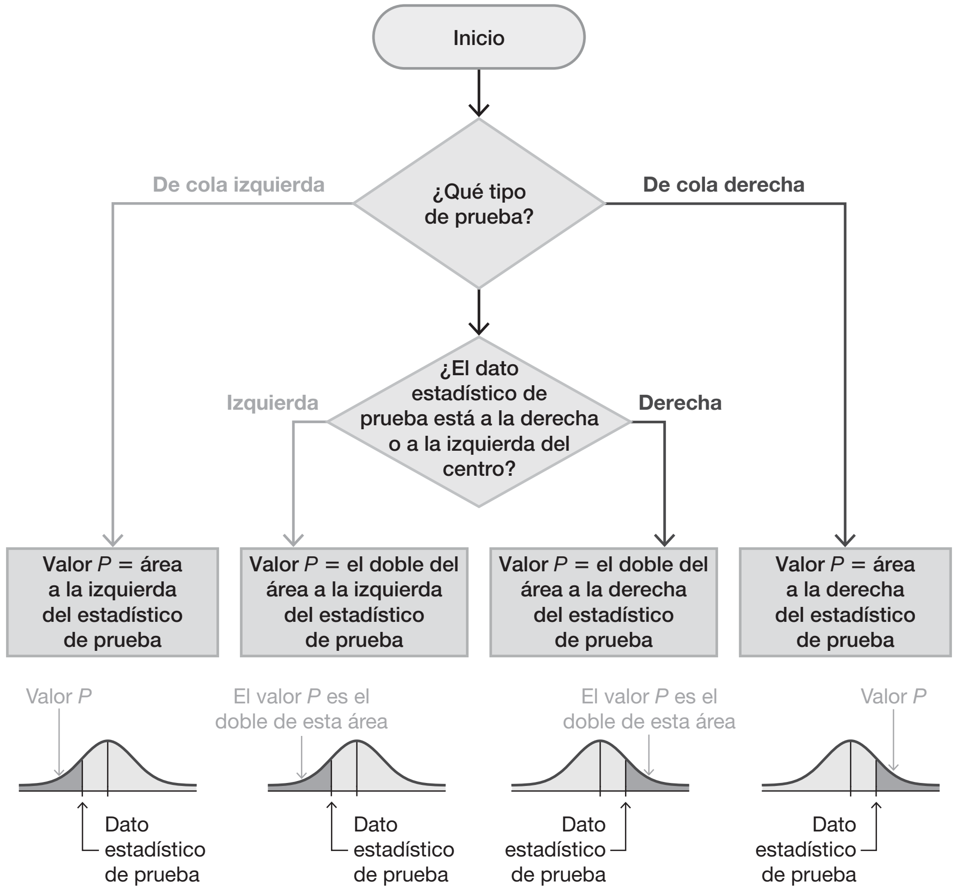


FIGURA 8-3 Determinación de valores P

PRECAUCIÓN No confunda un valor P con el parámetro p o el estadístico $\hat{p}$. Conozca la siguiente notación:

Valor P = probabilidad de que un estadístico de prueba sea al menos tan extremo como el obtenido
$\boldsymbol{p}$ = proporción poblacional
$\boldsymbol{\hat{p}}$ = proporción muestral

Controversia sobre el valor P y las pruebas de hipótesis

El método estándar para probar hipótesis y el uso de valores P tienen una aceptación y un uso muy generalizados, pero no todo mundo está convencido de que tales métodos sean sólidos. Los editores de la revista *Basic and Applied Social Psychology* adoptaron una postura enfática cuando dijeron que ya no publicarían artículos que incluyeran valores P. Dijeron que los valores P son una excusa para investigaciones de baja calidad y que el criterio del valor P es demasiado fácil de pasar. En el pasado, los valores P han sido malinterpretados y mal utilizados, por lo que un análisis estadístico serio e importante no debe basarse únicamente en los resultados del valor P. En cambio, sería conveniente considerar otros aspectos, como los siguientes.

- *Tamaño de muestra*: Las muestras muy grandes podrían dar como resultado valores P pequeños, lo que sugiere que los resultados son significativos cuando en realidad no representan una gran diferencia práctica.
- *Potencia:* La parte 3 de esta sección analiza el concepto de *potencia*, y con frecuencia resulta útil estudiar esta noción como parte de un análisis.

- *Otros factores:* En lugar de basarse sólo en un resultado, como el valor P, generalmente es mejor considerar otros resultados, por ejemplo un intervalo de confianza, salidas de simulaciones, la significancia práctica, el diseño del estudio, la calidad de la muestra, las consecuencias de los errores tipo I y tipo II (analizados en la parte 2 de esta sección) y la repetibilidad de los resultados.

Este capítulo presenta los métodos para pruebas de hipótesis y el uso de los valores P que se emplean actualmente, pero una vez más, se debe destacar que las aplicaciones importantes también deben considerar otros factores, como los listados anteriormente.

Método del valor crítico

Con el **método del valor crítico** (o método tradicional) para probar hipótesis, tomamos una decisión al comparar el estadístico de prueba con el (los) valor(es) crítico(s).

DEFINICIÓN

En una prueba de hipótesis, el (los) **valor(es) crítico(s)** separan la región crítica (donde se rechaza 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 hipótesis nula, la distribución muestral y el nivel de significancia α.

Ejemplo: La región crítica en la figura 8-4 está sombreada en gris oscuro. La figura 8-4 muestra que, con un nivel de significancia de $\alpha = 0.05$, el valor crítico es $z = 1.645$.

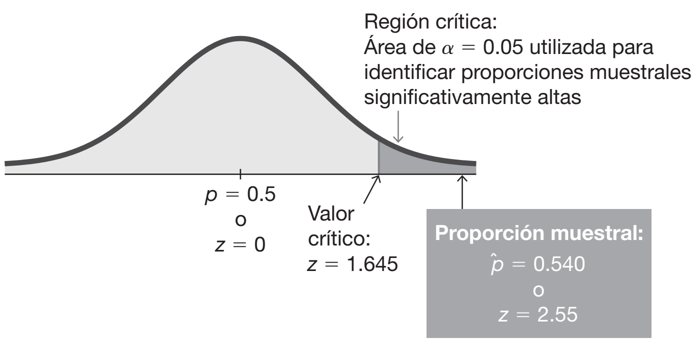


FIGURA 8-4 **Región crítica, valor crítico y estadístico de prueba**

Paso 7: Tome la decisión de rechazar o no rechazar H_0

Criterios de decisión para el método del valor *P*:

- Si el valor $P \leq \alpha$, rechace H_0. ("Si P es bajo, la nula debe irse").
- Si el valor $P > \alpha$, no rechace H_0.

Ejemplo: Con un nivel de significancia $\alpha = 0.05$ y un valor $P = 0.005$, tenemos la condición de que el valor $P \leq \alpha$, entonces rechace H_0. Recuerde, el valor P es la probabilidad de tener un resultado muestral al menos tan extremo como el obtenido, por lo que si el valor P es bajo (menor o igual a α), el estadístico muestral es significativamente bajo o significativamente alto.

Criterios de decisión para el método del valor crítico:

- Si el estadístico de prueba está en la región crítica, rechace H_0.
- Si el dato estadístico de prueba no está en la región crítica, no rechace H_0.

Ejemplo: Con el dato estadístico de prueba $z = 2.55$ y la región crítica desde $z = 1.645$ hasta el infinito, el dato estadístico de prueba cae dentro de la región crítica, por lo tanto, rechace H_0.

Paso 8: Reformule la decisión utilizando términos simples y no técnicos

Sin utilizar términos técnicos que la mayoría de las personas no entienden, establezca una conclusión final sobre la afirmación original con una redacción que pueda ser comprendida por quien no tiene conocimiento de los procedimientos estadísticos.

Ejemplo: Hay suficiente evidencia para respaldar la afirmación de que la mayoría de los consumidores no se sienten cómodos con las entregas mediante drones.

Redacción de la conclusión final Como una ayuda para redactar la conclusión final, consulte la tabla 8-3, que lista las cuatro circunstancias posibles y sus conclusiones correspondientes. Observe que sólo el primer caso conduce a una redacción que indica el *respaldo* a la conclusión original. Si desea respaldar alguna afirmación, indíquelo de modo que se convierta en la hipótesis alternativa, y luego espere que la hipótesis nula sea rechazada.

TABLA 8-3 Redacción de la conclusión final

Condición	Conclusión
La afirmación original no incluye igualdad, y se rechaza H_0.	"Hay suficiente evidencia para *respaldar* la afirmación de que... (afirmación original)".
La afirmación original no incluye igualdad, y no se rechaza H_0.	"No hay evidencia suficiente para respaldar la afirmación de que... (afirmación original)".
La afirmación original incluye igualdad, y se rechaza H_0.	"Hay evidencia suficiente para justificar el *rechazo* de la afirmación de que... (afirmación original)".
La afirmación original incluye igualdad, y no se rechaza H_0.	"No hay evidencia suficiente para justificar el rechazo de la afirmación de que... (afirmación original)".

¿Aceptar o no rechazar? Se recomienda decir que "no se rechaza la hipótesis nula" en vez de decir que "se acepta la hipótesis nula". El término *aceptar* es engañoso, porque implica incorrectamente que la hipótesis nula ha sido probada, pero nunca podemos probar una hipótesis nula. La frase "no se rechaza" dice más correctamente que la evidencia disponible no es suficientemente fuerte como para justificar el rechazo de la hipótesis nula.

Negativos múltiples Las conclusiones finales pueden incluir hasta tres términos negativos. (*Ejemplo:* "*No* hay evidencia suficiente para justificar el *rechazo* de la afirmación de que *no* hay diferencia entre 0.5 y la proporción poblacional"). Para conclusiones tan confusas, es mejor volver a expresarlas para que sean comprensibles. En lugar de decir que "no hay evidencia suficiente para justificar el rechazo de la afirmación de que no hay diferencia entre 0.5 y la proporción poblacional", un mejor enunciado sería este: "Hasta que se obtenga evidencia más sólida, se continúa asumiendo que la proporción poblacional es igual a 0.5".

PRECAUCIÓN Nunca concluya una prueba de hipótesis con una afirmación de "se rechaza la hipótesis nula" o "no se rechaza la hipótesis nula". Siempre dé sentido a la conclusión con un enunciado que use una redacción no técnica y simple, que aborde la afirmación original.

Intervalos de confianza para pruebas de hipótesis

En esta sección, hemos descrito los componentes individuales utilizados en una prueba de hipótesis, pero las siguientes secciones combinarán esos componentes en procedimientos completos. Podemos probar hipótesis sobre los parámetros poblacionales utilizando el método del valor P o el método del valor crítico resumido en la figura 8-1, o bien podemos utilizar intervalos de confianza.

Una estimación de intervalo de confianza para un parámetro poblacional contiene los valores probables de ese parámetro. Si un intervalo de confianza no incluye un valor declarado de un parámetro poblacional, rechace esa afirmación. Para las pruebas de hipótesis de dos colas, elabore 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, elabore un intervalo de confianza con un nivel de confianza de $1 - 2\alpha$. (Vea la tabla 8-1 en la página 360 para conocer los casos comunes). (Para una prueba de cola izquierda o una prueba de cola derecha, también podríamos usar un intervalo de confianza unilateral; consulte el ejercicio 40 de la sección 7-1). Después de elaborar el intervalo de confianza, use este criterio:

Una estimación del intervalo de confianza para un parámetro poblacional contiene los valores probables de ese parámetro. Por lo tanto, deberíamos rechazar una afirmación de que el parámetro poblacional tiene un valor que no está incluido en el intervalo de confianza.

Métodos equivalentes

En algunos casos, una conclusión basada en un intervalo de confianza puede ser diferente de una conclusión basada en una prueba de hipótesis. El método del valor P y el método del valor crítico son equivalentes en el sentido de que siempre conducen a la misma conclusión. La siguiente tabla muestra que, para los métodos incluidos en este capítulo, una estimación del intervalo de confianza para una proporción puede llevar a una conclusión diferente de la obtenida con una prueba de hipótesis.

Parámetro	¿Es un intervalo de confianza *equivalente* a una prueba de hipótesis en el sentido de que siempre conducen a la misma conclusión?
Proporción	No
Media	Sí
Desviación estándar o varianza	Sí

PARTE 2 Errores tipo I y tipo II

Cuando se prueba una hipótesis nula, llegamos a la conclusión de rechazarla o no rechazarla. Nuestras conclusiones a veces son correctas y en ocasiones son incorrectas (incluso si aplicamos todos los procedimientos de manera adecuada). La tabla 8-4 en la página siguiente incluye dos tipos de errores y los distinguimos llamándolos errores tipo I y tipo II, como se describe a continuación:

- **Error tipo I:** El error de rechazar la hipótesis nula cuando en realidad es verdadera. El símbolo α (alfa) se usa para representar la probabilidad de un error tipo I.

$$\boldsymbol{\alpha = P(\textbf{error tipo I}) = P(\textbf{rechazar } H_0 \textbf{ cuando } H_0 \textbf{ es verdadera})}$$

- **Error tipo II:** El error de no rechazar la hipótesis nula cuando en realidad es falsa. El símbolo β (beta) se usa para representar la probabilidad de un error tipo II.

$$\boldsymbol{\beta = P(\textbf{error tipo II}) = P(\textbf{no rechazar } H_0 \textbf{ cuando } H_0 \textbf{ es falsa})}$$

SUGERENCIA MNEMOTÉCNICA PARA ERRORES TIPO I Y TIPO II Recuerde “renueva Nora no fue”, y use las consonantes de esas palabras para recordar que un error tipo I es RNV: Rechazar Nula Verdadera (hipótesis), y un error tipo II es NRNF: No Rechazar Nula Falsa (hipótesis).

La aspirina no es útil para las personas de los signos Géminis y Libra

El médico Richard Peto envió un artículo a *Lancet*, una revista científica británica. El artículo afirmaba que los pacientes tenían mayores probabilidades de sobrevivir a un ataque cardiaco si se les trataba con aspirina pocas horas después de sufrir el infarto. Los editores de *Lancet* pidieron a Peto que separara sus resultados en subgrupos para ver si la recuperación era mejor o peor para ciertos grupos, como hombres y mujeres. Peto consideró que le pedían utilizar demasiados grupos, pero los editores insistieron. Peto accedió, pero sustentó sus objeciones mostrando que, cuando los pacientes se clasificaban según el signo del zodiaco, la aspirina resultaba inútil para los pacientes cardiacos de los signos Géminis y Libra, mientras que salvaba la vida a los nacidos bajo cualquier otro signo. Esto demuestra que cuando se realizan múltiples pruebas de hipótesis con muchos grupos distintos, hay una gran probabilidad de obtener algunos resultados incorrectos.

TABLA 8-4 Errores tipo I y tipo II

		Estado de la naturaleza verdadero	
		La hipótesis nula es verdadera	**La hipótesis nula es falsa**
Conclusión preliminar	**Rechazar H_0**	**Error tipo I:** Rechazar una H_0 verdadera. P (error tipo I) $= \alpha$	Decisión correcta
	No se puede rechazar H_0	Decisión correcta	**Error tipo II:** No se rechaza una H_0 falsa. P (error tipo II) $= \beta$

SUGERENCIA PARA DESCRIBIR ERRORES TIPO I Y TIPO II Las descripciones de un error de tipo I y un error de tipo II hacen referencia a que la *hipótesis nula* es verdadera o falsa, pero al redactar una afirmación que representa un error tipo I o un error tipo II, *asegúrese de que la conclusión aborde la afirmación original* (que puede ser o no ser la hipótesis nula). Vea el Ejemplo 2.

EJEMPLO 2 Descripción de errores tipo I y tipo II

Considere la afirmación de que un procedimiento médico diseñado para aumentar la probabilidad de tener una niña es efectivo, de modo que la probabilidad de tener una niña es $p >$ 0.5. Dadas las siguientes hipótesis nula y alternativa, escriba afirmaciones que describan (a) un error tipo I, y (b) un error tipo II.

H_0: $p = 0.5$

H_1: $p > 0.5$ (afirmación original que se abordará en la conclusión final)

SOLUCIÓN

a. Error tipo I: Un error tipo I es el error de rechazar una hipótesis nula verdadera, por lo que el siguiente es un error tipo I: en realidad $p = 0.5$, pero la evidencia muestral nos lleva a concluir que $p > 0.5$. (En este caso, un error tipo I es concluir que el procedimiento médico es efectivo cuando en realidad no tiene ningún efecto).

b. Error tipo II: Un error tipo II es el error de no rechazar la hipótesis nula cuando es falsa, por lo que el siguiente es un error tipo II: en realidad, $p > 0.5$, pero no logramos respaldar esa conclusión. (En este caso, un error tipo II es concluir que el procedimiento médico no tiene ningún efecto, cuando en realidad es efectivo para aumentar la probabilidad de tener una niña).

SU TURNO Resuelva el ejercicio 31 "Errores tipo I y tipo II".

Control de errores tipo I y tipo II El paso 4 de nuestro procedimiento estándar para pruebas de hipótesis es seleccionar un nivel de significancia α (como 0.05), que es la probabilidad de un error tipo I. Los valores de α, β y el tamaño de muestra n están relacionados, por lo que si elige cualesquiera dos, el tercero se determina automáticamente (aunque β no se puede establecer hasta que se haya especificado un valor alternativo del parámetro poblacional junto con α y n). Una práctica común consiste en seleccionar el nivel de significancia α y luego seleccionar un tamaño de muestra que sea práctico, de modo que se determine el valor de β. Generalmente, trate de usar el mayor α que pueda tolerar, pero para los errores tipo I con consecuencias más serias, seleccione valores más pequeños de α. Luego elija un tamaño de muestra n tan grande como sea razonable, con base en consideraciones de tiempo, costo y otros factores relevantes. Otra práctica común es seleccionar α y β para que el tamaño de muestra requerido n se determine automáticamente. (Consulte el ejemplo 4 "Determinación del tamaño de muestra requerido para lograr una potencia del 80%" en la parte 3 de esta sección).

PARTE 3 Potencia de una prueba de hipótesis

Usamos β para expresar la probabilidad de no rechazar una hipótesis nula falsa, entonces P(error tipo II) $= \beta$. De ahí que $1 - \beta$ es la probabilidad de rechazar una hipótesis nula falsa, entonces $1 - \beta$ es una probabilidad que representa una medida de la eficacia de una prueba de hipótesis.

DEFINICIÓN

La **potencia** de una prueba de hipótesis es la probabilidad $1 - \beta$ de rechazar una hipótesis nula falsa. El valor de la potencia se calcula utilizando un nivel de significancia particular α y un *valor particular* del parámetro poblacional que es una alternativa al valor supuesto como verdadero en la hipótesis nula.

Debido a que la determinación de la potencia requiere un valor particular que sea una alternativa al valor supuesto en la hipótesis nula, una prueba de hipótesis puede tener diferentes valores de potencia, dependiendo de los valores particulares del parámetro poblacional elegidos como alternativas a la hipótesis nula.

EJEMPLO 3 Potencia de una prueba de hipótesis

Considere los siguientes resultados preliminares del método XSORT para la selección de género: hubo 13 niñas entre los 14 bebés nacidos de parejas que usan el método XSORT. Si queremos probar la hipótesis de que las niñas son más probables ($p > 0.5$) con el método XSORT, tenemos las siguientes hipótesis nula y alternativa:

$$H_0: p = 0.5 \quad H_1: p > 0.5$$

Usemos un nivel de significancia de $\alpha = 0.05$. Además de todos los componentes de prueba dados, la determinación de la potencia requiere que seleccionemos un valor particular de p que sea una alternativa al valor supuesto en la hipótesis nula $H_0: p = 0.5$. Encuentre los valores de potencia correspondientes a los siguientes valores alternativos de p: 0.6, 0.7, 0.8 y 0.9.

SOLUCIÓN

Los valores de potencia en la siguiente tabla se encontraron utilizando Minitab y se usan cálculos exactos en vez de una aproximación normal a la distribución binomial.

Valor alternativo específico de p	β	Potencia de la prueba $= 1 - \beta$
0.6	0.820	0.180
0.7	0.564	0.436
0.8	0.227	0.773
0.9	0.012	0.988

INTERPRETACIÓN

Con base en los valores de potencia listados anteriormente, vemos que esta prueba de hipótesis tiene una potencia de 0.180 (o 18.0%) de rechazar $H_0: p = 0.5$ cuando la proporción poblacional p es en realidad 0.6. Es decir, si la proporción real de la población es en realidad 0.6, hay un 18.0% de posibilidades de llegar a la conclusión correcta al rechazar la hipótesis nula falsa de que $p = 0.5$. Esta baja potencia de 18.0% no es muy buena.

continúa

Proceso de aprobación de un fármaco

Lograr la aprobación de la Administración de Drogas y Alimentos de Estados Unidos para un fármaco nuevo es costoso y requiere de mucho tiempo. Las diferentes etapas para lograr la aprobación de un nuevo fármaco son:

- **Estudio de fase I:** Se prueba la seguridad del fármaco con un grupo pequeño de voluntarios (de 20 a 100).
- **Fase II:** Se prueba la eficacia del fármaco en ensayos aleatorios con un grupo más grande de sujetos (entre 100 y 300). Esta fase a menudo incluye sujetos asignados al azar a un grupo de tratamiento o a un grupo de placebo.
- **Fase III:** La meta consiste en comprender mejor la eficacia del fármaco, así como sus efectos adversos. En la fase III generalmente participan de 1,000 a 3,000 sujetos, y suele requerir varios años de pruebas.

Lisa Gibbs escribió en la revista *Money*: "la industria (farmacéutica) afirma que por cada 5,000 tratamientos que se someten a prueba, sólo cinco llegan a los ensayos clínicos y sólo 1 termina en las farmacias". Las estimaciones del costo total varían desde $40 millones hasta $1,500 millones.

Hay una probabilidad de 0.436 de rechazar $p = 0.5$ cuando el verdadero valor de p es en realidad 0.7. Tiene sentido que esta prueba sea más efectiva al rechazar la afirmación de $p = 0.5$ cuando la proporción poblacional es en realidad 0.7 que cuando dicha proporción es realmente 0.6. (Al identificar animales que se supone que son caballos, hay una mayor probabilidad de rechazar un elefante como caballo —por la mayor diferencia— que rechazar una mula como un caballo). En general, el aumento de la diferencia entre el valor supuesto y el valor real del parámetro produce un incremento en la potencia, como se muestra en la tabla de la página anterior.

SU TURNO Resuelva el ejercicio 34 "Cálculo de la potencia".

Debido a que los cálculos de la potencia son bastante complicados, se recomienda el uso de la tecnología. (En esta sección, sólo los ejercicios 33, 34 y 35 implican potencias).

Potencia y diseño de experimentos

Del mismo modo que 0.05 es una opción común para el nivel de significancia, una potencia de al menos 0.80 es un requisito común para determinar que una prueba de hipótesis es efectiva. (Algunos profesionales de la estadística argumentan que la potencia debería ser mayor, por ejemplo 0.85 o 0.90). Al diseñar un experimento, podríamos considerar cuál diferencia entre el valor declarado de un parámetro y su valor real es un tamaño importante de diferencia. Si se prueba la efectividad del método de selección de género XSORT, un cambio en la proporción de niñas de 0.5 a 0.501 no es muy importante, mientras que un cambio en esa proporción de 0.5 a 0.9 sería muy importante. Tales magnitudes de diferencias afectan la potencia. Al diseñar un experimento, con frecuencia puede usarse el objetivo de tener un valor de potencia de al menos 0.80 para determinar el tamaño de muestra mínimo requerido, como en el siguiente ejemplo.

EJEMPLO 4 Determinación del tamaño de muestra requerido para lograr una potencia del 80%

La siguiente es una afirmación similar a la de un artículo de la *Journal of the American Medical Association*: "El diseño del ensayo supuso que con un nivel de significancia de 0.05, se necesitarían 153 sujetos seleccionados aleatoriamente para lograr un 80% de potencia en la detección de una reducción en la tasa de enfermedades cardiacas coronarias de 0.5 a 0.4". Con base en esa afirmación, sabemos lo siguiente:

- Antes de realizar el experimento, los investigadores seleccionaron un nivel de significancia de 0.05 y una potencia de al menos 0.80.
- Los investigadores decidieron que una reducción en la proporción de enfermedades coronarias de 0.5 a 0.4 es una diferencia importante que quisieran detectar (mediante el rechazo correcto de una hipótesis nula falsa).
- Con un nivel de significancia de 0.05, potencia de 0.80 y proporción alternativa de 0.4, se usó tecnología como Minitab para encontrar que el tamaño de muestra mínimo requerido es 153.

Los investigadores pueden proceder a obtener una muestra de al menos 153 sujetos seleccionados al azar. Debido a factores como las tasas de deserción, es probable que los investigadores necesiten algo más de 153 sujetos. (Vea el ejercicio 35).

SU TURNO Resuelva el ejercicio 35 "Determinación del tamaño de muestra para lograr una potencia".

8-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Vitamina C y aspirina Una botella tiene una etiqueta que indica que contiene pastillas de Spring Valley con 500 mg de vitamina C, y otra tiene una etiqueta que indica que contiene tabletas de Bayer con 325 mg de aspirina. Al evaluar las afirmaciones sobre el contenido medio de las píldoras, ¿qué tendría implicaciones más serias: ¿el rechazo de la afirmación sobre la vitamina C de Spring Valley o el rechazo de la afirmación sobre la aspirina de Bayer? ¿Es aconsejable utilizar el mismo nivel de significancia para las pruebas de hipótesis sobre la cantidad media de vitamina C y la cantidad media de aspirina?

2. Estimaciones y pruebas de hipótesis El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye las temperaturas corporales de una muestra. Podríamos usar métodos del capítulo 7 para hacer una estimación, o podríamos usar esos valores para probar la creencia común de que la temperatura corporal media es de 98.6 °F. ¿Cuál es la diferencia entre la estimación y la prueba de hipótesis?

3. Estatura media de los hombres Se debe llevar a cabo una prueba de hipótesis formal utilizando la afirmación de que la estatura media de los hombres es igual a 174.1 cm.

a. ¿Cuál es la hipótesis nula y cómo se expresa?

b. ¿Cuál es la hipótesis alternativa y cómo se expresa?

c. ¿Cuáles son las posibles conclusiones que pueden obtenerse acerca de la hipótesis nula?

d. ¿Es posible concluir que "hay evidencia suficiente para respaldar la afirmación de que la estatura media de los hombres es igual a 174.1 cm"?

4. Interpretación del valor *P* El método Ericsson es una de varias técnicas de las que se afirma pueden aumentar la probabilidad de tener una niña. En un ensayo clínico, los resultados podrían analizarse con una prueba de hipótesis formal con la hipótesis alternativa de $p > 0.5$, que corresponde a la afirmación de que el método aumenta la probabilidad de tener una niña, por lo que la proporción de niñas es mayor a 0.5. Si tuviera interés en establecer el éxito del método, cuál de los siguientes valores *P* preferiría: 0.999, 0.5, 0.95, 0.05, 0.01, 0.001? ¿Por qué?

Identificación de H_0 y H_1. ***En los ejercicios 5 a 8, haga lo siguiente:***

a. Exprese la afirmación original en forma simbólica.

b. Identifique las hipótesis nula y alternativa.

5. Afirmación sobre **datos en línea:** La mayoría de los adultos borrarían toda su información personal en línea si pudieran. Una encuesta de GFI Software aplicada a 565 adultos seleccionados al azar mostró que 59% de ellos borraría toda su información personal en línea si pudieran.

6. Afirmación sobre el **teléfono celular:** Menos de 95% de los adultos tienen un teléfono celular. En una encuesta marista aplicada a 1128 adultos, 87% dijo que tienen un teléfono celular.

7. Afirmación sobre el **pulso:** El pulso medio (en latidos por minuto, o lpm) de los hombres adultos es igual a 69 lpm. Para la muestra aleatoria de 153 varones adultos en el conjunto de datos 1 "Datos corporales" del apéndice B, el pulso medio es de 69.6 lpm y la desviación estándar es de 11.3 lpm.

8. Afirmación sobre el **pulso:** La desviación estándar del pulso en los hombres adultos es mayor que 11 lpm. Para la muestra aleatoria de 153 hombres adultos en el conjunto de datos 1 "Datos corporales" del apéndice B, los pulsos tienen una desviación estándar de 11.3 lpm.

Conclusiones. ***En los ejercicios 9 a 12, refiérase al ejercicio identificado. Haga estimaciones subjetivas para decidir si los resultados son significativamente bajos o significativamente altos, luego formule una conclusión sobre la afirmación original. Por ejemplo, si la afirmación es que una moneda favorece las caras y los resultados de la muestra consisten en 11 caras en 20 lanzamientos, concluya que no hay evidencia suficiente para respaldar la afirmación de que la moneda favorece a las caras (porque es fácil obtener por casualidad 11 caras en 20 lanzamientos utilizando una moneda justa).***

9. Ejercicio 5 "Datos en línea"

10. Ejercicio 6 "Teléfono celular"

11. Ejercicio 7 "Pulsos"

12. Ejercicio 8 "Pulsos"

Estadísticos de prueba. ***En los ejercicios 13 a 16, refiérase al ejercicio identificado y encuentre el valor del estadístico de prueba. (Consulte la tabla 8-2 en la página 362 a fin de seleccionar la expresión correcta para evaluar el estadístico de prueba).***

13. Ejercicio 5 "Datos en línea"

14. Ejercicio 6 "Teléfono celular"

15. Ejercicio 7 "Pulsos"

16. Ejercicio 8 "Pulsos"

Valores *P* ***En los ejercicios 17 a 20, haga lo siguiente:***

a. Identifique la prueba de hipótesis como de dos colas, de cola izquierda o de cola derecha.

b. Encuentre el valor *P* (vea la figura 8-3 en la página 364).

c. Con un nivel de significancia $\alpha = 0.05$, ¿deberíamos rechazar H_0 o no deberíamos rechazarla?

17. El dato estadístico de prueba $z = 1.00$ se obtiene al probar la hipótesis de que $p > 0.3$.

18. El dato estadístico de prueba $z = -2.50$ se obtiene al probar la hipótesis de que $p < 0.75$.

19. El dato estadístico de prueba $z = 2.01$ se obtiene al probar la hipótesis de que $p \neq 0.345$.

20. El dato estadístico de prueba $z = -1.94$ se obtiene al probar la hipótesis de que $p = 3/8$.

Valores críticos. ***En los ejercicios 21 a 24, refiérase a la información del ejercicio dado y haga lo siguiente.***

a. Encuentre los valores críticos.

b. Con un nivel de significancia de $\alpha = 0.05$, ¿deberíamos rechazar H_0 o no deberíamos rechazarla?

21. Ejercicio 17

22. Ejercicio 18

23. Ejercicio 19

24. Ejercicio 20

Conclusiones finales. ***En los ejercicios 25 a 28, use un nivel de significancia $\alpha = 0.05$ y utilice la información dada para lo siguiente:***

a. Establezca una conclusión sobre la hipótesis nula. (Rechazar H_0 o no rechazarla).

b. Sin emplear términos o símbolos técnicos, establezca una conclusión final que aborde la afirmación original.

25. Afirmación original: Más de 58% de los adultos borrarían toda su información personal en línea si pudieran. La prueba de hipótesis da como resultado un valor *P* de 0.3257.

26. Afirmación original: Menos de 90% de los adultos tienen un teléfono celular. La prueba de hipótesis da como resultado un valor *P* de 0.0003.

27. Afirmación original: El pulso medio (en latidos por minuto) de los hombres adultos es de 72 lpm. La prueba de hipótesis da como resultado un valor *P* de 0.0095.

28. Afirmación original: La desviación estándar de los pulsos de los hombres adultos es mayor que 11 ppm. La prueba de hipótesis da como resultado un valor *P* de 0.3045.

Errores tipo I y tipo II. ***En los ejercicios 29 a 32, proporcione afirmaciones que identifiquen el error tipo I y el error tipo II que corresponden a la afirmación dada. (Aunque las conclusiones suelen expresarse en forma verbal, aquí las respuestas pueden expresarse con afirmaciones que incluyan expresiones simbólicas como $p = 0.1$).***

29. La proporción de personas que escriben con la mano izquierda es igual a 0.1.

30. La proporción de personas con ojos azules es igual a 0.35.

31. La proporción de adultos que usan Internet es mayor que 0.87.

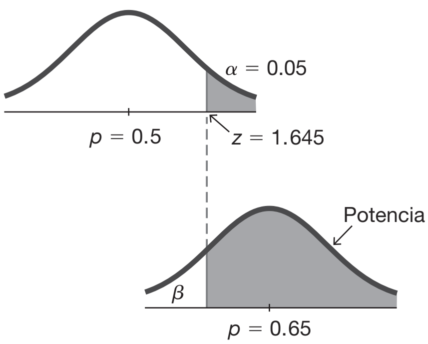
32. La proporción de personas que no requieren corrección de la vista es inferior a 0.25.

8-1 Más allá de lo básico

33. Interpretación de la potencia Las tabletas de Chantix (vareniclina) se usan como una ayuda para que las personas dejen de fumar. En un ensayo clínico, 129 sujetos fueron tratados con Chantix dos veces al día durante 12 semanas, y 16 sujetos experimentaron dolor abdominal (según los datos de Pfizer, Inc.). Si alguien establece que más de 8% de los usuarios de Chantix experimentan dolor abdominal, esa afirmación se respalda con una prueba de hipótesis realizada con un nivel de significancia de 0.05. Si se toma 0.18 como valor alternativo de p, la potencia de la prueba es 0.96. Interprete este valor de la potencia de la prueba.

34. Cálculo de la potencia Considere una prueba de hipótesis de la afirmación de que el método de selección de género Ericsson es eficaz para aumentar la probabilidad de tener una niña, por lo que la afirmación es $p > 0.5$. Suponga que se usa un nivel de significancia de $\alpha = 0.05$, y que la muestra es una muestra aleatoria simple de tamaño $n = 64$.

a. Si se supone que la proporción poblacional real es 0.65, encuentre la potencia de la prueba, que es la probabilidad de rechazar la hipótesis nula cuando es falsa. (*Sugerencia:* Con un nivel de significancia de 0.05, el valor crítico es $z = 1.645$, por lo que cualquier dato estadístico de prueba en la cola derecha de la gráfica superior adjunta estará en la región de rechazo donde se admite la afirmación. Encuentre la proporción muestral $\hat{p}$ en la gráfica superior, y úsela para determinar la potencia que se muestra en la gráfica inferior).



b. Explique por qué la región sombreada en gris de la gráfica inferior representa la potencia de la prueba.

35. Determinación del tamaño de muestra para lograr una potencia Algunos investigadores planean realizar una prueba de un método para la selección de género. Piensan usar la hipótesis alternativa H_1: $p > 0.5$ y un nivel de significancia $\alpha = 0.05$. Encuentre el tamaño de muestra requerido para lograr al menos 80% de potencia al detectar un aumento en p de 0.50 a 0.55. (Este es un ejercicio muy difícil. *Sugerencia:* Vea el ejercicio 34).

8-2 Prueba de una hipótesis respecto a una proporción

Concepto clave En esta sección se describe un procedimiento completo para probar una hipótesis hecha sobre una proporción poblacional p. Ilustramos las pruebas de hipótesis con el método del valor P, el método del valor crítico y el uso de intervalos de confianza. Los métodos de esta sección pueden usarse con afirmaciones sobre proporciones poblacionales, probabilidades o los equivalentes decimales de los porcentajes.

Existen diferentes métodos para probar una hipótesis sobre una proporción poblacional. La parte 1 de esta sección se basa en el uso de una aproximación normal a una distribución binomial, y este método sirve como una introducción a los conceptos básicos, pero no es un método usado por los estadísticos profesionales. En la parte 2 se analizan otros métodos que pueden requerir el uso de la tecnología.

PARTE 1 Método de la aproximación normal

El siguiente recuadro incluye los elementos clave utilizados para probar una hipótesis sobre una proporción poblacional mediante el uso de una distribución normal como una aproximación a una distribución binomial.

El dato estadístico de prueba anterior no incluye una corrección para la continuidad (como la descrita en la sección 6-6) porque su efecto tiende a ser muy pequeño en las muestras grandes.

ELEMENTOS CLAVE

Prueba de una hipótesis sobre una proporción poblacional (método de aproximación normal)

Objetivo

Realizar una prueba de hipótesis formal de una afirmación sobre una proporción poblacional p.

Notación

n = tamaño de muestra o número de ensayos

$\hat{p} = \frac{x}{n}$ (proporción *muestral*)

p = proporción poblacional (p es el valor usado en el enunciado de la hipótesis nula)

$q = 1 - p$

Requisitos

1. Las observaciones muestrales son una muestra aleatoria simple.
2. Se satisfacen las condiciones para una *distribución binomial*:
- Hay un número fijo de ensayos.
- Los ensayos son independientes.
- Cada ensayo tiene dos categorías: "éxito" y "fracaso".
- La probabilidad de éxito es la misma en todos los ensayos.

3. Se satisfacen las condiciones $np \geq 5$ y $nq \geq 5$, por lo que **la distribución binomial de las proporciones muestrales se puede aproximar mediante una distribución normal con** $\mu = np$ y $\sigma = \sqrt{npq}$ (como se describió en la sección 6-6). Note que p se usa aquí como la proporción *supuesta* utilizada en la afirmación, no como la proporción muestral $\hat{p}$.

Dato estadístico de prueba para probar una hipótesis sobre una proporción

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Valores P: Los valores P son proporcionados automáticamente por la tecnología. Si no dispone de una tecnología, use la distribución normal estándar (tabla A-2) y consulte la figura 8-3 en la página 364.

Valores críticos: Utilice la distribución normal estándar (tabla A-2).

Métodos equivalentes

Cuando se prueban hipótesis sobre proporciones, el método del intervalo de confianza no es equivalente a los métodos del valor P y del valor crítico, por lo que el método del intervalo de confianza podría dar lugar a una conclusión diferente. (Tanto el método del valor P como el método del valor crítico utilizan la misma desviación estándar basada en la *proporción p declarada*, por lo que son equivalentes entre sí; pero el método del intervalo de confianza utiliza una desviación estándar estimada que se basa en la *proporción muestral*). *Recomendación:* Use un intervalo de confianza para *estimar* una proporción poblacional, pero use el método de valor P o el método del valor crítico para *probar una hipótesis* sobre una proporción. Vea el ejercicio 34.

Afirmación: La mayoría de los consumidores se sienten incómodos con las entregas mediante drones

El problema del capítulo citó una encuesta de Pitney Bowes en la que se preguntó a 1009 consumidores si se sentían cómodos con que los drones entregaran sus compras, y 54% (o 545) de ellos respondieron “no”. Use estos resultados para probar la hipótesis de que la mayoría de los consumidores se sienten incómodos con las entregas mediante drones. Interpretamos que “la mayoría” significa “más de la mitad” o “mayor que 0.5”.

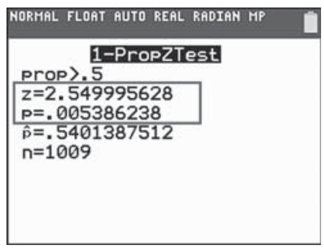
VERIFICACIÓN DE REQUISITOS Primero verificamos los tres requisitos.

1. Los 1009 consumidores se seleccionan al azar.
2. Hay un número fijo (1009) de ensayos independientes con dos categorías (el sujeto no se siente cómodo con las entregas mediante drones o se siente cómodo con esas entregas).
3. Los requisitos $np \geq 5$ y $nq \geq 5$ se satisfacen con $n = 1009$, $p = 0.5$ y $q = 0.5$. [El valor de $p = 0.5$ proviene de la afirmación. Obtenemos $np = (1009)(0.5) = 504.5$, que es mayor o igual que 5; y obtenemos $nq = (1009)(0.5) = 504.5$, que también es mayor o igual que 5].

Se satisfacen los tres requisitos.

Solución: Método del valor *P*

TI83/84 Plus



Tecnología: Los programas de computadora y las calculadoras suelen proporcionar un valor P, por lo que se utiliza el método del valor P. Consulte los resultados adjuntos de la calculadora TI-83/84 Plus que muestran la hipótesis alternativa “prop $>$ 0.5”, el dato estadístico de prueba $z = 2.55$ (redondeado) y el valor P de 0.0054 (redondeado).

Tabla A-2: Si no dispone de una tecnología, la figura 8-1 en la página 360 de la sección anterior lista los pasos para usar el método del valor P. Con base en los pasos de la figura 8-1, podemos probar la hipótesis de la siguiente manera.

Paso 1: La afirmación original es que la mayoría de los consumidores no se sienten cómodos con las entregas mediante drones, y esa afirmación se puede expresar en forma simbólica como $p > 0.5$.

Paso 2: El opuesto a la afirmación 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 de la siguiente manera:

$$H_0: p = 0.5$$

$$H_1: p > 0.5 \text{ (afirmación original)}$$

continúa

¿0.05 es una mala elección?

El valor de 0.05 es una opción muy común para servir como el punto de corte que separa los resultados considerados significativos de aquellos que no lo son. El escritor científico John Timmer escribió en *Ars Technica* que algunos problemas con las conclusiones científicas se deben al hecho de que, en ocasiones, la estadística es débil debido al uso común de 0.05 como nivel de significancia. Proporcionó ejemplos de experimentos de física de partículas y genética en los que los valores *P* deben ser mucho menores que 0.05. Citó un estudio del estadístico Valen Johnson, quien sugirió que deberíamos elevar los estándares al exigir que los experimentos usen un valor *P* de 0.005 o menor. Sabemos que la elección de 0.05 es arbitraria en gran medida, y que la disminución del nivel de significancia dará lugar a menos conclusiones significativas, junto con menos conclusiones erróneas.

Paso 4: Para el nivel de significancia, seleccionamos $\alpha = 0.05$, que es una opción muy común.

Paso 5: Puesto que estamos probando una afirmación sobre una proporción poblacional p, el estadístico muestral $\hat{p}$ es relevante para esta prueba. La distribución muestral de las proporciones muestrales $\hat{p}$ puede aproximarse mediante una distribución normal en este caso (como se describió en la sección 6-3).

Paso 6: El dato estadístico de prueba $z = 2.55$ se puede encontrar mediante el uso de tecnología o puede calcularse usando $\hat{p} = 545/1009$ (proporción muestral), $n = 1009$ (tamaño de muestra), $p = 0.5$ (supuesta en la hipótesis nula) y $q = 1 - 0.5 = 0.5$.

$$z = \frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} = \frac{\dfrac{545}{1009} - 0.5}{\sqrt{\dfrac{(0.5)(0.5)}{1009}}} = 2.55$$

El valor *P* se puede encontrar mediante una tecnología o puede determinarse utilizando el siguiente procedimiento, que se muestra en la figura 8-3 de la página 364:

Prueba de cola izquierda:	Valor P = área a la izquierda del dato estadístico de prueba z
Prueba de cola derecha:	Valor P = área a la derecha del dato estadístico de prueba z
Prueba de dos colas:	Valor P = el *doble* del área de la región extrema limitada por el dato estadístico de prueba z

Debido a que esta prueba de hipótesis está en la cola derecha con un dato estadístico de prueba $z = 2.55$, el valor *P* es el área a la derecha de $z = 2.55$. En la tabla A-2 vemos que el área acumulada a la izquierda de $z = 2.55$ es 0.9946, por lo que el área a la derecha de ese dato estadístico de prueba es $1 - 0.9946 = 0.0054$. Obtenemos el valor $P = 0.0054$. La figura 8-5 muestra el dato estadístico de prueba y el valor *P* para este ejemplo.

Paso 7: Debido a que el valor *P* de 0.0054 es menor o igual que el nivel de significancia $\alpha = 0.05$, rechazamos la hipótesis nula.

Paso 8: Debido a que rechazamos H_0: $p = 0.5$, respaldamos la hipótesis alternativa de $p > 0.5$. Concluimos que hay suficiente evidencia muestral para respaldar el argumento de que más de la mitad de los consumidores se sienten incómodos con las entregas mediante drones.

(Vea la tabla 8-3 en la página 366 para obtener ayuda con la redacción de esta conclusión final).

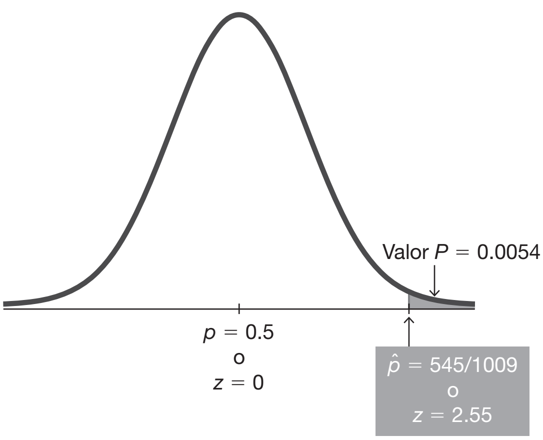


FIGURA 8-5 Método del valor *P*

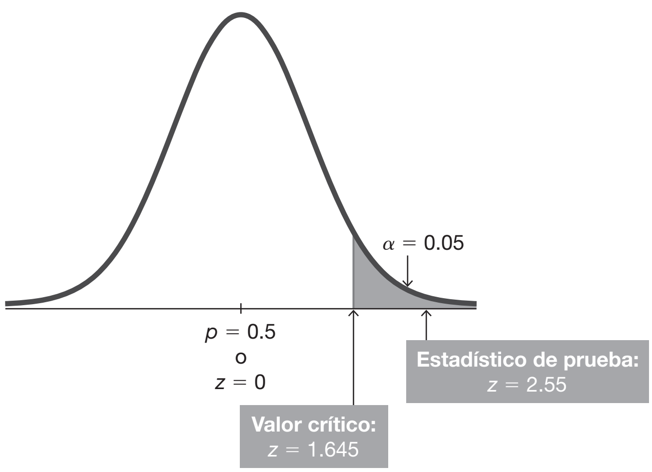


FIGURA 8-6 Método del valor crítico

PRECAUCIÓN No confunda la siguiente notación.

- Valor P = probabilidad de obtener un dato estadístico de prueba al menos tan extremo como el que representa los datos muestrales, suponiendo que la hipótesis nula H_0 es verdadera.
- p = proporción poblacional
- $\hat{p}$ = proporción muestral

Solución: Método del valor crítico

El método del valor crítico para probar hipótesis se resume en la figura 8-1 de la página 360. Cuando se usa el método del valor crítico con la afirmación dada en el ejemplo 1 de la sección 8-1, "La mayoría de los consumidores no están cómodos con las entregas mediante drones", los pasos 1 a 5 son iguales a los pasos correspondientes en el método del valor P descrito en las páginas anteriores. Por lo anterior, procedemos con el paso 6 del método del valor crítico.

Paso 6: El dato estadístico de prueba se calcula como $z = 2.55$ de acuerdo con el método anterior del valor P. Con el método del valor crítico, ahora encontramos los valores críticos (en lugar del valor P). Esta es una prueba de cola derecha, por lo que el área de la región crítica es la correspondiente a $\alpha = 0.05$ en la cola derecha. Con referencia a la tabla A-2 y mediante la aplicación de los métodos de la sección 6-1, encontramos que el valor crítico es $z = 1.645$, que está en el límite de la región crítica, como se muestra en la figura 8-6.

Paso 7: Dado que el dato estadístico de prueba cae dentro de la región crítica, rechazamos la hipótesis nula.

Paso 8: Puesto que rechazamos H_0: $p = 0.5$, concluimos que hay suficiente evidencia muestral para respaldar la afirmación de que la mayoría (más que 0.5) de los consumidores no se sienten cómodos con las entregas mediante drones. (Consulte la tabla 8-3 de la página 366 para obtener ayuda con la redacción de esta conclusión final).

Solución: Método del intervalo de confianza

La afirmación dada en el ejemplo 1 de la sección 8-1, "La mayoría de los consumidores no se sienten cómodos con las entregas mediante drones", puede probarse con un nivel de significancia de 0.05 elaborando un intervalo de confianza del 90%. (Consulte la tabla 8-1 en la página 360 para ver por qué el nivel de significancia 0.05 corresponde a un intervalo de confianza del 90%).

La estimación del intervalo de confianza del 90% para la proporción poblacional p se encuentra utilizando los datos muestrales que consisten en $n = 1009$ y $\hat{p} = 545/1009$. Mediante el uso de los métodos de la sección 7-1 obtenemos: $0.514 < p < 0.566$. El rango completo de valores en este intervalo de confianza es mayor que 0.5. Debido a que tenemos 90% de confianza en que los límites de 0.514 y 0.566 contienen el valor real de p, los datos muestrales parecen respaldar la afirmación de que la mayoría de los consumidores (más que 0.5) no se sienten cómodos con las entregas mediante drones. En este caso, la conclusión es la misma que con el método del valor P y el método del valor crítico, pero no siempre es así. Es posible que una conclusión basada en el intervalo de confianza pueda diferir de otra basada en el método del valor P o el método del valor crítico.

Determinación del número de éxitos *x*

Al usar software para realizar pruebas de hipótesis de proporciones, generalmente debemos ingresar el tamaño de muestra n y el número de éxitos x, pero es común que en las aplicaciones reales se dé la proporción muestral $\hat{p}$ en lugar de x. El número de éxitos x se puede encontrar evaluando $x = n\hat{p}$, como se ilustra en el ejemplo 1. Observe que en ese ejemplo el resultado de 5587.712 adultos debe redondearse al número entero más próximo de 5588.

EJEMPLO 1 Determinación del número de éxitos *x*

En la revista *Neurology* se describió un estudio de sonambulismo o "deambulación nocturna" e incluía información de que 29.2% de 19,136 adultos estadounidenses han experimentado sonambulismo, ¿cuál es la cantidad real de adultos que han experimentado sonambulismo?

SOLUCIÓN

El número de adultos que han sufrido sonambulismo es 29.2% de 19,136, o $0.292 \times 19{,}136 = 5587.712$, pero el resultado debe ser un número entero, por lo que redondeamos el producto al número entero más cercano, es decir 5588.

SU TURNO Resuelva el inciso (a) del ejercicio 1 "Número y proporción".

PRECAUCIÓN Cuando se realizan pruebas de hipótesis para afirmaciones sobre proporciones, es posible obtener resultados ligeramente diferentes al calcular el dato estadístico de prueba usando una proporción muestral dada, en vez de usar un valor redondeado de x al considerar $x = n\hat{p}$.

EJEMPLO 2 ¿Menos del 30% de los adultos ha experimentado sonambulismo?

Con base en los mismos datos de sonambulismo del ejemplo 1 ($n = 19{,}136$ y $\hat{p} = 29.2\%$), ¿un periodista podría afirmar de manera justificada que "menos del 30% de los adultos han experimentado sonambulismo"? Use un nivel de significancia de 0.05 para probar la hipótesis de que, para la población adulta, la proporción de aquellos que han padecido sonambulismo es menor que 0.30.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) La muestra es una muestra aleatoria simple. (2) Hay un número fijo (19,136) de ensayos independientes con dos categorías (un sujeto ha sufrido sonambulismo o no). (3) Los requisitos $np \geq 5$ y $nq \geq 5$ se satisfacen con $n = 19{,}136$ y $p = 0.30$. [Obtenemos $np = (19{,}136)(0.30) = 5740.8$, que es mayor o igual que 5, y también obtenemos $nq = (19{,}136)(0.70) = 13{,}395.2$, que es mayor o igual que 5]. Los tres requisitos se satisfacen.

Paso 1: La afirmación original se expresa en forma simbólica como $p < 0.30$.

Paso 2: El opuesto a la afirmación original es $p \geq 0.30$.

Paso 3: Como $p < 0.30$ no contiene igualdad, se convierte en H_1. Obtenemos

H_0: $p = 0.30$ (hipótesis nula)

H_1: $p < 0.30$ (hipótesis alternativa y afirmación original)

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Dado que la afirmación involucra la proporción p, el estadístico relevante para esta prueba es la proporción muestral $\hat{p}$ y la distribución muestral de las proporciones muestrales se puede aproximar mediante la distribución normal.

Paso 6: **Tecnología** Si utiliza tecnología, ésta le proporcionará el dato estadístico de prueba y el valor P. Vea los resultados adjuntos de StatCrunch que muestran que el dato estadístico de prueba es $z = -2.41$ (redondeado) y el valor $P = 0.008$.

StatCrunch

Hypothesis test results:
p : Proportion of successes
H_0 : p = 0.3
H_A : p < 0.3

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	5588	19136	0.29201505	0.0033127149	-2.4103945	0.008

Tabla A-2 Si no dispone de una tecnología, proceda de la siguiente manera para realizar la prueba de hipótesis usando el método del valor *P* resumido en la figura 8-1 de la página 360.

El dato estadístico de prueba $z = -2.41$ se calcula de la siguiente forma:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{5588}{19{,}136} - 0.30}{\sqrt{\frac{(0.30)(0.70)}{19{,}136}}} = -2.41$$

Consulte la figura 8-3 en la página 364 con el fin de conocer el procedimiento para encontrar el valor *P*. Para esta prueba de cola izquierda, el valor *P* es el área a la izquierda del estadístico de prueba. En la tabla A-2 vemos que el área a la izquierda de $z = -2.41$ es 0.0080, entonces el valor *P* es 0.0080.

Paso 7: Dado que el valor *P* de 0.0080 es menor o igual que el nivel de significancia de 0.05, rechazamos la hipótesis nula.

INTERPRETACIÓN

Debido a que rechazamos la hipótesis nula, respaldamos la hipótesis alternativa. Por lo tanto, concluimos que hay evidencia suficiente para respaldar la afirmación de que menos de 30% de los adultos ha tenido sonambulismo.

SU TURNO Resuelva el ejercicio 9 "Mickey D's".

Método de valor crítico Si tuviéramos que repetir el ejemplo 2 usando el método del valor crítico para pruebas de hipótesis, veríamos que en el paso 6 el valor crítico es $z = -1.645$, que se puede encontrar mediante el uso de la tecnología o con la tabla A-2. En el paso 7 rechazaríamos la hipótesis nula porque el dato estadístico de prueba $z = -2.41$ caería dentro de la región crítica limitada por $z = -1.645$. Entonces, llegaríamos a la misma conclusión dada en el ejemplo 2.

Método del intervalo de confianza Si tuviéramos que repetir el ejemplo 2 usando el método del intervalo de confianza, usaríamos un nivel de confianza del 90% porque tenemos una prueba de cola izquierda. (Consulte la tabla 8-1). Obtendríamos el siguiente intervalo de confianza del 90%: $0.287 < p < 0.297$. Debido a que todo el rango del intervalo de confianza cae por debajo de 0.30, hay suficiente evidencia para respaldar la afirmación de que menos del 30% de los adultos ha experimentado sonambulismo.

PARTE 2 Métodos exactos para probar hipótesis sobre una proporción poblacional *p*

En lugar de utilizar la distribución normal como una *aproximación* a la distribución binomial, podemos obtener resultados exactos empleando la propia distribución de probabilidad binomial. El cálculo manual de probabilidades binomiales son una verdadera molestia, pero la tecnología hace que este enfoque sea bastante simple. Además, este método exacto no requiere que $np \geq 5$ ni que $nq \geq 5$, por lo que tenemos un método aplicable aun cuando ese requisito no se cumple. Para probar hipótesis usando el método exacto, encuentre los valores *P* de la siguiente manera:

Método exacto Identifique el tamaño de muestra *n*, el número de éxitos *x* y el valor declarado de la proporción poblacional *p* (utilizado en la hipótesis nula); luego encuentre el valor *P* usando tecnología para encontrar probabilidades binomiales de la siguiente manera:

Prueba de cola izquierda:	Valor *P* =	*P*(*x* o menos éxitos entre *n* ensayos)
Prueba de cola derecha:	Valor *P* =	*P*(*x* o más éxitos entre *n* ensayos)
Prueba de dos colas:	Valor *P* =	El doble del menor de los valores anteriores de cola izquierda y cola derecha

Detectores de mentiras y la ley

¿Por qué no se exige que todos los sospechosos de un crimen sean sometidos a la prueba del detector de mentiras para así prescindir de los juicios? El Council of Scientific Affairs de la American Medical Association afirma que "está establecido que la clasificación de los sujetos con base en la culpabilidad puede realizarse con una precisión del 75 al 97%". Sin embargo, incluso con una precisión tan alta como el 97%, el porcentaje de resultados falsos positivos puede ser del 50%, de manera que la mitad de los sujetos inocentes aparecerían erróneamente como culpables. Una posibilidad tan alta de falsos positivos descarta el uso de pruebas con polígrafos como el único criterio para determinar la culpabilidad.

Nota: No existe un método universalmente aceptado para el caso exacto de dos colas, por lo que este caso se puede tratar con otros enfoques diferentes, algunos de los cuales son bastante complejos. Por ejemplo, Minitab usa una "prueba de razón de verosimilitud" que es diferente del enfoque anterior utilizado comúnmente.

EJEMPLO 3 Uso del método exacto

Al probar un método de selección de género, 10 parejas elegidas al azar son tratadas con ese método; cada una tiene un bebé y 9 de los bebés son niñas. Use un nivel de significancia de 0.05 para probar la hipótesis de que, con este método, la probabilidad de que un bebé sea niña es mayor a 0.75.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS El método de aproximación normal descrito en la parte 1 de esta sección requiere que $np \geq 5$ y $nq \geq 5$, pero $nq = (10)(0.25) = 2.5$, por lo que se viola el requisito. El método exacto sólo tiene los requisitos de ser una muestra aleatoria simple y satisfacer las condiciones para la distribución binomial, los cuales se satisfacen.

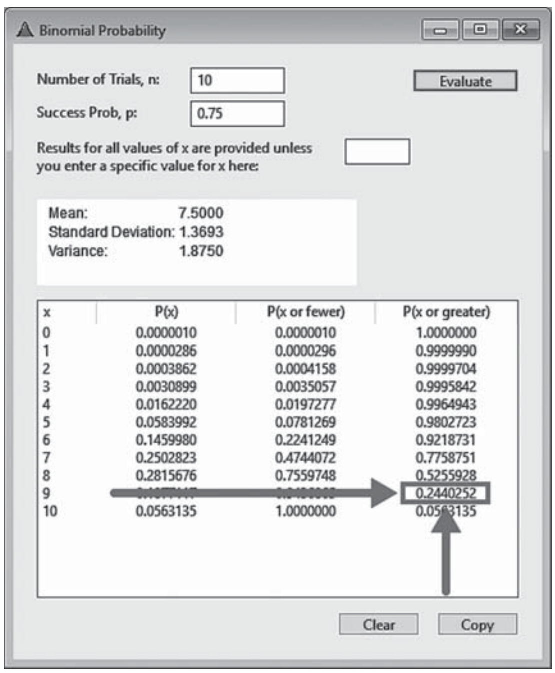
Las hipótesis nula y alternativa son:

$$H_0: p = 0.75 \text{ (hipótesis nula)}$$
$$H_1: p > 0.75 \text{ (hipótesis alternativa y afirmación original)}$$

En lugar de utilizar la distribución normal, usamos la tecnología para encontrar probabilidades en una distribución binomial con $p = 0.75$. Debido a que esta es una prueba de cola derecha, el valor P es la probabilidad de 9 o más éxitos entre 10 ensayos, suponiendo que $p = 0.75$. Vea la pantalla de Statdisk adjunta con las probabilidades exactas de la distribución binomial. Esta pantalla de Statdisk muestra que la probabilidad de 9 o más éxitos es 0.2440252 cuando se redondea a siete lugares decimales, por lo que el valor P es 0.2440252. El valor P es alto (mayor que 0.05), por lo que no podemos rechazar la hipótesis nula. No hay evidencia suficiente para respaldar la afirmación de que con el método de selección de género, la probabilidad de una niña es mayor que 0.75.

SU TURNO Resuelva el ejercicio 33 "Método exacto".

Statdisk



Mejora del método exacto Del método exacto se critica que es *demasiado conservador* en el sentido de que la probabilidad real de un error tipo I es siempre menor o igual a α, y podría ser mucho menor que α.

Con el método exacto, la probabilidad *real* de un error tipo I es menor o igual que α, que es la probabilidad *deseada* de un error tipo I.

Una **corrección de continuidad simple** mejora el comportamiento conservador del método exacto con un ajuste del valor P que se obtiene al restarle el valor que es la mitad de la probabilidad binomial en el límite, como se muestra a continuación. (Vea el ejercicio 33 "Método exacto"). Este método es fácil de aplicar si se dispone de tecnología para encontrar probabilidades binomiales.

Corrección de continuidad simple al método exacto

Prueba de cola izquierda: Valor $P = P(x \text{ o menos}) - \frac{1}{2}P(\text{exactamente } x)$

Prueba de cola derecha: Valor $P = P(x \text{ o más}) - \frac{1}{2}P(\text{exactamente } x)$

Prueba de dos colas: Valor $P =$ El doble del menor de los valores anteriores de cola izquierda y cola derecha

La "corrección de continuidad simple" anterior se describe en "Modifying the Exact Test for a Binomial Proportion and Comparisons with Other Approaches", de Alan Huston, *Journal of Applied Statistics*, vol. 33, núm. 7. Para consultar otra mejora que usa áreas de cola ponderadas con base en una medida de asimetría, vea el artículo anterior de Alan Huston.

Gane $ 1,000,000 por su percepción extra sensorial

El mago James Randi instituyó una fundación educativa que ofrece un premio de $1 millón a quien pueda demostrar poderes paranormales, sobrenaturales u ocultos. Cualquiera que posea un poder como el de adivinar el futuro, capacidad de percepción extrasensorial (PES) o la habilidad para comunicarse con los muertos puede ganar el premio si pasa ciertos procedimientos de prueba. Primero se realiza una prueba preliminar y después una formal, pero hasta ahora nadie ha salido con éxito de la prueba preliminar. La prueba formal se diseñaría con métodos estadísticos sólidos, y probablemente incluiría un análisis con una prueba de hipótesis formal. Según la fundación, se consulta a "especialistas competentes en estadística cuando se necesita evaluar los resultados o diseñar experimentos".

CENTRO DE TECNOLOGÍA

Prueba de hipótesis: Proporción

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y seleccione **Proportion One Sample** del submenú.
3. En *Alternative Hypothesis*, seleccione el formato utilizado para la hipótesis alternativa, ingrese el nivel de significancia, la proporción declarada (en la hipótesis nula), el tamaño de muestra y el número de éxitos.
4. Haga clic en **Evaluate.**

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **1 Proportion** del submenú.
3. Seleccione **Summarized data** del menú desplegable e ingrese el número de eventos y el número de ensayos (n).
4. Marque la casilla de **Perform Hypothesis Test** e ingrese la proporción utilizada en la hipótesis nula.
5. Haga clic en el botón **Options** e ingrese el nivel de confianza. (Ingrese 95.0 para un nivel de significancia de 0.05). En *Alternative Hypothesis* seleccione el formato utilizado para la hipótesis alternativa.
6. Para *Method*, seleccione **Normal approximation** para usar el mismo método de esta sección y haga clic en **OK** dos veces.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Proportion Stats** en el menú desplegable, luego seleccione **One Sample—With Summary** en el submenú.
3. Ingrese el número de éxitos y el número de observaciones (n).
4. Seleccione **Hypothesis Test for *p*** y para H_0 ingrese el valor declarado de la proporción poblacional (en la hipótesis nula). Para H_A, seleccione el formato utilizado para la hipótesis alternativa.
5. Haga clic en **Compute!**

continúa

CENTRO DE TECNOLOGÍA *continuación*

Prueba de hipótesis: Proporción

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **1-PropZTest** en el menú y presione ENTER
3. Ingrese la proporción de población declarada p_0, la cantidad de éxitos *x* y el tamaño de muestra *n*. Para *prop*, seleccione el formato utilizado para la hipótesis alternativa.
4. Seleccione **Calculate** y presione ENTER.

Excel

Complemento XLSTAT (requerido)

1. Haga clic en la pestaña **XLSTAT** en la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **Tests for one proportion** en el menú desplegable.
3. En *Data format*, seleccione **Frecuencia** si conoce el número de éxitos *x* o seleccione **Proportion** si conoce la proporción muestral $\hat{p}$.
4. Ingrese la frecuencia o proporción muestral, el tamaño de muestra y el valor declarado para la proporción poblacional (*proporción de prueba*).
5. Marque **z test** y desmarque **Continuity correction** para los métodos de esta sección.
6. Haga clic en la pestaña **Options**.
7. En *Alternative hypothesis* seleccione el formato utilizado para la hipótesis alternativa. Para *Hypothetisized Difference*, ingrese **0** e introduzca el nivel de significancia deseado (ingrese **5** para un nivel de significancia de 0.05). En *Variance* (*intervalo de confianza*) seleccione **Test Proportion** de prueba y en *Confidence Interval*, seleccione **Wald**.
8. Haga clic en **OK** para desplegar el resultado. El dato estadístico de prueba se etiqueta *z* (*valor observado*) y el valor *P* está debajo. También se mostrarán los valores críticos.

8-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

En los ejercicios 1 a 4, use los siguientes resultados de una encuesta de* USA Today *en la cual 510 personas eligieron responder a una pregunta que se publicó en el sitio web del diario: "¿Deberían los estadounidenses reemplazar las contraseñas por seguridad biométrica (huellas dactilares, etcétera)?" Entre los encuestados, 53% respondió "sí". Queremos probar la hipótesis de que más de la mitad de la población cree que las contraseñas deben reemplazarse por seguridad biométrica.

1. Número y proporción

a. Identifique el número real de encuestados que respondieron "sí".

b. Identifique la proporción muestral y el símbolo usado para representarla.

2. Hipótesis nula y alternativa Identifique la hipótesis nula y la hipótesis alternativa.

3. Equivalencia de métodos Si usamos el mismo nivel de significancia para realizar la prueba de hipótesis mediante el método del valor *P*, el método del valor crítico y un intervalo de confianza, ¿qué método no es equivalente a los otros dos?

4. Requisitos y conclusiones

a. ¿Se viola alguno de los tres requisitos? ¿Pueden los métodos de esta sección usarse para evaluar la afirmación?

b. Se estableció que podemos recordar fácilmente cómo interpretar los valores *P* con lo siguiente: "Si *P* es bajo, la nula debe irse". ¿Qué significa esto?

c. Otro truco mnemotécnico usado comúnmente es este: "Si *P* es alto, la nula volará". Dado que una prueba de hipótesis nunca da como resultado la conclusión de probar o respaldar una hipótesis nula, ¿de qué manera es engañoso este truco de memoria?

d. Los niveles de significancia comunes son 0.01 y 0.05. ¿Por qué no sería prudente usar un nivel de significancia con un número como 0.0483?

Uso de tecnología. ***En los ejercicios 5 a 8, identifique los valores indicados o interprete la pantalla dada. Use la distribución normal como una aproximación a la distribución binomial, según se describe en la parte I de esta sección, use un nivel de significancia de 0.05 y responda lo siguiente:***

a. ¿La prueba es de dos colas, de cola izquierda o de cola derecha?

b. ¿Cuál es el estadístico de prueba?

c. ¿Cuál es el valor *P*?

d. ¿Cuál es la hipótesis nula y qué concluye sobre ella?

e. ¿Cuál es la conclusión final?

5. Reacciones adversas a fármaco El medicamento Lipitor (atorvastatina) se usa para tratar el colesterol alto. En un ensayo clínico de Lipitor, 47 de 863 sujetos tratados experimentaron dolores de cabeza (según los datos de Pfizer). La pantalla adjunta de la calculadora TI-83/84 Plus muestra los resultados de una prueba a la afirmación de que menos del 10% de los sujetos tratados experimentan dolores de cabeza.

TI-83/84 Plus

NORMAL FLOAT AUTO REAL RADIAN MP
1-PropZTest
prop<.1
z=-4.45929186
p=4.1151493E-6
p̂=.0544611819
n=863

6. Vehículos auto-conducidos o autónomos En una encuesta de TE Connectivity aplicada a 1000 adultos, 29% dijo que se sentirían cómodos en un vehículo auto-conducido. La pantalla adjunta de StatCrunch es el resultado de probar la hipótesis de que más de 114 de los adultos se sentirían cómodos en un vehículo auto-dirigido.

StatCrunch

Hypothesis test results:
p : Proportion of successes
H_0 : p = 0.25
H_A : p > 0.25

Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	290	1000	0.29	0.013693064	2.921187	0.0017

7. Propiedad de teléfono celular Una encuesta del Pew Research Center aplicada a 2076 adultos seleccionados al azar mostró que 91% de ellos tienen teléfonos celulares. La siguiente pantalla de Minitab resulta de probar la hipótesis de que 92% de los adultos poseen teléfonos celulares.

Minitab

Test of p = 0.92 vs p ≠ 0.92

Sample	X	N	Sample p	95% CI	Z-Value	P-Value
1	1889	2076	0.909923	(0.897608, 0.922238)	−1.69	0.091

8. Seguridad biométrica En una encuesta realizada por *USA Today* y aplicada a 510 personas, 53% dijeron que los estadounidenses deberían reemplazar las contraseñas por seguridad biométrica, por ejemplo con huellas dactilares. La pantalla adjunta de Statdisk muestra los resultados de una prueba a la afirmación de que la mitad de los estadounidenses piensan que deben reemplazar las contraseñas por seguridad biométrica.

Statdisk

Sample proportion: 0.5294118
Test Statistic, z: 1.3284
Critical z: ±1.9600
P-Value: 0.1840

Pruebas de afirmaciones sobre proporciones. ***En los ejercicios 9 a 32, pruebe la afirmación dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor P o los valores críticos, luego establezca la conclusión sobre la hipótesis nula, así como una conclusión final que aborde la afirmación original. Use el método del valor P a menos que su profesor le indique lo contrario. Use la distribución normal como una aproximación a la distribución binomial, como se describe en la parte 1 de esta sección.***

9. Mickey D's En un estudio de la precisión de los pedidos de comida rápida, McDonald's tuvo 33 pedidos que no fueron precisos entre los 362 pedidos observados (según datos de la revista *QSR*). Use un nivel de significancia de 0.05 para probar la hipótesis de que la tasa de pedidos imprecisos es igual a 10%. ¿La tasa de precisión parece ser aceptable?

10. Eliquis El medicamento Eliquis (apixaban) se usa para ayudar a prevenir los coágulos en la sangre de ciertos pacientes. En ensayos clínicos, entre 5924 pacientes tratados con Eliquis, 153 desarrollaron la reacción adversa de náuseas (según datos de Bristol-Myers Squibb Co.). Use un nivel de significancia de 0.05 para evaluar la afirmación de que 3% de los usuarios de Eliquis desarrollan náuseas. ¿Las náuseas parecen ser una reacción adversa problemática?

11. Encuesta de células madre Se seleccionaron adultos al azar para una encuesta de *Newsweek*. Se les preguntó si "están a favor o en contra del uso de los impuestos federales para financiar la investigación médica utilizando células madre obtenidas de embriones humanos". De los encuestados, 481 estuvieron a favor, 401 se opusieron y 120 no estuvieron seguros. Un político afirma que las personas realmente no entienden el problema de las células madre y que sus respuestas a estas preguntas son respuestas aleatorias equivalentes a lanzar una moneda. Excluya a los 120 sujetos que dijeron que no estaban seguros y use un nivel de significancia de 0.01 para probar la hipótesis de que la proporción de sujetos que responden a favor es igual a 0.5. ¿Qué sugiere el resultado sobre la afirmación del político?

12. M&Ms El conjunto de datos 27 "Pesos de M&Ms" en el apéndice B lista datos de 100 M&Ms, y 27% de ellos son azules. La compañía de dulces Mars afirma que el porcentaje de M&Ms azules es igual a 24%. Use un nivel de significancia 0.05 para probar la hipótesis. ¿Debería la compañía Mars tomar medidas correctivas?

13. OxyContin El medicamento OxyContin (oxicodona) se usa para tratar el dolor, pero es peligroso porque resulta adictivo y puede ser letal. En ensayos clínicos, 227 sujetos fueron tratados con OxyContin y 52 de ellos experimentaron náuseas (en base a datos de Purdue Pharma L.P.). Use un nivel de significancia de 0.05 para probar la hipótesis de que más de 20% de los usuarios de OxyContin desarrollan náuseas. ¿La tasa de náuseas parece ser demasiado alta?

14. Negligencia médica En un estudio de 1228 demandas por negligencia médica seleccionadas al azar, se descubrió que 856 de ellas fueron desechadas o descartadas (según los datos de Physicians Insurers Association of America). Use un nivel de significancia de 0.01 para evaluar la afirmación de que la mayoría de las demandas por negligencia médica son desechadas o descartadas. ¿Esto debería ser reconfortante para los médicos?

15. Tasa de retorno de la encuesta En un estudio sobre el uso del teléfono celular y el dominio del hemisferio cerebral, se envió por correo electrónico una encuesta a 5000 sujetos seleccionados aleatoriamente de un grupo en línea relacionado con los oídos. Fueron devueltas 717 encuestas. Use un nivel de significancia de 0.01 para probar la hipótesis de que la tasa de retorno es menor a 15%.

16. Detección de drogas La compañía Drug Test Success ofrece una prueba "1-Panel-THC" para el consumo de marihuana. Entre 300 sujetos evaluados, los resultados de 27 sujetos fueron incorrectos (ya sea un falso positivo o un falso negativo). Use un nivel de significancia de 0.05 para probar la hipótesis de que menos de 10% de los resultados de la prueba son incorrectos. ¿La prueba parece ser buena para la mayoría de los propósitos?

17. Nacimientos Una muestra aleatoria de 860 nacimientos en el estado de Nueva York incluyó a 426 varones. Use un nivel de significancia de 0.05 para evaluar la afirmación de que 51.2% de los bebés recién nacidos son varones. ¿Los resultados respaldan la creencia de que 51.2% de los bebés recién nacidos son varones?

18. Genética mendeliana Cuando Mendel realizó sus famosos experimentos genéticos con chícharos, una muestra de descendientes consistió en 428 chícharos verdes y 152 chícharos amarillos. Use un nivel de significancia de 0.01 para probar la hipótesis de Mendel bajo las mismas circunstancias, 25% de los chícharos descendientes serán amarillos. ¿Qué podemos concluir sobre esta afirmación mendeliana?

19. Detectores de mentiras Los ensayos en un experimento con un polígrafo arrojan 98 resultados que incluyen 24 casos de resultados erróneos y 74 casos de resultados correctos (según datos de experimentos realizados por los investigadores Charles R. Honts de Boise State University y Gordon H. Barland del Instituto del Polígrafo del Departamento de Defensa). Use un nivel de significancia de 0.05 para probar la hipótesis de que los resultados del polígrafo son correctos menos del 80% de las veces. Con base en los resultados, ¿deberían prohibirse los resultados de las pruebas de polígrafo como evidencia en los juicios?

20. Repetición instantánea en el tenis El sistema electrónico Ojo de Halcón se usa en el tenis para desplegar una repetición instantánea que muestra si una pelota está dentro o fuera de los límites para que los jugadores puedan desafiar las decisiones tomadas por los árbitros. En un Abierto de EUA reciente, los jugadores individuales realizaron 879 desafíos y 231 de ellos tuvieron éxito al revocar la decisión del árbitro. Use un nivel de significancia de 0.01 para evaluar la afirmación de que menos de 1/3 de los desafíos son exitosos. ¿Qué sugieren los resultados sobre la capacidad de los tenistas para ver las jugadas mejor que los árbitros?

21. Terapia táctil Cuando tenía 9 años de edad, Emily Rosa hizo un experimento en una feria de ciencias en el que probó a terapeutas táctiles profesionales para ver si podían sentir su campo de energía. Lanzó una moneda para seleccionar su mano derecha o su mano izquierda, y luego pidió a los terapeutas que identificaran la mano seleccionada colocando su mano justo debajo de la mano de Emily sin verla y sin tocarla. Entre los 280 ensayos, los terapeutas táctiles fueron correctos 123 veces (según datos de "A Close Look at Therapeutic Touch", *Journal of the American Medical Association*, vol. 279, núm. 13). Use un nivel de significancia de 0.10 para evaluar la afirmación de que los terapeutas táctiles usan un método equivalente a las suposiciones al azar. ¿Los resultados sugieren que los terapeutas táctiles son efectivos?

22. Terapia táctil Repita el ejercicio anterior usando un nivel de significancia de 0.01. ¿La conclusión cambia?

23. Teléfonos celulares y cáncer En un estudio de 420,095 usuarios de teléfonos celulares daneses, 135 sujetos desarrollaron cáncer cerebral o del sistema nervioso (según datos de *Journal of the National Cancer Institute*, según se informó en *USA Today*). Pruebe la afirmación de una creencia bastante común de que tales cánceres se ven afectados por el uso del teléfono celular. Es decir, evalúe la afirmación de que los usuarios de teléfonos celulares desarrollan cáncer del cerebro o del sistema nervioso a una tasa diferente al índice de 0.0340% para las personas que no usan teléfonos celulares. Debido a que este problema tiene tanta importancia, use un nivel de significancia 0.005. Con base en los resultados, ¿los usuarios de teléfonos celulares deberían preocuparse por el cáncer cerebral o del sistema nervioso?

24. Exactitud de escáneres de caja en tiendas En un estudio de los escáneres de caja en tiendas, se revisaron 1234 artículos para determinar la precisión de los precios; se descubrió que 20 artículos tuvieron sobrecargos, y 1214 no los tuvieron (con base en datos de "UPC Scanner Pricing Systems: Are they Accurate?" de Goodstein, *Journal of Marketing*, vol. 58). Use un nivel de significancia de 0.05 para probar la hipótesis de que, con los escáneres, el 1% de las ventas tienen sobrecargos. (Antes de que se utilizaran los escáneres, la tasa de sobrecargos se estimó en aproximadamente 1%). Con base en estos resultados, ¿parecen los escáneres ayudar a los consumidores a evitar los sobrecargos?

25. Victorias en el Super Bowl De acuerdo con la muestra de los primeros 49 Super Bowls, 28 de ellos fueron ganados por equipos en la Conferencia Nacional de Fútbol Americano (NFC). Use un nivel de significancia de 0.05 para probar la hipótesis de que la probabilidad de una victoria en el Super Bowl para un equipo de la NFC es mayor a un medio.

26. Prueba de efectividad de los parches de nicotina En un estudio realizado con fumadores que intentaban dejar de fumar con la terapia de parches de nicotina, 39 fumaban un año después del tratamiento y 32 no fumaban en ese mismo periodo posterior (según datos de "High-Dose Nicotin Patches Therapy", de Dale *et al*., *Journal of the American Medical Association*, vol. 274, núm. 17). Use un nivel de significancia de 0.05 para evaluar la afirmación de que entre los fumadores que intentan dejar de fumar con la terapia de parches de nicotina, la mayoría fuma un año después del tratamiento. ¿Estos resultados sugieren que la terapia con parches de nicotina no es efectiva?

27. Regla del tiempo extra en el fútbol americano Antes de que la regla del tiempo extra en la Liga Nacional de Fútbol Americano se cambiara en 2011, entre 460 juegos con tiempo extra, 252 fueron ganados por el equipo que ganó el lanzamiento de la moneda al comienzo del tiempo extra. Utilizando un nivel de significancia de 0.05, pruebe la afirmación de que el lanzamiento de la moneda es justo en el sentido de que ninguno de los equipos tiene una ventaja al ganarlo. ¿El lanzamiento de la moneda parece ser justo?

28. Posposición de la muerte Una hipótesis interesante y popular es que las personas pueden posponer temporalmente la muerte para sobrevivir a una festividad o evento importante, como un cumpleaños. En un estudio, se encontró que hubo 6062 muertes en la semana anterior al Día de Acción de Gracias y 5938 muertes la semana posterior al Día de Acción de Gracias (según datos de "Holydays, Birthdays and Postponement of Cancer Death", de Young and Hade, *Journal of The American Medical Association*, vol. 292, núm. 24). Si las personas pueden posponer la muerte hasta después del Día de Acción de Gracias, entonces la proporción de muertes en la semana anterior debe ser menor a 0.5. Use un nivel de significancia de 0.05 para probar la hipótesis de que la proporción de muertes en la semana antes del Día de Acción de Gracias es menor a 0.5. De acuerdo con el resultado, ¿parece haber alguna indicación de que las personas puedan posponer temporalmente la muerte para sobrevivir las festividades de Acción de Gracias?

29. ¿Es Nessie Real? Esta pregunta fue publicada en el sitio web de America Online: ¿Cree usted que el monstruo de Loch Ness existe? Entre 21,346 respuestas, 64% fueron "sí". Use un nivel de significancia de 0.01 para probar la hipótesis de que la mayoría de las personas cree que existe el monstruo de Loch Ness. ¿Cómo se ve afectada la conclusión por el hecho de que los usuarios de Internet que vieron la pregunta podrían decidir si responder?

30. Dejar de fumar y cuidados constantes En un programa diseñado para ayudar a los pacientes a dejar de fumar, 198 sujetos recibieron atención *constante*, y 82.8% de ellos dejaron de fumar después de un mes (según datos de "Sustained Care Intervention and Postdischarge Smoking Cessation Among Hospitalized Adults", de Rigotti *et al*., *Journal of the American Medical Association*, vol. 312, núm. 7). Use un nivel de significancia de 0.01 para evaluar la afirmación de que 80% de los pacientes dejan de fumar cuando reciben cuidados constantes. ¿El cuidado constante parece ser efectivo?

31. Sesgo en la selección del jurado En el caso de *Casteneda vs Partida*, se encontró que durante un período de 11 años en el Condado de Hidalgo, Texas, se seleccionaron 870 personas para el deber de un gran jurado y 39% de ellos eran estadounidenses de ascendencia mexicana. Entre las personas elegibles para el deber de gran jurado, el 79.1% eran estadounidenses de ascendencia mexicana. Use un nivel de significancia de 0.01 para evaluar la afirmación de que el proceso de selección es parcial contra los estadounidenses de ascendencia mexicana. ¿El sistema de selección del jurado parece estar sesgado?

32. Uso de medicamentos En una encuesta a 3005 adultos de 57 a 85 años, se encontró que 81.7% de ellos usaba al menos un medicamento recetado (de acuerdo con datos de "Use of Prescription and Over-the-Counter Medications and Dietary Supplements Among Older Adults in the United States" de Qato *et al*., *Journal of the American Medical Association*, vol. 300, núm. 24). Use un nivel de significancia de 0.01 para evaluar la afirmación de que más de 3/4 de los adultos usan al menos un medicamento recetado. ¿El índice de uso de medicamentos recetados entre los adultos parece ser alto?

8-2 Más allá de lo básico

33. Método exacto Para cada uno de los tres métodos diferentes de pruebas de hipótesis (identificados en la columna de la izquierda), ingrese los valores *P* correspondientes a la hipótesis alternativa dada y los datos muestrales. Tenga en cuenta que las entradas en la última columna corresponden al problema del capítulo. ¿Qué tanto concuerdan los resultados con el tamaño de muestra grande?

	H_1: $p \neq 0.5$ $n = 10$, $x = 9$	H_1: $p \neq 0.4$ $n = 10$, $x = 9$	H_1: $p > 0.5$ $n = 1009$, $x = 545$
Aproximación normal			
Exacto			
Exacto con corrección de continuidad simple			

34. Uso de intervalos de confianza para probar hipótesis Al analizar los últimos dígitos de los números de teléfono en Port Jefferson, se encuentra que entre los 1000 dígitos seleccionados al azar, 119 son ceros. Si los dígitos se seleccionan al azar, la proporción de ceros debería ser 0.1.

a. Use el método del valor crítico con un nivel de significancia de 0.05 para probar la hipótesis de que la proporción de ceros es igual a 0.1.

b. Use el método del valor *P* con un nivel de significancia de 0.05 para probar la hipótesis de que la proporción de ceros es igual a 0.1.

c. Use los datos muestrales para elaborar una estimación del intervalo de confianza del 95% para la proporción de ceros. ¿Qué sugiere el intervalo de confianza sobre la afirmación de que la proporción de ceros es igual a 0.1?

d. Compare los resultados del método del valor crítico, el método del valor *P* y el método del intervalo de confianza. ¿Todos conducen a la misma conclusión?

35. Potencia Para una prueba de hipótesis con un nivel de significancia α, la probabilidad de un error de tipo I es α, mientras que la probabilidad β de un error de tipo II depende del valor particular de p que se use como alternativa a la hipótesis nula.

a. A partir de una hipótesis alternativa de $p < 0.4$, usando un tamaño de muestra de $n = 50$, y suponiendo que el valor verdadero de p es 0.25, encuentre la potencia de la prueba. Vea el ejercicio 34 "Cálculo de la potencia" en la sección 8-1. [*Sugerencia:* Use los valores $p = 0.25$ y $pq/n = (0.25)(0.75)/50$].

b. Encuentre el valor de β, la probabilidad de cometer un error tipo II.

c. Dadas las condiciones citadas en el inciso (a), encuentre la potencia de la prueba. ¿Qué nos dice la potencia sobre la efectividad de la prueba?

8-3 Prueba de una hipótesis respecto a una media

Concepto clave La prueba de una hipótesis sobre la media de una población es uno de los métodos más importantes presentados en este libro. La parte 1 de la presente sección trata sobre el caso muy realista y comúnmente utilizado en el cual no se conoce la desviación estándar de la población σ. La parte 2 incluye un breve estudio del procedimiento utilizado cuando se conoce σ, lo cual es muy raro.

PARTE 1 Prueba de una hipótesis acerca de μ con σ desconocida

En realidad, es muy raro que probemos una afirmación sobre un valor desconocido de una media poblacional μ, pero que de alguna manera conozcamos el valor de la desviación estándar poblacional σ. La situación realista es que probemos una afirmación sobre una media poblacional μ y el valor de la desviación estándar poblacional σ no se conozca. Cuando σ no se conoce, lo estimamos con la desviación estándar muestral s. A partir del teorema del límite central (sección 6-4), sabemos que la distribución de las medias muestrales $\bar{x}$ es aproximadamente una distribución normal con media $\mu_{\bar{x}} = \mu$ y desviación estándar $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, pero si se desconoce σ, lo estimamos con $s/\sqrt{n}$, que se ha utilizado en el dato estadístico de prueba para una "prueba *t*". Este dato estadístico de prueba tiene una distribución llamada distribución *t* de Student. Los requisitos, el estadístico de prueba, el valor *P* y los valores críticos se resumen en el cuadro de elementos clave que se presenta a continuación.

Métodos equivalentes

Para la prueba *t* descrita en esta sección, el método del valor *P*, el método del valor crítico y el método del intervalo de confianza son todos equivalentes en el sentido de que todos conducen a las mismas conclusiones.

ELEMENTOS CLAVE

Pruebas de hipótesis acerca de una media poblacional con σ desconocida

Objetivo

Usar una prueba de hipótesis formal para evaluar una afirmación sobre una población con media μ.

Notación

n = tamaño de muestra

s = desviación estándar *muestral*

$\bar{x}$ = media *muestral*

$\mu_{\bar{x}}$ = media *poblacional* (este valor se toma de la afirmación y se usa en el enunciado de la hipótesis nula H_0)

Requisitos

1. La muestra es una muestra aleatoria simple.
2. Se cumple al menos una de las siguientes condiciones: la población se distribuye normalmente o $n > 30$.

Dato estadístico de prueba para probar una hipótesis sobre una media

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$ Redondee t a tres lugares decimales, como en la tabla A-3).

Valores *P*: use software o la distribución *t* de Student (tabla A-3) con los grados de libertad dados por gl $= n - 1$. (La figura 8-3 en la página 364 resume el procedimiento para encontrar valores *P*).

Valores críticos: use la distribución *t* de Student (tabla A-3) con los grados de libertad dados por gl $= n - 1$. (Cuando la tabla A-3 no incluya el número de grados de libertad, usted puede ser conservador usando el siguiente número más bajo de grados de libertad que se encuentre en la tabla, puede emplear el número más cercano de grados de libertad en la tabla, o bien puede interpolar).

En cifras

En la actualidad hay 2.7 zettabytes (10^{21} bytes) de datos en nuestro universo digital.

Requisito de normalidad o $n > 30$ Esta prueba t es *robusta* contra una desviación de la normalidad, lo que significa que la prueba funciona razonablemente bien si la desviación de la normalidad no es demasiado extrema. Verifique que no haya valores atípicos y que el histograma o la gráfica de puntos tengan una forma que no esté muy alejada de una distribución normal.

Si la población original no se distribuye normalmente, usamos la condición $n > 30$ para justificar el uso de la distribución normal, pero no existe un tamaño de muestra mínimo específico exacto que funcione para todos los casos. Los tamaños de muestra de 15 a 30 son suficientes si la población tiene una distribución que no está alejada de la normal, pero algunas poblaciones tienen distribuciones muy desviadas de la normalidad, y pueden requerirse tamaños de muestra superiores a 30. En este texto usamos el criterio simplificado de $n > 30$ como justificación para tratar la distribución de las medias muestrales como una distribución normal, independientemente de qué tanto se desvíe de la normalidad.

Propiedades importantes de la distribución *t* de Student

A continuación se proporciona una breve reseña de las propiedades importantes de la distribución t de Student presentada por primera vez en la sección 7-2:

1. La distribución t de Student es diferente para distintos tamaños de muestra (vea la figura 7-4 en la sección 7-2).
2. La distribución t de Student tiene la misma forma de campana general que la distribución normal estándar; su forma más amplia refleja la mayor variabilidad que se espera cuando se usa s para estimar σ.
3. La distribución t de Student tiene una media de $t = 0$ (igual que 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 y es mayor que 1 (a diferencia de la distribución normal estándar, que tiene $\sigma = 1$).
5. A medida que el tamaño de muestra n aumenta, la distribución t de Student se acerca más a la distribución normal estándar.

Método del valor *P* con tecnología

Si se dispone de una tecnología adecuada, el método del valor P para pruebas de hipótesis es el camino a seguir.

EJEMPLO 1 Sueño en adultos: método del valor *P* con software

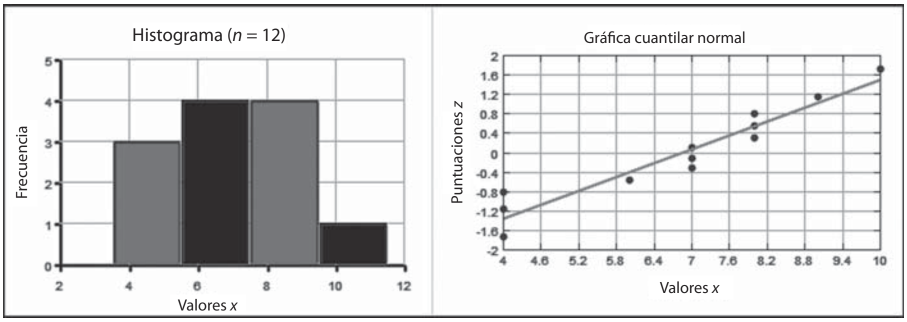
El autor obtuvo las horas de sueño de las personas adultas seleccionadas al azar que se incluyeron en el Estudio Nacional de Examen de Salud y Nutrición, y los tiempos (en horas) se listan a continuación. Los estadísticos no redondeados para esta muestra son $n = 12$, $\bar{x} = 6.83333333$ horas, $s = 1.99240984$ horas. Una recomendación común es que los adultos deben dormir entre 7 y 9 horas cada noche. Use el método del valor P con un nivel de significancia de 0.05 para evaluar la afirmación de que la cantidad promedio de sueño para los adultos es inferior a 7 horas.

4 8 4 4 8 6 9 7 7 10 7 8

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) La muestra es una muestra aleatoria simple. (2) El segundo requisito es que “la población se distribuya normalmente o $n > 30$”. El tamaño de muestra es $n = 12$, que no excede 30, por lo que debemos determinar si los datos de la muestra parecen ser de una población distribuida normalmente. El histograma y la gráfica cuantilar normal que se adjuntan, además de la aparente ausencia de valores atípicos, indican que la muestra parece provenir de una población con una distribución que es aproximadamente normal. Ambos requisitos se satisfacen.

Statdisk



Los pasos que siguen el procedimiento resumido en la figura 8-1 de la página 360 son:

Paso 1: La afirmación de que "el tiempo medio de sueño en los adultos es inferior a 7 horas" se convierte en $\mu < 7$ horas cuando se expresa en forma simbólica.

Paso 2: La alternativa (en forma simbólica) a la afirmación original es $\mu \geq 7$ horas.

Paso 3: Debido a que el enunciado $\mu < 7$ horas no contiene la condición de igualdad, se convierte en la hipótesis alternativa H_1. La hipótesis nula H_0 es la afirmación de que $\mu = 7$ horas.

H_0: $\mu = 7$ horas (hipótesis nula)

H_1: $\mu < 7$ horas (hipótesis alternativa y afirmación original)

Paso 4: Como se especifica en el enunciado del problema, el nivel de significancia es $\alpha = 0.05$.

Paso 5: Como la afirmación se hace sobre la *media poblacional* μ, el estadístico muestral más relevante para esta prueba es la *media muestral* $\bar{x}$, y usamos la distribución t.

Paso 6: Los estadísticos muestrales de $n = 12$, $\bar{x} = 6.83333333$ horas, $s = 1.99240984$ horas se utilizan para calcular el dato estadístico de prueba de la siguiente manera, pero las tecnologías proporcionan el dato estadístico de prueba de $t = -0.290$. En cálculos como el siguiente, es bueno conservar lugares decimales adicionales y no redondear.

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{6.83333333 - 7}{\frac{1.99240984}{\sqrt{12}}} = -0.290$$

Valor *P* con tecnología La tecnología podría usarse para obtener el valor P. En la página siguiente se muestran los resultados de varias tecnologías, y podemos ver que el valor P es 0.3887 (redondeado). (SPSS muestra un valor P de dos colas de 0.777, por lo que debe reducirse a la mitad para esta prueba de una cola).

Paso 7: Debido a que el valor P de 0.3887 es mayor que el nivel de significancia $\alpha = 0.05$, no podemos rechazar la hipótesis nula.

INTERPRETACIÓN

Paso 8: Puesto que no rechazamos la hipótesis nula, concluimos que no hay evidencia suficiente para respaldar la afirmación de que el tiempo medio de sueño en los adultos es inferior a 7 horas.

SU TURNO Resuelva el ejercicio 13 "Evaluaciones de cursos".

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
T-Test
μ<7
t=-.2897748534
p=.3886888459
x̄=6.833333333
Sx=1.99240984
n=12
```

Statdisk

t Test
Test Statistic, t: -0.2898
Critical t: -1.7959
P-Value: 0.3887

90% Confidence interval:
5.800414 < μ < 7.866252

Excel (XLSTAT)

Difference	-0.1667
t (Observed value)	-0.2898
t (Critical value)	-1.7959
DF	11
p-value (one-tailed)	0.3887
alpha	0.05

Minitab

Test of μ = 7 vs < 7

Variable	N	Mean	StDev	SE Mean	95% Upper Bound	T	P
Sleep	12	6.833	1.992	0.575	7.866	−0.29	0.389

JMP

Hypothesized Value	7
Actual Estimate	6.83333
DF	11
Std Dev	1.99241
t Test	
Test Statistic	-0.2898
Prob > |t|	0.7774
Prob > t	0.6113
Prob < t	0.3887

StatCrunch

Hypothesis test results:
μ : Mean of variable
H_0 : μ = 7
H_A : μ < 7

Variable	Sample Mean	Std. Err.	DF	T-Stat	P-value
Sleep	6.8333333	0.57515918	11	-0.28977485	0.3887

SPSS

	Test Value = 7					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
SLEEP	-.290	11	.777	-.16667	-1.4326	1.0993

Examine las pantallas de las tecnologías para ver que sólo dos de ellas incluyen valores críticos, pero todas contienen valores *P*. Esta es una razón importante por la cual el método del valor *P* para pruebas de hipótesis se ha utilizado ampliamente en los últimos años.

Método del valor *P* sin tecnología

Si no se dispone de una tecnología adecuada, podemos usar la tabla A-3 para identificar un rango de valores que contenga el valor *P*. Al utilizar la tabla A-3, tenga en cuenta que está diseñada sólo para valores positivos de *t* y áreas de la cola derecha, pero las áreas de la cola izquierda corresponden a los mismos valores de *t* con signos negativos.

EJEMPLO 2 Sueño en adultos: Método del valor *P* sin tecnología

El ejemplo 1 es una prueba de cola izquierda con un dato estadístico de prueba de $t = -0.290$ (redondeado) y un tamaño de muestra $n = 12$, por lo que el número de grados de libertad es gl $= n - 1 = 11$. Usando el dato estadístico de prueba $t = -0.290$ con la tabla A-3, examine los valores de *t* en la fila para gl $= 11$ y observe que 0.290 es menor que todos los valores *t* listados en la fila, lo que indica que el área en la cola izquierda debajo del dato estadístico de prueba $t = -0.290$ es mayor que 0.10. En este caso, la tabla A-3 nos permite concluir que el valor $P > 0.10$; por su parte la tecnología proporcionó el valor *P* de 0.3887. Con el valor $P > 0.10$, las conclusiones son las mismas que en el ejemplo 1.

SU TURNO Resuelva el ejercicio 21 “Plomo en la medicina”.

SUGERENCIA Debido a que el uso de la tabla A-3 para encontrar un rango de valores que contenga el valor *P* puede ser un poco complicado, el método del valor crítico (vea el ejemplo 3) podría ser más fácil de aplicar que el método *P* si la tecnología adecuada no estuviera disponible.

Método del valor crítico

EJEMPLO 3 Sueño en adultos: Método del valor crítico

El ejemplo 1 es una prueba de cola izquierda con dato estadístico de prueba $t = -0290$ (redondeado). El tamaño muestral es $n = 12$, por lo que el número de grados de libertad es gl $= n - 1 = 11$. Dado el nivel de significancia de $\alpha = 0.05$, consulte la fila de la tabla A-3 correspondiente a 11 grados de libertad, y consulte la columna que identifica un "área en una cola" de 0.05 (el nivel de significancia). La intersección de la fila y la columna arroja el valor crítico de $t = 1.796$, pero esta prueba es de cola izquierda, por lo que el valor crítico real es $t = -1.796$. La figura 8-7 muestra que el dato estadístico de prueba $t = -0.290$ no cae dentro de la región crítica limitada por el valor crítico $t = -1.796$, por lo que no podemos rechazar la hipótesis nula. Las conclusiones son las mismas que las dadas en el ejemplo 1.

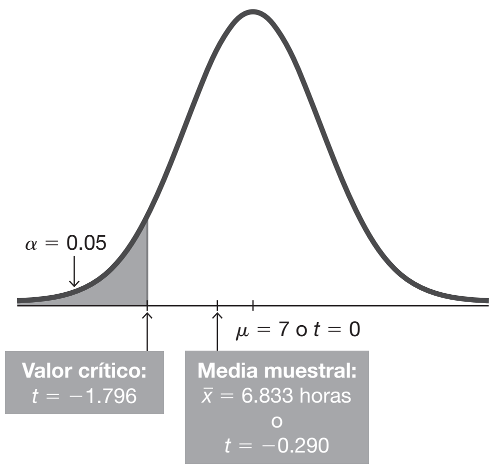


FIGURA 8-7 Prueba *t*: Método del valor crítico

Método del intervalo de confianza

EJEMPLO 4 Sueño en adultos: método del intervalo de confianza

El ejemplo 1 es una prueba de cola izquierda con un nivel de significancia $\alpha = 0.05$, por lo que deberíamos usar un 90% como nivel de confianza (como se indica en la tabla 8-1 de la página 360). Para los datos muestrales dados en el ejemplo 1, a continuación se presenta la estimación del intervalo de confianza del 90% para μ: 5.80 horas $< \mu <$ 7.87 horas. Al probar la hipótesis de que $\mu = 7$ horas, usamos H_0: $\mu = 7$ horas, pero el valor supuesto de $\mu = 7$ horas está contenido dentro de los límites del intervalo de confianza, lo que nos dice que 7 horas podrían ser el valor de μ. No tenemos evidencia suficiente para rechazar H_0: $\mu = 7$ horas, es decir, no se rechaza la hipótesis nula y se obtienen las mismas conclusiones dadas en el ejemplo 1.

EJEMPLO 5 ¿La temperatura corporal media es realmente 98.6 °F?

El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye temperaturas corporales medidas con los siguientes estadísticos para las 12 AM del día 2: $n = 106$, $\bar{x} = 98.20$ °F, $s = 0.62$ °F. (Este es el conjunto de datos favorito del autor). Use un nivel de significancia de 0.05 para probar la creencia común de que la media poblacional es 98.6 °F.

continúa

Detectores de mentiras humanos

Algunos investigadores examinaron la capacidad de 13,000 personas para determinar si alguien miente. Encontraron a 31 individuos con habilidades excepcionales para identificar mentiras. Estos detectores de mentiras humanos tenían tasas de éxito de alrededor del 90%. Además, descubrieron que los agentes federales y los alguaciles eran bastante hábiles para detectar mentiras, con tasas de éxito de alrededor del 80%. La profesora de psicología Maureen O'Sullivan entrevistó a todas las personas hábiles para identificar mentiras, y comentó que "todas ellas prestan atención a indicios no verbales y a los matices en el uso de palabras; además, los aplican de forma diferente a distintas personas. Estos individuos son capaces de decir ocho cosas sobre una persona después de observar un video de dos segundos. Es increíble todo lo que estas personas pueden observar". Es posible usar métodos estadísticos para distinguir entre las personas que son capaces de detectar mentiras y aquellas que no lo son.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Con el diseño de estudio utilizado, podemos tratar la muestra como una muestra aleatoria simple. (2) El segundo requisito es que "la población se distribuya normalmente o $n > 30$". El tamaño de muestra es $n = 106$, por lo que se cumple el segundo requisito y no es necesario investigar la normalidad de los datos. Ambos requisitos se satisfacen.

Los pasos que siguen el procedimiento resumido en la figura 8-1 son:

Paso 1: La afirmación de que "la media poblacional es 98.6 °F" se convierte en $\mu = 98.6$ °F cuando se expresa en forma simbólica.

Paso 2: La alternativa (en forma simbólica) a la afirmación original es $\mu \neq 98.6$ °F.

Paso 3: Debido a que el enunciado $\mu \neq 98.6$ °F no contiene la condición de igualdad, se convierte en la hipótesis alternativa H_1. La hipótesis nula H_0 es la afirmación de que $\mu = 98.6$ °F.

H_0: $\mu = 98.6$ °F (hipótesis nula y afirmación original)
H_1: $\mu \neq 98.6$ °F (hipótesis alternativa)

Paso 4: Como se especifica en el enunciado del problema, el nivel de significancia es $\alpha = 0.05$.

Paso 5: Debido a que la afirmación se hace acerca de la *media poblacional* μ, el estadístico muestral más relevante para esta prueba es la *media muestral* $\bar{x}$. Usamos la distribución t porque el dato estadístico muestral relevante es $\bar{x}$ y los requisitos para usar la distribución t están satisfechos.

Paso 6: Los datos estadísticos muestrales se usan para calcular el dato estadístico de prueba, pero el software usa valores no redondeados para proporcionar el dato estadístico de prueba $t = -6.61$.

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{98.20 - 98.6}{\frac{0.62}{\sqrt{106}}} = -6.64$$

Valor *P* El valor *P* es 0.0000 o 0+ (o "menos que 0.01" si se usa la tabla A-3).

Valores críticos: Los valores críticos son ± 1.983 (o ± 1.984 si se usa la tabla A-3).

Intervalo de confianza: El intervalo de confianza del 95% es $98.08\text{ °F} < \mu < 98.32\text{ °F}$.

Paso 7: Los tres enfoques conducen a la misma conclusión: Rechazar H_0.

- **Valor *P*:** El valor *P* de 0.0000 es menor que el nivel de significancia $\alpha = 0.05$.
- **Valores críticos:** El dato estadístico de prueba $t = -6.64$ cae en la región crítica limitada por ± 1.983.
- **Intervalo de confianza:** La media declarada de 98.6 °F no está dentro del intervalo de confianza de $98.08\text{ °F} < \mu < 98.32\text{ °F}$.

INTERPRETACIÓN

Paso 8: Hay suficiente evidencia para justificar el rechazo de la creencia común de que la media poblacional es 98.6 °F.

Métodos alternativos utilizados cuando la población no es normal y $n \leq 30$

Los métodos de esta sección incluyen dos requisitos: (1) La muestra es una muestra aleatoria simple; (2) la población se distribuye normalmente o $n > 30$. Si tenemos datos muestrales que no se recopilan de forma adecuada, como lo es una muestra de respuesta voluntaria, es probable que no haya nada que se pueda hacer para salvar los datos, y los métodos de esta sección no deberían usarse. Si los datos son una muestra aleatoria simple pero se infringe la segunda condición, existen métodos que podrían utilizarse, como los siguientes tres métodos alternativos:

- **Remuestreo bootstrap** Use el método del intervalo de confianza para pruebas de hipótesis, pero obtenga el intervalo de confianza usando el remuestreo bootstrap, como se describe en la sección 7-4. Tenga cuidado de usar el nivel de confianza apropiado, como se indica en la tabla 8-1 de la página 360. Rechace la hipótesis nula si los límites del intervalo de confianza no contienen el valor de la media declarada en la hipótesis nula. Vea el ejemplo 6.
- **Prueba del signo** Vea la sección 13-2.
- **Prueba de rangos con signo de Wilcoxon** Vea la sección 13-3.

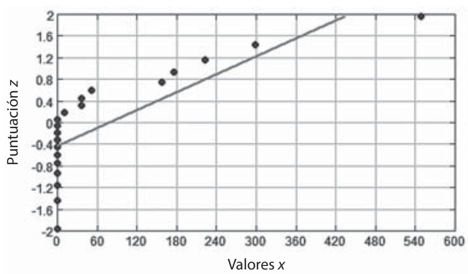
EJEMPLO 6 Remuestreo Bootstrap

A continuación se lista una muestra aleatoria de los tiempos (en segundos) en los que se consume tabaco en películas animadas para niños (del conjunto de datos 11 “Alcohol y tabaco en películas” del apéndice B). Use un nivel de significancia de 0.05 para evaluar la afirmación de que la muestra proviene de un población con una media mayor a 1 minuto o 60 segundos.

0 223 0 176 0 548 0 37 158 51 0 0 299 37 0 11 0 0 0 0

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS La prueba *t* descrita en la parte 1 de esta sección requiere que la población se distribuya normalmente o que $n > 30$, pero tenemos $n = 20$ y la gráfica cuantilar normal que la acompaña muestra que la muestra no parece proceder de una población normalmente distribuida. La prueba *t no* se debe utilizar.



En lugar de utilizar incorrectamente la prueba *t*, utilizamos el método de remuestreo bootstrap descrito en la sección 7-4. Después de obtener 1000 muestras bootstrap y encontrar la media de cada muestra, ordenamos las medias. Debido a que la prueba tiene una cola correcta con un nivel de significancia de 0.05, usamos las 1000 medias muestrales ordenadas para encontrar los límites del intervalo de confianza del 90% para $P_5 = 29.9$ segundos y $P_{95} = 132.9$ segundos. (Estos valores pueden variar un poco). El intervalo de confianza del 90% es 29.9 segundos $< \mu <$ 132.9 segundos. Como la media supuesta de 60 segundos está dentro de esos límites del intervalo de confianza, no podemos rechazar H_0: $\mu = 60$ segundos. No hay suficiente evidencia para respaldar H_1: $\mu > 60$ segundos.

INTERPRETACIÓN

No hay evidencia suficiente para respaldar la afirmación de que la muestra dada procede de una población con una media superior a 60 segundos.

PARTE 2 Prueba de una hipótesis acerca de μ cuando se conoce σ

En realidad, es muy raro probar una hipótesis sobre una media poblacional desconocida mientras se conoce de alguna manera la desviación estándar poblacional. Para este caso, el procedimiento es esencialmente el mismo que en la parte 1 de esta sección; los requisitos son los mismos, pero el estadístico de prueba, el valor *P* y los valores críticos se obtienen de la siguiente manera.

ELEMENTOS CLAVE

Prueba de una hipótesis sobre una media (cuando se conoce $\boldsymbol{\sigma}$)

Estadístico de prueba

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Valor *P*: Lo proporciona el software, o utilice la distribución normal estándar (tabla A-2) con el procedimiento de la figura 8-3 en la página 364.

Valores críticos: Use la distribución normal estándar (tabla A-2).

Si repetimos el ejemplo 1 “Sueño en adultos” con el supuesto de que se conoce el valor de $\sigma = 1.99240984$ horas, el dato estadístico de prueba es

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{6.83333333 - 7}{\frac{1.99240984}{\sqrt{12}}} = -0.29$$

Mediante el uso de este dato estadístico de prueba $z = -0.29$, podemos proceder a encontrar el valor *P*. (Consulte la figura 8-3 en la página 364 para ver el diagrama de flujo que resume el procedimiento para encontrar valores *P*). El ejemplo 1 se refiere a una prueba de cola izquierda, por lo que el valor *P* es el área a la *izquierda* de $z = -0.29$, que es 0.3859 (encontrada en la tabla A-2). Debido a que el valor *P* de 0.3859 es mayor que el nivel de significancia de $\alpha = 0.05$, no rechazamos la hipótesis nula, como lo hicimos en el ejemplo 1. Las conclusiones del ejemplo 1 son las mismas aquí.

Con σ conocida, obtenemos el valor $P = 0.3859$, pero con σ desconocida, el valor $P = 0.3887$. Con σ conocida, la distribución normal (z) proporciona un valor *P* más pequeño, por lo que los datos muestrales poseen una mayor significancia.

CENTRO DE TECNOLOGÍA

Prueba de hipótesis: Media

Acceda a complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y elija **Mean-One Sample** en el submenú.
3. Seleccione el formato deseado para la *hipótesis alternativa*, ingrese el nivel de significancia deseado y la media declarada (de la hipótesis nula).
4. *Si se usan estadísticos de resumen*: Seleccione la pestaña **Use Summary Statistics** e ingrese el tamaño de muestra, la media muestral y la desviación estándar muestral.

   Si se *usan datos muestrales*: Seleccione la pestaña **Use Data** y seleccione la columna de datos deseada.
5. Haga clic en **Evaluate**.

CENTRO DE TECNOLOGÍA *continuación*

Prueba de hipótesis: Media

Acceda a los complementos técnicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **1-Sample t** del submenú.
3. *Si se usan estadísticos de resumen:* Seleccione **Summarized Data** del menú desplegable e ingrese el tamaño de muestra, la media muestral y la desviación estándar muestral.

   Si se usan datos muestrales: Seleccione **One or more samples, each in a column** del menú desplegable y seleccione la columna de datos deseada.
4. Marque la casilla **Perform Hypothesis Test** e ingrese la media utilizada en la hipótesis nula.
5. Haga clic en el botón **Options** e ingrese el nivel de confianza. (Introduzca 95.0 para un nivel de significancia de 0.05). Para *Alternative Hypothesis*, seleccione el formato utilizado para la hipótesis alternativa.
6. Haga clic en **OK** dos veces.

SUGERENCIA: Otro procedimiento es hacer clic en **Assistant** en el menú superior, luego seleccione **Hypothesis Tests** y **1-Sample t**. Complete el cuadro de diálogo para obtener resultados, incluido el valor *P* y otra información útil.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **T Stats** del menú desplegable, luego elija **One Sample** en el submenú.
3. *Si se usan estadísticos de resumen:* Seleccione **With Summary** en el submenú e ingrese la media muestral, la desviación estándar muestral y el tamaño de muestra.

   Si se usan datos muestrales: Seleccione **With Data** en el submenú y elija la columna de datos deseada.
4. Seleccione **Hypothesis test for $\boldsymbol{\mu}$**, y para H_0 ingrese el valor declarado de la media poblacional (a partir de la hipótesis nula). Para H_A, seleccione el formato utilizado para la hipótesis alternativa.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **T-Test** en el menú y presione ENTER.
3. Introduzca el valor declarado de la media poblacional μ_0.
4. *Si se usan estadísticos de resumen:* Seleccione **Stats**, presione ENTER e ingrese la media muestral $\overline{x}$, la desviación estándar muestral *Sx* y el tamaño de muestra *n*.

   Si se usan datos muestrales: Seleccione **Data**, presione ENTER e ingrese el nombre de la lista que contiene los datos muestrales. *Freq* se debe establecer en 1.
5. Para μ, seleccione el formato utilizado para la hipótesis alternativa.
6. Seleccione **Calculate** y presione ENTER.

Excel

Complemento XLSTAT (Requerido)

Requiere los datos muestrales originales; no funciona con datos resumidos.

1. Haga clic en la pestaña **XLSTAT** en la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **One-Sample t-test and z-test** en el menú desplegable.
3. En *Data* ingrese el rango de celdas que contiene los datos muestrales. Para *Data format*, seleccione **One Sample**. Si la primera fila de datos contiene una etiqueta, también marque la casilla de **Column labels**.
4. Marque la casilla de **Student's t test**.
5. Haga clic en la pestaña **Options**.
6. En *Alternative hypothesis* seleccione el formato deseado ($\neq$ para pruebas de dos colas, $<$ para pruebas de cola izquierda, $>$ para pruebas de cola derecha).
7. Para *Theoretical Mean*, ingrese el valor declarado de la media poblacional (a partir de la hipótesis nula). Ingrese el nivel de significancia deseado (introduzca **5** para un nivel de significancia de 0.05).
8. Haga clic en **OK** para desplegar los resultados. El dato estadístico de prueba se identifica como *t* (valor observado) o *z* (valor observado). El valor *P* y el(los) valor(es) crítico(s) también se muestra(n).

8-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Videojuegos: verificación de requisitos Se observaron doce videojuegos diferentes que muestran uso de alcohol. Se registraron los tiempos (en segundos) de duración del consumo de alcohol que se detallan a continuación (según datos de "Content and Ratings of Teen-Rated Video Games", de Haninger y Thompson, *Journal of the American Medical Association*, vol. 291, núm. 7). ¿Qué requisitos deben cumplirse para probar la hipótesis de que la muestra proviene de una población con una media superior a 90 segundos? ¿Se satisfacen todos los requisitos?

84 14 583 50 0 57 207 43 178 0 2 57

2. gl Si utilizamos los datos muestrales del ejercicio 1 para una prueba t de la hipótesis de que la media poblacional es mayor que 90 segundos, ¿qué expresa gl y cuál es su valor?

3. Prueba *t* El ejercicio 2 se refiere a una prueba t. ¿Qué es una prueba t? ¿Por qué se usa la letra t? ¿Qué es poco realista acerca de los métodos de prueba z en la parte 2 de esta sección?

4. Intervalo de confianza Suponga que usaremos los datos muestrales del ejercicio 1 "Videojuegos" con un nivel de significancia de 0.05 en una prueba de la hipótesis de que la media poblacional es mayor que 90 segundos. Si queremos elaborar un intervalo de confianza que se utilizará para probar esa hipótesis, ¿qué nivel de confianza se debe emplear para el intervalo de confianza? Si se encuentra que el intervalo de confianza es 21.1 seg. $< \mu <$ 191.4 seg., ¿qué deberíamos concluir sobre la afirmación?

Determinación de valores *P*. ***En los ejercicios 5 a 8, utilice la tecnología para encontrar el valor P o use la tabla A-3 para determinar un rango de valores para el valor P.***

5. Velocidades de transferencia de datos en aeropuertos Se afirma que la media de las velocidades de datos de Verizon en los aeropuertos es $\mu = 14.00$ mbps. El tamaño de muestra es $n = 13$ y el dato estadístico de prueba es $t = -1.625$.

6. Temperaturas corporales Se afirma que para las temperaturas corporales de las 12 AM, la media es $\mu < 98.6$ ° F. El tamaño de muestra es $n = 4$ y el dato estadístico de prueba es $t = -2.503$.

7. Old Faithful Se afirma que el tiempo medio de duración (en segundos) de las erupciones del géiser Old Faithful es $\mu = 249$ seg. El tamaño de muestra es $n = 6$ y el dato estadístico de prueba es $t = 1.340$.

8. Tornados Se afirma que la anchura media (en yardas) de los tornados es $\mu < 140$ yd. El tamaño de muestra es $n = 21$ y el dato estadístico de prueba es $t = -0.024$.

Tecnología. ***En los ejercicios 9 a 12, pruebe la afirmación dada utilizando la pantalla proporcionada por una tecnología. Use un nivel de significancia de 0.05. Identifique las hipótesis nula y alternativa, el estadístico de prueba, el valor P (o el rango de valores P), o los valores críticos, y establezca una conclusión final que aborde la afirmación original.***

Para el ejercicio 9

```
NORMAL FLOAT AUTO REAL RADIAN MP
T-Test
μ<4
t=-.3662917532
p=.3578621222
x̄=3.71
Sx=5.598296024
n=50
```

9. Velocidades de transferencia de datos en aeropuertos El conjunto de datos 32 "Velocidades de transferencia datos en aeropuertos" en el apéndice B incluye velocidades de datos de Sprint (en mbps). La pantalla adjunta de una calculadora TI-83/84 Plus resulta del uso de esos datos para probar la hipótesis de que provienen de una población con una media inferior a 4.00 Mbps. Lleve a cabo la prueba de hipótesis empleando estos resultados.

Para el ejercicio 10

Difference	-0.476
t (Observed value)	-7.102
|t| (Critical value)	1.986
DF	92
p-value (Two-tailed)	< 0.0001
alpha	0.05

10. Temperaturas corporales El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye 93 temperaturas corporales medidas a las 12 AM. del día 1 de un estudio, y la pantalla adjunta de XLSTAT resulta del uso de esos datos para probar la hipótesis de que la temperatura corporal media es igual a 98.6 °F . Lleve a cabo la prueba de hipótesis empleando estos resultados.

Para el ejercicio 11

t Test	
Test Statistic, t:	−8.7201
Critical t:	±1.9695
P-Value:	0.0000

11. Old Faithful El conjunto de datos 23 "Old Faithful" en el apéndice B incluye información de 250 erupciones aleatorias del géiser Old Faithful. El Servicio Nacional de Parques hace predicciones de los tiempos para las próximas erupciones, y el conjunto de datos incluye los errores (en minutos) de esas predicciones. La pantalla adjunta de Statdisk resulta del uso de los errores de predicción para probar la hipótesis de que el error promedio es igual a cero. Comente sobre la precisión de las predicciones.

12. Tornados El conjunto de datos 22 "Tornados" en el apéndice B incluye datos de 500 tornados aleatorios. La pantalla adjunta de StatCrunch resulta del uso de las longitudes de los tornados (en millas) para probar la hipótesis de que la longitud media de los tornados es mayor a 2.5 millas.

Hypothesis test results:
μ : Mean of variable
$H_0 : \mu = 2.5$
$H_A : \mu > 2.5$

Variable	Sample Mean	Std. Err.	DF	T-Stat	P-value
Length	2.72424	0.29557683	499	0.75865215	0.2242

Pruebas de hipótesis. ***En los ejercicios 13 a 24, suponga que se ha seleccionado una muestra aleatoria simple y pruebe la afirmación dada. (A menos que su profesor indique otra cosa, use el método del valor P o el método del valor crítico para pruebas de hipótesis. Identifique las hipótesis nula y alternativa, el dato estadístico de prueba, el valor P (o el rango de valores P) o los valores críticos, y redacte una conclusión final que aborde la afirmación original.***

13. Evaluaciones de cursos El conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B incluye datos de evaluaciones de estudiantes a sus cursos. Los estadísticos de resumen son $n = 93$, $\bar{x} = 3.91$, $s = 0.53$. Use un nivel de significancia de 0.05 para probar la hipótesis de que la población de las evaluaciones de estudiantes a sus cursos tiene una media igual a 4.00.

14. Citas rápidas El conjunto de datos 18 "Citas rápidas" en el apéndice B incluye las calificaciones de "atractivo" para las citas masculinas expresadas por las citas femeninas. Los estadísticos de resumen son $n = 199$, $\bar{x} = 6.19$, $s = 1.99$. Use un nivel de significancia de 0.01 para probar la hipótesis de que la media poblacional de dichas calificaciones es menor a 7.00.

15. Ajo para reducir el colesterol En una prueba de la efectividad del ajo para reducir el colesterol, 49 sujetos fueron tratados con ajo crudo. Los niveles de colesterol se midieron antes y después del tratamiento. Los cambios (antes menos después) en sus niveles de colesterol LDL (en mg/dL) tienen una media de 0.4 y una desviación estándar de 21.0 (según datos de "Effect of Raw Garlic *vs* Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults with Moderate Hypercholesterolemia", de Gardner *et al*., *Archives of Internal Medicine*, vol. 167, núm. 4). Use un nivel de significancia de 0.05 para probar la hipótesis de que, mediante el tratamiento con ajo, el cambio promedio en el colesterol LDL es mayor que 0. ¿Qué sugieren los resultados sobre la efectividad del tratamiento con ajo?

16. Profundidades de epicentros de terremotos El conjunto de datos 21 "Terremotos" en el apéndice B lista las profundidades de terremotos, y los estadísticos de resumen son $n = 600$, $\bar{x} = 5.82$ km, $s = 4.93$ km. Use un nivel de significancia de 0.01 para evaluar la afirmación de un sismólogo de que estos terremotos son una población con una media igual a 5.00 km.

17. ¿Es la dieta práctica? Cuando 40 personas usaron la dieta Weight Watchers durante un año, su pérdida de peso promedio fue de 3.0 lb y la desviación estándar fue de 4.9 lb (según datos de "Comparison of the Atkins, Ornish, Weight Watchers y Zone Diets for Weight Loss and Heart Disease Reduction", de Dansingel *et al*., *Journal of the American Medical Association*, vol. 293, núm. 1). Use un nivel de significancia de 0.01 para evaluar la afirmación de que la pérdida de peso promedio es mayor que 0. De acuerdo con estos resultados, ¿la dieta parece tener significancia estadística? ¿La dieta parece tener una significancia práctica?

18. ¿Cuántas palabras en inglés? Se obtiene una muestra aleatoria simple de 10 páginas del *Merriam-Webster's Collegiate Dictionary*. Se determinan las cantidades de palabras definidas en esas páginas, con los siguientes resultados: $n = 10$, $\bar{x} = 53.3$ palabras, $s = 15.7$ palabras. Dado que este diccionario tiene 1459 páginas llenas de definiciones de palabras, la afirmación de que hay más de 70,000 palabras definidas es equivalente a la afirmación de que la cantidad media de palabras por página es mayor que 48.0. Suponga una población distribuida normalmente. Use un nivel de significancia de 0.01 para probar la hipótesis de que el número medio de palabras por página es mayor que 48.0. ¿Qué sugiere el resultado sobre la afirmación de que hay más de 70,000 palabras definidas?

19. Latas de Coca-Cola El conjunto de datos 26 "Pesos y volúmenes de bebidas de cola" en el apéndice B incluye los volúmenes (onzas) de una muestra de latas de Coca-Cola regular. Los estadísticos de resumen son $n = 36$, $\bar{x} = 12.19$ oz, $s = 0.11$ oz. Use un nivel de significancia de 0.05 para probar la hipótesis de que las latas de Coca-Cola tienen un volumen promedio de 12.00 onzas. ¿Parece que los consumidores están siendo engañados?

20. Tratamiento del insomnio Se realizó un ensayo clínico para evaluar la efectividad del medicamento zopiclona para tratar el insomnio en sujetos mayores. Antes del tratamiento con zopiclona, 16 sujetos tenían un tiempo de vigilia medio de 102.8 minutos. Después del tratamiento con zopiclona, los 16 sujetos tuvieron un tiempo de vigilia medio de 98.9 minutos y una desviación estándar de 42.3 minutos (según datos de "Cognitive Behavioral Therapy *vs* Zoplicone for Treatment of Chronic Primary Insomnia in Older Adults", de Sivertsen *et al*., *Journal of the American Medical Association*, vol. 295, núm. 24). Suponga que los 16 valores muestrales parecen provenir de una población distribuida normalmente, y pruebe la afirmación de que después del tratamiento con zopiclona, los sujetos tienen un tiempo de vigilia medio menor que 102.8 min. ¿Parece que la zopiclona es efectiva?

21. Plomo en medicinas A continuación se listan las concentraciones de plomo (en μg/g) medidas en diferentes medicamentos ayurvédicos. Ayurveda es un sistema médico tradicional comúnmente utilizado en la India. Las concentraciones de plomo listadas aquí provienen de medicamentos fabricados en Estados Unidos (según datos de "Lead, Mercury, and Arsenic in US and Indian Manufactured Ayurvedic Medicines Sold via the Internet", de Saper *et al*., *Journal of the American Medical Association*), vol. 300, núm. 8). Use un nivel de significancia de 0.05 para evaluar la afirmación de que la concentración media de plomo para todos estos medicamentos es menor a 14 μg/g.

3.0 6.5 6.0 5.5 20.5 7.5 12.0 20.5 11.5 17.5

22. ¿Tienes un minuto? Algunos alumnos del autor estimaron la duración de un minuto sin usar la referencia de un reloj y los tiempos (en segundos) se listan a continuación. Use un nivel de significancia de 0.05 para probar la hipótesis de que estos tiempos provienen de una población con una media igual a 60 segundos. ¿Parece que los estudiantes son razonablemente buenos para estimar un minuto?

69 81 39 65 42 21 60 63 66 48 64 70 96 91 65

23. Asientos para niños en el automóvil La Administración Nacional de Seguridad del Tráfico en Carreteras realizó pruebas de choque para los asientos para niños en los automóviles. A continuación se listan los resultados de esas pruebas, con las medidas indicadas en clc (unidades estándar de condición de lesión en la cabeza). El requisito de seguridad es que la medida de clc sea inferior a 1000 clc. Use un nivel de significancia de 0.01 para probar la hipótesis de que la muestra proviene de una población con una media inferior a 1000 clc. ¿Los resultados sugieren que todos los asientos para niños cumplen con el requisito especificado?

774 649 1210 546 431 612

24. Estaturas de supermodelos A continuación se listan las estaturas (en cm) para la muestra aleatoria simple de las supermodelos Lima, Bundchen, Ambrosio, Ebanks, Iman, Rubik, Kurkova, Kerr, Kroes, Swanepoel, Prinsloo, Hosk, Kloss, Robinson, Heatherton y Refaeli. Use un nivel de significancia de 0.01 para probar la hipótesis de que las supermodelos tienen alturas con una media mayor que la altura promedio de 162 cm para las mujeres de la población general. Dado que sólo hay 16 alturas representadas, ¿podemos realmente concluir que las supermodelos son más altas que la mujer típica?

178 177 176 174 175 178 175 178 178 177 180 176 180 178 180 176

Conjuntos de datos grandes del apéndice B. ***En los ejercicios 25 a 28, use el conjunto de datos del apéndice B para probar la hipótesis dada. Use el método del valor P a menos que su profesor especifique lo contrario.***

25. Pulsos Use los pulsos de mujeres adultas listados en el conjunto de datos 1 "Datos corporales" para probar si la media es inferior a 75 lpm. Use un nivel de significancia de 0.05.

26. Magnitudes de terremotos Use las magnitudes de terremotos listadas en el conjunto de datos 21 "Terremotos" y pruebe la afirmación de que la muestra proviene de una población con una media superior a 2.50. Utilice un nivel de significancia de 0.01.

27. Presión sanguínea diastólica en mujeres Use las medidas de presión arterial diastólica para mujeres adultas listadas en el conjunto de datos 1 "Datos corporales" y pruebe la afirmación de que la población de mujeres adultas tiene un nivel de presión arterial diastólica promedio menor que 90 mm Hg. Una presión arterial diastólica superior a 90 se considera hipertensión. Use un nivel de significancia de 0.05. Con base en el resultado, ¿podemos concluir que ninguna de las mujeres adultas en la muestra tiene hipertensión?

28. Presión sanguínea diastólica en hombres Repita el ejercicio anterior pero ahora utilice los datos de los hombres adultos.

8-3 Mas allá de lo básico

29. Prueba de hipótesis con σ conocida ¿Cómo cambian los resultados del ejercicio 13 "Evaluaciones de cursos" si se sabe que $\sigma = 0.53$? ¿El conocimiento de σ tiene mucho efecto?

30. Prueba de hipótesis con σ conocida ¿Cómo cambian los resultados del ejercicio 14 "Citas rápidas" si se sabe que σ es 1.99? ¿El conocimiento de σ tiene mucho efecto?

31. Determinación de valores *t* críticos Al encontrar valores críticos, a menudo necesitamos niveles de significancia distintos a los disponibles en la tabla A-3. Algunos programas de computadora aproximan los valores t críticos calculando $t = \sqrt{\text{gl} \cdot (e^{A^2/\text{gl}} - 1)}$ donde gl $= n - 1$, $e = 2.718$, $A = z(8 \cdot \text{gl} + 3)/(8 \cdot \text{gl} + 1)$, y z es la puntuación z crítica. Utilice esta aproximación para encontrar la puntuación t crítica para el ejercicio 12 "Tornados", usando un nivel de significancia de 0.05. Compare los resultados con la puntuación t crítica de 1.648 encontrada mediante el uso de tecnología. ¿Esta aproximación parece funcionar razonablemente bien?

32. Interpretación de la potencia Para los datos muestrales del ejemplo 1 "Sueño en adultos" de esta sección, Minitab y StatCrunch muestran que la prueba de hipótesis tiene una potencia de 0.4943 que respalda la afirmación de que $\mu < 7$ horas de sueño, cuando la media real de la población es 6.0 horas de sueño. Interprete este valor de la potencia, luego identifique el valor de β e interprételo. (Para la prueba t en esta sección, un "parámetro de no centralidad" complica los cálculos de potencia nucho más que el proceso descrito en la sección 8-1, por ello se recomienda usar software para realizar cálculos de potencia)

8-4 Prueba de una hipótesis respecto a una desviación estándar o varianza

Concepto clave Esta sección presenta métodos para llevar a cabo una prueba de hipótesis formal de una afirmación hecha sobre una desviación estándar σ o una varianza σ^2 poblacionales. Los métodos de esta sección usan la distribución ji cuadrada que se introdujo por primera vez en la sección 7-3. Los supuestos, el dato estadístico de prueba, el valor P y los valores críticos se resumen de la siguiente manera.

ELEMENTOS CLAVE

Pruebas de hipótesis sobre σ o σ^2

Objetivo

Realizar una prueba de hipótesis de una afirmación hecha sobre una desviación estándar σ o una varianza σ^2 poblacionales

Notación

n = tamaño de muestra

s = desviación estándar *muestral* — s^2 = varianza *muestral*

σ = desviación estándar *poblacional* — σ^2 = varianza *poblacional*

Requisitos

1. La muestra es una muestra aleatoria simple.
2. La población tiene una distribución normal. (Este es un requisito bastante estricto).

Dato estadístico de prueba

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$ (redondee a tres lugares decimales, como en la tabla A-4)

Valores *P*: Utilice la tecnología o la tabla A-4 con los grados de libertad: gl $= n - 1$.

Valores críticos: Use la tabla A-4 con los grados de libertad gl $= n - 1$.

PRECAUCIÓN La prueba χ^2 (ji cuadrada) de esta sección no es *sólida* frente a una desviación de la normalidad, lo que significa que la prueba no funciona bien si la población tiene una distribución que está lejos de ser normal. Por lo tanto, la condición de una población normalmente distribuida es un requisito mucho más estricto cuando se prueban hipótesis sobre σ o σ^2 que cuando se hace esto para una media poblacional μ.

Métodos equivalentes

Al probar hipótesis sobre σ o σ^2, el método del valor P, el método del valor crítico y el método del intervalo de confianza son equivalentes en el sentido de que siempre conducirán a la misma conclusión.

Propiedades de la distribución ji cuadrada

La distribución ji cuadrada se introdujo en la sección 7-3, donde notamos las siguientes propiedades importantes.

1. Todos los valores de χ^2 son no negativos, y la distribución no es simétrica (vea la figura 8-8).
2. Hay una distribución χ^2 diferente para cada número de grados de libertad (vea la figura 8-9).
3. Los valores críticos se encuentran en la tabla A-4 usando

$$\textbf{grados de libertad} = n - 1$$

La siguiente es una nota importante si se usa la tabla A-4 para encontrar valores críticos:

En la tabla A-4, cada valor crítico de χ^2 en el cuerpo de la tabla corresponde a un área dada en la fila superior de la tabla, y cada área en esa fila es un *área acumulada a la derecha* del valor crítico.

PRECAUCIÓN La tabla A-4 para la distribución ji cuadrada utiliza áreas acumuladas desde la *derecha* (a diferencia de la tabla A-2 para la distribución normal estándar, que proporciona áreas acumuladas desde la *izquierda*). Consulte el ejemplo 1 en la sección 7-3.

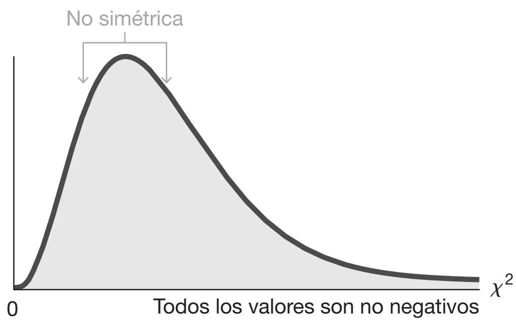


FIGURA 8-8 Propiedades de la distribución ji cuadrada

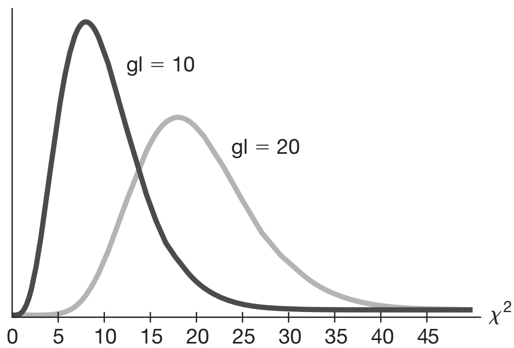


FIGURA 8-9 Distribución ji cuadrada para gl = 10 y gl = 20

EJEMPLO 1 Método del valor *P*: ¿Las estaturas varían menos entre las supermodelos?

A continuación se listan las estaturas (en cm) de la muestra aleatoria simple de las supermodelos Lima, Bundchen, Ambrosio, Ebanks, Iman, Rubik, Kurkova, Kerr, Kroes, Swanepoel, Prinsloo, Hosk, Kloss, Robinson, Heatherton y Refaeli. Use un nivel de significancia de 0.01 para probar la hipótesis de que las supermodelos tienen estaturas con una desviación estándar menor que $\sigma = 7.5$ cm (desviación estándar de la población femenina general). ¿Parece que las alturas de las supermodelos varían menos que las alturas de las mujeres en general?

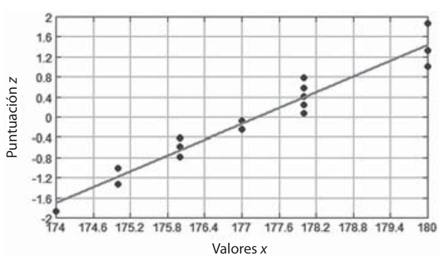
178 177 176 174 175 178 175 178 178 177 180 176 180 178 180 176

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) La muestra es una muestra aleatoria simple. (2) Al verificar la normalidad, vemos que la muestra no tiene valores atípicos, la gráfica cuantilar normal adjunta muestra puntos que están razonablemente cerca de un patrón en línea recta, y no hay otro patrón que no esté en esa misma disposición. Se satisfacen ambos requisitos.

Tecnología La tecnología capaz de realizar esta prueba generalmente mostrará el valor *P*. Es posible usar StatCrunch como se describe al final de esta sección, y el resultado será el que se muestra en la pantalla adjunta. (En vez de usar el valor supuesto de σ para H_0 y H_1, StatCrunch usa σ^2. Para la hipótesis nula, $\sigma = 7.5$ es equivalente a $\sigma^2 = 7.5^2 = 56.25$). La pantalla muestra que el dato estadístico de prueba es $\chi^2 = 0.907$ (redondeado) y el valor *P* es menor que 0.0001.

Statdisk



StatCrunch

Hypothesis test results:
σ^2 : Variance of population
$H_0 : \sigma^2 = 56.25$
$H_A : \sigma^2 < 56.25$

Variance	Sample Var.	DF	Chi-Square Stat	P-value
σ2	3.4000004	15	0.90666677	<0.0001

Paso 1: La afirmación de que "la desviación estándar es inferior a 7.5 cm" se expresa en forma simbólica como $\sigma < 7.5$ cm.

Paso 2: Si la afirmación original es falsa, entonces $\sigma \geq 7.5$ cm.

Paso 3: La expresión $\sigma < 7.5$ cm no contiene igualdad, por lo que se convierte en la hipótesis alternativa. La hipótesis nula es la afirmación de que $\sigma = 7.5$ cm.

$$H_0: \sigma = 7.5 \text{ cm}$$

$$H_1: \sigma < 7.5 \text{ cm (afirmación original)}$$

Paso 4: El nivel de significancia es $\alpha = 0.01$.

Paso 5: Dado que la afirmación se realiza sobre σ, usamos la distribución χ^2 (ji cuadrada).

Paso 6: La pantalla de StatCrunch muestra el dato estadístico de prueba $\chi^2 = 0.907$ (redondeado) y muestra que el valor *P* es menor que 0.0001.

Paso 7: Como el valor *P* es menor que el nivel de significancia $\alpha = 0.01$, rechazamos H_0.

continúa

INTERPRETACIÓN

Paso 8: Hay evidencia suficiente para respaldar la afirmación de que las supermodelos tienen estaturas con una desviación estándar inferior a los 7.5 cm para la población general de mujeres. Parece que las alturas de las supermodelos varían menos que las alturas de las mujeres en la población general.

SU TURNO Resuelva el ejercicio 7 "Temperatura corporal".

Método del valor crítico

Por lo general, la tecnología proporciona un valor P, por lo que se usa el método del valor P. Si no dispone de una tecnología, el método del valor P para pruebas de hipótesis es un poco desafiante porque la tabla A-4 nos permite encontrar sólo un rango de valores para el valor P. En cambio, podríamos usar el método del valor crítico. Los pasos 1 a 5 en el ejemplo 1 serían iguales. En el paso 6, el dato estadístico de prueba se calcula usando $\sigma = 7.5$ cm (como se supone en la hipótesis nula del ejemplo 1), $n = 16$ y $s = 1.843909$ cm, que es la desviación estándar no redondeada calculada a partir de la lista original de 16 estaturas. Obtenemos el siguiente estadístico de prueba:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(16-1)\ (1.843909)^2}{7.5^2} = 0.907$$

El valor crítico $\chi^2 = 5.229$ se encuentra en la tabla A-4, y corresponde a 15 grados de libertad y un "área a la derecha" de 0.99 (con base en el nivel de significancia de 0.01 para una prueba de cola izquierda). Vea la figura 8-10. En el paso 7, rechazamos la hipótesis nula porque el dato estadístico de prueba $\chi^2 = 0.907$ cae en la región crítica, como se muestra en la figura 8-10. Concluimos que hay suficiente evidencia para respaldar la afirmación de que las supermodelos tienen estaturas con una desviación estándar menor que los 7.5 cm para la población general de mujeres.

Método del intervalo de confianza

Como se indicó anteriormente, cuando se prueban hipótesis sobre σ o σ^2, el método del valor P, el método del valor crítico y el método del intervalo de confianza son todos equivalentes en el sentido de que siempre conducirán a la misma conclusión. Vea el ejemplo 2.

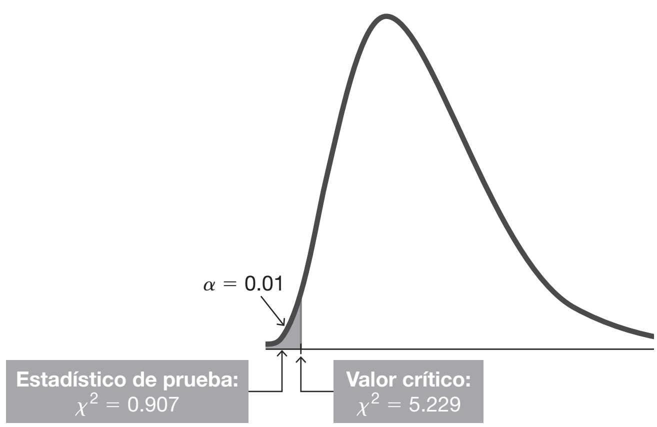


FIGURA 8-10 Prueba de la hipótesis $\sigma < 7.5$ cm

EJEMPLO 2 Estaturas de supermodelos: método del intervalo de confianza

Repita la prueba de hipótesis del ejemplo 1 elaborando un intervalo de confianza adecuado.

SOLUCIÓN

Primero, debemos tener cuidado de seleccionar el nivel de confianza correcto. Debido a que la prueba de hipótesis es de cola izquierda y el nivel de significancia es 0.01, deberíamos usar un nivel de confianza del 98% o 0.98. (Consulte la tabla 8-1 en la página 360 para obtener ayuda en la selección del nivel de confianza correcto).

Mediante el uso de los métodos descritos en la sección 7-3, podemos usar los datos muestrales listados en el ejemplo 1 para elaborar una estimación del intervalo de confianza del 98% para σ. Usamos $n = 16$, $s = 1.843909$ cm, $\chi_L^2 = 5.229$ y $\chi_R^2 = 30.578$. (Los valores críticos χ_L^2 y χ_R^2 se encuentran en la tabla A-4. Use la fila con gl $= n - 1 = 15$. El nivel de confianza de 0.98 corresponde a $\alpha = 0.02$, y dividimos esa área de 0.02 equitativamente entre las dos colas, de modo que las áreas a la *derecha* de los valores críticos sean 0.99 y 0.01. Consulte la tabla A-4 y use las columnas con áreas de 0.99 y 0.01, así como la fila 15).

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(16-1)(1.843909^2)}{30.578}} < \sigma < \sqrt{\frac{(16-1)(1.843909^2)}{5.229}}$$

$$1.3 \text{ cm} < \sigma < 3.1 \text{ cm}$$

Con este intervalo de confianza, podemos respaldar la afirmación de que $\sigma < 7.5$ cm porque todos los valores del intervalo de confianza son menores a 7.5 cm. Llegamos a la misma conclusión encontrada con los métodos del valor P y del valor crítico.

Método alternativo utilizado cuando la población no es normal

Los métodos de esta sección incluyen dos requisitos: (1) La muestra es una muestra aleatoria simple; (2) la población se distribuye normalmente. Si los datos muestrales no se recolectan de forma aleatoria, los métodos de esta sección no son aplicables. Si la muestra parece ser de una población que no tiene una distribución normal, podríamos usar el método del intervalo de confianza para probar las hipótesis, pero obtendríamos el intervalo de confianza mediante el remuestreo bootstrap, como se describe en la sección 7-4. Tenga cuidado de usar el nivel de confianza apropiado, como lo indica la tabla 8-1 en la página 360. Rechace la hipótesis nula si los límites del intervalo de confianza no contienen el valor de la media declarada en la hipótesis nula. Vea el proyecto de tecnología al final de este capítulo.

CENTRO DE TECNOLOGÍA

Prueba de hipótesis: Desviación estándar o varianza

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y seleccione **Standard Deviation One Sample** en el submenú.
3. Seleccione el formato deseado para la *Hipótesis Alternativa*, luego ingrese el nivel de significancia deseado y la desviación estándar declarada (a partir de la hipótesis nula).
4. *Si se usan estadísticos de resumen:* Seleccione la pestaña **Use Summary Statistics** e ingrese el tamaño de muestra y la desviación estándar muestral.

   Si se usan datos muestrales: Seleccione la pestaña **Use Data** y elija la columna de datos deseada.
5. Haga clic en **Evaluate.**

CENTRO DE TECNOLOGÍA *continuación*

Prueba de hipótesis: Desviación estándar o varianza

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **1 Variance** en el submenú.
3. *Si se usan estadísticos de resumen:* Seleccione **Sample standard deviation** (o **variance**) del menú desplegable e ingrese el tamaño de muestra y la desviación estándar (o varianza) muestral.

   Si se usan datos muestrales: Seleccione **One or more samples, each in a column** del menú desplegable y elija la columna de datos deseada.
4. Marque la casilla **Perform hypothesis test** e ingrese la desviación estándar (o varianza) utilizada en la hipótesis nula.
5. Haga clic en el botón **Options** e ingrese el nivel de confianza. (Introduzca 95.0 para un nivel de significancia de 0.05). Para *Alternative Hypothesis*, seleccione el formato utilizado para la hipótesis alternativa.
6. Haga clic en **OK** dos veces.

SUGERENCIA: Otro procedimiento es hacer clic en **Assistant** en el menú superior, luego seleccionar **Hypothesis Tests** y después **1-Sample Standard Deviation**. Complete el cuadro de diálogo para obtener los resultados, incluyendo el valor *P* y otra información útil.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Variance Stats** en el menú desplegable, luego elija **One Sample** en el submenú.
3. *Si se usan estadísticos de resumen:* Seleccione **With Summary** en el submenú e ingrese la varianza muestral y el tamaño de muestra.

   Si se usan datos muestrales: Seleccione **With Data** en el submenú y elija la columna de datos deseada.
4. Seleccione **Hypothesis test for σ^2** y para H_0 ingrese el valor declarado de la varianza poblacional. Para H_A, seleccione el formato utilizado para la hipótesis alternativa.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Descargue e instale los programas de Michael Lloyd *S2TEST* y *ZZINEWT* (disponibles en www.pearsonenespañol.com/triola) en su calculadora TI-83/84 Plus.
2. Presione el botón PRGM, luego seleccione **S2TEST** en el menú y presione ENTER dos veces.
3. Ingrese la varianza poblacional σx^2, la varianza muestral Sx^2 y el tamaño muestral n. Presione ENTER después de cada entrada.
4. Para SIGMA² seleccione el formato deseado para la hipótesis alternativa ($\neq$ para prueba de dos colas, $<$ para la prueba de cola izquierda, $>$ para la prueba de cola derecha).
5. Presione ENTER para obtener el dato estadístico de prueba y el valor *P*.

Excel

Complemento XLSTAT (requerido)

Requiere los datos muestrales originales; no funciona con datos resumidos.

1. Haga clic en la pestaña **XLSTAT** de la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **One-sample variance test** en el menú desplegable.
3. En *Data* ingrese el rango de celdas que contienen los datos muestrales. Para *Data format* seleccione **One column per sample**. Si la primera fila de datos contiene una etiqueta, también marque la casilla de **Column labels**.
4. Para *Range*, ingrese una ubicación de celda como D5 donde se mostrarán los resultados.
5. Haga clic en la pestaña **Options**.
6. En *Alternative hypothesis*, seleccione el formato deseado ($\neq$ para pruebas de dos colas, $<$ para pruebas de cola izquierda, $>$ para pruebas de cola derecha).
7. Para *Theoretical Variance*, ingrese el valor declarado de la varianza poblacional. Ingrese el nivel de significancia deseado (introduzca **5** para un nivel de significancia de 0.05).
8. Haga clic en **OK** para mostrar los resultados. El dato estadístico de prueba se identifica como ji cuadrada. También se muestra el valor *P*.

SUGERENCIA: El procedimiento anterior se basa en probar una hipótesis sobre una *varianza* poblacional; para probar una hipótesis sobre una desviación estándar poblacional, use el mismo procedimiento e ingrese σ^2 como la *varianza teórica*.

8-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Latas de Coca-Cola El conjunto de datos 26 "Pesos y volúmenes de bebidas de cola" en el apéndice B incluye los volúmenes (en oz) de la Coca-Cola regular. Con base en ese conjunto de datos, suponga que las latas se producen de manera que los volúmenes tengan una desviación estándar de 0.115 oz. Se está probando un nuevo proceso de llenado para las latas de Coca-Cola, y una muestra aleatoria de volúmenes se detalla a continuación. La muestra tiene los siguientes estadísticos de resumen: $n = 10$, $\bar{x} = 12.0004$ oz, $s = 0.268$ oz. Si queremos usar los datos muestrales para probar la hipótesis de que la muestra proviene de una población con una desviación estándar igual a 0.115 oz, ¿qué requisitos deben cumplirse? ¿De qué manera el requisito de normalidad para una prueba de hipótesis de una afirmación sobre una desviación estándar difiere del requisito de normalidad para una prueba de hipótesis de una afirmación sobre una media?

12.078 11.851 12.108 11.760 12.142 11.779 12.397 11.504 12.147 12.238

2. Latas de Coca-Cola Use los datos y la afirmación dados en el ejercicio 1 para identificar las hipótesis nula y alternativa, así como el estadístico de prueba. ¿Cuál es la distribución muestral del estadístico de prueba?

3. Latas de Coca-Cola Para los datos muestrales del ejercicio 1, obtenemos un "valor $P < 0.01$" al probar la hipótesis de que el nuevo proceso de llenado resulta en volúmenes con la misma desviación estándar de 0.115 oz.

a. ¿Qué deberíamos concluir sobre la hipótesis nula?

b. ¿Qué deberíamos concluir sobre la afirmación original?

c. ¿Qué sugieren estos resultados sobre el nuevo proceso de llenado?

4. Latas de Coca-Cola: intervalo de confianza Si usamos los datos proporcionados en el ejercicio 1, obtenemos la siguiente estimación del intervalo de confianza del 95% para la desviación estándar de los volúmenes con el nuevo proceso de llenado: 0.1846 oz $< \sigma <$ 0.4900 oz. ¿Qué nos dice este intervalo de confianza sobre el nuevo proceso de llenado?

Pruebas de hipótesis sobre la variación. ***En los ejercicios 5 a 16, pruebe la afirmación dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor P o los valores críticos, luego establezca una conclusión sobre la hipótesis nula, así como una conclusión final que aborde la afirmación original. Suponga que se selecciona una muestra aleatoria simple de una población distribuida normalmente.***

5. Pulsos de hombres Una muestra aleatoria simple de 153 hombres da como resultado una desviación estándar de 11.3 latidos por minuto (según el conjunto de datos 1 "Datos corporales" en el apéndice B). Comúnmente, el rango normal de los pulsos en adultos está entre 60 y 100 latidos por minuto. Si se aplica la regla práctica del rango a ese rango normal, el resultado es una desviación estándar de 10 latidos por minuto. Use los resultados muestrales con un nivel de significancia de 0.05 para probar la hipótesis de que los pulsos de los hombres tienen una desviación estándar de 10 latidos por minuto; vea la pantalla adjunta de StatCrunch para esta prueba. ¿Qué indican los resultados sobre la efectividad del uso de la regla práctica del rango con el "rango normal" de 60 a 100 latidos por minuto para estimar σ en este caso?

One sample variance hypothesis test:
σ^2 : Variance of variable
$H_0 : \sigma^2 = 100$
$H_A : \sigma^2 \neq 100$

Hypothesis test results:

Variable	Sample Var.	DF	Chi-Square Stat	P-value
PULSE	128.40282	152	195.17229	0.0208

6. Pulsos de mujeres Repita el ejercicio anterior utilizando los pulsos de mujeres listados en el conjunto de datos 1 "Datos corporales" del apéndice B. Para la muestra de los pulsos de las mujeres, $n = 147$ y $s = 12.5$. Consulte la pantalla adjunta de JMP que resulta del uso de la lista original de pulsos en lugar de los estadísticos de resumen. (*Sugerencia:* Las tres filas inferiores de la pantalla proporcionan valores P para una prueba de dos colas, una prueba de cola izquierda y una prueba de cola derecha, respectivamente). ¿Qué indican los resultados sobre la efectividad del uso de la regla práctica del rango con el "rango normal" de 60 a 100 latidos por minuto para estimar σ en este caso?

Test Standard Deviation

Hypothesized Value	10
Actual Estimate	12.5436
DF	146
Test	**ChiSquare**
Test Statistic	229.7176
Min PValue	<.0001*
Prob < ChiSq	1.0000
Prob > ChiSq	<.0001*

7. Temperatura corporal El ejemplo 5 en la sección 8-3 involucró una prueba de la hipótesis de que los humanos tienen temperaturas corporales con una media igual a 98.6 °F. La muestra de 106 temperaturas corporales tiene una desviación estándar de 0.62 °F. La conclusión en ese ejemplo cambiaría si la desviación estándar de la muestra fuera de 2.08 °F o mayor. Use un nivel de significancia de 0.01 para probar la hipótesis de que la muestra de 106 temperaturas corporales proviene de una población con una desviación estándar de menos de 2.08 °F. ¿Qué nos dice el resultado sobre la validez de la prueba de hipótesis en el ejemplo 5 de la sección 8-3?

8. Pesos al nacer Una muestra aleatoria simple de los pesos al nacer de 30 niñas tiene una desviación estándar de 829.5 hg. Use un nivel de significancia de 0.01 para probar la hipótesis de que el peso al nacer de las niñas tiene la misma desviación estándar que el peso al nacer de los niños, que es de 660.2 hg (con base en el conjunto de datos 4 "Nacimientos" del apéndice B).

9. Latas de bebida de cola Se selecciona una muestra aleatoria de 20 latas de aluminio para bebidas de cola con un espesor de 0.0109 pulgadas, se miden las cargas axiales y la desviación estándar es de 18.6 lb. La carga axial es la presión aplicada en la parte superior que hace que la lata se aplaste. Use un nivel de significancia de 0.05 para probar la hipótesis de que las latas con un espesor de 0.0109 pulgadas tienen cargas axiales con la misma desviación estándar que las cargas axiales de las latas que tienen un grosor de 0.0111 pulgadas. Las latas más gruesas tienen cargas axiales con una desviación estándar de 27.8 lb (según el conjunto de datos 30 "Latas de aluminio" en el apéndice B). ¿El espesor de las latas parece afectar la variación de las cargas axiales?

10. Resultados de exámenes de estadística Los exámenes en las clases de estadística del autor tienen calificaciones con una desviación estándar igual a 14.1. Una de sus últimas clases tuvo 27 calificaciones con una desviación estándar de 9.3. Use un nivel de significancia de 0.01 para probar la hipótesis de que esta clase tiene menos variación que otras clases en el pasado. ¿Una menor desviación estándar sugiere que esta última clase está mejorando?

11. Máquinas expendedoras de café La máquina expendedora Brazil dispensa café, y una muestra aleatoria de 27 tazas llenas tiene un contenido con una media de 7.14 oz y una desviación estándar de 0.17 oz. Use un nivel de significancia de 0.05 para probar la hipótesis de que la máquina dispensa cantidades con una desviación estándar mayor que la desviación estándar de 0.15 oz especificada en el diseño de la máquina.

12. Palabras habladas Se reclutaron parejas para realizar un estudio de cuántas palabras pronuncian las personas cada día. Una muestra aleatoria de 56 hombres resultó en una media de 16,576 palabras y una desviación estándar de 7871 palabras. Use un nivel de significancia de 0.01 para probar la hipótesis de que los varones tienen una desviación estándar que es mayor que la desviación estándar de 7460 palabras para las mujeres (según el conjunto de datos 24 "Las palabras cuentan").

13. Altímetros de aviones La empresa Skytek Avionics utiliza un nuevo método de producción para fabricar altímetros de aviones. Una muestra aleatoria simple de los nuevos altímetros resultó en los siguientes errores. Use un nivel de significancia de 0.05 para probar la hipótesis de que el nuevo método de producción tiene errores con una desviación estándar mayor a 32.2 pies, que fue la desviación estándar para el método de producción anterior. Si parece que la desviación estándar es mayor, ¿el nuevo método de producción parece ser mejor o peor que el método anterior? ¿Debería la empresa tomar alguna medida?

$$-42 \quad 78 \quad -22 \quad -72 \quad -45 \quad 15 \quad 17 \quad 51 \quad -5 \quad -53 \quad -9 \quad -109$$

14. Filas en los bancos Durante algún tiempo el banco Jefferson Valley tuvo una fila de espera de clientes en cada ventanilla de sus cajeros, pero ahora tiene una sola línea de espera que alimenta las ventanillas a medida que se desocupan. La desviación estándar de los tiempos de espera de los clientes con la configuración de línea múltiple anterior era de 1.8 minutos. A continuación se lista una muestra aleatoria simple de tiempos de espera (en minutos) con la línea de espera única. Use un nivel de significancia de 0.05 para probar la hipótesis de que, con una sola línea de espera, los tiempos tienen una desviación estándar de menos de 1.8 minutos. ¿Qué mejora ocurrió cuando los bancos cambiaron de varias colas a una sola línea de espera?

6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7

15. Tiempos de servicio para la comida rápida en auto A continuación se listan los tiempos de servicio en auto (segundos) registrados en McDonald's durante la hora de la cena (de acuerdo con el conjunto de datos 25 "Comida rápida" en el apéndice B). Si se supone que los tiempos de servicio durante la cena en Wendy's tienen una desviación estándar de $\sigma = -55.93$ segundos, use un nivel de significancia de 0.01 para probar la hipótesis de que los tiempos de servicio en McDonald's tienen la misma variación que los tiempos de servicio en Wendy's. ¿Debería McDonald's tomar alguna acción?

121 119 146 266 333 308 333 308

16. Especificaciones Mint A continuación se listan los pesos (en gramos) de una muestra aleatoria simple de centavos estadounidenses con el diseño del "trigo" (según el conjunto de datos 29 "Pesos de monedas" en el apéndice B). Las especificaciones Mint de EE.UU. ahora requieren una desviación estándar de 0.0230 gr para los pesos de los centavos. Use un nivel de significancia de 0.01 para probar la hipótesis de que los "centavos de trigo" se fabrican de modo que sus pesos tengan una desviación estándar igual a 0.0230 gr. ¿La especificación Mint parece cumplirse?

2.5024 2.5298 2.4998 2.4823 2.5163 2.5222 2.4900 2.4907 2.5017

Grandes conjuntos de datos del apéndice B. ***En los ejercicios 17 y 18, use el conjunto de datos del apéndice B para probar la hipótesis dada. Identifique la hipótesis nula, la hipótesis alternativa, el estadístico de prueba, el valor P o los valores críticos, luego establezca una conclusión sobre la hipótesis nula, así como una conclusión final que aborde la afirmación original.***

17. Tiempos de servicio para la comida rápida en auto Repita el ejercicio 15 usando los 50 tiempos de servicio en McDonald's durante las horas de cena en el conjunto de datos 25 "Comida rápida".

18. Especificaciones Mint Repita el ejercicio 16 usando los pesos de los 37 centavos posteriores a 1983 incluidos en el conjunto de datos 29 "Pesos de monedas" del apéndice B.

8-4 Más allá de lo básico

19. Determinación de valores críticos de χ^2 Para grandes cantidades de grados de libertad, podemos aproximar los valores críticos de χ^2 de la siguiente manera:

$$\chi^2 = \frac{1}{2}(z + \sqrt{2k - 1})^2$$

Aquí k es el número de grados de libertad y z es el valor crítico encontrado mediante la tecnología o con base en la tabla A-2. En el ejercicio 12 "Palabras habladas" tenemos gl = 55, por lo que la tabla A-4 no muestra un valor crítico exacto. Si queremos aproximar un valor crítico de χ^2 en la prueba de hipótesis de cola derecha con $\alpha = 0.01$ y un tamaño de muestra de 56, consideramos que $k = 55$ con $z = 2.33$ (o el valor más preciso de $z = 2.326348$ encontrado mediante la tecnología). Use esta aproximación para estimar el valor crítico de χ^2 en el ejercicio 12. ¿Qué tan cerca está del valor crítico de $\chi^2 = 82.292$ obtenido mediante el uso de Statdisk y Minitab?

20. Determinación de valores críticos de χ^2 Repita el ejercicio 19 usando la siguiente aproximación (con k y z como se describieron en el ejercicio 19):

$$\chi^2 = k\left(1 - \frac{2}{9k} + z\sqrt{\frac{2}{9k}}\right)^3$$

Examen rápido del capítulo

1. Distribuciones Usando los métodos de este capítulo, identifique la distribución que se debe usar para probar una hipótesis sobre el parámetro poblacional dado.

a. Media

b. Proporción

c. Desviación estándar

2. Colas Determine si la afirmación dada implica una prueba de hipótesis de cola izquierda, de dos colas o de cola derecha.

a. $p \neq 0.5$

b. $\mu < 98.6$ °F

c. $\sigma > 15$ cm

3. Encuesta de Instagram En una encuesta del Pew Research Center aplicada a usuarios de Internet con edades entre 18 y 29 años, 53% dijo que usaba Instagram. Queremos utilizar un nivel de significancia de 0.05 para probar la hipótesis de que la mayoría de los usuarios de Internet de entre 18 y 29 años de edad usan Instagram.

a. Identifique las hipótesis nula y alternativa.

b. Con un tamaño de muestra de 532, determine el valor del estadístico de prueba.

c. Se usa la tecnología para encontrar que el valor P para la prueba es 0.0827. ¿Qué deberíamos concluir sobre la hipótesis nula?

d. ¿Qué deberíamos concluir sobre la afirmación original?

4. Valor *P* Encuentre el valor P para una prueba de hipótesis de que el ingreso anual medio de un agente de la CIA es superior a \$81,623 (según datos de payscale.com) dado que el dato estadístico de prueba es $t = 1.304$ para una muestra de 40 agentes de la CIA.

5. Conclusiones Verdadero o falso: En las pruebas de hipótesis, nunca es válido llegar a una conclusión de respaldar la hipótesis nula.

6. Conclusiones Verdadero o falso: La conclusión de "no se rechaza la hipótesis nula" tiene exactamente el mismo significado que "se acepta la hipótesis nula".

7. Incertidumbre Verdadero o falso: Si se utilizan los métodos correctos para probar hipótesis con una gran muestra aleatoria simple que satisfaga los requisitos de la prueba, la conclusión siempre será verdadera.

8. Prueba ji cuadrada En una prueba de hipótesis de que $\sigma = 15$ para la población de puntuaciones de IQ de atletas profesionales, encontramos que el valor crítico más a la derecha es $\chi^2_R = 40.646$. ¿El valor crítico más a la izquierda χ^2_L es igual a -40.646?

9. Robustez Explique qué significan los siguientes enunciados: "La prueba t para una afirmación sobre μ es robusta, pero la prueba χ^2 para una afirmación sobre σ^2 no lo es".

10. Métodos equivalentes ¿Cuáles de las siguientes afirmaciones son verdaderas?

a. Cuando se prueba una hipótesis sobre una media poblacional μ, el método del valor P, el método del valor crítico y el método del intervalo de confianza son todos equivalentes en el sentido de que siempre arrojan las mismas conclusiones.

b. Cuando se prueba una hipótesis sobre una proporción poblacional p, el método del valor P, el método del valor crítico y el método del intervalo de confianza son todos equivalentes en el sentido de que siempre arrojan las mismas conclusiones.

c. Cuando se prueba una hipótesis sobre cualquier parámetro poblacional, el método del valor P, el método del valor crítico y el método del intervalo de confianza son todos equivalentes en el sentido de que siempre arrojan las mismas conclusiones.

Ejercicios de repaso

1. Verdadero/falso Caracterice cada una de las siguientes afirmaciones como verdaderas o falsas.

a. En una prueba de hipótesis, un valor P muy alto indica un fuerte respaldo a la hipótesis alternativa.

b. La distribución t de Student se puede usar para probar una hipótesis sobre una media poblacional siempre que los datos muestrales se seleccionen aleatoriamente de una población distribuida normalmente.

c. Cuando se usa una distribución χ^2 para probar una hipótesis acerca de una desviación estándar poblacional, existe un requisito muy poco estricto de que los datos muestrales provengan de una población que tiene una distribución normal.

d. Al realizar una prueba de hipótesis acerca de la proporción declarada de adultos que tienen pasaportes actualizados, los problemas con una muestra de conveniencia pueden superarse mediante el uso de un tamaño de muestra más grande.

e. Al repetir la misma prueba de hipótesis con diferentes muestras aleatorias del mismo tamaño, las conclusiones serán todas iguales.

2. Política Un empleado del condado de Essex, Nueva Jersey, seleccionó candidatos para puestos en las boletas electorales. Los demócratas fueron seleccionados primero en 40 de 41 boletas. Porque se suponía que debía usar un método de selección aleatorio, los republicanos afirmaron que, en lugar de utilizar la aleatoriedad, usó un método que favorecía a los demócratas. Emplee un nivel de significancia de 0.01 para probar la hipótesis de que el método de selección de boletas favorece a los demócratas.

3. Actrices ganadoras de un Oscar El conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B lista las edades de las actrices cuando ganaron el Oscar, y los estadísticos de resumen son $n = 87$, $\bar{x} = 36.2$ años, y $s = 11.5$ años. Use un nivel de significancia de 0.01 para evaluar la afirmación de que la edad promedio de las actrices cuando ganan los Oscar es mayor de 30 años.

4. Conteo de glóbulos rojos Se obtiene una muestra aleatoria simple de 40 hombres adultos y se mide el recuento de glóbulos rojos (en células por microlitro) para cada uno de ellos, con los siguientes resultados: $n = 40$, $\bar{x} = 4.932$ millones de células por microlitro, $s = 0.504$ millones de células por microlitro (según el conjunto de datos 1 "Datos corporales" del apéndice B). Use un nivel de significancia de 0.01 para probar la hipótesis de que la muestra proviene de una población con una media menor a 5.4 millones de células por microlitro, que a menudo se usa como el límite superior del rango de valores normales. ¿El resultado sugiere que cada uno de los 40 hombres tiene un recuento de glóbulos rojos por debajo de 5.4 millones de células por microlitro?

5. Percepción y realidad En una elección presidencial, 308 de los 611 votantes encuestados dijeron que votaron por el candidato que ganó (según datos del ICR Survey Research Group). Use un nivel de significancia de 0.05 para probar la hipótesis de que entre todos los votantes, el porcentaje que cree que votaron por el candidato ganador es igual a 43%, que es el porcentaje real de votos para el candidato ganador. ¿Qué sugiere el resultado sobre las percepciones de los votantes?

6. IMC de Miss Estados Unidos Una tendencia reivindicada de ganadoras más delgadas de Miss Estados Unidos ha generado cargos de que el concurso fomenta hábitos de dieta poco saludables entre las mujeres jóvenes. A continuación se listan los índices de masa corporal (IMC) para las ganadores recientes de Miss Estados Unidos. Use un nivel de significancia de 0.01 para probar la hipótesis de que las ganadoras recientes provienen de una población con un IMC promedio menor a 20.16, que fue el IMC para las ganadoras en las décadas de 1920 y 1930. Dado que el IMC es una medida de las cantidades relativas de grasa corporal y estatura, ¿las ganadoras recientes parecen ser significativamente más delgadas que las de las décadas de 1920 y 1930?

19.5 20.3 19.6 20.2 17.8 17.9 19.1 18.8 17.6 16.8

7. IMC de Miss Estados Unidos Use los mismos índices de IMC dados en el ejercicio 6. Use un nivel de significancia 0.01 para probar la hipótesis de que las ganadoras recientes de Miss Estados Unidos provienen de una población con una desviación estándar igual a 1.34, que fue la desviación estándar del IMC para las ganadoras en las décadas de 1920 y 1930. ¿Las ganadoras recientes parecen tener una variación diferente a la de las décadas de 1920 y 1930?

8. Error tipo I y error tipo II

a. En general, ¿qué es un error tipo I y qué es un error tipo II?

b. Para la prueba de hipótesis del ejercicio 6 "IMC de Miss Estados Unidos", escriba un enunciado que sería un error tipo I y otro que sería un error tipo II.

Ejercicios de repaso acumulado

1. Muertes por rayo A continuación se listan las muertes anuales por rayo en Estados Unidos, durante una secuencia de 14 años recientes y consecutivos. Encuentre los valores de los estadísticos indicados.

51 44 51 43 32 38 48 45 27 34 29 26 28 23

a. Media **b.** Mediana **c.** Desviación estándar **d.** Varianza **e.** Rango

f. ¿Qué característica importante de los datos no se revela a partir de un examen de los estadísticos, y qué herramienta sería útil para revelarla?

2. Muertes por rayo Consulte los datos muestrales del primer ejercicio de repaso acumulado.

a. ¿Cuál es el nivel de medición de los datos (nominal, ordinal, de intervalo, de razón)?

b. ¿Los valores son discretos o continuos?

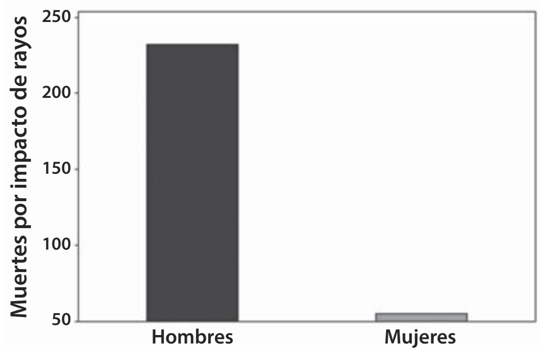
c. ¿Los datos son categóricos o cuantitativos?

d. ¿La muestra es una muestra aleatoria simple?

3. Intervalo de confianza para las muertes por rayo Use los valores muestrales dados en el primer ejercicio de repaso acumulado para elaborar una estimación del intervalo de confianza del 99% para la media poblacional. Suponga que la población tiene una distribución normal. Escriba un enunciado breve que interprete el intervalo de confianza.

4. Prueba de hipótesis para muertes por rayo Consulte los datos muestrales dados en el primer ejercicio de repaso acumulado y considere que esos datos son una muestra aleatoria de las muertes anuales por rayo en los últimos años. Use los datos con un nivel de significancia de 0.01 para evaluar la afirmación de que el número medio de muertes anuales por rayo es menor que el promedio de 72.6 muertes en la década de 1980. Si la media ahora es menor que en el pasado, identifique uno de los varios factores que podrían explicar la disminución.

5. Muertes por rayo La gráfica de barras adjunta muestra el número de muertes por rayo, clasificadas según el sexo, durante un período reciente de nueve años. ¿Qué es incorrecto en la gráfica?



6. Muertes por rayo La gráfica del ejercicio de repaso acumulado 5 se creó utilizando datos que constan de 232 muertes de hombres y 55 muertes de mujeres a consecuencia de rayos. Suponga que estos datos son muertes por rayo seleccionadas al azar y procedemos a evaluar la afirmación de que la proporción de muertes masculinas es mayor a 1/2. Use un nivel de significancia de 0.01. ¿Hay alguna explicación para el resultado?

7. Muertes por rayo La gráfica del ejercicio de repaso acumulado 5 se creó utilizando datos que constan de 232 muertes de hombres y 55 muertes de mujeres a consecuencia de rayos. Suponga que estos datos son muertes por rayo seleccionadas al azar y procedemos a elaborar una estimación del intervalo de confianza del 95% para la proporción de hombres entre todas las muertes por rayo. Con base en el resultado, ¿parece factible que los hombres y las mujeres tengan las mismas posibilidades de morir a consecuencia de un rayo?

8. Muertes por rayo Con base en los resultados dados en el ejercicio de repaso acumulado 6, suponga que para una muerte por rayo seleccionada al azar, hay una probabilidad de 0.8 de que la víctima sea un hombre.

a. Encuentre la probabilidad de que tres personas al azar muertas por rayo sean todos hombres.

b. Encuentre la probabilidad de que tres personas al azar muertas por rayo sean todas mujeres.

c. Encuentre la probabilidad de que entre tres personas muertas por rayo, al menos uno sea un hombre.

d. Si cinco personas muertas por rayo se seleccionan al azar, encuentre la probabilidad de que exactamente tres de ellos sean hombres.

e. Un estudio implica la selección aleatoria de diferentes grupos de 50 personas muertas por rayo. Para esos grupos, encuentre la media y la desviación estándar del número de víctimas masculinas.

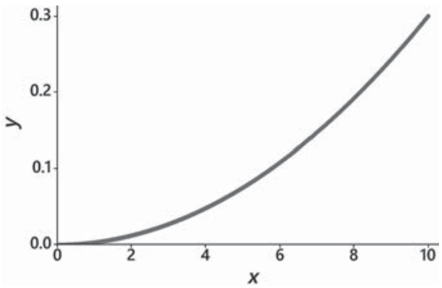
f. Para los mismos grupos descritos en el inciso (e), ¿46 sería un número significativamente alto de hombres en un grupo? Explique.

Proyectos de tecnología

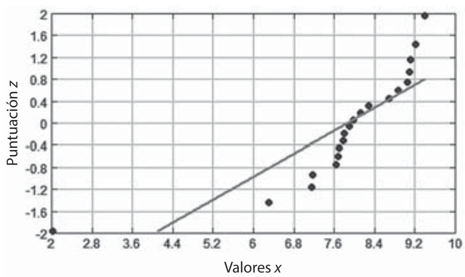
Bootstrap y robustez Considere la distribución de probabilidad definida por la fórmula $P(x) = \frac{3x^2}{1000}$ donde x puede ser cualquier valor entre 0 y 10 inclusive (no sólo los enteros). La gráfica adjunta de esta distribución de probabilidad muestra que su forma se aleja mucho de la forma de campana para una distribución normal. Esta distribución de probabilidad tiene parámetros $\mu = 7.5$ y $\sigma = 1.93649$. A continuación se lista una muestra aleatoria simple de valores de esta distribución, y se presenta la gráfica cuantilar normal para esa muestra. Dada la forma muy distinta a la normal de la distribución, no es sorprendente ver la gráfica cuantilar normal con puntos que quedan muy alejados de un patrón en línea recta, lo que confirma que la muestra no parece proceder de una población normalmente distribuida.

8.69	2.03	9.09	7.15	9.05	9.40	6.30	7.89	7.98	7.67
7.77	7.17	8.86	8.29	9.21	7.80	7.70	8.12	9.11	7.64

$P(x) = \frac{3x^2}{1000}$ **definida en [0, 10]**



Gráfica cuantilar normal de 20 valores muestrales



a. Media Pruebe la afirmación de que los 20 valores muestrales dados son de una población que tiene una media igual a 7.5, que es la media poblacional conocida. Debido a que la muestra no proviene de una población distribuida normalmente y dado que $n = 20$ no satisface el requisito de $n > 30$, no deberíamos usar los métodos de la sección 8-3. En lugar de ello, pruebe la afirmación utilizando el método del intervalo de confianza basado en una muestra bootstrap de tamaño 1000 (consulte la sección 7-4). Use un nivel de significancia de 0.05. ¿El intervalo de confianza bootstrap contiene la media poblacional conocida de 7.5? ¿El método bootstrap es efectivo para esta prueba? ¿Qué sucede si realizamos esta prueba sin ninguna precaución y elaboramos el intervalo de confianza del 95% usando la distribución t como se describió en la sección 7-2?

b. Desviación estándar Pruebe la afirmación de que los 20 valores muestrales provienen de una población con una desviación estándar igual a 1.93649, que es la desviación estándar poblacional conocida. Use el método del intervalo de confianza basado en una muestra bootstrap de tamaño 1000. (Vea la sección 7-4). Use un nivel de significancia de 0.05. ¿El intervalo de confianza bootstrap contiene la desviación estándar poblacional conocida de 1.93649? ¿El método bootstrap es efectivo para esta prueba? ¿Qué sucede si realizamos la prueba sin ninguna precaución y elaboramos el intervalo de confianza del 95% mediante el uso de la distribución χ^2 como se describe en la sección 7-3?

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: prueba de la vacuna Salk

El experimento de salud más grande jamás llevado a cabo incluyó una prueba de la vacuna Salk diseñada para proteger a los niños de los efectos devastadores de la polio. La prueba incluyó a 201,229 niños que recibieron la vacuna Salk, y 33 de ellos desarrollaron polio. La afirmación de que la vacuna Salk es efectiva es equivalente a la afirmación de que la proporción de niños vacunados que desarrollan polio es inferior a 0.0000573, que era la tasa de polio entre los niños que no recibieron la vacuna Salk. (*Nota:* El experimento real de la vacuna Salk involucró a otro grupo de 200,745 niños a los que se les inyectó una solución salina ineficaz en lugar de la vacuna Salk). El diseño del estudio con un grupo de tratamiento y un grupo placebo es muy común y efectivo. Los métodos para comparar dos proporciones se presentan en el capítulo 9).

continúa

Análisis de los resultados

a. Pruebe la afirmación dada usando un nivel de significancia de 0.05. ¿La vacuna Salk parece ser efectiva?

b. Para la prueba de hipótesis del inciso (a), considere los siguientes dos errores:

- Concluir que la vacuna Salk es efectiva cuando no lo es.
- Concluir que la vacuna Salk no es efectiva cuando sí lo es.

Determine cuál de los dos errores anteriores es un error tipo I y cuál es un error tipo II. ¿Qué error tendría peores consecuencias? ¿Cómo podría realizarse la prueba de hipótesis para reducir la posibilidad de cometer el error más grave?

Actividades de cooperación en grupo

1. Actividad fuera de clase Aquí se presenta el desglose de los colores de auto más comunes de PPG Industries: 23% son blancos, 18% son negros, 76% son grises, 15% son plateados y 10% son rojos. Después de seleccionar uno de los colores indicados, grupos de tres o cuatro estudiantes deben ir al estacionamiento de la universidad y seleccionar al azar los automóviles para probar la hipótesis de que el porcentaje del color elegido es el indicado.

2. Actividad fuera de clase En Estados Unidos, 40% de las personas tienen ojos de color marrón, según el Dr. P Sorita Soni de la Universidad de Indiana. Grupos de tres o cuatro estudiantes deben seleccionar personas al azar e identificar el color de sus ojos. Se puede evaluar la afirmación de que 40% de las personas tienen ojos marrones.

3. Actividad en clase Sin utilizar ningún dispositivo de medición, cada alumno debe trazar una línea que crea mide 3 pulgadas de largo y otra que crea mide 3 cm de largo. Luego use reglas para medir y registrar las longitudes de las líneas dibujadas. Encuentre las medias y las desviaciones estándar de los dos conjuntos de longitudes. Pruebe la afirmación de que las líneas que se estiman en 3 pulgadas tienen una longitud media igual a 3 pulgadas. Pruebe la afirmación de que las líneas que se estiman en 3 cm tienen una longitud media igual a 3 cm. Compare los resultados. ¿Las estimaciones de 3 pulgadas parecen ser más precisas que las de 3 cm? ¿Qué sugieren estos resultados?

4. Actividad en clase Suponga que un método de selección de género puede afectar la probabilidad de que un bebé sea niña, de modo que la probabilidad se convierta en 1/4. Cada alumno debe simular 20 nacimientos extrayendo 20 naipes de un mazo barajado. Reemplace cada naipe después de que haya sido extraído, luego vuelva a barajar. Considere que los corazones son niñas y que todos los demás naipes son niños. Después de hacer 20 selecciones y registrar los "géneros" de los bebés, use un nivel de significancia de 0.10 para probar la hipótesis de que la proporción de niñas es igual a 1/4. ¿Cuántos estudiantes se espera que obtengan resultados que lleven a la conclusión equivocada de que la proporción no es 1/4)? ¿Cómo se relaciona eso con la probabilidad de un error tipo I? ¿Este procedimiento parece ser efectivo para identificar la efectividad del método de selección de género? (Si no dispone de naipes, use otra forma de simular los nacimientos, por ejemplo un generador de números aleatorios en una calculadora o los dígitos de los números de teléfono o del Seguro Social).

5. Actividad fuera de clase Grupos de tres o cuatro estudiantes deben ir a la biblioteca y recolectar una muestra que conste de las edades de los libros (en función de las fechas de copyright). Planifique y describa el plan de muestreo, ejecute el procedimiento y luego use los resultados para probar la hipótesis de que la edad promedio de los libros en la biblioteca es mayor a 20 años.

6. Actividad en clase Cada estudiante debe escribir un estimado de la edad del actual presidente de Estados Unidos. Se deben recolectar todas las estimaciones, y calcular la media muestral y la desviación estándar. Luego pruebe la hipótesis de que la media de tales estimaciones es igual a la edad actual del presidente.

7. Actividad en clase Se debe diseñar un proyecto en clase para realizar una prueba en la que a cada alumno se le dé a probar Coca-Cola y Pepsi. Luego se le pedirá al estudiante que identifique cuál muestra es de Coca-Cola. Después de recopilar todos los resultados, pruebe la afirmación de que la tasa de éxito es mejor que la tasa que se esperaría con conjeturas aleatorias.

8. Actividad en clase Cada estudiante debe estimar la longitud del aula. Los valores deben basarse en estimaciones visuales, sin que se tomen medidas reales. Una vez que se hayan recopilado las estimaciones, mida la longitud del aula y luego pruebe la afirmación de que la media muestral es igual a la longitud real del aula. ¿Existe una "sabiduría colectiva", según la cual la media de la clase es aproximadamente igual a la longitud real del aula?

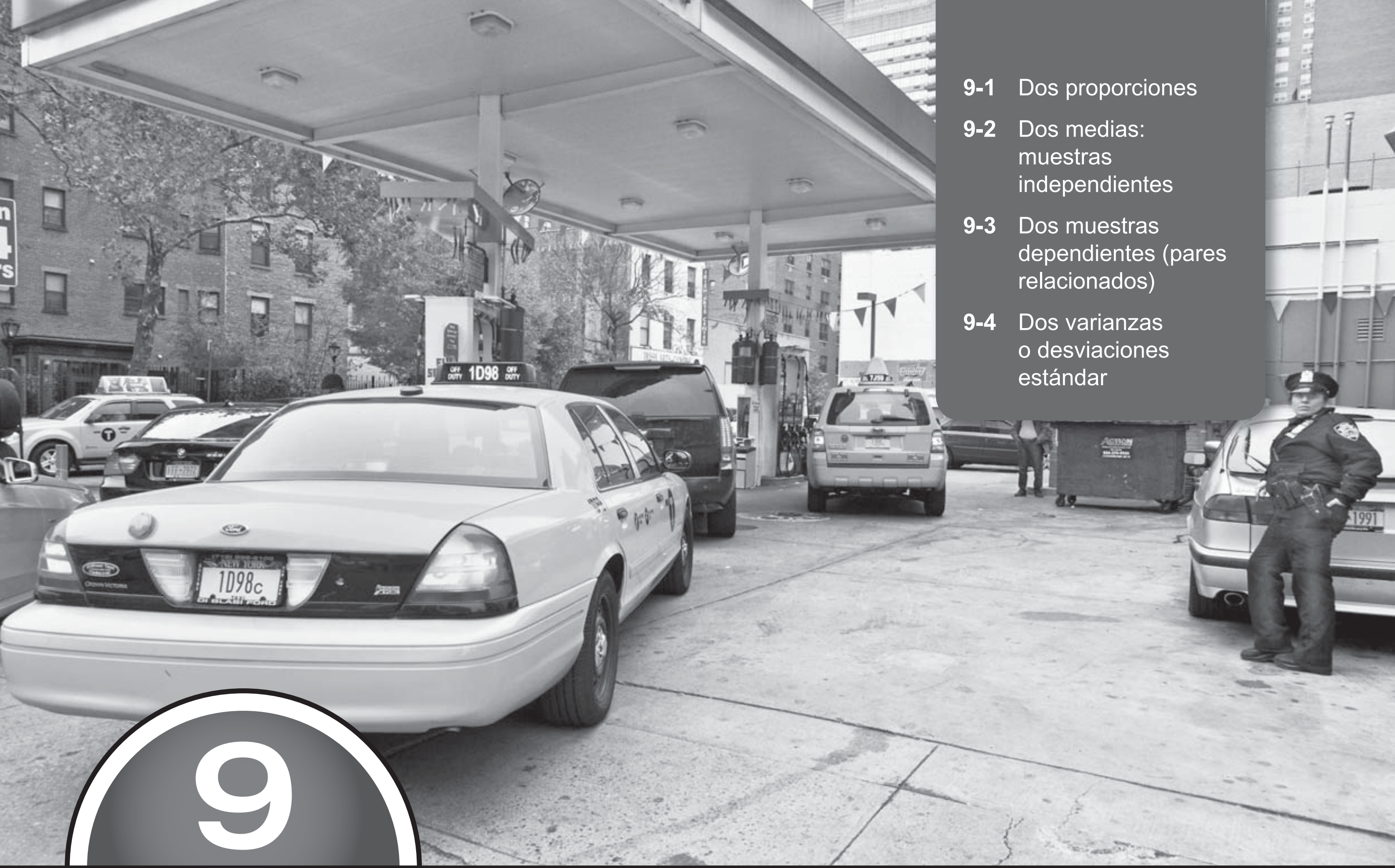
9. Actividad fuera de clase Usando un reloj de pulsera que sea razonablemente preciso, establezca la hora de modo que sea exacta. Visite www.time.gov para establecer la hora exacta. Si no puede establecer la hora al segundo más cercano, registre el error del reloj que está usando. Ahora compare el tiempo en este reloj con el tiempo en otros relojes que no se han configurado a la hora exacta. Registre los errores con signos negativos para los relojes que están adelantados a la hora real y con signos positivos para los relojes que están atrasados con respecto a la hora real. Use los datos para probar la hipótesis de que el error promedio de todos los relojes de pulsera es igual a 0. ¿Estamos colectivamente a tiempo, o estamos atrasados o adelantados? Pruebe también la afirmación de que la desviación estándar de los errores es inferior a 1 minuto. ¿Cuáles son las implicaciones prácticas de una desviación estándar que sea excesivamente grande?

10. Actividad en clase En un grupo de tres o cuatro personas, realice un experimento de percepción extrasensorial (PES) seleccionando a uno de los miembros del grupo como el sujeto. Dibuje un círculo en una pequeña hoja de papel y dibuje un cuadrado en otra hoja del mismo tamaño. Repita el siguiente experimento 20 veces: Seleccione al azar el círculo o el cuadrado y colóquelo en la mano del sujeto detrás de su espalda para que no lo pueda ver, luego pídale que identifique la forma (sin verla); registre si la respuesta es correcta. Pruebe la afirmación de que el sujeto tiene PES debido a que la proporción de respuestas correctas es significativamente mayor que 0.5.

11. Actividad en clase Después de dividirse en grupos de entre 10 y 20 personas, cada miembro del grupo debe registrar la cantidad de sus latidos en un minuto. Después de calcular la media y la desviación estándar muestrales, cada grupo debe proceder a evaluar la afirmación de que la media es mayor a 48 latidos por minuto, que es el resultado del autor. (Cuando las personas hacen ejercicio, tienden a tener tasas de pulso más bajas, y el autor corre 5 millas varias veces por semana. ¡Vaya tipo!)

12. Actividad fuera de clase En grupos de tres o cuatro estudiantes, reúna datos para determinar si los sujetos tienen una página de Facebook, luego combine los resultados y pruebe la afirmación de que más de 3/4 de los estudiantes tienen una página de Facebook.

13. Actividad fuera de clase Cada alumno debe encontrar un artículo en una revista profesional que incluya una prueba de hipótesis del tipo discutido en este capítulo. Redacte un informe breve que describa la prueba de hipótesis y su papel en el contexto del artículo.



9

INFERENCIAS A PARTIR DE DOS MUESTRAS

PROBLEMA DEL CAPÍTULO

Reglamentos de las placas de automóvil: ¿se aplican de igual forma en diferentes estados?

Muchos conductores han aprendido que las reglas del límite de velocidad se aplican discrecionalmente y con variaciones de un estado a otro en cuanto al rigor de la aplicación. Nadie recibe una multa por ir tres millas por encima del límite de velocidad. Algunos policías parecen operar con la regla de “9, estás bien, 10 eres mío”. En la práctica, las leyes del límite de velocidad suelen tratarse más como directrices.

Cuando el autor compró recientemente un automóvil nuevo, se sorprendió cuando el concesionario le dijo que le darían dos placas, pero sólo le instalarían la placa trasera. El vendedor afirmó que la ley estatal exige placas en las partes delantera y trasera del automóvil; pero la policía no se molesta en multar a los conductores que llevan sólo la placa trasera. ¿Los reglamentos de las placas se aplican como leyes o se les trata como directrices? ¿El cumplimiento de las leyes es diferente en los distintos estados? Para ayudar a responder estas preguntas, el autor recopiló los datos muestrales que se presentan en la tabla 9-1; los automóviles incluidos en dicha tabla son vehículos no comerciales.

TABLA 9-1 Número de automóviles con placas trasera y delantera

	Connecticut	Nueva York
Automóviles con placa trasera solamente	239	9
Automóviles con placa delantera y trasera	1810	541
Total	**2049**	**550**

Connecticut y Nueva York son estados contiguos, ambos tienen reglamentos que exigen placas delanteras y traseras.

Análisis La proporción de autos "ilegales" en Connecticut con placas traseras solamente es 239/2049, o 11.7%. La proporción de automóviles "ilegales" en Nueva York sólo con placas traseras es 9/550, o 1.6%. Los porcentajes muestrales de 11.7% y 1.6% son bastante diferentes, pero ¿son significativamente diferentes? Esa pregunta se tratará en la sección 9-1, que incluye métodos para comparar dos proporciones muestrales.

Debido a que hay tantos estudios con tantas aplicaciones diferentes que implican una comparación de dos muestras, los métodos de este capítulo son realmente muy importantes.

OBJETIVOS DEL CAPÍTULO

La estadística inferencial implica la formación de conclusiones (o inferencias) sobre un parámetro poblacional. Dos actividades cruciales de la estadística inferencial son la estimación de los valores de parámetros poblacionales utilizando intervalos de confianza (como en el capítulo 7) y las pruebas de hipótesis sobre parámetros poblacionales (como en el capítulo 8). Los capítulos 7 y 8 involucraron métodos para tratar una muestra de una población, y este capítulo extiende esos métodos a situaciones que involucran a dos poblaciones. Los objetivos del capítulo son:

9-1 Dos proporciones

- Realizar una prueba de hipótesis formal de una afirmación hecha sobre dos proporciones poblacionales.
- Elaborar una estimación del intervalo de confianza para la diferencia entre dos proporciones poblacionales.

9-2 Dos medias: muestras independientes

- Distinguir entre una situación que involucra dos muestras independientes y una situación que involucra dos muestras que no son independientes.
- Realizar una prueba de hipótesis formal de una afirmación hecha sobre dos medias poblacionales independientes.
- Elaborar una estimación del intervalo de confianza para la diferencia entre dos medias poblacionales.

9-3 Dos muestras dependientes (pares relacionados)

- Identificar datos muestrales que constan de pares relacionados.
- Realizar una prueba de hipótesis formal de una afirmación hecha sobre la media de las diferencias entre pares relacionados.
- Elaborar una estimación del intervalo de confianza para la media de las diferencias entre pares relacionados.

9-4 Dos varianzas o desviaciones estándar

- Desarrollar la capacidad de realizar una prueba de hipótesis formal de una afirmación hecha sobre dos desviaciones estándar o varianzas poblacionales.

9-1 Dos proporciones

Concepto clave En esta sección presentamos métodos para (1) probar una afirmación hecha sobre dos proporciones poblacionales y (2) elaborar una estimación del intervalo de confianza para la diferencia entre dos proporciones poblacionales. Los métodos de este capítulo también se pueden usar con probabilidades o con los equivalentes decimales de los porcentajes.

ELEMENTOS CLAVE

Inferencias sobre dos proporciones

Objetivos

1. **Prueba de hipótesis:** Realizar una prueba de hipótesis de una afirmación sobre dos proporciones poblacionales.
2. **Intervalo de confianza:** Elaborar una estimación del intervalo de confianza para la diferencia entre dos proporciones poblacionales.

Notación para dos proporciones

Para la población 1, sea

p_1 = proporción *poblacional* $\qquad \hat{p}_1 = \dfrac{x_1}{n_1}$ (proporción *muestral*)

n_1 = tamaño de la primera muestra $\qquad \hat{q}_1 = 1 - \hat{p}_1$ (complemento de $\hat{p}_1$)

x_1 = número de éxitos en la primera muestra

Las notaciones correspondientes p_2, n_2, x_2, $\hat{p}_2$ y $\hat{q}_2$ son aplicables a la población 2.

Proporción muestral agrupada

La **proporción muestral agrupada** se expresa con $\overline{p}$ y combina las dos proporciones muestrales en una proporción, como se muestra a continuación:

$$\overline{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\overline{q} = 1 - \overline{p}$$

Requisitos

1. Las proporciones muestrales son de dos muestras aleatorias simples.
2. Las dos muestras son *independientes*. (Las muestras serán independientes si los valores muestrales seleccionados de una población no están relacionados o de alguna forma naturalmente pareados con los valores muestrales de la otra población).
3. Para cada una de las dos muestras, hay al menos 5 éxitos y al menos 5 fracasos. (Es decir, $n\hat{p} \geq 5$ y $n\hat{q} \geq 5$ para cada una de las dos muestras).

Dato estadístico de prueba para dos proporciones (con H_0: $p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\overline{p}\,\overline{q}}{n_1} + \dfrac{\overline{p}\,\overline{q}}{n_2}}} \qquad \text{donde } p_1 - p_2 = 0 \text{ (supuesto en la hipótesis nula)}$$

Donde $\overline{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ (proporción muestral agrupada) y $\overline{q} = 1 - \overline{p}$

Valor P: Los valores P son proporcionados automáticamente por la tecnología. Si no dispone de una tecnología, use la tabla A-2 (distribución normal estándar) y encuentre el valor P mediante el procedimiento dado en la figura 8-3 de la página 364.

Valores críticos: Use la tabla A-2. (Con base en el nivel de significancia α, encuentre los valores críticos mediante los mismos procedimientos presentados en la sección 8-1).

Estimación del intervalo de confianza para $p_1 - p_2$

La estimación del intervalo de confianza para 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 $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$.

Redondeo: Redondee los límites del intervalo de confianza a tres dígitos significativos.

Métodos equivalentes

Al probar una afirmación sobre dos proporciones poblacionales:

- El método del valor P y el método del valor crítico son equivalentes.
- El método del intervalo de confianza *no* es equivalente al método del valor P o al método del valor crítico.

Recomendación: Si desea probar una afirmación sobre dos proporciones poblacionales, use el método del valor P o el método del valor crítico; si desea estimar la diferencia entre dos proporciones poblacionales, utilice el método del intervalo de confianza.

Pruebas de hipótesis

Para las pruebas de hipótesis hechas sobre dos proporciones poblacionales, consideramos sólo las que tienen una hipótesis nula del tipo $p_1 = p_2$ (por lo que la hipótesis nula es H_0: $p_1 = p_2$). Bajo el supuesto que $p_1 = p_2$, las estimaciones de $\hat{p}_1$ y $\hat{p}_2$ se combinan para proporcionar la mejor estimación del valor común de $\hat{p}_1$ y $\hat{p}_2$, y ese valor combinado es la proporción muestral agrupada $\bar{p}$ dada en el recuadro de elementos clave anterior. El siguiente ejemplo ayudará a aclarar los roles de x_1, n_1, $\hat{p}_1$, $\bar{p}$, etcétera. Tenga en cuenta que, en el supuesto de proporciones poblacionales iguales, la mejor estimación de la proporción poblacional común se obtiene agrupando ambas muestras en una muestra grande, de modo que $\bar{p}$ es el estimador de la proporción poblacional común.

Método del valor P

EJEMPLO 1 Proporciones de automóviles con placas traseras solamente: ¿son las proporciones iguales en Connecticut y Nueva York?

La tabla 9-1 del problema del capítulo incluye dos proporciones muestrales de automóviles con matrículas traseras solamente:

Connecticut: $\hat{p}_1 = 239/2049 = 0.117$

Nueva York: $\hat{p}_2 = 9/550 = 0.016$

Use un nivel de significancia de 0.05 y el método del valor P para probar la afirmación de que Connecticut y Nueva York tienen la misma proporción de automóviles con matrículas traseras solamente.

continúa

Margen de error en el liderazgo

Los autores Stephen Ansolabehere y Thomas Belin escribieron lo siguiente en su artículo "Poll Faulting" (revista *Chance*): "Nuestra mayor crítica al reporte de los resultados de las encuestas es para el margen de error de una proporción individual (por lo general $\pm 3\%$), cuando la atención de los medios de comunicación está claramente dirigida al liderazgo de un candidato". Señalan que el liderazgo es realmente la diferencia entre dos proporciones ($p_1 - p_2$) y explican cómo desarrollaron la siguiente regla práctica: el liderazgo es aproximadamente $\sqrt{3}$ veces más grande que el margen de error para cualquier proporción individual. Para una encuesta típica de preelección, un margen de error reportado de $\pm 3\%$ se convierte en alrededor de $\pm 5\%$ para el liderazgo de un candidato sobre otro. Los autores señalan que debe informarse el margen de error en el liderazgo para este tipo de encuestas.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Primero verificamos que se cumplan los tres requisitos necesarios. (1) Las dos muestras son muestras aleatorias simples (¡confíe en el autor!). (2) Las dos muestras son independientes porque los automóviles en las muestras no están relacionados o emparejados de ninguna manera. (3) Consideremos un "éxito" como un automóvil con una placa trasera solamente. Para Connecticut, la cantidad de éxitos es de 239 y la cantidad de fracasos (automóviles con placas delantera y trasera) es 1810, por lo que ambos son al menos 5. Para Nueva York, hay 9 éxitos y 541 fracasos, y ambos son al menos 5. Los requisitos se satisfacen.

Los siguientes pasos son del método del valor *P* para pruebas de hipótesis, que se resume en la figura 8-1 de la página 360.

Paso 1: La afirmación de que "Connecticut y Nueva York tienen la misma proporción de automóviles con placas traseras solamente" puede expresarse como $p_1 = p_2$.

Paso 2: Si $p_1 = p_2$ es falso, entonces $p_1 \neq p_2$.

Paso 3: Debido a que la afirmación de que $p_1 \neq p_2$ no contiene igualdad, se convierte en la hipótesis alternativa. La hipótesis nula es la declaración de igualdad, entonces tenemos

$$H_0: p_1 = p_2 \qquad H_1: p_1 \neq p_2$$

Paso 4: El nivel de significancia se especificó como $\alpha = 0.05$, por lo que usaremos $\alpha = 0.05$. (¿Quiénes somos nosotros para discutir?)

Paso 5: Este paso y el siguiente pueden evitarse usando tecnología; vea la pantalla que se presenta después de este ejemplo. Si no se utiliza software, se emplea la distribución normal (con el dato estadístico de prueba dado en el cuadro de elementos clave) como una aproximación a la distribución binomial. Estimamos el valor común de p_1 y p_2 con la estimación muestral agrupada $\overline{p}$, que se calcula de la manera mostrada a continuación; se emplean decimales adicionales para minimizar los errores de redondeo en cálculos posteriores.

$$\overline{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{239 + 9}{2049 + 550} = 0.09542132$$

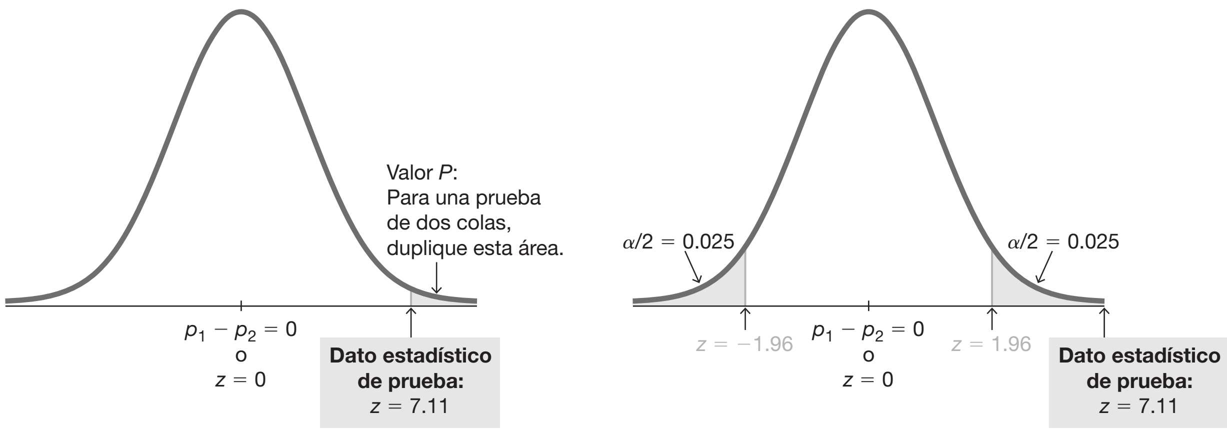
Con $\overline{p} = 0.09542132$, se deduce que $\overline{q} = 1 - 0.09542132 = 0.90457868$.

Paso 6: Debido a que suponemos en la hipótesis nula que $p_1 = p_2$, el valor de $p_1 - p_2$ es 0 en el siguiente cálculo del dato estadístico de prueba:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\overline{p}\,\overline{q}}{n_1} + \dfrac{\overline{p}\,\overline{q}}{n_2}}}$$

$$= \frac{\left(\dfrac{239}{2049} - \dfrac{9}{550}\right) - 0}{\sqrt{\dfrac{(0.09542132)(0.90457868)}{2049} + \dfrac{(0.09542132)(0.90457868)}{550}}} = 7.11$$

Esta es una prueba de dos colas, por lo que el valor *P* es el doble del área a la derecha del dato estadístico de prueba $z = 7.11$ (como lo indica la figura 8-3 en la página 364). Consulte la tabla A-2 y descubra que el área a la derecha del dato estadístico de prueba $z = 7.11$ es 0.0001, por lo que el valor *P* es 0.0002. La tecnología proporciona un valor *P* más preciso de 0.00000000000119, que a menudo se expresa como 0.0000 o "valor $P < 0.0001$". El dato estadístico de prueba y el valor *P* se muestran en la figura 9-1(a).



(a) Método del valor *P*   **(b) Método del valor crítico**

FIGURA 9-1 Prueba de hipótesis con dos proporciones

Paso 7: Debido a que el valor P de 0.0000 es menor que el nivel de significancia $\alpha = 0.05$, rechazamos la hipótesis nula de $p_1 = p_2$. ("Si P es bajo, la nula debe irse").

INTERPRETACIÓN

Debemos abordar la afirmación original de que "Connecticut y Nueva York tienen la misma proporción de automóviles con matrículas traseras solamente". Como rechazamos la hipótesis nula, concluimos que hay evidencia suficiente para justificar el rechazo de la afirmación de que $p_1 = p_2$. Es decir, hay evidencia suficiente para concluir que Connecticut y Nueva York tienen *diferentes* proporciones de automóviles con matrículas traseras solamente. Es razonable especular que la aplicación del reglamento de las placas es mucho más estricta en Nueva York que en Connecticut, y es por eso que es menos probable que los propietarios de automóviles de Connecticut instalen la placa delantera.

SU TURNO Resuelva el ejercicio 7 "Reglamentos de placas".

Tecnología El software y las calculadoras suelen proporcionar un valor P, por lo que generalmente se usa el método del valor P para probar una afirmación sobre dos proporciones. Vea los resultados de Statdisk adjuntos al ejemplo 1 que muestran el dato estadístico de prueba $z = 7.11$ (redondeado) y el valor P de 0.0000.

Statdisk

```
Pooled proportion: 0.0954213
Test Statistic, z: 7.1074
Critical z: ±1.9600
P-Value: 0.0000

95% Confidence interval:
0.0827974 < p1-p2 < 0.1177599
```

Método del valor crítico

El método del valor crítico para pruebas de hipótesis (vea la figura 8-1 en la página 360) también se puede usar para el ejemplo 1. En el paso 6, en lugar de encontrar el valor P, determine los valores críticos. Con un nivel de significancia de $\alpha = 0.05$ en una prueba de dos colas basada en la distribución normal, la tabla A-2 muestra que un área de $\alpha = 0.05$ dividida equitativamente entre las dos colas corresponde a los valores críticos de $z = \pm 1.96$. En la figura 9-1(b) podemos ver que el dato estadístico de prueba $z = 7.11$ cae dentro de la región crítica más allá del valor crítico de 1.96. Nuevamente rechazamos la hipótesis nula. Las conclusiones son las mismas que en el ejemplo 1.

Experimento sobre la poliomielitis

En 1954 se realizó un experimento para probar la eficacia de la vacuna Salk como protección contra los devastadores efectos de la poliomielitis. Aproximadamente 200,000 niños recibieron una inyección de solución salina inocua, y otros 200,000 una inyección de la vacuna. El experimento fue "doblemente ciego" porque los niños inyectados no sabían si estaban recibiendo la vacuna real o el placebo, en tanto que los doctores que aplicaban las inyecciones y evaluaban los resultados tampoco lo sabían. Sólo 33 de los 200,000 niños vacunados contrajeron poliomielitis paralítica tiempo después, mientras que 115 de los 200,000 inyectados con la solución salina contrajeron posteriormente la enfermedad. Un análisis estadístico de estos y otros resultados llevó a la conclusión de que la vacuna Salk realmente era eficaz contra la poliomielitis paralítica.

Intervalos de confianza

Mediante el uso del formato dado en el recuadro de elementos clave anterior, podemos elaborar una estimación del intervalo de confianza de la diferencia entre las proporciones poblacionales ($p_1 - p_2$). Si una estimación del intervalo de confianza de $p_1 - p_2$ no incluye 0, tenemos evidencia que sugiere que p_1 y p_2 tienen valores diferentes. El intervalo de confianza utiliza una desviación estándar basada en los valores *estimados* de las proporciones poblacionales, mientras que una prueba de hipótesis utiliza una desviación estándar basada en el *supuesto* de que las dos proporciones poblacionales son iguales. En consecuencia, una conclusión basada en un intervalo de confianza puede ser diferente de otra conclusión basada en una prueba de hipótesis. Vea la advertencia que sigue al ejemplo 2.

EJEMPLO 2 Intervalo de confianza para una afirmación sobre dos proporciones

Utilice los datos muestrales dados en el ejemplo 1 para elaborar un cálculo del intervalo de confianza del 95% para la diferencia entre las dos proporciones poblacionales. ¿Qué sugiere el resultado sobre la afirmación de que "Connecticut y Nueva York tienen la misma proporción de automóviles con placas traseras solamente"?

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Estamos usando los datos del ejemplo 1, por lo que aquí se aplica la misma verificación de requisitos y, entonces, los requisitos se cumplen. ✓

El intervalo de confianza se puede encontrar usando tecnología; vea la pantalla de Statdisk anterior. Si no emplea una tecnología, proceda de la siguiente manera.

Con un nivel de confianza del 95%, $z_{\alpha/2} = 1.96$ (a partir de la tabla A-2). Primero calculamos el valor del margen de error E como se muestra a continuación.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 1.96\sqrt{\frac{\left(\frac{239}{2049}\right)\left(\frac{1810}{2049}\right)}{2049} + \frac{\left(\frac{9}{550}\right)\left(\frac{541}{550}\right)}{550}} = 0.017482$$

Con $\hat{p}_1 = 239/2049 = 0.116642$ y $\hat{p}_2 = 9/550 = 0.016364$, obtenemos $\hat{p}_1 - \hat{p}_2 = 0.100278$. Con $\hat{p}_1 - \hat{p}_2 = 0.100278$ y $E = 0.017482$, el intervalo de confianza se evalúa de la siguiente manera, con los límites del intervalo de confianza redondeados a tres dígitos significativos:

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$
$$0.100278 - 0.017482 < (p_1 - p_2) < 0.100278 + 0.017482$$
$$0.0828 < (p_1 - p_2) < 0.118$$

Vea la pantalla de Statdisk anterior que muestra el mismo intervalo de confianza obtenido aquí.

INTERPRETACIÓN

Los límites del intervalo de confianza no contienen 0, lo que sugiere que hay una diferencia significativa entre las dos proporciones. El intervalo de confianza sugiere que el valor de p_1 es mayor que el de p_2, por lo que parece haber suficiente evidencia para justificar el rechazo de la afirmación de que "Connecticut y Nueva York tienen la misma proporción de automóviles con placas traseras solamente". Esta "sugerencia" puede ser respaldada con la prueba de hipótesis formal en el ejemplo 1.

PRECAUCIÓN Uso de un intervalo de confianza. No pruebe la igualdad de dos proporciones poblacionales dadas si existe una superposición entre dos estimaciones de intervalos de confianza individuales para las dos proporciones poblacionales particulares. Cuando se compara con la estimación del intervalo de confianza de $p_1 - p_2$, el análisis de superposición entre 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 $p_1 \neq p_2$). Vea el ejercicio 25 "Superposición de intervalos de confianza".

¿Qué podemos hacer cuando los requisitos no se satisfacen?

Muestras malas Si incumplimos el requisito de tener dos muestras aleatorias simples, podríamos enfrentar un gran problema. Por ejemplo, si tenemos dos muestras de conveniencia, probablemente no haya nada que pueda hacerse para salvarlas.

Menos de 5 éxitos o menos de 5 fracasos en una prueba de hipótesis Si incumplimos el requisito de que cada una de las dos muestras tenga al menos 5 éxitos y al menos 5 fracasos, podemos usar la *prueba exacta de Fisher*, que proporciona un valor *P* exacto en lugar de usar el método basado en una aproximación de la distribución normal. La prueba exacta de Fisher implica cálculos muy complicados, por lo que se recomienda ampliamente el uso de la tecnología. Statdisk, Minitab, XLSTAT y StatCrunch tienen la capacidad de realizar la prueba exacta de Fisher. (Vea la sección 11-2).

Menos de 5 éxitos o menos de 5 fracasos en un intervalo de confianza Si incumplimos el requisito de que cada una de las dos muestras tenga al menos 5 éxitos y al menos 5 fracasos, podemos usar métodos de remuestreo bootstrap para elaborar un intervalo de confianza. Vea la sección 7-4.

Justificación: ¿por qué funcionan los procedimientos de esta sección?

Pruebas de hipótesis Con $n_1\hat{p}_1 \geq 5$ y $n_1\hat{q}_1 \geq 5$, la distribución de $\hat{p}_1$ se puede aproximar mediante una distribución normal con media p_1, la desviación estándar $\sqrt{p_1q_1/n_1}$, y la varianza p_1q_1/n_1 (de acuerdo con las secciones 6-6 y 7-1). Esto también se aplica a la segunda muestra. Las distribuciones de $\hat{p}_1$ y $\hat{p}_2$ se aproximan cada una mediante una distribución normal, por lo que la diferencia $\hat{p}_1 - \hat{p}_2$ también se aproximará por medio de una distribución normal con la media $p_1 - p_2$ y la 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 anterior se basa en la siguiente propiedad: la varianza de las *diferencias* entre dos variables aleatorias independientes es la *suma* de sus varianzas individuales).

La estimación agrupada del valor común de p_1 y p_2 es $\overline{p} = (x_1 + x_2)/(n_1 + n_2)$. Si remplazamos p_1 y p_2 por $\overline{p}$ y sustituimos q_1 y q_2 por $\overline{q} = 1 - \overline{p}$, la varianza anterior conduce a la siguiente desviación estándar:

$$\sigma_{(\hat{p}_1-\hat{p}_2)} = \sqrt{\frac{\overline{p}\,\overline{q}}{n_1} + \frac{\overline{p}\,\overline{q}}{n_2}}$$

Sabemos que la distribución de $\hat{p}_1 - \hat{p}_2$ es aproximadamente normal, con media $p_1 - p_2$ y desviación estándar como la mostrada anteriormente, por lo que el dato estadístico de prueba z tiene la forma dada en el recuadro de elementos clave al principio de esta sección.

Intervalo de confianza La forma del intervalo de confianza requiere una expresión para la varianza diferente de la dada previamente. Al elaborar una estimación del intervalo de con-

continúa en la página 423

El autor como testigo

El autor fue requerido para testificar ante la Suprema Corte del estado de Nueva York en el caso de un ex alumno que impugnaba una reelección perdida ante la oficina del ayuntamiento del condado Dutchess. El autor testificó utilizando la estadística para demostrar que el comportamiento de votación en el distrito impugnado fue significativamente diferente del comportamiento en todos los demás distritos. Cuando el abogado de la oposición se refirió a los resultados de un intervalo de confianza, preguntó si el 5% de error (de un intervalo de confianza del 95%) podría añadirse a los tres puntos porcentuales del margen de error para obtener un error total del 8%; de esa forma, reveló que no entendía el concepto básico de un intervalo de confianza. El juez citó el testimonio del autor, apoyó la demanda del ex alumno y ordenó una nueva elección en el distrito impugnado. Después, este juicio fue anulado por la corte de apelación con base en que las irregularidades de la votación debían haberse impugnado antes de la elección, no después.

CENTRO DE TECNOLOGÍA

Inferencias con dos proporciones

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Pruebas de hipótesis

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y **Proportion Two Samples** en el submenú.
3. Seleccione el formato deseado para la *hipótesis alternativa* e ingrese el nivel de significancia. Para ambas muestras, introduzca el tamaño de la muestra y la cantidad de éxitos.
4. Haga clic en **Evaluate**.

Intervalos de confianza

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Confidence Intervals** en el menú desplegable y **Proportion Two Samples** en el submenú.
3. Ingrese el nivel de confianza deseado. Para ambas muestras, introduzca el tamaño de la muestra y la cantidad de éxitos.
4. Haga clic en **Evaluate**.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y elija **2 Proportions** en el submenú.
3. Seleccione **Summarized Data** en el menú desplegable e ingrese el número de eventos y el número de ensayos para ambas muestras.
4. Haga clic en el botón **Options** e ingrese el nivel de confianza. Si prueba una hipótesis, ingrese **0** para *Hypothesized Difference* y seleccione el formato deseado para *Alternative Hypothesis*.
5. Para *Test Method* seleccione **Use the pooled estimate of the proportion**.
6. Haga clic en **OK** dos veces.

SUGERENCIA: Otro procedimiento consiste en hacer clic en **Assistant** en el menú superior, luego seleccionar **Hypothesis Tests** y **2-Sample % Defective**. Complete el cuadro de diálogo para obtener los resultados, incluyendo el valor *P* y otra información útil.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Proportions Stats** en el menú desplegable, luego seleccione **Two Sample—With Summary** en el submenú.
3. Ingrese el número de éxitos y el número de observaciones para ambas muestras.
4. *Pruebas de hipótesis:* Seleccione **Hypothesys test for p_1-p_2**. Ingrese **0** para la diferencia hipotética (H_0) y seleccione el formato deseado para la hipótesis alternativa (H_A).
   Intervalos de confianza: Seleccione **Confidence interval for p_1-p_2** e ingrese el nivel de confianza.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

Pruebas de hipótesis:

1. Presione **STAT**, luego seleccione **TESTS** en el menú superior.
2. Seleccione **2-PropZTest** en el menú y presione **ENTER**.
3. Ingrese el número de éxitos (x) y el número de observaciones (n) para ambas muestras. Seleccione el formato deseado para la hipótesis alternativa (*p1*).
4. Seleccione **Calculate** y presione **ENTER**.

Intervalos de confianza:

1. Presione **STAT** después seleccione **TESTS** en el menú superior.
2. Seleccione **2-PropZInt** en el menú y presione **ENTER**
3. Ingrese el número de éxitos (x) y el número de observaciones (n) para ambas muestras. Introduzca el nivel de confianza deseado (*C-Level*)
4. Seleccione **Calculate** y presione **ENTER**

Excel

Complemento XLSTAT (requerido)

1. Haga clic en la pestaña **XLSTAT** de la cinta de opciones y luego haga clic en **Parametric Tests**.
2. Seleccione **Tests for two proportions** en el menú desplegable.
3. En *Data Format* seleccione **Frequency** si conoce el número de éxitos x o seleccione **Proportion** si conoce la proporción muestral $\hat{p}$.
4. Ingrese la frecuencia o la proporción muestral y el tamaño de ambas muestras.
5. Marque **z test**, desmarque **Continuity Correction** y desmarque **Monte Carlo method**.
6. Haga clic en la pestaña **Options**.
7. *Prueba de hipótesis:*
   En *Alternative hypothesis*, seleccione el formato deseado ($\neq$ para una prueba de dos colas, $<$ para una prueba de cola izquierda $>$, para una prueba de cola derecha). Para *Hypothesized difference* ingrese **0**. Introduzca el nivel de significancia deseado (ingrese **5** para el nivel de significancia 0.05). En *Variance*, seleccione **pq(1/n1 + 1/n2)**.
   Intervalos de confianza:
   En *Alternative hypothesis*, seleccione $\neq$ **D** para una prueba de dos colas. Para *Hypothesized difference* ingrese **0**. Introduzca el nivel de significancia deseado (ingrese **5** para un nivel de confianza del 95%). En *Variance*, seleccione **p1q1/n1 + p2q2/n2**.
8. Haga clic en **OK** para desplegar los resultados que incluyen el dato estadístico de prueba etiquetado como z (valor observado), el valor *P* y el intervalo de confianza.

fianza para la diferencia entre dos proporciones, no suponemos que las dos proporciones son iguales, y estimamos la desviación estándar como

$$\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

En el dato estadístico de prueba

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

use los valores positivos y negativos de z (para dos colas) y despeje $p_1 - p_2$. Los resultados son los límites del intervalo de confianza dados en el recuadro de elementos clave cerca del inicio de esta sección.

9-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Verificación de requisitos En el ensayo clínico más grande que se haya realizado, 401,974 niños fueron asignados aleatoriamente a dos grupos. El grupo de tratamiento consistió en 201,229 niños que recibieron la vacuna Salk para la poliomielitis, y 33 de esos niños desarrollaron polio. Los otros 200,745 niños recibieron un placebo, y 115 de esos niños desarrollaron poliomielitis. Si queremos utilizar los métodos de esta sección para evaluar la afirmación de que la tasa de poliomielitis es menor para los niños que reciben la vacuna Salk, ¿se satisfacen los requisitos para una prueba de hipótesis? Explique.

2. Notación Para los datos muestrales dados en el ejercicio 1, considere que el grupo de tratamiento de la vacuna Salk es la primera muestra. Identifique los valores de n_1, $\hat{p}_1$, $\hat{q}_1$, n_2, $\hat{p}_2$, $\hat{q}_2$, $\overline{p}$ y $\overline{q}$. Redondee todos los valores a seis dígitos significativos.

3. Hipótesis y conclusiones Consulte la prueba de hipótesis descrita en el ejercicio 1.

a. Identifique la hipótesis nula y la hipótesis alternativa.

b. Si el valor P para la prueba se reporta como "menor que 0.001". ¿Qué deberíamos concluir sobre la afirmación original?

4. Uso de intervalos de confianza

a. Suponga que queremos usar un nivel de significancia de 0.05 para probar la afirmación de que $p_1 < p_2$. ¿Qué es mejor: una prueba de hipótesis o un intervalo de confianza?

b. En general, cuando se trata con inferencias de dos proporciones poblacionales, ¿Cuáles dos de las siguientes opciones son equivalentes: Método del intervalo de confianza; método del valor P; método del valor crítico?

c. Si queremos utilizar un nivel de significancia de 0.05 para probar la afirmación de que $p_1 < p_2$, ¿qué nivel de confianza deberíamos usar?

d. Si probamos la afirmación del inciso (c) usando los datos muestrales del ejercicio 1, obtenemos el siguiente intervalo de confianza: $-0.000508 < p_1 - p_2 < -0.000309$. ¿Qué sugiere este intervalo de confianza sobre la afirmación?

Interpretación de pantallas. ***En los ejercicios 5 y 6, use los resultados de las pantallas dadas.***

5. Prueba de guantes de laboratorio El *New York Times* publicó un artículo sobre un estudio en el que la profesora Denise Korniewicz y otros investigadores del Johns Hopkins sometieron guantes de laboratorio a esfuerzo. Entre 240 guantes de vinilo, 63% sufrieron filtración de virus. Entre 240 guantes de látex, 7% tuvieron filtración de virus. Vea la pantalla adjunta con los resultados de Statdisk. Utilizando un nivel de significancia de 0.01, pruebe la afirmación de que los guantes de vinilo tienen una mayor tasa de filtración de virus que los guantes de látex.

Statdisk

```
Pooled proportion: 0.35
Test Statistic, z: 12.8231
Critical z: 2.3264
P-Value: 0.0000

98% Confidence interval:
0.4762035 < p1-p2 < 0.6404632
```

continúa

6. Tratamiento del síndrome del túnel carpiano El síndrome del túnel carpiano es una molestia común en la muñeca que resulta de un nervio comprimido, y a menudo es el resultado del uso prolongado de movimientos repetitivos de la muñeca, como los asociados con el uso de un teclado o el ratón. En un ensayo aleatorio controlado, 73 pacientes fueron tratados con cirugía y el tratamiento de 67 resultó exitoso. Entre 83 pacientes tratados con férulas, 60 tuvieron éxito en su tratamiento (según datos de "Splinting *vs* Surgery in the Treatment of Carpal Tunnel Syndrome", de Gerritsen *et al*., *Journal of the American Medical Association*, vol. 288, núm. 10). Use la pantalla adjunta de StatCrunch con un nivel de significancia de 0.01 para evaluar la afirmación de que la tasa de éxito es mejor con cirugía.

StatCrunch

Difference	0.1949
z (Observed value)	3.1226
z (Critical value)	2.3263
p-value (one-tailed)	0.0009
alpha	0.01

Pruebas de hipótesis sobre proporciones. ***En los ejercicios 7 a 22, pruebe la afirmación dada. Identifique la hipótesis nula, la hipótesis alternativa, el dato estadístico de prueba, el valor P o el(los) valor(es) crítico(s), luego formule una conclusión sobre la hipótesis nula, así como una conclusión final que aborde la afirmación original.***

7. Reglamentos de placas El problema del capítulo involucró a los automóviles de pasajeros en Connecticut y en Nueva York, pero aquí consideramos tanto los autos de pasajeros, como los camiones comerciales. Entre 2049 automóviles de pasajeros en Connecticut, 239 tenían sólo placas traseras. Entre 334 camiones en Connecticut, 45 tenían sólo placas traseras (según muestras recogidas por el autor). Una hipótesis razonable es que los propietarios de automóviles de pasajeros violan los reglamentos de placas en una proporción mayor que los propietarios de camiones comerciales. Use un nivel de significancia de 0.05 para probar esa hipótesis.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza apropiado.

8. Precisión de los pedidos de comida rápida en auto En un estudio de pedidos en auto en Burger King, se encontró que 264 pedidos fueron precisos y 54 no lo fueron. Para McDonald's, se encontraron 329 pedidos precisos, mientras que 33 pedidos no fueron precisos (según datos de la revista *QSR*). Use un nivel de significancia de 0.05 para probar la afirmación de que Burger King y McDonald's tienen las mismas tasas de precisión.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. En relación con la precisión de los pedidos, ¿alguna de estas cadenas de restaurantes parece ser mejor?

9. Programas para dejar de fumar Entre 198 fumadores que se sometieron a un programa de "atención sostenida", 51 ya no fumaban después de seis meses. Entre 199 fumadores que se sometieron a un programa de "atención estándar", 30 ya no fumaban después de seis meses (según datos de "Sustained Care Intervention and Postdischarge Smoking Cessation Among Hospitalized Adults", de Rigotti *et al*., *Journal of the American Medical Association*, vol. 312, núm. 7). Queremos usar un nivel de significancia de 0.01 para evaluar la afirmación de que la tasa de éxito para dejar de fumar es mayor con el programa de atención sostenida.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. ¿La diferencia entre los dos programas tiene un significado práctico?

10. Desafíos en el tenis Desde que se presentó el sistema de repetición instantánea "ojo de halcón" para el tenis en el Abierto de Estados Unidos en 2006, los tenistas varones han desafiado 2441 decisiones de los árbitros, con el resultado de que 1027 de las decisiones fueron anuladas. Las mujeres han desafiado 1273 decisiones de los árbitros, y 509 de éstas fueron anuladas. Queremos usar un nivel de significancia de 0.05 para evaluar la afirmación de que los hombres y las mujeres tienen igual éxito en los desafíos a los árbitros.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. De acuerdo con los resultados, ¿parece que los hombres y las mujeres tienen el mismo éxito al desafiar las decisiones de los árbitros?

11. Sueños en blanco y negro Se realizó un estudio para determinar la proporción de personas que sueñan en blanco y negro en lugar de a color. Entre 306 personas mayores de 55 años, 68 sueñan en blanco y negro, y entre 298 personas menores de 25 años, 13 sueñan en blanco y negro (según datos de "Do We Dream In Color?" de Eva Murzyn, *Consciousness and Cognition*, vol. 17, núm. 4). Queremos utilizar un nivel de significancia de 0.01 para evaluar la afirmación de que la proporción de personas mayores de 55 años que sueñan en blanco y negro es mayor que la proporción de menores de 25 años.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Una explicación dada por los resultados es que aquellos mayores de 55 años crecieron expuestos a medios desplegados principalmente en blanco y negro. ¿Se pueden usar los resultados de los incisos (a) y (b) para verificar esa explicación?

12. Ensayos clínicos de OxyContin El OxyContin (la oxicodona) es un medicamento utilizado para tratar el dolor, pero es bien conocido por su adictividad y peligrosidad. En un ensayo clínico, entre sujetos tratados con OxyContin, 52 desarrollaron náuseas y 175 no lo hicieron. Entre otros sujetos que recibieron placebos, 5 desarrollaron náuseas y 40 no lo hicieron (según datos de Purdue Pharma L.P.). Use un nivel de significancia de 0.05 para evaluar la diferencia entre las tasas de náuseas de las personas tratadas con OxyContin y las que recibieron un placebo.

a. Use una prueba de hipótesis.

b. Use un intervalo de confianza adecuado.

c. ¿Las náuseas parecen ser una reacción adversa que resulta del uso del OxyContin?

13. ¿Son efectivos los cinturones de seguridad? Se obtiene una muestra aleatoria simple de ocupantes del asiento delantero involucrados en accidentes automovilísticos. Entre 2823 ocupantes que no usaban cinturones de seguridad, 31 murieron. Entre 7765 ocupantes que usaban cinturones de seguridad, 16 murieron (según datos de "Who Wants Airbags?" de Meyer y Finney, *Chance*, vol. 18, núm. 2). Queremos usar un nivel de significancia de 0.05 para probar la afirmación de que los cinturones de seguridad son efectivos para reducir la mortalidad.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. ¿Qué sugiere el resultado sobre la efectividad de los cinturones de seguridad?

14. Paro cardíaco de día y de noche Un estudio investigó las tasas de supervivencia para pacientes hospitalizados que sufrieron un paro cardíaco. Entre los 58,593 pacientes que tuvieron un paro cardíaco durante el día, 11,604 sobrevivieron y fueron dados de alta. Entre 28,155 pacientes que sufrieron un paro cardíaco por la noche, 4139 sobrevivieron y fueron dados de alta (según datos de "Survival from In-Hospital Cardiac Arrest During Nights and Weekends", de Peberdy *et al*., *Journal of the American Medical Association*, vol. 299, núm. 7). Queremos usar un nivel de significancia de 0.01 para probar la afirmación de que las tasas de supervivencia son iguales para el día y para la noche.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Según los resultados, ¿parece que para los pacientes hospitalizados que sufren un paro cardíaco, la tasa de supervivencia es igual de día que de noche?

15. ¿Es la equinácea efectiva para los resfriados? Por lo general, los rinovirus causan resfriados comunes. En una prueba de la efectividad de la equinácea, 40 de los 45 sujetos tratados con esta planta desarrollaron infecciones por rinovirus. En un grupo placebo, 88 de los 103 sujetos desarrollaron infecciones por rinovirus (con base en datos de "An Evaluation of Echinacea Angustifolia in Experimental Rhinovirus Infections" de Turner *et al*., *New England Journal of Medicine*, vol. 353, núm. 4). Queremos usar un nivel de significancia de 0.05 para evaluar la afirmación de que la equinácea tiene un efecto en las infecciones por rinovirus.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. De acuerdo con los resultados, ¿parece que la equinácea tiene algún efecto sobre la tasa de infección?

16. Mosquiteros para reducir la malaria En un ensayo controlado y aleatorizado en Kenia, se probaron mosquiteros tratados con insecticida como una forma de reducir la malaria. Entre 343 niños que usaron mosquiteros, 15 desarrollaron malaria. Entre 294 niños que no usaron mosquiteros, 27 desarrollaron malaria (según datos de "Sustainability of Reductions in Malaria Transmission and Infant Mortality in Western Kenya with Use of Insecticide-Treated Bednets", de Lindblade *et al*., *Journal of the American Medical Association*, vol. 291, núm. 21). Queremos usar un nivel de significancia de 0.01 para evaluar la afirmación de que la incidencia de malaria es menor en los niños que usan mosquiteros.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Según los resultados, ¿los mosquiteros parecen ser efectivos?

17. Teléfonos celulares y destreza Se realizó un estudio para investigar la asociación entre el uso del teléfono celular y el dominio hemisférico del cerebro. Entre 216 sujetos que prefieren usar su oído izquierdo para el teléfono celular, 166 eran diestros. Entre 452 sujetos que prefieren usar su oído derecho para el teléfono celular, 436 eran diestros (según datos de "Hemispheric Dominance and Cell Phone Use", de Seidman *et al*., *JAMA Otolaryngology-Head & Neck Surgery*, vol. 139 , núm. 5). Queremos usar un nivel de significancia de 0.01 para probar la afirmación de que la proporción de destreza en aquellos que prefieren usar su oído izquierdo para el teléfono celular es menor que la tasa de destreza en aquellos que prefieren usar su oído derecho para el celular. (Trate de no confundirse demasiado aquí).

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

18. Efecto de la denominación Se realizó un ensayo en China con 75 mujeres a quienes se les dio un billete de 100 yuanes, mientras que otras 75 mujeres recibieron 100 yuanes en forma de billetes más pequeños (un billete de 50 yuanes más dos billetes de 20 yuanes y dos de 5 yuanes). Entre quienes recibieron un solo billete, 60 gastaron parte o todo el dinero. Entre quienes recibieron los billetes más pequeños, 68 gastaron parte o la totalidad del dinero (según datos de "The Denomination Effect", de Raghubir y Srivastava, *Journal of Consumer Research*, vol. 36). Queremos usar un nivel de significancia de 0.05 para probar la afirmación de que cuando se les da un solo billete grande, una proporción menor de mujeres en China gasta una parte o la totalidad del dinero en comparación con la proporción de mujeres en China que reciben el mismo monto en billetes pequeños.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Si el nivel de significación se cambia a 0.01, ¿cambia la conclusión?

19. Tratamiento del dolor de cabeza En un estudio de los tratamientos para los dolores de cabeza "en racimo", 150 pacientes fueron tratados con oxígeno y otros 148 pacientes recibieron un placebo consistente en aire ordinario. Entre los 150 pacientes en el grupo de tratamiento con oxígeno, 116 estuvieron libres de dolores de cabeza 15 minutos después del tratamiento. Entre los 148 pacientes que recibieron el placebo, 29 estuvieron libres de dolores de cabeza 15 minutos después del tratamiento (según datos de "High-Flow Oxygen for Treatment of Cluster Headache", de Cohen, Burns y Goadsby, *Journal of the American Medical Association*, vol. 302, núm. 22). Queremos usar un nivel de significancia de 0.01 para probar la afirmación de que el tratamiento con oxígeno es efectivo.

continúa

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Según los resultados, ¿el tratamiento con oxígeno es efectivo?

20. ¿La aspirina previene enfermedades del corazón? En un ensayo diseñado para evaluar la efectividad de la aspirina en la prevención de enfermedades cardíacas, 11,037 médicos varones fueron tratados con aspirina y 11,034 médicos varones recibieron placebos. Entre los sujetos en el grupo de tratamiento con aspirina, 139 experimentaron infartos al miocardio (ataques cardíacos). Entre los sujetos que recibieron placebos, 239 experimentaron infartos al miocardio (según datos del "Final Report on the Aspirin Component of the Ongoing Physicians' Health Study", *New England Journal of Medicine*, vol. 321: 129-135). Use un nivel de significancia de 0.05 para evaluar la afirmación de que la aspirina no tiene efecto sobre los infartos al miocardio.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Según los resultados, ¿parece que la aspirina es efectiva?

21. Zurdos En una muestra aleatoria de hombres, se encontró que 23 escriben con la mano izquierda y 217 no. En una muestra aleatoria de mujeres, se encontró que 65 escriben con la mano izquierda y 455 no (según datos de "The Left-Handed: Their Sinister History", de Elaine Fowler Costas, *Education Resources Information Center*, artículo 399519). Queremos usar un nivel de significancia de 0.01 para probar la afirmación de que la tasa de zurdera entre los hombres es menor que entre las mujeres.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. Con base en los resultados, ¿la tasa de zurdera entre los hombres es menor que entre las mujeres?

22. Vehículo terrestre contra helicóptero para lesiones graves Un estudio investigó las tasas de mortalidad en pacientes con lesiones traumáticas graves. Entre 61,909 pacientes transportados en helicóptero, murieron 7813. Entre 161,566 pacientes transportados en vehículos terrestres, 17,775 murieron (según datos de "Association Between Helicopter *vs* Ground Emergency Medical Services and Survival for Adults With Major Trauma", de Galvagno *et al*., *Journal of the American Medical Association*, vol. 307, núm. 15). Use un nivel de significancia de 0.01 para probar la afirmación de que la tasa de mortalidad es mayor para los pacientes transportados en helicóptero.

a. Pruebe la afirmación usando una prueba de hipótesis.

b. Pruebe la afirmación elaborando un intervalo de confianza adecuado.

c. De acuerdo con los resultados de la prueba y las tasas muestrales reales, ¿hay un modo de transporte mejor que el otro? ¿Hay otros factores importantes a considerar?

9-1 Más allá de lo básico

23. Determinación del tamaño de muestra. El tamaño de 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 puede encontrar utilizando la siguiente expresión:

$$E = z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

continúa

Reemplace n_1 y n_2 con n en la fórmula anterior (suponiendo que ambas muestras tienen el mismo tamaño) y reemplace p_1, q_1, p_2 y q_2 por 0.5 (porque no se conocen sus valores). La solución para n resulta en la siguiente expresión:

$$n = \frac{z_{\alpha/2}^2}{2E^2}$$

Use esta expresión para encontrar el tamaño de cada muestra si desea estimar la diferencia entre las proporciones de hombres y mujeres que poseen teléfonos inteligentes. Suponga que desea un 95% de confianza de que el error no será mayor que 0.03.

24. Bostezo y la prueba exacta de Fisher En un segmento de la serie de televisión "Cazadores de mitos", se llevó a cabo un experimento para poner a prueba la creencia común de que es más probable que las personas bostecen cuando ven a otros bostezar. En un grupo, 34 sujetos fueron expuestos a bostezos, y 10 de ellos bostezaron. En otro grupo, 16 sujetos no fueron expuestos a bostezos, y 4 de ellos bostezaron. Queremos probar la creencia de que las personas son más propensas a bostezar cuando están expuestas al bostezo.

a. ¿Por qué no podemos probar la afirmación utilizando los métodos de esta sección?

b. Si ignoramos los requisitos y usamos los métodos de esta sección, ¿cuál es el valor P? ¿Cómo se compara con el valor P de 0.5128 que se obtendría al usar la prueba exacta de Fisher?

c. Comente la conclusión del segmento de "Cazadores de mitos" de que bostezar es contagioso.

25. Superposición de intervalos de confianza En el artículo "Juicio de la importancia de las diferencias examinando la superposición entre intervalos de confianza", de Schenker y Gentleman (*American Statistician*, vol. 55, núm. 3), los autores consideran los datos muestrales en la siguiente afirmación: "Se han obtenido dos muestras aleatorias simples e independientes, cada una de tamaño 200, y en la primera muestra 112 personas tienen el atributo, mientras que en la segunda muestra 88 personas lo tienen".

a. Use los métodos de esta sección para elaborar una estimación del intervalo de confianza del 95% para la diferencia $p_1 - p_2$. ¿Qué sugiere el resultado sobre la igualdad de p_1 y p_2?

b. Use los métodos de la sección 7-1 para elaborar estimaciones individuales del intervalo de confianza del 95% para cada una de las dos proporciones poblacionales. Después de comparar la superposición entre los dos intervalos de confianza, ¿qué concluye usted acerca de la igualdad de p_1 y p_2?

c. Use un nivel de significancia de 0.05 para probar la afirmación de que las dos proporciones poblacionales son iguales. ¿Qué concluye usted?

d. Con base en los resultados anteriores, ¿qué se debería concluir acerca de la igualdad de p_1 y p_2? ¿Cuál de los tres métodos anteriores es menos efectivo para evaluar la igualdad de p_1 y p_2?

26. Equivalencia de la prueba de hipótesis y el intervalo de confianza Se toman dos muestras aleatorias simples diferentes de dos poblaciones distintas. La primera muestra consiste en 20 personas; 10 de ellas tienen un atributo común. La segunda muestra consta de 2000 personas, 1404 de las cuales tienen el mismo atributo común. Compare los resultados de una prueba de hipótesis de $p_1 = p_2$ (con un nivel de significancia de 0.05) y una estimación del intervalo de confianza del 95% para $p_1 - p_2$.

9-2 Dos medias: muestras independientes

Concepto clave En esta sección se presentan métodos para usar datos de dos muestras independientes, con el fin de probar hipótesis hechas sobre dos medias poblacionales o para elaborar estimaciones de intervalo de confianza para la diferencia entre dos medias poblacionales. En la parte 1 analizamos situaciones en las que las desviaciones estándar de las dos poblaciones son desconocidas y no se supone que sean iguales. En la parte 2, estudiamos brevemente otras dos situaciones: (1) Las dos desviaciones estándar poblacionales son desconocidas, pero se supone que son iguales; (2) el caso poco realista en el que se conocen las dos desviaciones estándar poblacionales.

PARTE 1 Muestras independientes: σ_1 y σ_2 desconocidas y que no se suponen iguales

Esta sección involucra dos muestras *independientes* y la siguiente sección trata con muestras que son dependientes. Es importante conocer la diferencia entre muestras independientes y muestras dependientes.

> **DEFINICIONES**
>
> Dos muestras son **independientes** si los valores muestrales de una población no están relacionados o de alguna manera naturalmente emparejados o combinados con los valores muestrales de otra población.
>
> Dos muestras son **dependientes** (o constan de **pares relacionados**) si los valores muestrales se corresponden de alguna manera, donde la correspondencia se basa en una relación inherente (es decir, cada par de valores muestrales consiste en dos medidas del mismo sujeto, como datos de antes y después, o cada par de valores muestrales consiste en pares coincidentes, como datos de marido y mujer, y donde la coincidencia se basa en alguna relación significativa). *Precaución*: la "dependencia" no requiere una relación de causa/efecto directa.

> **SUGERENCIA:** Si las dos muestras tienen diferentes tamaños de muestra sin datos faltantes, deben ser independientes. Si las dos muestras tienen el mismo tamaño de muestra, las muestras pueden ser independientes o no.

A continuación se muestra un ejemplo de muestras independientes y otro de muestras dependientes:

- ***Muestras independientes:* estaturas de hombres y mujeres** El conjunto de datos 1 "Datos corporales" en el apéndice B incluye las siguientes estaturas (en cm) muestrales de hombres y mujeres, y las dos muestras no se corresponden de acuerdo con alguna relación inherente. En realidad, son dos muestras independientes que, por casualidad, se enumeran de una manera que podría hacernos pensar incorrectamente que se corresponden.

Estaturas (en cm) de hombres	172	154	156	158	169
Estaturas (en cm) de mujeres	186	161	179	167	179

- ***Muestras dependientes:* estaturas de esposos y esposas** Algunos estudiantes del autor reunieron datos que consisten en las estaturas (en cm) de esposos y las estaturas (en cm) de sus esposas. Cinco de esos pares de estaturas se enumeran a continuación. Estas dos muestras son dependientes, porque la estatura de cada esposo se corresponde con la estatura de su esposa.

Estatura (cm) del esposo	175	180	173	176	178
Altura (cm) de la esposa	160	165	163	162	166

Para realizar inferencias sobre las medias de dos poblaciones independientes, el siguiente recuadro resume los elementos clave de una prueba de hipótesis y una estimación del intervalo de confianza para la diferencia entre las medias poblacionales.

¿Los agentes inmobiliarios consiguen los mejores precios para sus clientes?

Cuando un agente inmobiliario vende una casa, ¿consigue el mejor precio para el vendedor? Steven Levitt y Stephen Dubner exploraron esta pregunta en *Freakonomics*. Los autores reunieron datos de miles de viviendas cerca de Chicago, incluyendo las casas de los propios agentes, y escribieron lo siguiente: "Existe una forma de descubrirlo: medir la diferencia entre los datos de ventas de las casas que pertenecen a los agentes inmobiliarios y las casas que venden a nombre de los clientes. Al utilizar los datos de ventas de 100,000 casas de Chicago, y controlando distintas variables (ubicación, antigüedad y calidad de la casa, características estéticas, etcétera), resultó que un agente inmobiliario mantiene su propia casa en el mercado un promedio de 10 días más y la vende por 3% más, lo que representa casi $10,000 más en el caso de una residencia de $300,000". Una conclusión como ésta se puede obtener utilizando los métodos que se estudian en la presente sección.

ELEMENTOS CLAVE

Inferencias sobre dos medias: muestras independientes

Objetivos

1. **Prueba de hipótesis:** Realizar una prueba de hipótesis de una afirmación sobre dos medias poblacionales independientes.
2. **Intervalo de confianza:** Elaborar una estimación del intervalo de confianza para la diferencia entre dos medias poblacionales independientes.

Notación

Para la población 1 sea

μ_1 = media *poblacional*

$\bar{x}_1$ = media *muestral*

σ_1 = desviación estándar *poblacional*

s_1 = desviación estándar *muestral*

n_1 = tamaño de la primera muestra

Las notaciones correspondientes μ_2, σ_2, $\bar{x}_2$, s_2 y n_2 se aplican a la población 2

Requisitos

1. Los valores de σ_1 y σ_2 son desconocidos y no suponemos que sean iguales.
2. Las dos muestras son *independientes*.
3. Ambas muestras son *muestras aleatorias simples*.
4. Se cumple al menos una de las siguientes condiciones: 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. (Los métodos utilizados aquí son robustos contra las desviaciones de la normalidad, por lo que para muestras pequeñas, el requisito de normalidad es flexible en el sentido de que los procedimientos funcionan bien siempre que no haya valores atípicos y las desviaciones de la normalidad no sean demasiado extremas).

Dato estadístico de prueba de hipótesis para dos medias: muestras independientes (con H_0: $\mu_1 = \mu_2$)

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

(donde a menudo se supone que $\mu_1 - \mu_2$ es 0)

Grados de libertad Al determinar valores críticos o valores P, use lo siguiente para encontrar el número de grados de libertad, expresados por gl. (Aunque estos dos métodos suelen resultar en diferentes cantidades de grados de libertad, la conclusión de una prueba de hipótesis rara vez se ve afectada por la elección).

1. Use esta estimación simple y conservadora:

$$\textbf{gl = el menor de } n_1 - 1 \textbf{ y } n_2 - 1$$

2. Por lo general las tecnologías usan la estimación más precisa pero más difícil dada en la fórmula 9-1.

FÓRMULA 9-1

$$\text{gl} = \frac{(A + B)^2}{\dfrac{A^2}{n_1 - 1} + \dfrac{B^2}{n_2 - 1}}$$

$$\text{donde} \quad A = \frac{s_1^2}{n_1} \quad \text{y} \quad B = \frac{s_2^2}{n_2}$$

Nota: Las respuestas en el apéndice D incluyen resultados de tecnologías basadas en la fórmula 9-1, junto con respuestas de "tabla" basadas en el uso de la tabla A-3 con la estimación simple de gl dada en la opción 1 anterior.

Valores *P*: Los valores P son proporcionados automáticamente por la tecnología. Si usted no dispone de una tecnología, consulte la distribución t en la tabla A-3. Use el procedimiento resumido en la figura 8-3 de la página 364.

Valores críticos: Consulte la distribución t en la tabla A-3.

Estimación del intervalo de confianza para $\mu_1 - \mu_2$: muestras independientes

La estimación del intervalo de confianza para la diferencia $\mu_1 - \mu_2$ es

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

donde

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

y el número de grados de libertad gl es el descrito para las pruebas de hipótesis. (En este libro, usamos gl = el menor de $n_1 - 1$ y $n_2 - 1$).

Métodos equivalentes

El método del valor *P* y el método del valor crítico para pruebas de hipótesis, así como los intervalos de confianza utilizan la misma distribución y el mismo error estándar, por lo que todos son equivalentes en el sentido de que dan como resultado las mismas conclusiones.

Método del valor *P*

EJEMPLO 1 **¿Los profesores y profesoras son calificados de manera diferente por los estudiantes?**

A continuación se listan las puntuaciones otorgadas por estudiantes para evaluar los cursos impartidos por profesoras y profesores (del conjunto de datos 17 “Evaluaciones de cursos” en el apéndice B). Use un nivel de significancia de 0.05 para probar la afirmación de que las dos muestras provienen de poblaciones con la misma media. ¿Parece haber una diferencia en las puntuaciones para los cursos impartidos por profesoras y profesores?

Profesoras	4.3	4.3	4.4	4.0	3.4	4.7	2.9	4.0	4.3	3.4	3.4	3.3			
Profesores	4.5	3.7	4.2	3.9	3.1	4.0	3.8	3.4	4.5	3.8	4.3	4.4	4.1	4.2	4.0

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) No se conocen los valores de las desviaciones estándar poblacionales y no suponemos que sean iguales. (2) Las dos muestras son independientes porque las profesoras y los profesores no están relacionados o emparejados de ninguna manera. (3) Las muestras son aleatorias simples. (4) Ambas muestras son pequeñas (30 o menos), por lo que debemos determinar si provienen de poblaciones con distribuciones normales. Las gráficas cuantilares normales de las dos muestras sugieren que provienen de poblaciones con distribuciones no muy alejadas de la normal. Todos los requisitos se satisfacen. ✓

Con base en el método del valor *P* resumido en la figura 8-1 de la página 360, podemos probar la afirmación de la siguiente manera.

Paso 1: La afirmación de que “las dos muestras provienen de poblaciones con la misma media” se puede expresar como $\mu_1 = \mu_2$.

Paso 2: Si la afirmación original es falsa, entonces $\mu_1 \neq \mu_2$.

Paso 3: La hipótesis alternativa es la expresión que no contiene igualdad, y la hipótesis nula es una expresión de igualdad, por lo que tenemos

$$H_0: \mu_1 = \mu_2 \qquad H_1: \mu_1 \neq \mu_2$$

continúa

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 eficaces para bajar de peso. Los anuncios afirmaban que después de tomar las cápsulas, la grasa sería bloqueada y las calorías se quemarían, incluso sin realizar ejercicio. Puesto que la Federal Trade Commission identificó hipótesis que parecían no tener fundamento, se multó a la compañía con \$10 millones por publicidad engañosa.

La eficacia de tratamientos como éstos puede determinarse con experimentos en los cuales se administra el tratamiento a un grupo de sujetos seleccionados al azar, mientras que se da un placebo a otro grupo de sujetos seleccionados al azar. Las pérdidas de peso resultantes se comparan utilizando métodos estadísticos, como los descritos en esta sección.

Ahora procedemos con el supuesto de que $\mu_1 = \mu_2$, o $\mu_1 - \mu_2 = 0$.

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Dado que tenemos dos muestras independientes y estamos probando una afirmación sobre las dos medias poblacionales, usamos una distribución t con el dato estadístico de prueba dado anteriormente en esta sección.

Paso 6: El dato estadístico de prueba se calcula utilizando los siguientes datos estadísticos (con lugares decimales adicionales) obtenidos a partir de los datos muestrales que se listan: profesoras: $n = 12$, $\bar{x} = 3.866667$, $s = 0.563001$; profesores: $n = 15$, $\bar{x} = 3.993333$, $s = 0.395450$.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{(3.866667 - 3.993333) - 0}{\sqrt{\dfrac{0.563001^2}{12} + \dfrac{0.395450^2}{15}}} = -0.660$$

Valor *P* Con el dato estadístico de prueba $t = -0.660$, se hace referencia a la tabla A-3 (Distribución t). El número de grados de libertad es el menor entre $n_1 - 1$ y $n_2 - 1$, o el menor de $(12 - 1)$ y $(15 - 1)$, que es 11. Con gl $= 11$ y una prueba de dos colas, la tabla A-3 indica que el valor P es mayor que 0.20. La tecnología proporcionará el valor P de 0.5172 si se utilizan los datos originales o los datos estadísticos muestrales no redondeados.

Paso 7: Como el valor P es mayor que el nivel de significancia de 0.05, no rechazamos la hipótesis nula. ("Si P es *bajo*, la nula debe irse").

INTERPRETACIÓN

Paso 8: No hay evidencia suficiente para justificar el rechazo de la afirmación de que las profesoras y los profesores tienen la misma puntuación media en la evaluación de sus cursos.

SU TURNO Resuelva el ejercicio 7 "Color y creatividad".

Tecnología La parte truculenta en el método del valor P presentado anteriormente es que la tabla A-3 puede dar sólo un rango para el valor P, y la determinación de ese rango suele ser algo difícil. La tecnología proporciona automáticamente el valor P, por lo que la tecnología hace que este método sea bastante sencillo. Vea la pantalla adjunta de XLSTAT que muestra el estadístico de prueba $t = -0.660$ (redondeado) y el valor P de 0.5172.

XLSTAT

Difference	-0.1267
t (Observed value)	-0.6599
|t| (Critical value)	2.0926
DF	19
p-value (Two-tailed)	0.5172
alpha	0.05

Método del valor crítico

Si no se dispone de una tecnología, el método del valor crítico para probar una afirmación sobre dos medias es generalmente más fácil que el método del valor P. El ejemplo 1 se puede resolver usando el método del valor crítico. Para encontrar los valores críticos en la tabla A-3, usamos gl = el menor de $n_1 - 1$ y $n_2 - 1$ como una manera relativamente fácil de evitar el uso del cálculo realmente intrincado que requiere la fórmula 9-1. En el ejemplo 1, con tamaños de muestra de $n_1 = 12$ y $n_2 = 15$, el número de grados de libertad es 11, que es el menor entre 11 y 14. En la tabla A-3 con gl $= 11$ y $\alpha = 0.05$ en dos colas, se obtienen los valores críticos de $t = \pm 2.201$. La tecnología se puede utilizar para encontrar los valores críticos más precisos de $t = \pm 2.093$. La figura 9-2 muestra el dato estadístico de prueba más preciso y los valores críticos que se encuentran con la tecnología. El dato estadístico de prueba de $t = -0.660$ cae entre los valores críticos, por lo que no se encuentra en la región crítica y no podemos rechazar la hipótesis nula, como hicimos en el ejemplo 1.

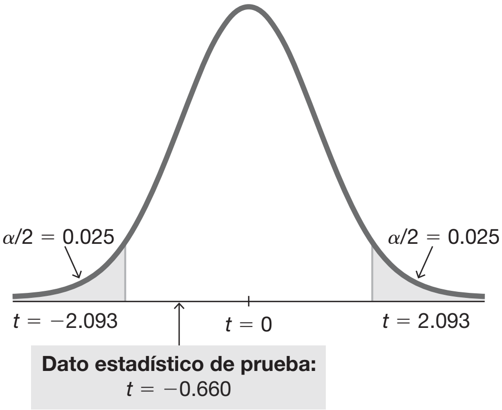


FIGURA 9-2 Prueba de hipótesis de las medias de dos poblaciones independientes

Intervalos de confianza

EJEMPLO 2 Intervalo de confianza para las puntuaciones dadas a cursos impartidos por profesores y profesoras

Con base en los datos muestrales dados en el ejemplo 1, elabore una estimación del intervalo de confianza del 95% para la diferencia entre la puntuación media de la evaluación de cursos impartidos por profesoras y la puntuación media de la evaluación de cursos impartidos por profesores.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Debido a que estamos utilizando los mismos datos del ejemplo 1, aquí se aplica la misma verificación de requisitos, por lo que se satisfacen todos los requisitos.

Primero encontramos el valor del margen de error E. En la tabla A-3 con gl = 11 y $\alpha = 0.05$ en dos colas, obtenemos los valores críticos de $t = \pm 2.201$. (Se puede usar la tecnología para encontrar los valores críticos más precisos de $t = \pm 2.093$).

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.201\sqrt{\frac{0.563001^2}{12} + \frac{0.395450^2}{15}} = 0.422452$$

Usando $E = 0.422452$, $\bar{x}_1 = 3.866667$ y $\bar{x}_2 = 3.993333$, ahora podemos determinar el intervalo de confianza de la siguiente manera:

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

$$-0.55 < (\mu_1 - \mu_2) < 0.30$$

Si usamos la tecnología para obtener resultados más precisos, determinamos el intervalo de confianza $-0.53 < (\mu_1 - \mu_2) < 0.27$, por lo que podemos ver que el intervalo de confianza anterior es bastante bueno, aunque utilizamos un método simplificado para encontrar el número de grados de libertad (en lugar de obtener resultados más precisos utilizando la fórmula 9-1).

INTERPRETACIÓN

Tenemos un 95% de confianza en que los límites de -0.53 y 0.27 realmente contienen la diferencia entre las dos medias poblacionales. Debido a que esos límites contienen a 0, este intervalo de confianza sugiere que no hay una diferencia significativa entre la puntuación media de las evaluaciones de cursos impartidos por profesoras y la puntuación media de las evaluaciones de cursos impartidos por profesores.

Súper Bowl

Se invitó a un grupo de estudiantes a un juego del Súper Bowl, y a la mitad de ellos se les dieron tazones grandes, con capacidad de 4 litros, con botanas, mientras que la otra mitad recibió tazones más pequeños, de 2 litros, con el mismo tipo de contenido. Los que recibieron los tazones grandes consumieron un 56% más botanas que los que utilizaron los tazones pequeños. (Vea “Super Bowls: Serving Bowl Size and Food Consumption”, de Wansink y Cheney, *Journal of the American Medical Association*, vol. 293, núm. 14).

Otro estudio demostró que existe “un incremento significativo de los accidentes automovilísticos fatales las horas posteriores a la transmisión del Súper Bowl en Estados Unidos”. Los investigadores analizaron 20,377 muertes en 27 domingos de Súper Bowl y en otros 54 domingos que se utilizaron como control. Encontraron un incremento del 41% en las muertes después de los juegos del Súper Bowl. (Vea “Do Fatal Crashes Increase Following A Super Bowl Telecast?”, de Redelmeier y Stewart, *Chance*, vol. 18, núm. 1).

PARTE 2 Métodos alternativos

La parte 1 de esta sección trató de situaciones en las que las dos desviaciones estándar poblacionales son desconocidas y no se supone que sean iguales. En la parte 2 abordamos otras dos situaciones:

1. Las dos desviaciones estándar poblacionales son desconocidas, pero se supone que son iguales.
2. Las dos desviaciones estándar poblacionales son conocidas.

Método alternativo: suponga que $\sigma_1 = \sigma_2$ y combine las varianzas muestrales

Incluso cuando no se conocen los valores específicos de σ_1 y σ_2, si se puede suponer que tienen *algún* valor, las varianzas muestrales s_2^1 y s_2^2 se pueden *combinar* para obtener una estimación de la varianza poblacional común σ^2. La **estimación combinada de σ^2** se expresa con s_p^1 y es un promedio ponderado de s_2^1 y s_2^2, que se usa en el dato estadístico de prueba para este caso:

$$\textbf{Dato estadístico de prueba} \quad t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$$

$$\text{donde } s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} \text{ (varianza muestral combinada)}$$

y el número de grados de libertad es gl $= n_1 + n_2 - 2$.

Los requisitos para este caso son los mismos que en la parte 1, excepto que el primer requisito es que σ_1 y σ_2 no se conocen, pero se supone que son iguales. Los intervalos de confianza se encuentran al evaluar $(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$ con el siguiente margen de error E.

$$\textbf{Margen de error para el intervalo de confianza} \quad E = t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

donde s_p^2 está dado en el estadístico de prueba anterior, y gl $= n_1 + n_2 - 2$.

¿Cuándo deberíamos suponer que $\sigma_1 = \sigma_2$? Si usamos la aleatoriedad para asignar sujetos a los grupos de tratamiento y placebo, sabemos que las muestras provienen de la misma población. Entonces, si realizamos una prueba de hipótesis suponiendo que dos medias poblacionales son iguales, no es irrazonable suponer que las muestras son de poblaciones con las mismas desviaciones estándar (pero aún debemos verificar ese supuesto).

Ventaja de combinar La ventaja de este método alternativo consistente en combinar las varianzas muestrales es que el número de grados de libertad es un poco más alto, por lo que las pruebas de hipótesis tienen más potencia y los intervalos de confianza son un poco más estrechos.

Método alternativo usado cuando se conocen σ_1 y σ_2

En realidad, las desviaciones estándar poblacionales σ_1 y σ_2 casi nunca se conocen, pero si de alguna manera son conocidas, el estadístico de prueba y el intervalo de confianza se basan en la distribución normal y no en la distribución t. Los requisitos son los mismos que los establecidos en la parte 1, a excepción del primer requisito que pide se conozcan σ_1 y σ_2. Los valores críticos y los valores P se encuentran usando la tecnología o la tabla A-2, y el dato estadístico de prueba para este caso es el siguiente:

$$\textbf{Dato estadístico de prueba} \quad z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

Los intervalos de confianza se encuentran evaluando
$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$, donde:

Margen de error para el intervalo de confianza $E = z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

¿Qué ocurre si se conoce una desviación estándar y la otra es desconocida? Si se conoce σ_1 pero σ_2 es desconocida, utilice los procedimientos de la parte 1 de esta sección con los siguientes cambios: Reemplace s_1 con el valor conocido de σ_1 y use el número de grados de libertad encontrado a partir de la expresión siguiente. (Vea "The Two-Sample *t* Test with One Variance Unknown", de Maity y Sherman, *The American Statistician*, vol. 60, núm. 2).

$$df = \frac{\left(\dfrac{\sigma_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Estrategia recomendada para dos medias independientes

La estrategia recomendada para los métodos de esta sección es:

Suponga que σ_1 y σ_2 son desconocidas, *no* suponga que $\sigma_1 = \sigma_2$, y use el dato estadístico de prueba y el intervalo de confianza dados en la parte 1 de esta sección.

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 sustraída. Los siguientes son algunos indicadores que se utilizan para ello. En periodos comparables, las muestras de ventas tienen medias que son significativamente diferentes. La cantidad media de ventas decrece de manera significativa. Existe un incremento significativo en la proporción de registros de "no venta" en la caja registradora. Existe una disminución significativa en la proporción de la recepción de efectivo y de cheques. Se pueden aplicar los métodos de prueba de hipótesis para identificar indicadores como estos. (Vea "How to Catch a Thief", de Manly y Thomson, *Chance*, vol. 11, núm. 4)

CENTRO DE TECNOLOGÍA

Inferencias con dos medias: muestras independientes

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Pruebas de hipótesis

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y **Mean-Two Independent Samples** del submenú.
3. Seleccione el formato deseado para la *hipótesis alternativa* e ingrese el nivel de significancia. Para ambas muestras, introduzca los datos estadísticos muestrales o haga clic en la pestaña **Use Data** para usar columnas de datos.
4. Para el *método de análisis*, seleccione **Unequal variances: No Pool.**
5. Haga clic en **Evaluate**.

Intervalos de confianza

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Confidence Intervals** en el menú desplegable y **Mean-Two Independent Samples** del submenú.
3. Ingrese el nivel de confianza deseado. Para ambas muestras, introduzca los datos estadísticos muestrales o haga clic en la pestaña **Use Data** para usar columnas de datos.
4. Para el *método de análisis*, seleccione **Unequal variances: No Pool**.
5. Haga clic en **Evaluate**.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y elija **2-sample t** del submenú.
3. *Si se usan datos estadísticos de resumen:* Seleccione **Summarized data** del menú desplegable e ingrese el tamaño de muestra, la media muestral y la desviación estándar muestral para cada muestra.
   Si se usan datos muestrales: Seleccione **Each sample is in its own column** del menú desplegable y elija las columnas de datos deseadas.
4. Haga clic en el botón **Options** e ingrese el nivel de confianza. Introduzca **0** para *diferencia hipotética* y seleccione el formato deseado para la *hipótesis alternativa*.
5. No marque **Assume equal variances**. Marque esta casilla sólo si desea suponer que las poblaciones tienen las mismas varianzas —lo cual no es recomendable.
6. Haga clic en **OK** dos veces.

Sugerencia: Otro procedimiento consiste en hacer clic en **Assistant** en el menú superior, seleccionar **Hypothesis Tests** y **2-Sample t**. Complete el cuadro de diálogo para obtener los resultados, incluido el valor *P* y otra información útil.

continúa

CENTRO DE TECNOLOGÍA *continuación*

Inferencias con dos medias: muestras independientes

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **T Stats** en el menú desplegable, luego elija **Two Sample** en el submenú.
3. *Si se usan datos estadísticos de resumen:* Seleccione **With Summary** en el submenú e ingrese la media muestral, la desviación estándar muestral y el tamaño de cada muestra.
   Si se usan datos muestrales: Seleccione **With Data** del submenú y elija la columna de datos deseada para cada muestra.
4. No marque **Pool variances**. Marque esta casilla sólo si desea suponer que las poblaciones tienen las mismas varianzas—lo cual no es recomendable.
5. Para *pruebas de hipótesis:* Seleccione **Hypothesis test for $\mu_1 - \mu_2$**. Para la diferencia hipotética (H_0) ingrese **0** y seleccione el formato deseado para la hipótesis alternativa (H_A).
   Intervalos de confianza: Seleccione **Confidence interval for $\mu_1 - \mu_2$** e ingrese el nivel de confianza.
6. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

Pruebas de hipótesis:

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **2-SampTTest** en el menú y presione ENTER.
3. Seleccione **Data** si tiene datos muestrales en listas o **Stats** si posee estadísticos de resumen. Presione ENTER e introduzca los nombres de la lista (establezca *Freq* = **1**) o los estadísticos de resumen.
4. Para μ_1, seleccione el formato deseado para la hipótesis alternativa.
5. Para pruebas combinadas seleccione **No**. Seleccione *Yes* sólo si se cree que las varianzas poblacionales son iguales.
6. Seleccione **Calculate** y presione ENTER.

Intervalos de confianza:

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **2-SampTInt** en el menú y presione ENTER.
3. Seleccione **Data** si tiene datos muestrales en listas o **Stats** si posee estadísticos de resumen. Presione **ENTER** e ingrese los nombres de la lista (establezca *Freq* = **1**) o los estadísticos de resumen.
4. Para *C-Level* introduzca el nivel de confianza deseado.
5. Para pruebas combinadas seleccione **No**. Seleccione *Yes* sólo si se cree que las varianzas poblacionales son iguales.
6. Seleccione **Calculate** y presione ENTER.

Excel

Prueba de hipótesis

Complemento XLSTAT

Requiere datos muestrales originales; no funciona con datos resumidos.

1. Haga clic en la pestaña **XLSTAT** en la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **Two-sample t-test and z-test** en el menú desplegable.
3. En las *muestras 1 y 2*, ingrese el rango de celdas que contiene los datos muestrales. Para el *formato de datos*, seleccione **One column per sample**.
4. Seleccione **Student's t test**.
5. Si la primera fila de datos contiene una etiqueta, marque la casilla de **Column labels**.
6. Haga clic en la pestaña **Options**.
7. En *hipótesis alternativa*, seleccione el formato deseado ($\neq$ para pruebas de dos colas, $<$ para pruebas de cola izquierda, $>$ para pruebas de cola derecha). Ingrese **0** para la *diferencia hipotética* (*D*) e introduzca el nivel de significancia deseado (ingrese **5** para un nivel de significancia de 0.05).
8. Desmarque la casilla **Assume equality** y desmarque la casilla **Cochran-Cox**. Desmarque la casilla **Use an *F*-test**.
9. Haga clic en **OK** para desplegar el dato estadístico de prueba (etiquetado como *t Observed value*) y el valor *P*.

CENTRO DE TECNOLOGÍA *continuación*

Inferencias con dos medias: muestras independientes

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Excel

Excel (Complemento de análisis de datos)

1. Haga clic en **Data** en la cinta, luego haga clic en la pestaña **Data Analysis**.
2. Seleccione **t-Test: Two-Sample Assuming Unequal Variances** y haga clic en **OK**.
3. Ingrese el rango de datos para cada variable en los cuadros de *rango de variables*. Si la primera fila contiene una etiqueta, marque la casilla **Labels**.
4. Ingrese **0** para la *diferencia hipotética de las medias*.
5. Ingrese el nivel de significancia deseado en la casilla *Alpha* y haga clic en **OK**. El dato estadístico de prueba se etiqueta *t Stat* y el valor *P* se etiqueta *P*.

Intervalo de confianza

Complemento XLSTAT (requerido)

Requiere datos muestrales originales; no funciona con datos resumidos.

1 a 5. Siga los pasos 1 a 5 para la **prueba de hipótesis** usando el complemento XLSTAT.

6. Haga clic en la pestaña **Options**.
7. En *hipótesis alternativa*, seleccione la opción de dos colas ≠. Ingrese **0** para la *diferencia hipotética* (*D*) e introduzca el nivel de significancia deseado (ingrese **5** para un nivel de confianza del 95%).
8. Desmarque la casilla **Assume equality** y desmarque la casilla **Cochran-Cox**. Desmarque la casilla **Use an *F*-test**.
9. Haga clic en **OK** para desplegar el intervalo de confianza.

9-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Muestras independientes y dependientes ¿Cuál de las siguientes opciones implica muestras independientes?

a. El conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B incluye pares de edades de actrices y actores en el momento en que ganaron un Oscar a las categorías de Mejor Actriz y Mejor Actor. Se lista el par de edades de los ganadores para cada año, y cada par consta de edades que coinciden con el año en que se ganaron los Oscar.

b. Conjunto de datos 15 "Presidentes" en el apéndice B se incluyen las estaturas de los presidentes electos junto con las estaturas de sus principales oponentes. El par de estaturas se lista para cada elección.

c. El conjunto de datos 26 "Pesos y volúmenes de bebidas de cola" en el apéndice B incluye los volúmenes de los contenidos en 36 latas de Coca-Cola regular y los volúmenes de los contenidos en 36 latas de Pepsi regular.

2. Intervalo de confianza para hemoglobina Se obtienen muestras grandes de mujeres y hombres, y se mide el nivel de hemoglobina en cada sujeto. A continuación se presenta el intervalo de confianza del 95% para la diferencia entre las dos medias poblacionales, donde las medidas de las mujeres corresponden a la población 1 y las medidas de los hombres corresponden a la población 2: $-1.76 \text{ g/dL} < \mu_1 - \mu_2 < -1.62 \text{ g/dL}$.

a. ¿Qué sugiere el intervalo de confianza sobre la igualdad del nivel medio de hemoglobina en las mujeres y el nivel medio de hemoglobina en los hombres?

b. Escriba un enunciado breve que interprete ese intervalo de confianza.

c. Exprese el intervalo de confianza con las medidas de los hombres como la población 1 y las medidas de las mujeres como la población 2.

3. Pruebas de hipótesis e intervalos de confianza para hemoglobina

a. El ejercicio 2 incluye un intervalo de confianza. Si se usa el método del valor P o el método del valor crítico de la parte 1 de esta sección para evaluar la afirmación de que las mujeres y los hombres tienen los mismos niveles medios de hemoglobina, ¿las pruebas de hipótesis y el intervalo de confianza darán la misma conclusión?

b. En general, si realiza una prueba de hipótesis utilizando los métodos de la parte 1 de esta sección, el método del valor P, el método del valor crítico, y el método del intervalo de confianza resultan en la misma conclusión?

c. Suponga que desea usar un nivel de significancia de 0.01 para probar la afirmación de que el nivel medio de hemoglobina en las mujeres es *menor* que el nivel medio de hemoglobina en los hombres. ¿Qué *nivel de confianza* debería usarse si desea probar esa afirmación utilizando un intervalo de confianza?

4. Grados de libertad Para el ejemplo 1 de la página 431, se usó gl = el menor de $n_1 - 1$ y $n_2 - 1$, de donde se obtuvo gl = 11 y los valores críticos correspondientes $t = \pm 2.201$. Si calculamos gl usando la fórmula 9-1, obtenemos gl = 19.063, y los valores críticos correspondientes son $t = \pm 2.093$. ¿De qué manera el uso de los valores críticos de $t = \pm 2.201$ es más "conservador" que el uso de los valores críticos $t = \pm 2.093$?

En los ejercicios 5 a 20, suponga que las dos muestras son muestras aleatorias simples e independientes, seleccionadas de poblaciones distribuidas normalmente, y no asuma que las desviaciones estándar poblacionales son iguales. (Nota: Las respuestas en el apéndice D consideran el uso de tecnologías basadas en la fórmula 9-1, junto con respuestas basadas en la tabla A-3 con gl igual al menor de $n_1 - 1$ y $n_2 - 1$).

5. Coca-Cola regular y Diet Coke El conjunto de datos 26 "Pesos y volúmenes de bebidas de Cola" en el apéndice B incluye los pesos (en lb) del contenido de latas de Diet Coke ($n = 36$, $\bar{x} = 0.78479$ lb, $s = 0.00439$ lb) y del contenido de latas de Coca-Cola regular ($n = 36$, $\bar{x} = 0.81682$ lb, $s = 0.00751$ lb).

a. Use un nivel de significancia de 0.05 para probar la afirmación de que el contenido de las latas de Diet Coke tienen pesos con una media menor que la media para la Coca-Cola regular.

b. Elabore el intervalo de confianza adecuado para la prueba de hipótesis del inciso (a).

c. ¿Puede usted explicar por qué las latas de Diet Coke pesarían menos que las latas de Coca-Cola regular?

6. Coca-Cola y Pepsi El conjunto de datos 26 "Pesos y volúmenes de bebidas de Cola" en el apéndice B incluye volúmenes contenidos en latas de Coca-Cola regular ($n = 36$, $\bar{x} = 12.19$ oz, $s = 0.11$ oz) y volúmenes contenidos en latas de Pepsi regular ($n = 36$, $\bar{x} = 12.29$ oz, $s = 0.09$ oz).

a. Use un nivel de significancia de 0.05 para probar la afirmación de que las latas de Coca-Cola regular y Pepsi regular tienen el mismo volumen medio.

b. Elabore el intervalo de confianza adecuado para la prueba de hipótesis del inciso (a).

c. ¿Qué se puede concluir? ¿Parece haber una diferencia? ¿Hay algún significado práctico?

7. Color y creatividad Investigadores de la University of British Columbia realizaron pruebas para investigar los efectos del color en la creatividad. A los sujetos con un fondo rojo se les pidió que pensaran en usos creativos para un ladrillo, a otros sujetos con un fondo azul se les dio la misma tarea. Las respuestas fueron calificadas por un panel de jueces y los resultados de las puntuaciones de creatividad se dan a continuación. Los puntajes más altos corresponden a más creatividad. Los investigadores afirman que "el azul mejora el rendimiento en una tarea creativa".

a. Use un nivel de significancia de 0.01 para probar la afirmación de que el azul mejora el rendimiento en una tarea creativa.

b. Elabore el intervalo de confianza adecuado para la prueba de hipótesis del inciso (a). ¿Qué indica el hecho de que este intervalo de confianza nos conduzca a la misma conclusión del inciso (a)?

Fondo rojo:	$n = 35$, $\bar{x} = 3.39$, $s = 0.97$
Fondo azul:	$n = 36$, $\bar{x} = 3.97$, $s = 0.63$

8. Color y cognición Algunos investigadores de la University of British Columbia realizaron un estudio para investigar los efectos del color en tareas cognitivas. Se mostraron palabras en una pantalla de computadora con colores de fondo en rojo y azul. A continuación se muestran los resultados de las puntuaciones en una prueba sobre la memorización de las palabras. Los puntajes más altos corresponden a un mayor recuerdo de las palabras.

a. Use un nivel de significancia de 0.05 para probar la afirmación de que las muestras provienen de poblaciones con la misma media.

b. Elabore un intervalo de confianza adecuado para la prueba de hipótesis del inciso (a). ¿Qué se puede decir del intervalo de confianza que conduce a la misma conclusión que en el inciso (a)?

c. ¿El color de fondo parece tener un efecto en las puntuaciones para el recuerdo de las palabras? Si es así, ¿qué color parece estar asociado con los puntajes más altos de la memorización de las palabras?

Fondo rojo	$n = 35, \bar{x} = 15.89, s = 5.90$
Fondo azul	$n = 36, \bar{x} = 12.31, s = 5.48$

9. Tratamiento del dolor con imanes Las personas gastan alrededor de 55 mil millones de dólares anuales en la compra de imanes utilizados para tratar una gran variedad de dolores. Los investigadores realizaron un estudio para determinar si los imanes son efectivos en el tratamiento del dolor de espalda. El dolor se midió utilizando la escala analógica visual, y los resultados mostrados a continuación se encuentran entre las conclusiones del estudio (basado en 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). Las puntuaciones más altas corresponden a mayores niveles de dolor.

a. Use un nivel de significancia de 0.05 para probar la afirmación de que las personas tratadas con imanes tienen una reducción media del dolor mayor que quienes reciben un tratamiento simulado (similar a un placebo).

b. Elabore el intervalo de confianza adecuado para la prueba de hipótesis del inciso (a).

c. ¿Parece que los imanes son efectivos para tratar el dolor de espalda? ¿Es válido argumentar que los imanes podrían parecer efectivos si los tamaños de muestra fueran más grandes?

Reducción en el nivel de dolor después del tratamiento con imanes: $n = 20, \bar{x} = 0.49, s = 0.96$

Reducción del nivel de dolor después del tratamiento simulado: $n = 20, \bar{x} = 0.44, s = 1.4$

10. Fumadores pasivos El conjunto de datos 12 "Fumadores pasivos y activos" en el apéndice B incluye los niveles de cotinina medidos en un grupo de no fumadores expuestos al humo de tabaco ($n = 40$, $\bar{x} = 60.58$ ng/mL, $s = 138.08$ ng/mL) y un grupo de no fumadores que no está expuesto al humo de tabaco ($n = 40$, $\bar{x} = 16.35$ ng/mL, $s = 62.53$ ng/mL). La cotinina es un metabolito de la nicotina, lo que significa que cuando el cuerpo absorbe nicotina, se produce la cotinina.

a. Use un nivel de significancia de 0.05 para evaluar la afirmación de que los no fumadores expuestos al humo de tabaco tienen un nivel medio de cotinina más alto que los no fumadores que no están expuestos al humo de tabaco.

b. Elabore el intervalo de confianza adecuado para la prueba de hipótesis del inciso (a).

c. ¿Qué se puede concluir sobre los efectos del humo en los fumadores pasivos?

11. IMC Sabemos que el peso medio de los hombres es mayor que el peso medio de las mujeres, y la altura media de los hombres es mayor que la altura media de las mujeres. El índice de masa corporal (IMC) de una persona se calcula dividiendo el peso (kg) por el cuadrado de la altura (m). A continuación se presentan las estadísticas de IMC para muestras aleatorias de mujeres y hombres tomadas del conjunto de datos 1 "Datos corporales" en el apéndice B.

a. Use un nivel de significancia de 0.05 para evaluar la afirmación de que las mujeres y los hombres tienen el mismo IMC promedio.

b. Elabore el intervalo de confianza adecuado para probar la afirmación del inciso (a).

c. ¿Las mujeres y los hombres parecen tener el mismo IMC medio?

IMC femenino: $n = 70, \bar{x} = 29.10, s = 7.39$

IMC masculino: $n = 80, \bar{x} = 28.38, s = 5.37$

12. IQ y exposición al plomo El Conjunto de datos 7 "IQ y plomo" en el apéndice B lista las puntuaciones de IQ completos para una muestra aleatoria de sujetos con bajos niveles de plomo en la sangre y otra muestra aleatoria de sujetos con niveles altos de plomo en la sangre. Los estadísticos se resumen a continuación.

a. Use un nivel de significancia de 0.05 para evaluar la afirmación de que la puntuación media de IQ de las personas con niveles bajos de plomo en la sangre es más alta que la puntuación media de IQ de las personas con niveles altos de plomo en la sangre.

b. Elabore un intervalo de confianza adecuado para la prueba de hipótesis en el inciso (a).

c. ¿La exposición al plomo parece tener un efecto en las puntuaciones de IQ?

Nivel bajo de plomo en la sangre: $n = 78, \bar{x} = 92.88462, s = 15.34451$

Nivel alto de plomo en la sangre: $n = 21, \bar{x} = 86.90476, s = 8.988352$

13. ¿Los profesores y las profesoras son calificados de manera diferente?

a. Use un nivel de significancia de 0.05 para probar la afirmación de que dos muestras de puntuaciones de evaluación de cursos provienen de poblaciones con la misma media. Utilice los siguientes estadísticos de resumen: profesoras: $n = 40, \bar{x} = 3.79, s = 0.51$; profesores: $n = 53, \bar{x} = 4.01, s = 0.53$. (El uso de los datos brutos en el conjunto de datos 17 "Evaluaciones de cursos" arrojará resultados diferentes).

b. Con base en los estadísticos de resumen dados en el inciso (a), elabore una estimación del intervalo de confianza del 95% para la diferencia entre la puntuación media de las evaluaciones de cursos para las profesoras y los profesores.

c. En el ejemplo 1 se usaron datos muestrales similares con muestras de tamaño 12 y 15, y se llegó a la conclusión de que no hay evidencia suficiente para justificar el rechazo de la hipótesis nula. ¿Las muestras más grandes en este ejercicio afectan mucho los resultados?

14. Cinturones de seguridad Un estudio sobre el uso del cinturón de seguridad involucró a niños que fueron hospitalizados después de choques automovilísticos. Para un grupo de 123 niños que usaban el cinturón de seguridad, el número de días en unidades de cuidados intensivos (UCI) tiene una media de 0.83 y una desviación estándar de 1.77. Para un grupo de 290 niños que no usaban el cinturón de seguridad, el número medio de días pasados en las UCI es de 1.39 y la desviación estándar de 3.06 (según datos 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).

a. Use un nivel de significancia de 0.05 para probar la afirmación de que los niños que usan cinturones de seguridad tienen un tiempo medio menor en una UCI que la media de los niños que no usan cinturones de seguridad.

b. Elabore un intervalo de confianza adecuado para la prueba de hipótesis en el inciso (a).

c. ¿Qué conclusión importante sugieren los resultados?

15. ¿Las monedas de ¢25 ahora son más ligeras? Los pesos de las monedas de ¢25 se consideran rigurosamente en el diseño de las máquinas expendedoras que todos hemos llegado a conocer y amar. El conjunto de datos 29 "Pesos de monedas" en el apéndice B incluye pesos de una muestra de monedas de ¢25 anteriores a 1964 ($n = 40, \bar{x} = 6.19267$ g, $s = 0.08700$ g) y pesos de una muestra de monedas de ¢25 posteriores a 1964 ($n = 40, \bar{x} = 5.63930$ g, $s = 0.06194$ g).

a. Use un nivel de significancia de 0.05 para probar la afirmación de que las monedas de ¢25 previas a 1964 tienen un peso medio que es mayor que el peso medio de las monedas de ¢25 posteriores a 1964.

b. Elabore un intervalo de confianza adecuado para la prueba de hipótesis en el inciso (a).

c. ¿Las monedas de ¢25 posteriores a 1964 parecen pesar menos que antes de 1964? Si es así, ¿por qué las máquinas expendedoras no se ven afectadas en gran medida por la diferencia?

16. Cosas malas en películas para niños El conjunto de datos 11 "Alcohol y tabaco en películas" del apéndice B incluye la cantidad de tiempo (en segundos) que se muestra consumo de tabaco en las películas animadas para niños. Para las películas de Disney, $n = 33, \bar{x} = 61.6$ seg, $s = 118.8$ seg; para otras películas, $n = 17, \bar{x} = 49.3$ seg, $s = 69.3$ seg. Los tiempos ordenados para las películas que no son de Disney se listan a continuación.

a. Use un nivel de significancia de 0.05 para probar la afirmación de que las películas animadas para niños de Disney y otras películas animadas para niños tienen el mismo tiempo medio que muestra consumo de tabaco.

b. Elabore un intervalo de confianza adecuado para la prueba de hipótesis del inciso (a).

c. Realice una inspección visual rápida de los tiempos listados para las películas que no son de Disney y comente sobre los requisitos de normalidad. ¿Cómo afecta la normalidad de los 17 tiempos que no son de Disney a los resultados?

0 0 0 0 0 0 1 5 6 17 24 55 91 117 155 162 205

17. ¿Los profesores y las profesoras son calificados de manera diferente? A continuación se listan puntuaciones de evaluación emitidos por estudiantes a sus profesoras y profesores, en el conjunto de datos 17 "Evaluaciones de cursos" del apéndice B. Pruebe la afirmación de que las profesoras y los profesores tienen las mismas puntuaciones medias de evaluación. ¿Parece haber una diferencia?

Profesoras	4.4	3.4	4.8	2.9	4.4	4.9	3.5	3.7	3.4	4.8
Profesores	4.0	3.6	4.1	4.1	3.5	4.6	4.0	4.3	4.5	4.3

18. Edades de automóviles y taxis Cuando el autor visitó Dublín, Irlanda (hogar del empleado William Gosset de Guinness Brewery, quien desarrolló por primera vez la distribución *t*), registró las edades de automóviles de pasajeros seleccionados al azar y taxis seleccionados aleatoriamente. Las edades se pueden encontrar en las placas. (La diversión de viajar con el autor no tiene límite). Las edades (en años) se detallan a continuación. Podríamos esperar que los taxis sean más nuevos, así que pruebe la afirmación de que la edad media de los autos es mayor que la edad media de los taxis.

Edades de autos	4	0	8	11	14	3	4	4	3	5	8	3	3	7	4	6	6	1	8	2	15	11	4	1	6	1	8
Edades de taxis	8	8	0	3	8	4	3	3	6	11	7	7	6	9	5	10	8	4	3	4							

19. ¿Old Faithful no es tan fiel? A continuación se listan los intervalos de tiempo (en minutos) entre las erupciones del géiser Old Faithful. Los tiempos "recientes" ocurrieron en los últimos años, y los tiempos "pasados" son de 1995. ¿Parece que el intervalo de tiempo medio ha cambiado? ¿La conclusión se ve afectada si el nivel de significancia es 0.05 o 0.01?

Recientes	78	91	89	79	57	100	62	87	70	88	82	83	56	81	74	102	61
Pasados	89	88	97	98	64	85	85	96	87	95	90	95					

20. Bloque en exámenes Muchos estudiantes han tenido la desagradable experiencia de entrar en pánico en un examen porque la primera pregunta era excepcionalmente difícil. Se estudió la disposición de las preguntas de examen en cuanto a su efecto sobre la ansiedad. Las siguientes puntuaciones son medidas de "ansiedad debilitante en exámenes", que la mayoría de nosotros llamamos pánico o bloqueo (según datos de "Item Arrangement, Cognitive Entry Characteristics. Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4). ¿Hay suficiente evidencia para apoyar la afirmación de que las dos poblaciones de puntuaciones tienen diferentes medias? ¿Hay suficiente evidencia para respaldar el argumento de que la disposición de las preguntas de examen tiene un efecto en la puntación? ¿La conclusión se ve afectada si el nivel de significancia es 0.05 o 0.01?

Preguntas dispuestas de 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

Preguntas dispuestas de 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

Conjuntos de datos grandes. ***En los ejercicios 21 a 24, use los conjuntos de datos indicados en el apéndice B. Los conjuntos de datos completos se pueden encontrar en www.pearsonenespañol.com/triola. Suponga que las dos muestras son muestras aleatorias simples independientes, seleccionadas de poblaciones normalmente distribuidas. No suponga que las desviaciones estándar de la población son iguales.***

21. ¿Los hombres hablan menos que las mujeres? Consulte el conjunto de datos 24 "Las palabras cuentan" y use los conteos de palabras de los hombres en la tercera columna y los conteos de palabras de las mujeres en la cuarta columna. Use un nivel de significancia de 0.05 para probar la afirmación de que los hombres hablan menos que las mujeres.

22. ¿Los hombres y las mujeres tienen la misma presión arterial diastólica media? Consulte el conjunto de datos 1 "Datos corporales" y use un nivel de significancia de 0.05 para probar la afirmación de que las mujeres y los hombres tienen la misma presión arterial diastólica media.

23. Pesos al nacer Consulte el conjunto de datos 4 "Nacimientos" y use los pesos al nacer de niños y niñas. Pruebe la afirmación de que, al nacer, las niñas tienen un peso medio menor que los niños.

24. Estadías después del parto Consulte el conjunto de datos 4 "Nacimientos" y use la "duración de la estadía" para niños y niñas. La duración de la estadía es la cantidad de días que el bebé permaneció en el hospital. Pruebe la afirmación de que los niños y niñas tienen la misma duración de estadía media.

9-2 Más allá de lo básico

25. Combinación Repita el ejercicio 12 "IQ y plomo" suponiendo que las dos desviaciones estándar poblacionales son iguales, por lo que $\sigma_1 = \sigma_2$. Use el método apropiado de la parte 2 de esta sección. ¿La combinación de las desviaciones estándar arroja resultados que muestran una mayor significancia?

26. Grados de libertad En el ejercicio 20 "Bloqueo en exámenes", el uso de "el menor entre $n_1 - 1$ y $n_2 - 1$" para la cantidad de grados de libertad da como resultado gl $= 15$. Determine el número de grados de libertad con la fórmula 9-1, ¿cómo se ven afectadas las pruebas de hipótesis y los intervalos de confianza con el uso de la fórmula 9-1 en lugar de "el menor entre $n_1 - 1$ y $n_2 - 1$"?

27. Sin variación en una muestra Se realizó un experimento para probar los efectos del alcohol. Los investigadores midieron los niveles de alcohol en el aliento de un grupo de personas bajo tratamiento que bebieron etanol y otro grupo que recibió un placebo. Los resultados se dan a continuación (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). Use un nivel de significancia de 0.05 para probar la afirmación de que los dos grupos muestrales provienen de poblaciones con la misma media.

Grupo de tratamiento: $n_1 = 22, \bar{x}_1 = 0.049, s_1 = 0.015$

Grupo placebo: $n_2 = 22, \bar{x}_2 = 0.000, s_2 = 0.000$

9-3 Dos muestras dependientes (pares relacionados)

Concepto clave En esta sección se presentan métodos para probar hipótesis y elaborar intervalos de confianza que involucran la media de las diferencias entre los valores de dos poblaciones dependientes (*dependientes* en el sentido de que los datos consisten en pares relacionados). Los pares se deben combinar de acuerdo con alguna relación, como las mediciones antes/después de los mismos sujetos o las puntuaciones de IQ de esposos y esposas.

Buen diseño experimental

Suponga que queremos probar la efectividad del curso preparatorio Kaplan para el SAT. Lo mejor sería utilizar puntuaciones del tipo antes/después de un solo grupo de estudiantes que tomaron el curso, y no utilizar puntuaciones de un grupo de estudiantes que no tomaron el curso Kaplan y otro grupo que sí lo tomo. La ventaja de usar pares relacionados (puntuaciones antes/después) es que reducimos la variación externa, lo que podría ocurrir con dos muestras independientes diferentes. Esta estrategia para diseñar un experimento se puede generalizar mediante el siguiente principio de diseño:

> **Cuando se diseña un experimento o se planifica un estudio observacional, el uso de muestras dependientes con pares relacionados suele ser mejor que el uso de dos muestras independientes.**

Déja Vu una y otra vez Los métodos de pruebas de hipótesis en esta sección son los *mismos métodos* para probar una afirmación sobre una media poblacional (parte 1 de la sección 8-3), excepto que aquí usamos las *diferencias* de los pares de datos muestrales relacionados.
No existen procedimientos exactos para tratar con muestras dependientes, pero los siguientes métodos de aproximación se usan con frecuencia.

ELEMENTOS CLAVE

Inferencias sobre diferencias de pares relacionados

Objetivos

1. **Prueba de hipótesis:** Usar las diferencias de dos muestras dependientes (pares relacionados) para probar una afirmación sobre la media poblacional de todas esas diferencias.
2. **Intervalo de confianza:** Usar las diferencias de dos muestras dependientes (pares relacionados) para elaborar una estimación del intervalo de confianza de la media poblacional de todas esas diferencias.

Notación para muestras dependientes

d = diferencia individual entre los dos valores en un solo par relacionado

μ_d = valor medio de las diferencias d para la *población* de todos los pares de datos relacionados

$\overline{d}$ = valor medio de las diferencias d para los datos *muestrales* relacionados

s_d = desviación estándar de las diferencias d para los datos *muestrales* relacionados

n = número de *pares* de datos muestrales

Requisitos

1. Los datos muestrales son dependientes (pares relacionados).
2. Los pares relacionados son una muestra aleatoria simple.
3. Se cumple al menos una de las siguientes condiciones: El número de pares de datos muestrales es grande (n / 30) o los pares de valores tienen diferencias que provienen de una población con distribución aproximadamente normal. Estos métodos son robustos contra las desviaciones de la normalidad, por lo que el requisito de normalidad es laxo.

Dato estadístico de prueba para muestras dependientes (con H_0: $\mu_d = 0$)

$$t = \frac{\overline{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}}$$

Valores P: Los valores P pueden obtenerse automáticamente con tecnología o puede usarse la distribución t de la tabla A-3. Use el procedimiento dado en la figura 8-3 de la página 364.

Valores críticos: Use la tabla A-3 (distribución t). Para los grados de libertad, use gl $= n - 1$.

Intervalos de confianza para muestras dependientes

$$\overline{d} - E < \mu_d < \overline{d} + E$$

$$\text{donde} \quad E = t_{\alpha/2}\frac{s_d}{\sqrt{n}} \quad \text{(Grados de libertad: gl} = n - 1\text{).}$$

Crest y muestras dependientes

A fines de la década de 1950, Procter & Gamble lanzó al mercado la pasta dental Crest como el primer dentífrico con fluoruro. Con la finalidad 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 utilizando una pasta dental ordinaria sin fluoruro. Se asumió que cada par de gemelos tendría hábitos semejantes de alimentación y de cepillado, así como características genéticas similares. Los resultados indicaron 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.

Procedimientos para inferencias con muestras dependientes

1. Verifique que los datos muestrales consisten en muestras dependientes (o pares relacionados) y verifique que se satisfacen los requisitos indicados en el recuadro de elementos clave anterior.
2. Encuentre la diferencia d para cada par de valores muestrales. (*Precaución:* Asegúrese de restar de manera consistente, por ejemplo "antes − después").
3. Determine el valor de $\overline{d}$ (media de las diferencias) y s_d (desviación estándar de las diferencias).
4. Para las pruebas de hipótesis y los intervalos de confianza, utilice los mismos procedimientos de prueba t utilizados para una media poblacional única (como se describió en la parte 1 de la sección 8-3).

Métodos equivalentes

Debido a que la prueba de hipótesis y el intervalo de confianza en esta sección usan la misma distribución y el mismo error estándar, son *equivalentes* en el sentido de que dan como resultado las mismas conclusiones. En consecuencia, una hipótesis nula de que la diferencia media es igual a 0 puede probarse determinando si el intervalo de confianza incluye a 0.

EJEMPLO 1 ¿Las mejores actrices suelen ser más jóvenes que los mejores actores?

A continuación consideraremos un aspecto de cómo tratamos de manera diferente a mujeres y hombres en función de su edad. El conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B lista las edades de las actrices cuando ganaron un Oscar en la categoría de Mejor Actriz, junto con las edades de los actores cuando ganaron un Oscar en la categoría de Mejor Actor. Las edades se corresponden de acuerdo con el año en que se otorgaron los premios. La tabla 9-2 incluye una pequeña selección aleatoria de los datos disponibles para que podamos ilustrar mejor los procedimientos de esta sección. Use los datos muestrales de la tabla 9-2 con un nivel de significancia de 0.05 para probar la afirmación de que, para la población de edades de las mejores actrices y los mejores actores, las diferencias tienen una media menor a 0 (lo que indica que las actrices suelen ser más jóvenes que los mejores actores).

TABLA 9-2 Edades de las mejores actrices y los mejores actores

Actriz (años)	28	28	31	29	35
Actor (años)	62	37	36	38	29
Diferencia d	−34	−9	−5	−9	6

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Abordamos los tres requisitos listados anteriormente en el recuadro de elementos clave. (1) Las muestras son dependientes porque los valores se relacionan por el año en que se otorgaron los premios. (2) Los pares de datos se seleccionan aleatoriamente. Consideraremos que los datos son una muestra aleatoria simple. (3) Debido a que el número de pares de datos es $n = 5$, que no es grande, deberíamos verificar la normalidad de las diferencias y verificar los valores atípicos. No hay valores atípicos, y una gráfica cuantilar normal mostraría que los puntos se aproximan a un patrón de línea recta sin ningún otro patrón, por lo que las diferencias satisfacen el requisito laxo de ser de una población distribuida normalmente. Todos los requisitos se satisfacen.

Seguiremos el mismo método para las pruebas de hipótesis que usamos para probar una afirmación sobre una media (vea la figura 8-1 en la página 360), pero utilizaremos las *diferencias* en lugar de los datos muestrales sin procesar.

Paso 1: La afirmación de que las diferencias tienen una media inferior a 0 se puede expresar como $\mu_d < 0$ años.

Paso 2: Si la afirmación original no es verdadera, tenemos $\mu_d \geq 0$ años.

Paso 3: La hipótesis nula debe expresar igualdad y la hipótesis alternativa no puede incluir igualdad, por lo que

$$H_0: \mu_d = 0 \text{ años} \qquad H_1: \mu_d < 0 \text{ años (afirmación original)}$$

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Usamos la distribución t de Student.

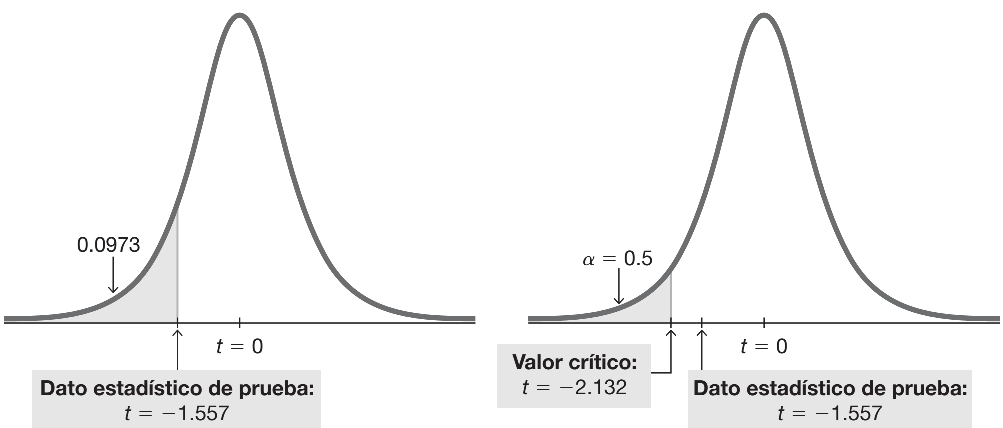
Paso 6: Antes de encontrar el valor del dato estadístico de prueba, primero debemos determinar los valores de $\overline{d}$ y s_d. Usamos las diferencias de la tabla 9-2 ($-34, -9, -5, -9, 6$) para encontrar los siguientes estadísticos muestrales: $\overline{d} = -10.2$ años y $s_d = 14.7$ años. Usando estos estadísticos muestrales y el supuesto de la hipótesis nula de que $\mu_d = 0$ años, ahora podemos determinar el valor del dato estadístico de prueba. (El valor de $t = -1.557$ se obtiene si se usan valores no redondeados de $\overline{d}$ y s_d; la tecnología proporcionará un dato estadístico de prueba $t = -1.557$).

$$t = \frac{\overline{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}} = \frac{-10.2 - 0}{\dfrac{14.7}{\sqrt{5}}} = -1.552$$

Método del valor *P* Debido a que estamos usando una distribución t, nos remitimos a la tabla A-3 en la fila con gl $= 4$ y vemos que el dato estadístico de prueba $t = -1.552$ corresponde a un "área en una cola" que es mayor que 0.05, entonces el valor $P > 0.05$. La tecnología proporcionaría un valor $P = 0.0973$. Vea la figura 9-3 (a).

Método del valor crítico Consulte la tabla A-3 para encontrar el valor crítico de $t = -2.132$ de la siguiente manera: Use la columna para 0.05 (área en una cola), y use la fila con $n - 1 = 4$ grados de libertad. El valor crítico $t = -2.132$ es negativo porque esta prueba es de cola izquierda donde todos los valores de t son negativos. Vea la figura 9-3(b).

Paso 7: Si usamos el método del valor P, no rechazamos H_0 porque el valor P es mayor que el nivel de significancia de 0.05. Si usamos el método del valor crítico, no rechazamos H_0 porque el dato estadístico de prueba no se encuentra en la región crítica.



(a) Método del valor *P* **(b) Método del valor crítico**

FIGURA 9-3 Prueba de hipótesis con muestras dependientes

INTERPRETACIÓN

Concluimos que no hay evidencia suficiente para respaldar $\mu_d < 0$. No hay evidencia suficiente para respaldar la afirmación de que, para la población de edades de las mejores actrices y los mejores actores, las diferencias tienen una media menor a 0. No hay evidencia suficiente para concluir que las mejores actrices suelen ser más jóvenes que los mejores actores.

SU TURNO Resuelva el ejercicio 5 "Prueba de hipótesis de los Oscar".

Gemelos en Twinsburg

Durante el primer fin de semana de agosto de cada año, Twinsburg, Ohio, celebra su festival anual "Día de los gemelos en Twinsburg". Miles de gemelos de todo el mundo han acudido a este festival en el pasado. Los científicos consideraron el festival como una oportunidad para estudiar a gemelos idénticos. Puesto que tienen la misma estructura genética básica, los gemelos idénticos son ideales para estudiar los distintos efectos de la herencia y del ambiente en una gran variedad de rasgos, como la calvicie masculina, las enfermedades cardiacas y la sordera, rasgos que se estudiaron recientemente en uno de los festivales de Twinsburg. Un estudio con gemelos demostró que la miopía se ve muy afectada por factores hereditarios y no por factores ambientales como ver televisión, navegar en Internet o participar en juegos de video o de computadora.

Brecha de género en las pruebas de medicamentos

Un estudio de la relación entre los ataques cerebrales y las dosis de aspirina consumidas involucró a 22,000 médicos varones. Este estudio, como muchos otros, excluyó a las mujeres. La Oficina de Contabilidad General criticó a los Institutos Nacionales de Salud por no incluir ambos sexos en muchos estudios debido a que los resultados de las pruebas médicas en hombres no se aplican necesariamente a las mujeres. Por ejemplo, las cabezas de las mujeres son diferentes de las de los hombres en muchas formas importantes. Al obtener conclusiones basadas en resultados muestrales, debemos desconfiar de una inferencia que se extienda a una población mayor que aquella de la que se extrajo la muestra.

Tecnología El software y las calculadoras normalmente proporcionan un valor P, por lo que usualmente se utiliza el método del valor P para probar hipótesis. Vea la pantalla adjunta de Statdisk que muestra el dato estadístico de prueba $t = -1.5566$ y el valor P de 0.0973. Debido a que el valor P es mayor que el nivel de significancia de 0.05, no rechazamos la hipótesis nula y concluimos que no hay evidencia suficiente para respaldar la afirmación de que, para la población de edades de las mejores actrices y los mejores actores, las diferencias tienen una media menor a 0.

Statdisk

Sample size, n: 5
Difference Mean, d: -10.2
Difference Standard Deviation, sd: 14.65264
Test Statistic, t: -1.5566
Critical t: -2.1318
P-Value: 0.0973

90% Confidence interval:
-24.16971 < μd < 3.769712

EJEMPLO 2 Intervalo de confianza para estimar la media de las diferencias de edad

Usando los mismos datos muestrales de la tabla 9-2, elabore una estimación del intervalo de confianza del 90% para μ_d, que es la media de las diferencias de edad. Al usar un nivel de confianza del 90%, obtenemos un resultado que podría usarse para la prueba de hipótesis en el ejemplo 1. (Debido a que la prueba de hipótesis es unilateral con un nivel de significancia de $\alpha = 0.05$, el nivel de confianza debe ser del 90%. Consulte la tabla 8-1 en la página 360).

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS La solución para el ejemplo 1 incluye la verificación de que se satisfacen los requisitos. ✓

La pantalla de Statdisk anterior muestra el intervalo de confianza del 90%. Se encuentra usando los valores de $\overline{d} = -10.2$ años, $s_d = 14.7$ años, y $t_{\alpha/2} = 2.132$ (encontrado en la tabla A-3 con $n - 1 = 4$ grados de libertad y un área de 0.10 dividida equitativamente entre las dos colas). Primero encontramos el valor del margen de error E.

$$E = t_{\alpha/2}\frac{s_d}{\sqrt{n}} = 2.132 \cdot \frac{14.7}{\sqrt{5}} = 14.015853$$

Ahora determinamos el intervalo de confianza.

$$\overline{d} - E < \mu_d < \overline{d} + E$$
$$-10.2 - 14.015853 < \mu_d < -10.2 + 14.015853$$
$$-24.2 \text{ años} < \mu_d < 3.8 \text{ años}$$

INTERPRETACIÓN

Tenemos un 90% de confianza en que los límites de −24.2 años y 3.8 años contienen el valor real de la media de las diferencias de edad. A largo plazo, el 90% de tales muestras dará lugar a límites del intervalo de confianza que en realidad contengan la media real de la población de las diferencias. Observe que el intervalo de confianza incluye el valor de 0 años, por lo que es muy posible que la media de las diferencias sea igual a 0 años, lo que indica que no hay una diferencia significativa entre las edades de las mejores actrices y de los mejores actores. Recuerde que esta conclusión se basa en la muy pequeña muestra incluida en la tabla 9-2.

Método alternativo utilizado cuando la población no es normal y $n \leq 30$

Bootstrap El recuadro de elementos clave al principio de esta sección incluye el siguiente requisito: El número de pares de datos muestrales es grande ($n > 30$) o los pares de valores tienen diferencias que provienen de una población con distribución aproximadamente normal. Si se infringe esa condición, podemos usar el "Procedimiento Bootstrap para una estimación del intervalo de confianza de un parámetro" incluido en la sección 7-4. Use ese procedimiento para encontrar la *media* de cada muestra bootstrap, luego ordene esas medias, después utilice percentiles para encontrar el intervalo de confianza que puede usarse en las pruebas de hipótesis. Vea el ejercicio 26 "Bootstrap".

CENTRO DE TECNOLOGÍA

Inferencias con dos medias: muestras dependientes

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Requiere datos muestrales relacionados introducidos en columnas.

Pruebas de hipótesis

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y **Mean-Matched Pairs** del submenú.
3. Seleccione el formato deseado para la *hipótesis alternativa*, ingrese el nivel de significancia y seleccione las columnas de datos a comparar.
4. Haga clic en **Evaluate**.

Intervalos de confianza

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Confidence Intervals** en el menú desplegable y **Mean-Matched Pairs** del submenú.
3. Ingrese el nivel de confianza deseado y seleccione las columnas de datos a comparar.
4. Haga clic en **Evaluate**.

Minitab

Requiere datos muestrales relacionados introducidos en columnas.

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y elija **Paired t** en el submenú.
3. Seleccione **Each simple is in a column** del menú desplegable y elija las columnas de datos deseadas.
4. Haga clic en el botón **Options** e ingrese el nivel de confianza. Introduzca **0** para la *diferencia hipotética* y seleccione el formato deseado para la *hipótesis alternativa*.
5. Haga clic en **OK** dos veces.

StatCrunch

Requiere datos muestrales relacionados introducidos en columnas.

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **T Stats** en el menú desplegable, luego elija **Paired** en el submenú.
3. Seleccione las columnas que contienen los datos muestrales relacionados.
4. *Pruebas de hipótesis:* Seleccione **Hypothesis test for** $\mu_D = \mu_1 - \mu_2$. Para la diferencia hipotética (H_0), ingrese **0** y seleccione el formato deseado para la hipótesis alternativa (H_A).

   Intervalos de confianza: Seleccione **Confidence interval for** $\mu_D = \mu_1 - \mu_2$ e ingrese el nivel de confianza.
5. Haga clic en **Compute!**

continúa

CENTRO DE TECNOLOGÍA *continuación*

Inferencias con dos medias: muestras dependientes

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

Precaución: No use el elemento del menú **2-SampTTest** porque se aplica sólo a *muestras independientes*.

1. Ingrese los datos para la primera variable en la lista *L1* y los datos para la segunda variable en la lista *L2*.
2. Cree una lista de diferencias y almacene la lista en *L3* ingresando **L1** – **L2** STO› **L3**.
3. Presione STAT, luego seleccione **TESTS** en el menú superior.

Pruebas de hipótesis

4. Elija **T-Test** y presione ENTER.
5. Seleccione **Data** y presione ENTER.
6. Para μ_0 ingrese **0**.
7. Para *List* introduzca **L3** (establezca *Freq* = **1**).
8. Para μ seleccione el formato deseado para la hipótesis alternativa.
9. Seleccione **Calculate** y presione ENTER.

Intervalo de confianza

4. Elija **TInterval** y presione ENTER.
5. Seleccione **Data**, presione ENTER e ingrese **L3** para el nombre de la lista (establezca *Freq* = **1**).
6. Ingrese el nivel de confianza deseado *C-Level*.
7. Seleccione **Calculate** y presione ENTER.

Excel

Requiere *datos muestrales relacionados* introducidos en columnas.

Prueba de hipótesis

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **Two sample t-test and z-test** en el menú desplegable.
3. En las *muestras 1 y 2*, ingrese el rango de celdas que contiene los datos muestrales relacionados. Para el *formato de datos*, seleccione **Paired samples**.
4. Seleccione **Student's t test**.
5. Si la primera fila de datos contiene una etiqueta, marque la casilla de **Column labels**.
6. Haga clic en la pestaña **Options**.
7. En *hipótesis alternativa*, seleccione el formato deseado ($\neq$ para pruebas de dos colas, $<$ para pruebas de cola izquierda, $>$ para pruebas de cola derecha). Ingrese **0** para *diferencia hipotética* (*D*) e introduzca el nivel de significancia deseado (ingrese **5** para nivel de significancia 0.05).
8. Haga clic en **OK** para visualizar el dato estadístico de prueba (etiquetado como *t Observed value*) y el valor *P*.

Excel (complemento de análisis de datos)

1. Haga clic en **Data** en la cinta, luego haga clic en la pestaña **Data Analysis**.
2. Seleccione **t-Test: Paired Two Sample for Means** y haga clic en **OK**.
3. Ingrese el rango de datos para cada variable en las casillas de *rango de variables*. Si la primera fila contiene una etiqueta, marque la casilla **Labels**.
4. Ingrese **0** para la *diferencia de medias hipotética*.
5. Introduzca el nivel de significancia deseado en la casilla *Alpha*.
6. Haga clic en **OK**. Los resultados incluyen el dato estadístico de prueba (etiquetado como *t Stat*), los valores *P* para una prueba de una cola y de dos colas, y los valores críticos para una prueba de una cola y de dos colas.

Intervalo de confianza

Complemento XLSTAT (requerido)

1 a 5. Siga los pasos 1 a 5 para **Pruebas de Hipótesis** usando el complemento XLSTAT.

6. Haga clic en la pestaña **Options**.
7. Para la *hipótesis alternativa*, seleccione la opción $\neq$ de dos colas. Ingrese **0** para la *diferencia hipotética* (*D*) e ingrese el nivel de significancia deseado (ingrese **5** para un nivel de confianza del 95%).
8. Haga clic en **OK** para visualizar el intervalo de confianza.

9-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. ¿Verdadera? Para los métodos de esta sección. ¿Cuál de las siguientes hipótesis es verdadera?

a. Al probar una afirmación con diez pares de estaturas relacionadas, las pruebas de hipótesis usando el método del valor P, el método del valor crítico y el método del intervalo de confianza darán como resultado la misma conclusión.

b. Los métodos de esta sección son *robustos* contra las desviaciones de la normalidad, lo que significa que la distribución de las diferencias muestrales debe ser muy cercana a una distribución normal.

c. Si queremos usar un intervalo de confianza para probar la afirmación de que $\mu_d < 0$ con un nivel de significancia de 0.01, el intervalo de confianza debe tener un nivel de confianza del 98%.

d. Los métodos de esta sección se pueden utilizar con los ingresos anuales de 50 abogados seleccionados al azar en Carolina del Norte y 50 abogados seleccionados al azar en Carolina del Sur.

e. Con diez pares de estaturas relacionadas, los métodos de esta sección requieren que usemos $n = 20$.

2. Notación A continuación se listan las temperaturas corporales de cinco sujetos diferentes, medidos a las 8 AM y nuevamente a las 12 AM (del conjunto de datos 3 "Temperaturas corporales" en el apéndice B). Encuentre los valores de $\overline{d}$ y s_d. En general. ¿Qué representa μ_d?

Temperatura (°F) a las 8 AM	97.8	99.0	97.4	97.4	97.5
Temperatura (°F) a las 12 AM	98.6	99.5	97.5	97.3	97.6

3. Unidades de medida Si los valores listados en el ejercicio 2 se modifican para ser expresados en grados Celsius y no en grados Fahrenheit. ¿Cómo se ven afectados los resultados de las pruebas de hipótesis?

4. Grados de libertad Si usamos los datos muestrales del ejercicio 2 para elaborar un intervalo de confianza del 99%, ¿cuál es el número de grados de libertad que se deben usar para encontrar el valor crítico de $t_{\alpha/2}$? ¿Cuál es el valor crítico $t_{\alpha/2}$?

En los ejercicios 5 a 16, use los datos muestrales relacionados que se presentan, suponga que las muestras son muestras aleatorias simples y que las diferencias tienen una distribución aproximadamente normal.

5. Prueba de hipótesis del Oscar

a. En el ejemplo 1 de la página 444 en esta sección, se usaron solo cinco pares de datos del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B. Repita la prueba de hipótesis del ejemplo 1 usando los datos que se presentan a continuación. Use un nivel de significancia de 0.05 como en el ejemplo 1.

b. Elabore el intervalo de confianza que podría usarse para la prueba de hipótesis descrita en el inciso (a). ¿Qué característica del intervalo de confianza conduce a la misma conclusión obtenida en ese inciso?

Actriz (años)	28	28	31	29	35	26	26	41	30	34
Actor (años)	62	37	36	38	29	34	51	39	37	42

6. Estaturas de presidentes Una teoría popular es que los candidatos presidenciales tienen una ventaja si son más altos que sus principales oponentes. Se listan las estaturas (en cm) de los presidentes junto con las estaturas de sus principales oponentes (del conjunto de datos 15 "Presidentes").

a. Utilice los datos muestrales con un nivel de significancia de 0.05 para probar la afirmación de que, para la población de estaturas de los presidentes y de sus principales oponentes, las diferencias tienen una media mayor a 0 cm.

b. Elabore el intervalo de confianza que podría usarse para la prueba de hipótesis descrita en el inciso (a). ¿Qué característica del intervalo de confianza conduce a la misma conclusión obtenida en ese inciso?

Estatura (en cm) del presidente	185	178	175	183	193	173
Estatura (en cm) del principal oponente	171	180	173	175	188	178

7. Temperaturas corporales A continuación se listan las temperaturas corporales de siete sujetos medidas en dos momentos distintos de un día (del conjunto de datos 3 "Temperaturas corporales" en el apéndice B).

a. Use un nivel de significancia de 0.05 para probar la afirmación de que no hay diferencia entre las temperaturas corporales medidas a las 8 AM y a las 12 AM.

b. Elabore el intervalo de confianza que podría usarse para la prueba de hipótesis descrita en el inciso (a). ¿Qué característica del intervalo de confianza conduce a la misma conclusión obtenida en ese inciso?

Temperatura (°F) a las 8 AM	96.6	97.0	97.0	97.8	97.0	97.4	96.6
Temperatura (°F) a las 12 AM	99.0	98.4	98.0	98.6	98.5	98.9	98.4

8. La palabra hablada A continuación se listan las cantidades de palabras que cada miembro de seis parejas diferentes habla en un día. Los datos se seleccionan aleatoriamente de las dos primeras columnas del conjunto de datos 24 "Las palabras cuentan" en el apéndice B.

a. Use un nivel de significancia de 0.05 para probar la afirmación de que, entre las parejas, los hombres hablan menos palabras en un día que las mujeres.

b. Elabore el intervalo de confianza que podría usarse para la prueba de hipótesis descrita en el inciso (a). ¿Qué característica del intervalo de confianza conduce a la misma conclusión obtenida en ese inciso?

Hombres	15,684	26,429	1,411	7,771	18,876	15,477	14,069	25,835
Mujeres	24,625	13,397	18,338	17,791	12,964	16,937	16,255	18,667

9. Estaturas de madres e hijas A continuación se listan las estaturas (en pulgadas) de madres y sus primeras hijas. Los datos provienen de una revista de Francis Galton. (Consulte el conjunto de datos 5 "Estaturas familiares" en el apéndice B). Use un nivel de significancia de 0.05 para evaluar la afirmación de que no hay diferencia en las estaturas entre las madres y sus primeras hijas.

Estatura de la madre	68.0	60.0	61.0	63.5	69.0	64.0	69.0	64.0	63.5	66.0
Estatura de la hija	68.5	60.0	63.5	67.5	68.0	65.5	69.0	68.0	64.5	63.0

10. Estaturas de padres e hijos A continuación se listan las alturas (en pulgadas) de padres y sus primeros hijos. Los datos provienen de una revista sostenida por Francis Galton. (Consulte el conjunto de datos 5 "Estaturas familiares" en el apéndice B). Use un nivel de significancia de 0.05 para evaluar la afirmación de que no hay diferencia en las estaturas entre los padres y sus primeros hijos.

Estatura del padre	72.0	66.0	69.0	70.0	70.0	70.0	70.0	75.0	68.2	65.0
Estatura del hijo	73.0	68.0	68.0	71.0	70.0	70.0	71.0	71.0	70.0	63.0

11. Citas rápidas: Atributos A continuación se listan las calificaciones de "atributos" otorgadas por los participantes en una sesión de citas rápidas. Cada calificación de atributos es la suma de las puntuaciones para cinco atributos (sinceridad, inteligencia, diversión, ambición, intereses compartidos). Las calificaciones presentadas provienen del conjunto de datos 18 "Citas rápidas" en el apéndice B. Use un nivel de significancia de 0.05 para probar la afirmación de que existe una diferencia entre las calificaciones de atributos femeninos y masculinos.

Calificación de hombre dada por mujer	29	38	36	37	30	34	35	23	43
Calificación de mujer dada por hombre	36	34	34	33	31	17	31	30	42

12. Citas rápidas: Atractivo A continuación se listan las calificaciones de "atractivo" otorgadas por los participantes en una sesión de citas rápidas. Cada calificación de atributo es la suma de las calificaciones de cinco atributos (sinceridad, inteligencia, ambición, diversión, intereses compartidos). Los valores presentados provienen del conjunto de datos 18 "Citas rápidas". Use un nivel de significancia de 0.05 para probar la afirmación de que existe una diferencia entre las calificaciones de atractivo femenino y masculino.

Calificación de hombre dada por mujer	4.0	8.0	7.0	7.0	6.0	8.0	6.0	4.0	2.0	5.0	9.5	7.0
Calificación de mujer dada por hombre	6.0	8.0	7.0	9.0	5.0	7.0	5.0	4.0	6.0	8.0	6.0	5.0

13. Viernes 13 Algunos investigadores recogieron datos sobre el número de admisiones hospitalarias resultantes de accidentes automovilísticos, y los resultados se dan a continuación para los viernes 6 de cada mes y los viernes 13 del mismo mes (según datos de "Is Friday the 13th Bad for Your Health?, de Scanlon *et al*., *British Medical Journal*, vol. 307, disponible en el recurso en línea de conjuntos de datos *Data and Story Line*). Elabore una estimación del intervalo de confianza del 95% para la media poblacional de las diferencias entre los ingresos hospitalarios en los días viernes 6 de un mes y los días viernes 13 de ese mismo mes. Use el intervalo de confianza para evaluar la afirmación de que, cuando el día 13 de un mes cae en viernes, el número de ingresos hospitalarios por choques de vehículos de motor no se ve afectado.

Viernes 6	9	6	11	11	3	5
Viernes 13	13	12	14	10	4	12

14. Dos cabezas son mejores que una A continuación se listan volúmenes cerebrales (en cm^3) de gemelos provenientes del conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B. Elabore una estimación del intervalo de confianza del 99% para la media de las diferencias entre los volúmenes cerebrales de los gemelos nacidos en primer lugar y los gemelos nacidos en segundo lugar. ¿Qué sugiere el intervalo de confianza?

Nacidos en primer lugar	1005	1035	1281	1051	1034	1079	1104	1439	1029	1160
Nacidos en segundo lugar	963	1027	1272	1079	1070	1173	1067	1347	1100	1204

15. Hipnotismo para reducir el dolor Se realizó un estudio con el propósito de investigar la efectividad del hipnotismo para reducir el dolor. Los resultados de los sujetos seleccionados al azar se dan en la tabla adjunta (basada en "An Analysis of Factors That Contribute to the Efficacy of Hypnotic Analgesia", de Price y Barber, *Journal of Abnormal Psychology*, vol. 96, núm. 1) antes y después de la hipnosis; las medidas se dan en centímetros en una escala de dolor. Los valores más altos corresponden a mayores niveles de dolor. Elabore un intervalo de confianza del 95% para la media de las diferencias "antes/después". ¿Parece que el hipnotismo es efectivo para reducir el 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

16. Estaturas masculinas auto-reportadas y medidas Como parte de la Encuesta Nacional de Examen de la Salud y la Nutrición, el Departamento de Salud y Servicios Humanos obtuvo estaturas reportadas por los mismos sujetos (en pulgadas) y estaturas medidas (en pulgadas) para varones entre 12 y 16 años. A continuación se listan los resultados de la muestra. Elabore una estimación del intervalo de confianza del 99% para la diferencia de medias entre las estaturas reportadas y las estaturas medidas. Interprete el intervalo de confianza resultante y comente las implicaciones del hecho que los límites del intervalo de confianza contengan a 0.

Reportadas	68	71	63	70	71	60	65	64	54	63	66	72
Medidas	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67.0	55.6	74.2	65.0	70.8

Conjuntos de datos grandes. ***En los ejercicios 17 a 24, use los conjuntos de datos indicados en el apéndice B. Los conjuntos de datos completos se pueden encontrar en*** **www.pearsonenespañol.com/triola*****. Suponga que los datos muestrales relacionados son muestras aleatorias simples y que las diferencias tienen una distribución aproximadamente normal.***

17. Oscar Repita el ejercicio 5 "Prueba de hipótesis del Oscar" utilizando todos los datos muestrales del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B. Tenga en cuenta que los pares de datos consisten en edades que se relacionan por el año en que se ganó el Oscar. Nuevamente use un nivel de significancia de 0.05.

18. Estaturas de presidentes Repita el ejercicio 6 "Estaturas de presidentes" usando todos los datos muestrales del conjunto de datos 15 "Presidentes" en el apéndice B.

19. Temperaturas corporales Repita el ejercicio 7 "Temperaturas corporales" usando todas las temperaturas corporales de las 8 AM y las 12 AM del día 2 que se listan en el conjunto de datos 3 "Temperaturas corporales" del apéndice B.

20. La palabra hablada Repita el ejercicio 8 "La palabra hablada" usando todos los datos contenidos en las dos primeras columnas del conjunto de datos 24 "Las palabras cuentan" en el apéndice B.

21. Estaturas de madres e hijas Repita el ejercicio 9 "Estaturas de madres e hijas" usando todas las estaturas de madres e hijas listadas en el conjunto de datos 5 "Estaturas familiares" del apéndice B.

22. Estaturas de padres e hijos Repita el ejercicio 10 "Estaturas de padres e hijos" usando todas las estaturas de padres e hijos listadas en el conjunto de datos 5 "Estaturas familiares" del apéndice B.

23. Citas rápidas: Atributos Repita el ejercicio 11 "Citas rápidas: Atributos" usando todas las calificaciones de atributos otorgadas por mujeres y hombres. Las calificaciones se listan en el conjunto de datos 18 "Citas rápidas" del apéndice B.

24. Citas rápidas: Atractivo Repita el ejercicio 12 "Citas rápidas: Atractivo" usando todas las calificaciones de atractivo otorgadas por mujeres y hombres. Las calificaciones se listan en el conjunto de datos 18 "Citas rápidas" del apéndice B.

9-3 Más allá de lo básico

25. Temperaturas corporales Consulte el conjunto de datos 3 "Temperaturas corporales" en el apéndice B y use todos los pares de temperaturas corporales relacionadas a las 8 AM y a las 12 AM del día 1. Si utiliza un nivel de significancia 0.05 para probar una afirmación sobre la diferencia entre las temperaturas a las 8 AM y a las 12 AM del día 1, ¿cómo se ven afectados los resultados de la prueba de hipótesis y del intervalo de confianza si las temperaturas se convierten de grados Fahrenheit a grados Celsius? ¿Cuál es la relación entre los límites del intervalo de confianza para las temperaturas corporales en grados Fahrenheit y los mismos límites para las temperaturas corporales en grados Celsius? *Sugerencia:* Tenga en cuenta que $C = \frac{5}{9}(F - 32)$.

26. Bootstrap

a. Si los datos muestrales relacionados (x, y) son tales que los valores de x no parecen ser de una población con una distribución normal, y los valores de y tampoco parecen ser de una población con distribución normal, ¿se deduce que los valores de $\bar{d}$ no parecerán ser de una población con distribución normal?

b. Para la prueba de hipótesis descrita en el ejercicio 25 "Temperaturas corporales", use las temperaturas en grados Fahrenheit y haga una estimación del intervalo de confianza del 95% para μ_d con base en 1000 muestras bootstrap. Genere las muestras bootstrap utilizando los valores de $\bar{d}$.

9-4 Dos varianzas o desviaciones estándar

Concepto clave En esta sección presentamos la prueba F para probar hipótesis sobre dos varianzas (o desviaciones estándar) poblacionales. La prueba F (llamada así en honor al estadístico Sir Ronald Fisher) usa la distribución F presentada en esta sección. La prueba F requiere que ambas poblaciones tengan distribuciones normales. En vez de ser robusta, esta prueba es muy sensible a desviaciones de las distribuciones normales, por lo que el requisito de normalidad es bastante estricto. La parte 1 describe el procedimiento de la prueba F para realizar una prueba de hipótesis, y la parte 2 proporciona una breve descripción de dos métodos alternativos para comparar la variación entre dos muestras.

PARTE 1 Prueba *F* con dos varianzas o desviaciones estándar

El siguiente recuadro de elementos clave incluye aspectos de una prueba de hipótesis sobre una afirmación acerca de dos varianzas o dos desviaciones estándar poblacionales. El procedimiento se basa en el uso de dos varianzas muestrales, pero el *mismo procedimiento* se usa para las hipótesis sobre dos desviaciones estándar poblacionales.

La prueba *F* real podría ser de dos colas, de cola izquierda o de cola derecha, pero podemos facilitar los cálculos al estipular que la mayor de las varianzas muestrales se expresa con s_1^2. Se deduce que la varianza muestral más pequeña se expresa como s_2^2. Este estipulación nos permite evitar el problema algo complicado de encontrar un valor crítico de *F* para la cola izquierda.

ELEMENTOS CLAVE

Prueba de hipótesis con dos varianzas o desviaciones estándar

Objetivo

Realizar una prueba de hipótesis de una afirmación sobre dos varianzas o dos desviaciones estándar poblacionales. (Cualquier afirmación hecha sobre dos desviaciones estándar poblacionales se puede expresar de nuevo con una afirmación equivalente acerca de dos varianzas poblacionales, por lo que el mismo procedimiento se usa para dos desviaciones estándar poblacionales o dos varianzas poblacionales).

Notación

s_1^2 = *mayor* de las dos varianzas muestrales
n_1 = tamaño de muestra con la varianza *mayor*
σ_1^2 = varianza poblacional de donde se obtuvo la muestra con la *mayor* varianza
Los símbolos s_2^2, n_2 y σ_2^2 se utilizan para la otra muestra y la otra población.

Requisitos

1. Las dos poblaciones son *independientes*.
2. Las dos muestras son muestras aleatorias simples.
3. Cada una de las dos poblaciones debe estar *distribuida normalmente*, independientemente de su tamaño de muestra. Esta prueba *F no es robusta* contra desviaciones de la normalidad. Por lo tanto, funciona mal si una o ambas poblaciones tienen una distribución que no es normal. El requisito de distribuciones normales es bastante estricto para esta prueba *F*.

Dato estadístico de prueba para pruebas de hipótesis con dos varianzas (con H_0: $\sigma_1^2 = \sigma_2^2$)

$$F = \frac{s_1^2}{s_2^2}$$ (donde s_1^2 es la *mayor* de las dos varianzas muestrales)

Valores *P*: Los valores *P* son proporcionados automáticamente por la tecnología. Si no dispone de una tecnología, use el valor calculado del dato estadístico de prueba *F* con la tabla A-5 a fin de encontrar un rango para el valor *P*.

Valores críticos: Use la tabla A-5 para encontrar los valores *F* críticos que se determinan de la siguiente manera:

1. El nivel de significancia α (la tabla A-5 incluye valores críticos para $\alpha = 0.025$ y $\alpha = 0.05$).
2. **Grados de libertad del numerador = $n_1 - 1$** (determina la *columna* de la tabla A-5)
3. **Grados de libertad del denominador = $n_2 - 1$** (determina la *fila* de la Tabla A-5) Para el nivel de significancia $\alpha = 0.05$, consulte la tabla A-5 y use el área de la cola derecha de 0.025 o 0.05, según el tipo de prueba. Como se muestra a continuación:

- *Prueba de dos colas:* Use la tabla A-5 con 0.025 en la cola derecha. (El nivel de significancia de 0.05 se divide entre las dos colas, por lo que el área en la cola derecha es 0.025).
- *Prueba de una cola:* Use la tabla A-5 con $\alpha = 0.05$ en la cola derecha.

Determinación del valor F crítico para la cola derecha: Dado que estamos estipulando que la varianza muestral más grande es $s^2{}_1$, todas las pruebas de una cola serán de cola derecha y todas las pruebas de dos colas requerirán que encontremos sólo el valor crítico ubicado en la derecha. (No hay necesidad de encontrar el valor crítico en la cola izquierda, lo cual no es muy difícil. Vea el ejercicio 19 “Determinación de valores críticos *F* inferiores”).

¡Explore los datos! Debido a que el requisito para la prueba F sobre las distribuciones normales es bastante estricto, asegúrese de examinar las distribuciones de las dos muestras usando histogramas y gráficas cuantilares normales, y confirme que no haya valores atípicos. (Vea "Evaluación de la normalidad" en la sección 6-5).

Distribución *F*

Para dos poblaciones normalmente distribuidas con varianzas iguales ($\sigma_1^2 = \sigma_2^2$), la distribución muestral del dato estadístico de prueba $F = s_1^2/s_2^2$ es la **distribución *F*** que se muestra en la figura 9-4 (dado que aún no imponemos la estipulación de que la varianza muestral más grande es s_1^2). Si repite el proceso de selección de muestras de dos poblaciones distribuidas normalmente con varianzas iguales, la distribución de la relación s_1^2/ s_2^2 es la distribución F.

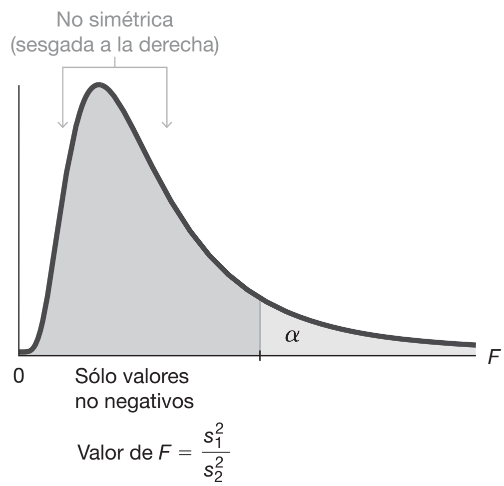


FIGURA 9-4 Distribución *F*

Existe una distribución F distinta para cada par diferente de grados de libertad en el numerador y el denominador. Vea la figura 9-4 y observe las siguientes propiedades de la distribución F:

- La distribución F no es simétrica.
- Los valores de la distribución F no pueden ser negativos.
- La forma exacta de la distribución F depende de los dos diferentes grados de libertad.

Interpretación del valor del dato estadístico de prueba *F*

Si las dos poblaciones tienen varianzas iguales, entonces la relación s_1^2/s_2^2 tenderá a estar cerca de 1. Debido a que estamos estipulando que s_1^2 es la varianza muestral más grande, la relación s_1^2/s_2^2 será un número *grande* cuando los valores de s_1^2 y s_2^2 sean muy diferentes. En consecuencia, un valor de F cercano a 1 será evidencia a favor de $\sigma_1^2 = \sigma_2^2$, pero un valor grande de F será evidencia contra $\sigma_1^2 = \sigma_2^2$.

Los valores *grandes* de *F* son evidencia en contra de $\sigma_1^2 = \sigma_2^2$.

EJEMPLO 1 Puntuaciones de evaluación de cursos

A continuación se listan las mismas puntuaciones de evaluación de cursos emitidas por estudiantes que se usaron en el ejemplo 1 de la sección 9-2, donde probamos la afirmación de que las dos muestras provienen de poblaciones con la misma *media*. Use los mismos datos con un nivel de significancia de 0.05 con el fin de probar la afirmación de que las puntuaciones de evaluación de cursos para las profesoras y los profesores tienen la misma *variación*.

Profesoras	4.3	4.3	4.4	4.0	3.4	4.7	2.9	4.0	4.3	3.4	3.4	3.3			
Profesores	4.5	3.7	4.2	3.9	3.1	4.0	3.8	3.4	4.5	3.8	4.3	4.4	4.1	4.2	4.0

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Las dos poblaciones son independientes entre sí. Las dos muestras no se corresponden de ninguna manera. (2) Dado el diseño para el estudio, suponemos que las dos muestras se pueden tratar como muestras aleatorias simples. (3) Una gráfica cuantilar normal de cada conjunto muestral de puntuaciones de evaluación de cursos indica que ambas muestras parecen ser de poblaciones con una distribución normal. Los requisitos se satisfacen.

Para las profesoras, obtenemos $s = 0.5630006$ y para los profesores $s = 0.3954503$. Podemos realizar la prueba utilizando varianzas o desviaciones estándar. Debido a que estipulamos en esta sección que la varianza más grande se expresa con s_1^2, sea $s_1^2 = 0.5630006^2$ y $s_2^2 = 0.3954503^2$.

Paso 1: La afirmación de que los profesores y las profesoras tienen la misma variación se puede expresar simbólicamente como $\sigma_1^2 = \sigma_2^2$ o como $\sigma_1 = \sigma_2$. Usaremos $\sigma_1 = \sigma_2$.

Paso 2: Si la afirmación original es falsa, entonces $\sigma_1 \neq \sigma_2$.

Paso 3: Dado que la hipótesis nula es la declaración de igualdad y la hipótesis alternativa no puede contener igualdad, tenemos

$$H_0: \sigma_1 = \sigma_2 \text{ (afirmación original)} \qquad H_1: \sigma_1 \neq \sigma_2$$

Paso 4: El nivel de significancia es $\alpha = 0.05$.

Paso 5: Como esta prueba involucra dos varianzas poblacionales, usamos la distribución F.

Paso 6: El dato estadístico de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{0.5630006^2}{0.3954503^2} = 2.0269$$

Método del valor *P* Debido al formato de la tabla A-5, el método del valor P es un poco complicado sin tecnología, pero aquí vamos. Para una prueba de dos colas con un nivel de significancia de 0.05, hay un área de 0.025 en la cola derecha, así que usamos las dos páginas de la distribución F (tabla A-5) con "0.025 en la cola derecha". Los grados de libertad en el numerador $= n_1 - 1 = 11$ y los grados de libertad en el denominador $= n_2 - 1 = 14$, por lo que el valor crítico de F está entre 3.1469 y 3.0502. El dato estadístico de prueba $F = 2.0269$ es menor que el valor crítico, por lo que sabemos que el área a la derecha del dato estadístico de prueba es mayor que 0.025, y se deduce que para esta prueba de dos colas, el valor $P > 0.05$.

Método del valor crítico Al igual que con el método del valor P, encontramos que el valor crítico está entre 3.1469 y 3.0502. El dato estadístico de prueba $F = 2.0269$ es menor que el valor crítico (que está entre 3.1469 y 3.0502), por lo que el dato estadístico de prueba no cae en la región crítica. Vea la figura 9-5.

continúa

En cifras

12,386,344: Número de diferentes jugadas posibles en cada juego de béisbol.
9: Número de diferentes formas posibles en las que un bateador puede llegar a primera base.
5 a 7: Número de lanzamientos que se usa una pelota de béisbol típica en un juego de Grandes Ligas.

Paso 7: La figura 9-5 muestra que el dato estadístico de prueba $F = 2.0269$ no cae dentro de la región crítica, por lo que no podemos rechazar la hipótesis nula de las varianzas iguales. No hay evidencia suficiente para respaldar el rechazo de la afirmación de que las desviaciones estándar son iguales.

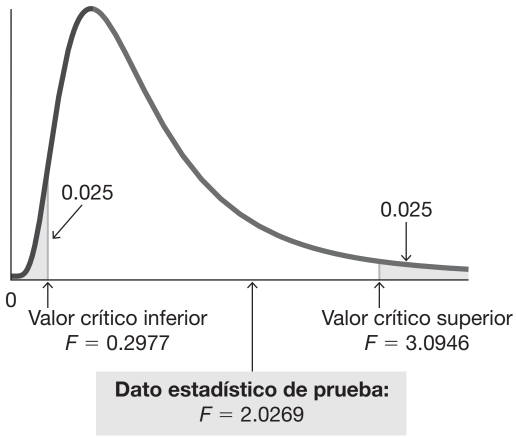


FIGURA 9-5 Prueba *F* de las desviaciones estándar iguales

INTERPRETACIÓN

Paso 8: No hay evidencia suficiente para respaldar el rechazo de la afirmación de que las dos desviaciones estándar son iguales. Las puntuaciones de evaluación de cursos para las profesoras y los profesores parecen tener la misma cantidad de variación.

SU TURNO Resuelva el ejercicio 5 "Color y creatividad".

XLSTAT

Ratio	2.0269
F (Observed value)	2.0269
F (Critical value)	3.0946
DF1	11
DF2	14
p-value (Two-tailed)	0.2139
alpha	0.05

Tecnología Por lo general, el software y las calculadoras proporcionan un valor *P*, por lo que aquí se usará el método del valor *P*. Vea los resultados adjuntos de XLSTAT, los cuales muestran el dato estadístico de prueba $F = 2.0269$ y el valor $P = 0.2139$. Debido a que el valor *P* de 0.2139 es mayor que el nivel de significancia de $\alpha = 0.05$, no podemos rechazar la hipótesis nula, y obtenemos las mismas conclusiones que en el ejemplo 1.

Precaución: La parte 2 de la sección 9-2 incluye métodos para probar hipótesis sobre dos medias poblacionales, y uno de esos métodos tiene un requisito de que $\sigma_1 = \sigma_2$. *No* se recomienda utilizar la prueba *F* como una forma de decidir si se cumple este requisito. Para la sección 9-2, el uso de la prueba *F* corre el riesgo de emplear diferencias que son demasiado pequeñas para tener un efecto en la prueba *t* de dos muestras independientes. Con frecuencia, ese método se describe como análogo a enviar a alguien al mar en un bote de remos (la prueba *F* preliminar) para determinar si el mar es seguro para un transatlántico (la prueba *t*).

PARTE 2 Métodos alternativos

La parte 1 de esta sección presentó la prueba *F* para probar hipótesis sobre desviaciones estándar (o varianzas) de dos poblaciones independientes. Debido a que la prueba *F* es muy sensible a desviaciones de la normalidad, ahora describiremos brevemente dos métodos alternativos que no son tan sensibles a tales desviaciones.

Conteo de cinco

El método del *conteo de cinco* es una alternativa relativamente simple a la prueba *F*, además no requiere poblaciones normalmente distribuidas. (Consulte "'A Quick, Compact, Two-Sample Dispersion Test: Count Five", de McGrath y Yeh, *American Statistician*, vol. 59, núm. 1). Si los dos tamaños de muestra son iguales, y si una muestra tiene al menos cinco de las mayores desviaciones absolutas medias (DAM), entonces concluimos que su población tiene una varianza mayor. Para ver el procedimiento específico, consulte el ejercicio 17 "Prueba del conteo de cinco".

Prueba Levene-Brown-Forsythe

La *prueba Levene-Brown-Forsythe* (o *prueba de Levene modificada*) es otra alternativa a la prueba F, y es mucho más robusta contra desviaciones de la normalidad. Esta prueba comienza con una transformación de cada conjunto de valores muestrales. Dentro de la primera muestra, reemplace cada valor x con $|x - \text{mediana}|$ y aplique la misma transformación a la segunda muestra. Con base en los valores transformados, realice una prueba t para la igualdad de medias con muestras independientes, como se describe en la parte 1 de la sección 9-2. Dado que los valores transformados son ahora desviaciones, la prueba t para la igualdad de medias es en realidad una prueba que compara la variación en las dos muestras. Vea el ejercicio 18 "Prueba Levene-Brown-Forsythe".

Existen otras alternativas a la prueba F, así como ajustes que mejoran el desempeño de dicha prueba. Consulte "Fixing the F Test for Equal Variances", de Shoemaker, *American Statistician*, vol. 57, núm. 2.

En cifras

Cada minuto hay más de 100,000 tweets, 170 millones de correos electrónicos y 12 millones de mensajes instantáneos.

CENTRO DE TECNOLOGÍA

Inferencias a partir de dos desviaciones estándar

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Hypothesis Testing** en el menú desplegable y **Standard Deviation Two Samples** en el submenú.
3. Seleccione el formato deseado para la *hipótesis alternativa* e ingrese el nivel de significancia.
4. *Si se usan datos estadísticos de resumen:* Seleccione la pestaña **Use Summary Statistics** e ingrese el tamaño de muestra y la desviación estándar muestral para cada muestra.

   Si se usan datos muestrales: Seleccione la pestaña **Use Data** y elija las columnas de datos deseadas.
5. Haga clic en **Evaluate**.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y seleccione **2 Variances** del submenú.
3. *Si se usan datos estadísticos de resumen:* Seleccione **Sample variances** o **Sample standard deviations** en el menú desplegable e ingrese los tamaños de muestra y las varianzas o desviaciones estándar muestrales.

   Si se usan datos muestrales: Seleccione **Each sample is in its own column** del menú desplegable y elija las columnas de datos deseadas.
4. Haga clic en el botón **Options** y seleccione la proporción deseada. Ingrese el nivel de confianza deseado e introduzca **1** para la *razón hipotética*. Seleccione el formato deseado para la *hipótesis alternativa*.
5. Marque la casilla etiquetada como **Use test and confidence intervals based on normal distribution**.
6. Haga clic en **OK** dos veces.

SUGERENCIA: Otro procedimiento es hacer clic en **Assistant** en el menú superior, luego seleccionar **Hypothesis Tests** y **2-Sample Standard Deviation**. Complete el cuadro de diálogo para obtener los resultados, incluyendo el valor *P* y otra información útil.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Variance Stats** en el menú desplegable, luego elija **Two Sample** en el submenú.
3. *Si se usan datos estadísticos de resumen:* Seleccione **With Summary** en el submenú e ingrese las varianzas muestrales y los tamaños de muestra.

   Si se usan datos muestrales: Seleccione **With Data** del submenú y elija la columna de datos deseada para cada muestra.
4. Seleccione **Hypothesis test for σ_1^2/σ_1^2**.Ingrese la razón hipotética (H_0) y seleccione el formato deseado para la hipótesis alternativa (H_A).
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **2-SampFTest** en el menú y presione ENTER.
3. Seleccione **Data** si tiene datos muestrales en listas o **Stats** si tiene estadísticos de resumen. Presione ENTER e ingrese los nombres de lista (establezca *Freq* = **1**) o los estadísticos de resumen.
4. Para σ_1, seleccione el formato deseado para la hipótesis alternativa.
5. Seleccione **Calculate** y presione ENTER.

continúa

CENTRO DE TECNOLOGÍA *continuación*

Inferencias a partir de dos desviaciones estándar

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Excel

Complemento XLSTAT

Requiere datos muestrales originales, no funciona con datos resumidos.

1. Haga clic en la pestaña **XLSTAT** en la cinta y luego haga clic en **Parametric Tests**.
2. Seleccione **Two-sample comparison of variances** en el menú desplegable.
3. Para las *muestras 1 y 2* ingrese el rango de celdas que contiene los datos muestrales. Para el *formato de datos*, seleccione **One column per sample**. Si la primera fila de datos contiene una etiqueta, también marque la casilla de **Column labels**.
4. Seleccione **Fisher's F test**.
5. Haga clic en la pestaña **Options**.
6. En *hipótesis alternativa*, seleccione el formato deseado ($\neq$ para pruebas de dos colas, $<$ para pruebas de cola izquierda, $>$ para pruebas de cola derecha). Ingrese la *razón hipotética* (generalmente **1**) e introduzca el nivel de significancia deseado (**5** para un nivel de significancia de 0.05).
7. Haga clic en **OK** para visualizar el dato estadístico de prueba (etiquetado como *F Observed value*) y el valor *P*.

Excel (complemento del análisis de datos)

1. Haga clic en **Data** en la cinta, luego haga clic en la pestaña **Data Analysis**.
2. Seleccione **F-Test Two-Sample for Variances** y haga clic en **OK**.
3. Ingrese el rango de datos para cada variable en los cuadros de *rango de variable*. Si la primera fila contiene una etiqueta, marque la casilla **Labels**.
4. Ingrese el nivel de significancia deseado en la casilla de *alfa* y haga clic en **OK**. Los resultados incluyen el dato estadístico de prueba *F*, el valor *P* para una prueba de una cola y el valor crítico para una prueba de una cola.

SUGERENCIA: Para efectuar una prueba de dos colas, haga dos ajustes: (1) Ingrese un valor que sea la mitad del nivel de significancia (Paso 4) y (2) duplique el valor *P* dado en los resultados.

9-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Dato estadístico de prueba *F*

a. Si s_1^2 representa la mayor de dos varianzas muestrales, ¿el dato estadístico de prueba F puede ser menor que 1?

b. ¿El dato estadístico de prueba F puede ser un número negativo?

c. Si se prueba la afirmación de que $\sigma_1^2 \neq \sigma_2^2$, ¿qué sabemos sobre las dos muestras si el dato estadístico de prueba F está muy cerca de 1?

d. ¿La distribución F es simétrica, sesgada hacia la izquierda o sesgada hacia la derecha?

2. Prueba *F* Si se utilizan los datos muestrales del conjunto de datos 1 “Datos corporales” en el apéndice B para una prueba sobre la afirmación de que las alturas de los hombres y las alturas de las mujeres tienen varianzas diferentes, encontramos que $s = 7.48296$ cm para las mujeres y $s = 7.10098$ cm para los hombres.

a. Encuentre los valores de s_1^2 y s_2^2 y exprésalos con las unidades de medición adecuadas.

b. Identifique las hipótesis nula y alternativa.

c. Encuentre el valor del dato estadístico de prueba F y redondéelo a cuatro decimales.

d. El valor P para esta prueba es 0.5225. ¿Qué se puede concluir sobre la afirmación declarada?

3. Prueba de normalidad Para la prueba de hipótesis descrita en el ejercicio 2, los tamaños de muestra son $n_1 = 147$ y $n_2 = 153$. Cuando se utiliza la prueba F con estos datos. ¿Es correcto razonar que no hay necesidad de verificar la normalidad porque $n_1 > 30$ y $n_2 > 30$?

4. Robustez ¿Qué significa cuando decimos que la prueba F descrita en esta sección no es *robusta* contra desviaciones de la normalidad?

En los ejercicios 5 a 16, pruebe la afirmación dada.

5. Color y creatividad Investigadores de la University of British Columbia realizaron pruebas para investigar los efectos del color en la creatividad. Se pidió a los sujetos, ante un fondo rojo, que pensaran en usos creativos para un ladrillo; otros sujetos, ante un fondo azul, recibieron la misma tarea. Las respuestas fueron calificadas por un panel de jueces y los resultados de las puntuaciones de creatividad se dan a continuación. Use un nivel de significancia de 0.05 para probar la afirmación de que las puntuaciones de las tareas creativas tienen la misma variación con un fondo rojo y un fondo azul.

Fondo rojo:	$n = 35, \bar{x} = 3.39, s = 0.97$
Fondo azul:	$n = 36, \bar{x} = 3.97, s = 0.63$

6. Color y memoria Investigadores de University of British Columbia realizaron ensayos para investigar los efectos del color en la precisión de la memoria. Los sujetos recibieron tareas que involucraban palabras mostradas en una pantalla de computadora con colores de fondo en rojo y en azul. Los sujetos estudiaron 36 palabras durante 2 minutos, y luego se les pidió que recordaran tantas palabras como pudieran después de esperar 20 minutos. Los resultados de las puntuaciones en la prueba de memorización de palabras se dan a continuación. Use un nivel de significancia de 0.05 para probar la afirmación de que la variación de las puntuaciones es la misma con el fondo rojo y con el fondo azul.

Puntuaciones de precisión

Fondo rojo:	$n = 35, \bar{x} = 15.89, s = 5.90$
Fondo azul:	$n = 36, \bar{x} = 12.31, s = 5.48$

7. Prueba de los efectos del alcohol Algunos investigadores realizaron un experimento para evaluar los efectos de a1coho1. Los errores se registraron en una prueba de habilidades visuales y motoras para un grupo de tratamiento de 22 personas que bebieron etanol y otro grupo de 22 personas que recibieron un placebo. Los errores para el grupo de tratamiento tienen una desviación estándar de 2.20 y los errores para el grupo placebo tienen una desviación estándar de 0.72 (de acuerdo con los 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). Use un nivel de significancia de 0.05 para probar la afirmación de que el grupo con tratamiento tiene errores que varían significativamente más que los errores del grupo placebo.

8. Fumadores pasivos El conjunto de datos 12 "Fumadores pasivos y activos" incluye los niveles de cotinina medidos en un grupo de fumadores ($n = 40$, $\bar{x} = 172.48$ ng/mL, $s = 119.50$ ng/mL) y un grupo de no fumadores que no están expuestos al humo del tabaco ($n = 40$, $\bar{x} = 16.35$ ng/mL, $s = 62.53$ ng/mL). La cotinina es un metabolito de la nicotina, lo que significa que cuando la nicotina es absorbida por el cuerpo, se produce la cotinina.

a. Use un nivel de significancia de 0.05 para evaluar la afirmación de que la variación de cotinina en los fumadores es mayor que la variación de cotinina en los no fumadores que no están expuestos al humo de tabaco.

b. Las 40 mediciones de cotinina del grupo no fumador consisten en los siguientes valores (todos en ng/mL): 1, 1, 90, 244, 309, y otros 35 valores que son todos 0. ¿Esta muestra parece ser de una población distribuida normalmente? Si no es así, ¿cómo se ven afectados los resultados del inciso (a)?

9. Coca-Cola y Diet Coke El conjunto de datos 26 "Pesos y volúmenes de bebidas de cola" en el apéndice B incluyen los pesos (en libras) para el contenido de una muestra de latas de Coca-Cola regular ($n = 36$, $\bar{x} = 0.81682$ lb, $s = 0.00751$ lb) y los pesos del contenido para una muestra de latas de Diet Coke ($n = 36$, $\bar{x} = 0.78479$ lb, $s = 0.00439$ lb). Use un nivel de significancia de 0.05 para probar la afirmación de que la variación es la misma para ambos tipos de bebida.

10. IQ y exposición al plomo El conjunto de datos 7 "IQ y plomo" en el apéndice B lista las puntuaciones de IQ completas para una muestra aleatoria de sujetos con bajos niveles de plomo en la sangre y otra muestra aleatoria de sujetos con niveles altos de plomo en la sangre. Los estadísticos se resumen al inicio de la página siguiente. Use un nivel de significancia de 0.05 para probar la afirmación de que las puntuaciones de IQ de las personas con niveles bajos de plomo varían más que las puntuaciones de IQ de las personas con niveles altos de plomo.

Bajo nivel de plomo: $n = 78, \bar{x} = 92.88462, s = 15.34451$

Alto nivel de plomo: $n = 21, \bar{x} = 86.90476, s = 8.988352$

11. Tratamiento del dolor con imanes Algunos investigadores realizaron un estudio para determinar si los imanes son efectivos en el tratamiento del dolor de espalda, con los resultados que se presentan a continuación (de acuerdo con datos de "Bipolar Permanent Magnets for the Treatment of Chronic Lower Back Pain: A Pilot Study", de Collacott, Zimmerman, White y Rindone, *Journal of the American Medical Association*, vol. 283, núm. 10). Los valores representan mediciones de dolor usando una escala analógica visual. Use un nivel de significancia de 0.05 para probar la afirmación de que aquellas personas que reciben un tratamiento simulado (similar a un placebo) tienen reducciones de dolor que varían más que las reducciones de dolor para las personas tratadas con imanes.

Reducción del nivel de dolor después del tratamiento simulado: $n = 20, \bar{x} = 0.44, s = 1.4.$

Reducción del nivel de dolor después del tratamiento con imanes: $n = 20, \bar{x} = 0.49, s = 0.96$

12. Edades de automóviles y taxis Cuando el autor visitó Dublín, Irlanda (hogar del empleado William Gosset de Guinness Brewery, quien desarrolló por primera vez la distribución *t*), registró edades de automóviles de pasajeros seleccionados al azar y de taxis seleccionados aleatoriamente. Las edades (en años) se listan a continuación. Use un nivel de significancia de 0.05 para probar la afirmación de que en Dublín, las edades de los automóviles y los taxis tienen la misma variación.

Edades de automóviles	4	0	8	11	14	3	4	4	3	5	8	3	3	7	4	6	6	1	8	2	15	11	4	1	6	1	8
Edades de taxis	8	8	0	3	8	4	3	3	6	11	7	7	6	9	5	10	8	4	3	4							

13. Puntuaciones para evaluar a profesores A continuación se listan las puntuaciones de evaluación otorgadas por estudiantes a sus profesoras y profesores, provenientes del conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B. Use un nivel de significancia de 0.05 para probar la afirmación de que las profesoras y los profesores tienen una puntuación de evaluación con la misma variación.

Profesoras	4.4	3.4	4.8	2.9	4.4	4.9	3.5	3.7	3.4	4.8
Profesores	4.0	3.6	4.1	4.1	3.5	4.6	4.0	4.3	4.5	4.3

Hombres	Mujeres
$n_1 = 11$	$n_2 = 59$
$\bar{x}_1 = 97.69°F$	$\bar{x}_2 = 97.45°F$
$s_1 = 0.89°F$	$s_2 = 0.66°F$

14. Temperaturas corporales de hombres y mujeres Si utilizamos las temperaturas corporales de las 8 AM del día 2, como se presentan en el conjunto de datos 3 "Temperaturas corporales" del apéndice B, obtenemos los estadísticos mostrados en la tabla adjunta. Utilice estos datos con un nivel de significancia de 0.05 para probar la afirmación de que los hombres tienen temperaturas corporales que varían más que las temperaturas corporales de las mujeres.

15. Old Faithful A continuación se listan los intervalos de tiempo (en minutos) entre las erupciones del géiser Old Faithful. Los tiempos "recientes" ocurrieron en los últimos años, y los tiempos "pasados" son de 1995. ¿Parece que la variación de los tiempos entre las erupciones ha cambiado?

Recientes	78	91	89	79	57	100	62	87	70	88	82	83	56	81	74	102	61
Pasados	89	88	97	98	64	85	85	96	87	95	90	95					

16. Bloqueo en exámenes Muchos estudiantes han tenido la desagradable experiencia de entrar en pánico en un examen porque la primera pregunta era excepcionalmente difícil. La disposición de las preguntas de examen fue estudiada por su efecto sobre la ansiedad. Las siguientes puntuaciones son medidas de "ansiedad debilitante en exámenes", que la mayoría de nosotros llamamos *pánico* o *bloqueo* (según datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4). Utilice un nivel de significancia de 0.05 para probar la afirmación de que las dos poblaciones de puntuaciones tienen diferentes cantidades de variación.

Preguntas dispuestas de 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

Preguntas dispuestas de 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

9-4 Más allá de lo básico

17. Prueba del conteo de cinco para comparar la variación en dos poblaciones Repita el ejercicio 16 "Bloqueo en exámenes", pero en lugar de usar la prueba F, use el siguiente procedimiento a fin de realizar la prueba del "conteo de cinco" para variaciones iguales (que no es tan complicado como podría parecer).

a. Para cada valor x en la primera muestra, encuentre la desviación absoluta $|x - \bar{x}|$, luego ordene los valores de desviación absoluta. Haga lo mismo para la segunda muestra.

b. Sea c_1 el conteo del número de valores de desviación absoluta en la primera muestra que son mayores que el valor de la desviación absoluta más grande en la segunda muestra. Además, sea c_2 es el conteo del número de valores de desviación absoluta en la segunda muestra que son mayores que el valor de la desviación absoluta más grande en la primera muestra. (Uno de estos conteos siempre será cero).

c. Si los tamaños de muestra son iguales ($n_1 = n_2$), use un valor crítico de 5. Si $n_1 \neq n_2$, calcule el valor crítico que se muestra a continuación.

$$\frac{\log(\alpha/2)}{\log\left(\dfrac{n_1}{n_1 + n_2}\right)}$$

d. Si $c_1 \geq$ valor crítico, entonces concluya que $\sigma_1^2 > \sigma_2^2$. Si $c_2 \geq$ valor crítico, entonces concluya que $\sigma_2^2 > \sigma_1^2$. En otro caso, no se puede rechazar la hipótesis nula de $\sigma_1^2 = \sigma_1^2$.

18. Prueba Levene-Brown-Forsythe Repita el ejercicio 16 "Bloqueo en exámenes" utilizando la prueba Levene-Brown-Forsythe.

19. Determinación de valores F inferiores Para las pruebas de hipótesis que son de dos colas, los métodos de la parte 1 requieren que encontremos sólo el valor crítico superior. Denotemos el valor crítico superior por F_R, donde el subíndice indica el valor crítico para la cola derecha. El valor crítico inferior F_L (para la cola izquierda) se puede encontrar de la siguiente manera: (1) Intercambie los grados de libertad utilizados para encontrar F_R, luego (2) mediante el uso de los grados de libertad encontrados en el paso 1, determine el valor F a partir de la tabla A-5; (3) obtenga el recíproco del valor F encontrado en el paso 2, y el resultado es F_L. Encuentre los valores críticos F_L y F_R para el ejercicio 16 "Bloqueo en exámenes".

Examen rápido del capítulo

En los ejercicios 1 a 5, use los siguientes resultados de encuesta: Se preguntó a los sujetos seleccionados al azar si sabían que la Tierra ha perdido la mitad de la población de su vida silvestre durante los últimos 50 años. Entre 1121 mujeres, 23% dijo que sí lo sabía. Entre 1084 hombres, 26% dijo que sí lo sabía (según datos de una encuesta de Harris).

1. Biodiversidad Identifique las hipótesis nula y alternativa que resultan de la afirmación de que, para las personas que conocían la declaración, la proporción de mujeres es igual a la proporción de hombres.

2. Biodiversidad Encuentre los valores de x_1 (el número de mujeres que conocían la declaración), x_2 (el número de hombres que conocían la declaración), $\hat{p}_1$, $\hat{p}_2$, y la proporción combinada $\bar{p}$ obtenidos al probar la afirmación dada en el ejercicio 1.

3. Biodiversidad Al probar la afirmación de que $p_1 = p_2$, se obtiene un dato estadístico de prueba $z = -1.64$. Encuentre el valor P para la prueba de hipótesis.

4. Biodiversidad Cuando se utilizan los datos muestrales dados a fin de elaborar una estimación del intervalo de confianza del 95% para la diferencia entre las dos proporciones poblaciones, se obtiene el resultado de $(-0.0659, 0.00591)$ mediante el uso de la tecnología.

a. Exprese ese intervalo de confianza en un formato que use el símbolo $<$.

b. ¿Qué característica del intervalo de confianza es una base para decidir si existe una diferencia significativa entre la proporción de mujeres que conocen la declaración y la proporción de hombres que la conocen?

5. Biodiversidad Suponga que se obtiene un valor P de 0.1 al probar la afirmación dada en el ejercicio 1 "Biodiversidad". ¿Qué se debe concluir sobre la hipótesis nula? ¿Cuál debería ser la conclusión final?

6. ¿Verdadera? Determine si la siguiente afirmación es verdadera: Cuando se obtienen muestras aleatorias de 50 hombres y 50 mujeres, y queremos probar la afirmación de que los hombres y las mujeres tienen diferentes ingresos anuales medios, no hay necesidad de confirmar que las muestras provienen de poblaciones con distribuciones normales.

7. ¿Verdadera? Al recolectar muestras aleatorias para evaluar la afirmación de que la proporción de mujeres agentes de la CIA ubicadas en Estados Unidos es igual a la proporción de mujeres agentes de la CIA ubicadas fuera de Estados Unidos, hay un requisito de que $np \geq 30$ y $nq \geq 30$.

8. ¿Dependiente o independiente? A continuación se listan las mediciones de presión arterial sistólica (mm Hg) tomadas de los brazos derecho e izquierdo de una mujer sana en diferentes momentos (según datos de "Consistency of Blood Pressure Differences Between the Left and Right Arms", de Eguchi *et al*., *Archives of Internal Medicine*, vol. 167). ¿Los datos son dependientes o independientes?

Brazo derecho	102	101	94	79	79
Brazo izquierdo	175	169	182	146	144

9. Hipótesis Identifique las hipótesis nula y alternativa con el fin de usar los datos muestrales del ejercicio 8, en una prueba de la afirmación de que existen diferencias entre las presiones arteriales sistólicas en el brazo derecho y en el brazo izquierdo; tales diferencias provienen de una población con una media igual a 0.

10. Estadísticos de prueba Identifique el dato estadístico de prueba que debe usarse para probar las siguientes hipótesis dadas.

a. La media de las diferencias entre las puntuaciones de IQ de los esposos y las puntuaciones de IQ de sus esposas es igual a 0.

b. La edad media de las agentes femeninas de la CIA es igual a la edad media de los agentes masculinos de la CIA.

c. La proporción de hombres zurdos es igual a la proporción de mujeres zurdas.

d. La variación entre los pulsos de las mujeres es igual a la variación entre los pulsos de los hombres.

Ejercicios de repaso

1. Efecto de la denominación En el artículo "The Denomination Effect" de Priya Raghubir y Joydeep Srivastava, *Journal of Consumer Research*, vol. 36, los investigadores reportaron los resultados obtenidos en sus estudios para determinar si las personas tienen diferentes características de gasto cuando poseen billetes más grandes, por ejemplo billetes de \$20, en vez de billetes más pequeños como veinte billetes de \$1. En un ensayo, 89 estudiantes de negocios de pregrado de dos universidades diferentes fueron asignados aleatoriamente a dos grupos diferentes. En el grupo "billete de un dólar", a 46 sujetos se les dieron billetes de un dólar; el grupo "un cuarto" constó de 43 sujetos que recibieron cuartos de dólar. A todos los sujetos de ambos grupos se les dio la opción de quedarse con el dinero o comprar goma de mascar o pastillas de menta. El artículo incluye la afirmación de que "es menos probable gastar el dinero en una denominación grande, en comparación con una cantidad equivalente en denominaciones más pequeñas". Pruebe esa afirmación usando un nivel de significancia de 0.05 con los siguientes datos muestrales del estudio.

	Grupo 1	Grupo 2
	Sujetos que recibieron un billete de \$1	Sujetos que recibieron 4 cuartos de dólar
Gastaron el dinero	$x_1 = 12$	$x_2 = 27$
Sujetos en el grupo	$n_1 = 46$	$n_2 = 43$

2. Efecto de la denominación Elabore el intervalo de confianza que podría usarse para probar la afirmación del ejercicio 1. ¿Qué característica del intervalo de confianza conduce a la misma conclusión que en el ejercicio 1?

3. Estaturas A continuación se listan las estaturas (en cm) seleccionadas al azar de una muestra de mujeres y las estaturas (en cm) seleccionadas al azar de una muestra de hombres (del conjunto de datos 1 "Datos corporales" en el apéndice B). Use un nivel de confianza del 95% para estimar la magnitud de la diferencia entre la estatura media de las mujeres y la estatura media de los hombres.

Mujeres:	160.3	167.7	166.9	153.3	160.0	177.3	169.1	134.5	163.3	171.1
Hombres:	190.3	169.8	179.8	179.8	177.0	178.5	173.5	178.7	179.0	181.3

4. Estaturas Use un nivel de significancia de 0.01 con los datos muestrales del ejercicio 3 para probar la afirmación de que las mujeres tienen estaturas con una media menor que la estatura media de los hombres.

5. Resultados antes/después del tratamiento Captopril es un medicamento diseñado para reducir la presión arterial sistólica. Cuando los sujetos fueron tratados con este medicamento, sus lecturas de presión arterial sistólica (en mm Hg) se midieron antes y después de tomar el medicamento. Los resultados se dan en la tabla adjunta (con base en datos de "Essential Hypertension: Effect of an Oral Inhibitor of Angiotensin-Converting Enzyme", de MacGregor *et al*., *British Medical Journal*, vol. 2). Si se usa un nivel de significancia de 0.01, ¿hay suficiente evidencia para respaldar la afirmación de que el Captopril es efectivo para disminuir la presión arterial 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

6. Precisión del testigo presencial de la policía ¿Afecta el estrés de los testigos presenciales de la policía su capacidad de recordar? Este tema fue estudiado en un experimento que probó la memoria de testigos oculares, una semana después de un interrogatorio no estresante de un sospechoso cooperativo y un interrogatorio estresante de un sospechoso no cooperativo y beligerante. A continuación se proporciona el número de detalles recordados una semana después de que se registró el incidente, así como los estadísticos de resumen (según datos de "Eyewitness Memory of Police Trainees for Realistic Role Plays", de Yuille *et al.*, *Journal of Applied Psichology*, vol. 79, núm. 6). Use un nivel de significancia de 0.01 para probar la afirmación del artículo de que "el estrés disminuye la cantidad de detalles recordados".

$$\text{Sin estrés: } n = 40, \overline{x} = 53.3, s = 11.6$$

$$\text{Con estrés: } n = 40, \overline{x} = 45.3, s = 13.2$$

7. ¿Los vuelos son más baratos cuando se programan con anticipación? A continuación se listan los costos (en dólares) de vuelos desde Nueva York (JFK) a Los Ángeles (LAX). Use un nivel de significancia de 0.01 para evaluar la afirmación de que los vuelos programados con un día de antelación cuestan más que los vuelos programados con 30 días de anticipación. ¿Qué estrategia parece ser efectiva para ahorrar dinero al volar?

	Delta	Jet Blue	American	Virgin	Alaska	United
1 día de anticipación	501	634	633	646	633	642
30 días de anticipación	148	149	156	156	252	313

8. Variación de estaturas Use los datos muestrales dados en el ejercicio 3 "Estaturas" y pruebe la afirmación de que las mujeres y los hombres tienen estaturas con la misma variación. Use un nivel de significancia de 0.05.

Ejercicios de repaso acumulado

Estaturas familiares. ***En los ejercicios 1 a 5, use las siguientes estaturas (en pulgadas). Los datos se relacionan de modo que cada columna contenga estaturas de la misma familia.***

Padre	68.0	68.0	65.5	66.0	67.5	70.0	68.0	71.0
Madre	64.0	60.0	63.0	59.0	62.0	69.0	65.5	66.0
Hijo	71.0	64.0	71.0	68.0	70.0	71.0	71.7	71.0

1. a. ¿Las tres muestras son independientes o dependientes? ¿Por qué?

b. Encuentre la media, la mediana, el rango, la desviación estándar y la varianza de las estaturas de los hijos.

c. ¿Cuál es el nivel de medición de los datos muestrales (nominal, ordinal, de intervalo, de razón)?

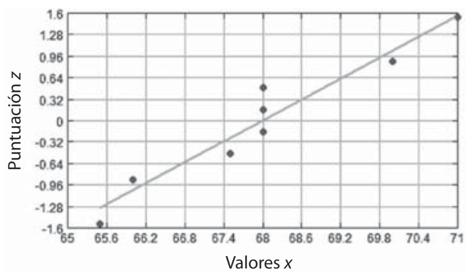
d. ¿Las estaturas originales sin redondear son datos discretos o datos continuos?

2. Diagrama de dispersión Elabore un diagrama de dispersión de las estaturas padre/hijo, luego interprételo.

3. Intervalo de confianza Elabore una estimación del intervalo de confianza del 95% para la estatura media de los hijos. Escriba un breve enunciado que interprete el intervalo de confianza.

4. Prueba de hipótesis Use un nivel de significancia de 0.05 para probar la afirmación de que las diferencias entre las estaturas de los padres y sus hijos tienen una media de 0 pulgadas.

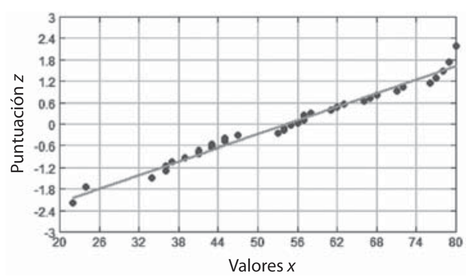
5. Evaluación de la normalidad Interprete la gráfica cuantilar normal de las estaturas de los padres.



6. Tiempos de reacción al frenar: Histograma A continuación se clasifican los tiempos de reacción al frenar (en 1/10,000 seg) para sujetos masculinos y femeninos (según datos del Comprobador de tiempo de reacción de frenado RT-2S). Elabore un histograma para los tiempos de reacción de los hombres. Use un ancho de clase de 8 y utilice 28 como el límite inferior de la primera clase. Para el eje horizontal, use los puntos medios de clase. ¿Parece que los datos provienen de una población con una distribución normal?

Hombres	28	30	31	34	34	36	36	36	36	38	39	40	40	40	40	41	41	41
	42	42	44	46	47	48	48	49	51	53	54	54	56	57	60	61	61	63
Mujeres	22	24	34	36	36	37	39	41	41	43	43	45	45	47	53	54	54	55
	56	57	57	57	58	61	62	63	66	67	68	71	72	76	77	78	79	80

7. Tiempos de reacción al frenar: ¿Normal? La gráfica cuantilar normal adjunta se obtiene utilizando los tiempos de reacción de frenado de las mujeres listados en el ejercicio 6. Interprete esta gráfica.



8. Tiempos de reacción al frenar: Diagramas de caja y bigotes Utilice los mismos datos del ejercicio 6 y use la misma escala para elaborar un diagrama de caja y bigotes de los tiempos de reacción al frenar de los hombres y otro diagrama de caja y bigotes para los tiempos de reacción al frenar de las mujeres. ¿Qué sugieren estos diagramas?

9. Tiempos de reacción al frenar: Prueba de hipótesis Utilice los datos muestrales del ejercicio 6 con un nivel de significancia de 0.01 para evaluar la afirmación de que los hombres y las mujeres tienen la misma media.

10. Tiempos de reacción al frenar: Intervalos de confianza

a. Elabore una estimación del intervalo de confianza del 99% para el tiempo medio de reacción al frenar de los hombres, elabore una estimación del intervalo de confianza del 99% para el tiempo medio de reacción al frenar de las mujeres, luego compare los resultados.

b. Elabore una estimación del intervalo de confianza del 99% para la diferencia entre el tiempo medio de reacción al frenar de los hombres y el tiempo medio de reacción al frenar de las mujeres.

c. ¿Con cuáles resultados se comparan de mejor manera los tiempos de reacción medios de hombres y mujeres, con los resultados del inciso (a) o con los del inciso (b)?

Proyecto de tecnología

Statdisk, Minitab, Excel, StatCrunch, la calculadora TI-83/84 Plus y muchas otras tecnologías son capaces de generar datos distribuidos normalmente, extraídos de una población con una media y una desviación estándar específicas. Las puntuaciones de las pruebas de densidad ósea se miden como puntuaciones z que tienen una distribución normal con media 0 y desviación estándar 1. Genere dos conjuntos de datos muestrales que representen puntuaciones de densidad ósea simulados, como se muestra a continuación.

- **Grupo de tratamiento:** Genere 10 valores muestrales de una población de puntuaciones de densidad ósea distribuida normalmente, con media 0 y desviación estándar 1.
- **Grupo placebo:** Genere 15 valores muestrales de una población de puntuaciones de densidad ósea distribuida normalmente con media 0 y desviación estándar 1.

Statdisk:	Seleccione **Data**, luego **Normal Generator**.
Minitab:	Seleccione **Calc**, **Random Data**, **Normal**
Excel:	Seleccione **Data Analysis**, **Randon Number Generation**.
TI-83/84 Plus:	Presione **MATH**, seleccione **PROB**, luego use la función **randNorm** con el formato de ($\overline{x}$, s, n).
StatCrunch:	Haga clic en **Data**, seleccione **Simulate**, seleccione **Normal**.

Debido a que cada una de las dos muestras consiste en selecciones aleatorias de una población normalmente distribuida con una media de 0 y una desviación estándar de 1, los datos se generan de manera que ambos conjuntos de datos realmente provengan de la misma población, entonces no debería haber diferencia entre las dos medias muestrales.

a. Después de generar los dos conjuntos de datos, use un nivel de significancia de 0.10 para probar la afirmación 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 conducirán a la conclusión de que las dos medias poblacionales son diferentes? ¿Cómo se relaciona esto con un error tipo I?

c. Si los datos generados lo llevan a la conclusión de que las dos medias poblacionales son diferentes. ¿Esta conclusión sería correcta o incorrecta en la realidad? ¿Cómo lo sabe?

d. Si el inciso (a) se repite 20 veces. ¿Cuál es la probabilidad de que ninguna de las pruebas de hipótesis conduzca al rechazo de la hipótesis nula?

e. Repita el inciso (a) 20 veces. ¿Con qué frecuencia se rechazó la hipótesis nula de la igualdad de medias? ¿Es este el resultado que usted esperaba?

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿El cambio en la regla de la NFL tuvo el efecto deseado?

Entre los 460 juegos con tiempos extra de la Liga Nacional de Fútbol Americano (NFL) entre 1974 y 2011, 252 de los equipos que ganaron el lanzamiento de moneda previo al tiempo extra ganaron el juego. Durante esos años, un equipo podía ganar el lanzamiento de la moneda y avanzar por el terreno de juego para ganar el juego mediante un gol de campo, y el otro equipo nunca tendría la posesión del balón. Eso simplemente no parecía justo. A partir de 2012, se cambiaron las reglas del tiempo extra. En los primeros tres años con las nuevas reglas de tiempo extra, 47 juegos se decidieron de esa manera, y el equipo que ganó el lanzamiento de la moneda resultó vencedor en 24 de esos juegos.

Análisis de los resultados

1. Primero *explore* las dos proporciones de victorias en tiempo extra. ¿Parece haber una diferencia? Si es así, ¿cuál es?

2. Cree una afirmación que deberá ser probada, luego pruébela. Use una prueba de hipótesis y un intervalo de confianza.

3. ¿Qué concluye usted sobre la eficacia del cambio en la regla del tiempo extra?

Actividades en equipo

1. Actividad fuera de clase El problema del capítulo se basa en observaciones de automóviles sólo con placas traseras en estados con leyes que requieren placas delanteras y traseras. Trabajen en grupos de tres o cuatro y reúnan datos en su estado. Usen una prueba de hipótesis para probar la afirmación de que, en su estado, la proporción de automóviles sólo con placas traseras es igual que la proporción de 239/2049 de Connecticut. (Los estudiantes de Connecticut pueden comparar la proporción que obtengan con la proporción de 239/2049 obtenida por el autor).

2. Actividad fuera de clase Realice una encuesta a parejas y registre el número de tarjetas de crédito que tiene cada persona. Analice los datos pareados para determinar si los hombres en relaciones de pareja tienen más tarjetas de crédito que las mujeres. Intente identificar las razones de cualquier discrepancia.

3. Actividad fuera de clase Mida y registre la estatura del hombre y la estatura de la mujer en cada una de varias parejas diferentes. Estime la media de las diferencias. Compare el resultado con la diferencia entre la estatura media de los hombres y la estatura media de las mujeres incluidas en el conjunto de datos 1 “Datos corporales” en el apéndice B. ¿Los resultados sugieren que la estatura es un factor cuando las personas seleccionan parejas?

4. Actividad fuera de clase ¿Las estimaciones están influidas por números anclados? Consulte la actividad en equipo relacionada del capítulo 3 en la página 129. En el capítulo 3 observamos que, según el autor John Rubin, cuando las personas deben estimar un valor, su estimación suele estar “anclada” (o influenciada por) un número anterior. En esa actividad del capítulo 3, se solicitó 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 se pidió a otros que estimaran rápidamente el valor de $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$. En el capítulo 3, podríamos comparar los dos conjuntos de resultados mediante el uso de estadísticos (como la media) y gráficas (como los diagramas de caja y bigotes). Los métodos de este capítulo ahora nos permiten comparar los resultados con una prueba de hipótesis formal. Específicamente, recopile sus propios datos muestrales y pruebe la afirmación de que cuando comenzamos con números más grandes (como en $8 \times 7 \times 6$), nuestras estimaciones tienden a ser mayores.

5. Actividad en clase Divídanse en grupos de acuerdo con su género, aproximadamente con 10 o 12 estudiantes en cada grupo. Cada miembro del grupo debe registrar su pulso contando el número de latidos en 1 minuto, luego se deben calcular los estadísticos grupales (n, $\bar{x}$, s). Los grupos deben probar la hipótesis nula de que no hay diferencia entre su pulso medio y la media de los pulsos de la población de donde se seleccionaron los sujetos del mismo sexo para el conjunto de datos 1 “Datos corporales” en el apéndice B.

6. Actividad fuera de clase Seleccione al azar una muestra de estudiantes varones y una muestra de estudiantes mujeres y haga una pregunta con respuesta “sí o no” a cada persona seleccionada; por ejemplo, si respaldan la pena de muerte para las personas condenadas por asesinato, o si creen que el gobierno federal debería financiar la investigación con células madre. Registre la respuesta, el sexo del encuestado, y el sexo de la persona que hace la pregunta. Use una prueba de hipótesis formal para determinar si existe una diferencia entre las proporciones de respuestas afirmativas de hombres y mujeres. Además, determine si las respuestas parecen estar influenciadas por el sexo del entrevistador.

7. Actividad fuera de clase Elabore una breve encuesta con sólo unas pocas preguntas, incluida una que le pida al sujeto reportar su estatura. Después de que el sujeto haya completado la encuesta, mida la estatura del sujeto (sin zapatos) usando un sistema de medición preciso. Registre el sexo, la estatura reportada, y la estatura medida de cada sujeto. ¿Los sujetos masculinos parecen exagerar sus alturas? ¿Las mujeres parecen exagerar sus alturas? ¿Los errores para los hombres parecen tener la misma media que los errores para las mujeres?

8. Actividad en clase Sin usar ningún dispositivo de medición. Pida a cada alumno que trace una línea de 3 pulgadas de largo y otra línea con 3 cm de longitud. Luego use reglas para medir y registrar las longitudes de las líneas dibujadas. Registre los errores junto con los géneros de los estudiantes que hicieron las estimaciones. Pruebe la afirmación de que al estimar la longitud de una línea de 3 pulgadas, el error medio de los hombres es igual al error medio de las mujeres. Además, ¿los resultados muestran que tenemos una mejor comprensión del sistema británico de medidas (pulgadas) que del sistema internacional (centímetros)?

9. Actividad fuera de clase Obtenga muestras aleatorias simples de automóviles en los estacionamientos de estudiantes y profesores, y pruebe la afirmación de que los estudiantes y el personal docente tienen las mismas proporciones de carros extranjeros.

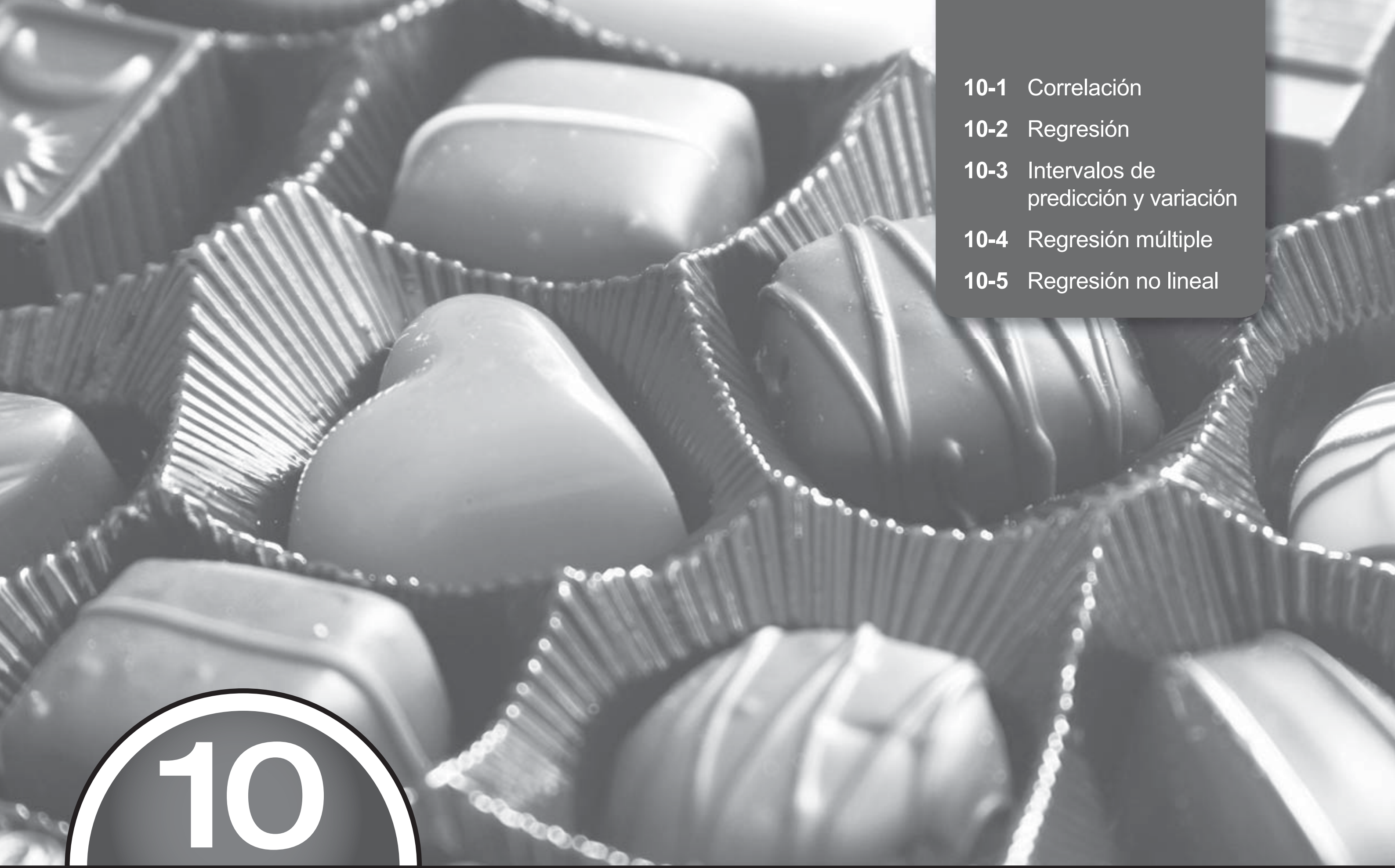
10. Actividad fuera de clase Obtenga datos muestrales para probar la afirmación de que, en la biblioteca de la universidad, los libros de ciencia tienen una antigüedad media menor que la antigüedad media de las novelas de ficción.

11. Actividad fuera de clase Realice experimentos y recopile datos para probar la afirmación de que no hay diferencias en el sabor entre el agua potable común y las diferentes marcas de agua embotellada.

12. Actividad fuera de clase Recopile datos muestrales y pruebe la afirmación de que las personas que hacen ejercicio tienden a tener pulsos más bajos que las que no hacen ejercicio.

13. Actividad fuera de clase Recopile datos muestrales y pruebe la afirmación de que la proporción de mujeres estudiantes que fuman es igual a la proporción de hombres estudiantes que fuman.

14. Actividad fuera de clase Recopile datos muestrales para probar la afirmación de que las mujeres llevan consigo más cambio (dinero fraccionario) que los hombres.



10

CORRELACIÓN Y REGRESIÓN

PROBLEMA DEL CAPÍTULO

¿Comer más chocolate para ganar un Premio Nobel?

Si usted quiere ganar un Premio Nobel, ¿debería estudiar materias como física, química y economía, o debería comer más chocolate? La tabla 10-1 lista el consumo de chocolate (kg *per cápita*) y el número de ganadores del Premio Nobel (por cada 10 millones de personas) para 23 países. Vea el conjunto de datos 16 "Premios Nobel y chocolate" en el apéndice B, donde se identifican los países. En la sección 2-4 se presentaron métodos para elaborar diagramas de dispersión, y la figura 10-1 es el diagrama de dispersión de los datos pareados de chocolate/Nobel. Con base en los métodos de este capítulo, podemos abordar preguntas como las siguientes:

- ¿Existe una *correlación* entre el consumo de chocolate y la tasa de ganadores del Premio Nobel?
- Si existe una correlación entre el consumo de chocolate y la tasa de ganadores del Premio Nobel, ¿podemos describirla con una ecuación para que sea posible predecir la tasa de ganadores del Premio Nobel dada la tasa de consumo de chocolate?

TABLA 10-1 Consumo de chocolate y premios Nobel

Chocolate	Nobel
4.5	5.5
10.2	24.3
4.4	8.6
2.9	0.1
3.9	6.1
0.7	0.1
8.5	25.3
7.3	7.6
6.3	9.0
11.6	12.7
2.5	1.9
8.8	12.7
3.7	3.3
1.8	1.5
4.5	11.4
9.4	25.5
3.6	3.1
2.0	1.9
3.6	1.7
6.4	31.9
11.9	31.5
9.7	18.9
5.3	10.8

- Si un país aumenta su consumo de chocolate, ¿es probable que experimente un aumento en la tasa de premios Nobel?

La pregunta anterior implica al menos tanto sentido común como conocimiento estadístico. Al igual que todos los temas en estadística, el sentido común del pensamiento crítico demuestra ser una herramienta indispensable.

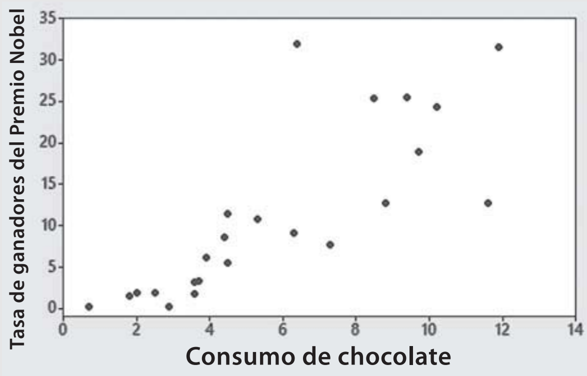


FIGURA 10-1 **Diagrama de dispersión del consumo de chocolate y la tasa de ganadores del Premio Nobel**

OBJETIVOS DEL CAPÍTULO

Uno de los métodos principales de este capítulo es analizar datos de muestras *pareadas*. En la sección 9-3 consideramos datos muestrales que consisten en pares relacionados, pero el objetivo en la sección 9-3 era hacer inferencias sobre la *media de las diferencias* de los pares relacionados. En este capítulo consideramos de nuevo los datos muestrales pareados, pero el objetivo es fundamentalmente diferente al de la sección 9-3. Aquí presentamos métodos para determinar si existe una correlación o asociación entre dos variables. Para las correlaciones lineales, podemos identificar la ecuación de una línea recta que mejor se ajuste a los datos, y podemos usar esa ecuación para predecir el valor de una variable dado el valor de la otra variable. Los objetivos del capítulo son los siguientes:

10-1 Correlación

- Usar datos pareados para encontrar el valor del coeficiente de correlación lineal *r*.
- Determinar si hay evidencia suficiente para respaldar la conclusión de que existe una correlación lineal entre dos variables.

10-2 Regresión

- Usar datos muestrales pareados para encontrar la ecuación de la línea de regresión.
- Encontrar el mejor valor predicho de una variable dado algún valor de la otra variable.

>>>

10-3 Intervalos de predicción y variación

- Usar datos muestrales pareados para determinar el valor del coeficiente de determinación r^2 e interpretar ese valor.
- Utilizar un valor dado de una variable para encontrar un intervalo de predicción para la otra variable.

10-4 Regresión múltiple

- Interpretar los resultados de la tecnología para determinar si una ecuación de regresión múltiple es adecuada para hacer predicciones.
- Comparar los resultados de diferentes combinaciones de variables predictoras e identificar la combinación que resulta en la mejor ecuación de regresión múltiple.

10-5 Regresión no lineal

- Usar datos pareados para identificar los modelos lineales, cuadráticos, logarítmicos, exponenciales y de potencia.
- Determinar qué modelo se ajusta mejor a los datos pareados.

10-1 Correlación

Concepto clave En la parte 1 presentamos el *coeficiente de correlación lineal r*, un número que mide qué tan bien se ajustan los datos muestrales pareados a un patrón de línea recta al ser graficados. Usamos la muestra de datos pareados (en ocasiones llamados **datos bivariados**) para encontrar el valor de r (por lo general empleando tecnología), y luego utilizamos ese valor para decidir si existe una correlación lineal entre las dos variables. En esta sección consideramos sólo relaciones *lineales*, lo que significa que cuando se representan en un diagrama de dispersión, los puntos se aproximan a un patrón en *línea recta*. En la parte 2, analizamos los métodos para realizar una prueba de hipótesis formal que puede usarse con el fin de decidir si existe una correlación lineal entre todos los valores poblacionales de las dos variables.

PARTE 1 Conceptos básicos de la correlación

Iniciamos con la definición básica de *correlación*, un término comúnmente utilizado en el contexto de una asociación entre dos variables.

DEFINICIONES

Existe una **correlación** entre dos variables cuando los valores de una variable están de alguna manera asociados con los valores de la otra variable.

Existe una **correlación lineal** entre dos variables cuando existe una correlación y los puntos graficados de los datos pareados dan como resultado un patrón que se puede aproximar mediante una línea recta.

La tabla 10-1, por ejemplo, incluye datos muestrales pareados que consisten en datos del consumo de chocolate y ganadores del Premio Nobel. Determinaremos si existe una correlación lineal entre la variable x (consumo de chocolate) y la variable y (ganadores del Premio Nobel). En lugar de saltar ciegamente sobre el cálculo del coeficiente de correlación lineal r, es aconsejable primero *explorar* los datos.

¡Explore!

Debido a que siempre conviene explorar los datos muestrales antes de aplicar un procedimiento estadístico formal, deberíamos usar un diagrama de dispersión para explorar visualmente los datos pareados. La figura 10-1 en el problema del capítulo muestra un diagrama de dispersión de los datos. Observamos que no hay valores atípicos (puntos de datos que estuvieran muy alejados de los demás puntos). Los puntos en la figura 10-1 parecen mostrar un patrón de valores crecientes de la tasa de premios Nobel correspondiente a valores crecientes de la tasa de consumo de chocolate.

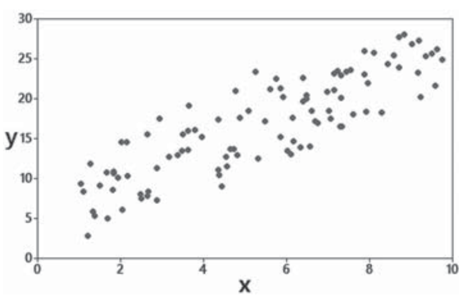
Interpretación de diagramas de dispersión

La figura 10-2 muestra cuatro diagramas de dispersión con diferentes características.

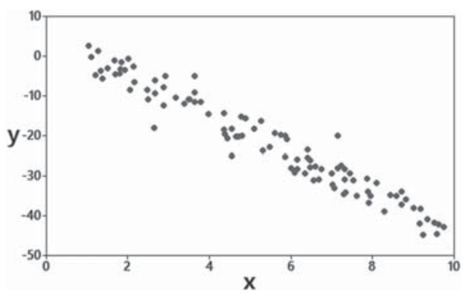
- Figura 10-2(a): Línea recta definida o patrón lineal. Decimos que existe una correlación lineal *positiva* entre x y y, puesto que a medida que los valores de x aumentan, los valores correspondientes de y también aumentan.
- Figura 10-2(b): Línea recta definida o patrón lineal. Decimos que existe una correlación lineal *negativa* entre x y y, puesto que a medida que los valores x aumentan, los valores correspondientes de y disminuyen.
- Figura 10-2(c): Sin patrón definido, lo cual sugiere que no hay correlación entre x y y.
- Figura 10-2(d): Patrón definido que sugiere una correlación entre x y y. Pero el patrón no es el de una línea recta.

Medición de la fuerza de la correlación lineal con *r*

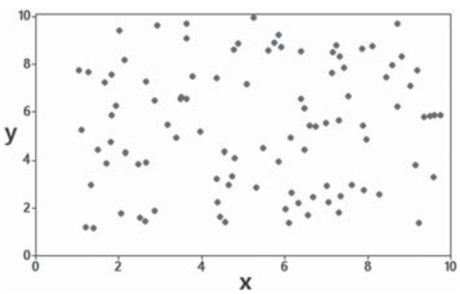
Debido a que las conclusiones basadas en exámenes visuales de diagramas de dispersión son subjetivas en gran medida, necesitamos mediciones más objetivas. Usaremos el coeficiente de correlación lineal r, que es un número que mide la fuerza de la asociación lineal entre las dos variables.



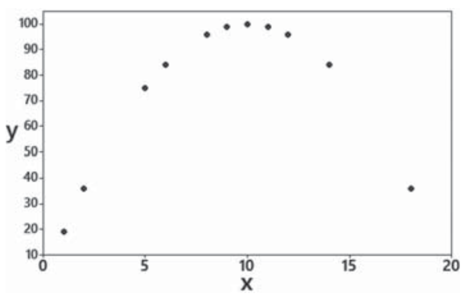
(a) Correlación positiva: r = 0.859



(b) Correlación negativa: r = –0.971



(c) Sin correlación: r = 0.074



(d) Relación no lineal: r = 0.330

FIGURA 10-2 **Diagramas de dispersión**

En cifras

15 mil millones: La cantidad de años que demoraría el reloj atómico en Boulder, Colorado, para retrasarse un segundo.

DEFINICIÓN

El **coeficiente de correlación lineal r** mide la fuerza de la correlación lineal entre los valores cuantitativos pareados x y y en una *muestra*. El coeficiente de correlación lineal r ($-1 \leq r < 0$ para correlación negativa y $0 \leq r \leq 1$ para correlación positiva), se calcula utilizando la fórmula 10-1 o la fórmula 10-2, incluidas en el siguiente cuadro de elementos clave. [El coeficiente de correlación lineal se denomina en ocasiones **coeficiente de correlación del momento del producto Pearson** en honor a Karl Pearson (1857-1936), quien lo desarrolló originalmente].

Debido a que el coeficiente de correlación lineal r se calcula usando datos muestrales, es un estadístico muestral utilizado para medir la fuerza de la correlación lineal entre x y y. Si tuviéramos todos los pares de valores x y y de toda una población, el resultado de la fórmula 10-1 o la fórmula 10-2 sería un parámetro poblacional, representado por ρ (letra griega rho).

ELEMENTOS CLAVE

Cálculo e interpretación del coeficiente de correlación lineal r

Objetivo

Determinar si existe una correlación lineal entre dos variables.

Notación para el coeficiente de correlación lineal

n	número de *pares* de datos muestrales.
Σ	expresa la suma de los elementos indicados
Σx	suma de todos los valores x.
Σx^2	indica que cada valor x debe elevarse al cuadrado y después esos cuadrados deben sumarse.
$(\Sigma x)^2$	indica que los valores de x deben sumarse y el total debe elevarse al cuadrado. Evite confundir Σx^2 y $(\Sigma x)^2$.
Σxy	indica que cada valor x debe multiplicarse por su correspondiente valor y. Después se debe obtener la suma de todos esos productos.
r	coeficiente de correlación lineal para los datos *muestrales*.
ρ	coeficiente de correlación lineal para una *población* de datos pareados.

Requisitos

Dada cualquier recopilación de datos muestrales cuantitativos pareados, siempre es posible calcular el coeficiente de correlación lineal r; pero los siguientes requisitos se deben cumplir cuando se utilizan los datos muestrales pareados para llegar a una conclusión sobre la correlación lineal en la población correspondiente de los datos pareados.

1. La muestra de datos pareados (x, y) es una muestra aleatoria simple de datos cuantitativos. (Es importante que los datos muestrales no se hayan recopilado utilizando algún método inadecuado, como el uso de una muestra de respuesta voluntaria).
2. El examen visual del diagrama de dispersión debe confirmar que los puntos se aproximan a un patrón en línea recta.*
3. Debido a que los resultados pueden verse fuertemente afectados por la presencia de valores atípicos, es necesario eliminar tales valores si se sabe que representan errores. Los efectos de cualquier otro valor atípico se deben considerar calculando r con y sin los valores atípicos incluidos.*

Nota:* Los requisitos 2 y 3 son intentos simplificados para verificar el siguiente requisito formal: Los pares de datos (x, y) deben tener una **distribución normal bivariada. Las distribuciones normales se analizaron en el capítulo 6, pero este supuesto requiere básicamente que para cualquier valor fijo de x, los valores correspondientes de y tengan una distribución que sea aproximadamente normal, y para cualquier valor fijo de y, los valores de x tengan una distribución que sea aproximadamente normal. Este requisito suele ser difícil de verificar; por lo tanto, por ahora utilizaremos los requisitos 2 y 3 como se indicó anteriormente.

continúa

Fórmulas para calcular *r*

FÓRMULA 10-1

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}} \quad \text{(Buen formato para los cálculos)}$$

FÓRMULA 10-2

$$r = \frac{\Sigma(z_x z_y)}{n - 1} \quad \text{(Buen formato para entender)}$$

donde z_x expresa la puntuación z para un valor muestral individual x y z_y es la puntuación z para el valor muestral correspondiente y.

Redondeo del coeficiente de correlación lineal *r*

Redondee el coeficiente de correlación lineal r a tres lugares decimales para que su valor pueda compararse directamente con los valores críticos en la tabla A-6.

Interpretación del coeficiente de correlación lineal *r*

- ***Uso del valor P de la tecnología para interpretar r:*** Utilice el valor P y el nivel de significancia α de la siguiente manera:

  Valor $P \leq \alpha$: Respalda la afirmación de una correlación lineal.
  Valor $P > \alpha$: No respalda la afirmación de una correlación lineal.

- ***Uso de la tabla A-6 para interpretar r:*** Considere que los valores críticos de la tabla A-6 o de la tecnología son tanto positivos como negativos, dibuje una gráfica similar a la figura 10-3 que acompaña al ejemplo 4 en la página 477 y luego use los siguientes criterios de decisión:

  Correlación Si el coeficiente de correlación lineal calculado r se encuentra en la cola izquierda más allá del valor crítico más a la izquierda o si se encuentra en la cola derecha más allá del valor crítico más a la derecha (es decir, $|r| \geq$ valor crítico), concluya que hay suficiente evidencia para respaldar la afirmación de una correlación lineal.

  Sin correlación Si el coeficiente de correlación lineal calculado se encuentra *entre* los dos valores críticos (es decir, $|r| <$ valor crítico), se concluye que no hay pruebas suficientes para respaldar la afirmación de una correlación lineal.

PRECAUCIÓN Recuerde que los métodos de esta sección se aplican a una correlación *lineal*. Si usted concluye que no parece haber una correlación lineal, es posible que haya alguna otra asociación que no sea lineal; como en la figura 10-2(d). Siempre elabore un diagrama de dispersión para ver las relaciones que pueden no ser lineales.

Propiedades del coeficiente de correlación lineal *r*

1. El valor de r siempre está entre -1 y 1 inclusive. Es decir, $-1 \leq r \leq 1$.
2. Si todos los valores de cualquiera de las variables se convierten a una escala diferente, el valor de r no cambia.
3. El valor de r no se ve afectado por la elección de x o y. Si se intercambian todos los valores de x y y, y el valor de r no cambiará.
4. r mide la fuerza de una relación *lineal*. No está diseñado para medir la fuerza de una relación que no sea lineal [como en la figura 10-2(d)].
5. r es muy sensible a los valores atípicos en el sentido de que un único valor de este tipo podría afectar dramáticamente su valor.

Cálculo del coeficiente de correlación lineal *r*

Los siguientes tres ejemplos ilustran tres métodos diferentes para encontrar el valor del coeficiente de correlación lineal *r*, pero sólo es necesario utilizar un método. *Se recomienda el uso de la tecnología (como en el ejemplo 1)*. Si los cálculos manuales son absolutamente necesarios, se recomienda emplear la fórmula 10-1 (como en el ejemplo 2). Si se desea una mejor comprensión de *r*, es recomendable el uso de la fórmula 10-2 (como en el ejemplo 3).

EJEMPLO 1 Determinación de *r* usando la tecnología

Para ilustrar mejor el cálculo de *r*, utilizamos la tabla 10-2, que lista cinco de los valores de datos pareados de la tabla 10-1 después de redondearlos. Use la tecnología para encontrar el valor del coeficiente de correlación *r* para los datos en la tabla 10-2.

TABLA 10-2 Consumo de chocolate y premios Nobel

Chocolate	5	6	4	4	5
Nobel	6	9	3	2	11

SOLUCIÓN

El valor de *r* se calculará automáticamente con un software o una calculadora. Vea las pantallas adjuntas de tecnología, las cuales muestran que $r = 0.795$ (redondeado).

Statdisk

Correlation Results:
Correlation coeff, r: 0.7949366
Critical r: ±0.8783393
P-value (two-tailed): 0.10798

Minitab

Correlation: Nobel, Chocolate

Pearson correlation of Nobel and Chocolate = 0.795
P-Value = 0.108

StatCrunch

Correlation between Chocolate and Nobel is:
0.7949366

XLSTAT

Variables	Chocolate	Nobel
Chocolate	1	0.7949
Nobel	0.7949	1

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
      LinRegTTest
 y=a+bx
 β≠0 and ρ≠0
↑df=3
 a=-11.28571429
 b=3.642857143
 s=2.68594224
 r²=.6319241983
 r=.7949366001
```

SU TURNO Resuelva el ejercicio 13 "Internet y premios Nobel"

EJEMPLO 2 Determinación de *r* usando la fórmula 10-1

Use la fórmula 10-1 para determinar el valor del coeficiente de correlación lineal *r* para los cinco pares de datos de chocolate/Nobel listados en la tabla 10-2.

SOLUCIÓN

Con base en la fórmula 10-1, el valor de r se calcula de la manera mostrada a continuación. Aquí, la variable x se usa para los valores del consumo de chocolate, y la variable y se usa para los valores de los premios Nobel. Debido a que hay cinco pares de datos, $n = 5$. Los valores requeridos restantes se calculan en la tabla 10-3.

TABLA 10-3 Cálculo de r con la fórmula 10-1

x (Chocolate)	y (Nobel)	x^2	y^2	xy
5	6	25	36	30
6	9	36	81	54
4	3	16	9	12
4	2	16	4	8
5	11	25	121	55
$\Sigma x = 24$	$\Sigma y = 31$	$\Sigma x^2 = 118$	$\Sigma y^2 = 251$	$\Sigma xy = 159$

Si se utiliza la fórmula 10-1 con los datos pareados de la tabla 10-3, r se calcula de la siguiente manera:

$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

$$= \frac{5(159) - (24)(31)}{\sqrt{5(118) - (24)^2}\sqrt{5(251) - (31)^2}}$$

$$= \frac{51}{\sqrt{14}\sqrt{294}} = 0.795$$

SU TURNO Resuelva el ejercicio 13 “Internet y premios Nobel”.

EJEMPLO 3 Determinación de r usando la fórmula 10-2

Use la fórmula 10-2 para encontrar el valor del coeficiente de correlación lineal r para los cinco pares de datos de chocolate/Nobel listados en la tabla 10-2.

SOLUCIÓN

Si los cálculos manuales son absolutamente necesarios, la fórmula 10-1 es mucho más fácil que la fórmula 10-2, pero la fórmula 10-2 tiene la ventaja de hacer que sea más fácil *entender* cómo funciona r. (Consulte la *justificación* para r que se analizará más adelante en esta sección). Como en el ejemplo 2, la variable x se usa para los valores del chocolate, y la variable y se utiliza para los valores del Nobel. En la fórmula 10-2, cada valor muestral se reemplaza por su puntuación z correspondiente. Por ejemplo, al usar números no redondeados, los valores del chocolate tienen una media de $\bar{x} = 4.8$ y una desviación estándar de $s_x = 0.836660$, así que el primer valor del chocolate 5 se convierte en una puntuación z de 0.239046 como se muestra aquí:

$$z_x = \frac{x - \bar{x}}{s_x} = \frac{5 - 4.8}{0.836660} = 0.239046$$

La tabla 10-4 lista las puntuaciones z para todos los valores del chocolate (vea la tercera columna) y las puntuaciones z para todos los valores del Nobel (vea la cuarta columna). La última columna de la tabla 10-4 lista los productos $z_x \cdot z_y$.

continúa

¿Los conductores foráneos que rebasan el límite de velocidad son multados con mayor frecuencia?

¿Es más probable que la policía multe a un conductor foráneo que a uno local? Michael Makowsky y Thomas Stratmann, investigadores de la Universidad George Mason, analizaron esta pregunta al examinar más de 60,000 advertencias y multas impuestas por la policía de Massachusetts durante un año. Los autores encontraron que los conductores que no eran de la ciudad tenían un 10% más de probabilidad de ser multados que los conductores locales, y la cifra del 10% llegó al 20% en el caso de los conductores que eran de otro estado. También encontraron una relación estadística entre la economía de la ciudad y las multas por exceso de velocidad. Los conductores foráneos, comparados con los conductores locales, tenían 37% más probabilidades de ser multados al exceder la velocidad límite en una ciudad donde los votantes habían rechazado la propuesta de aumentar los impuestos en una cantidad 2.5% mayor que lo permitido por las leyes estatales. Estos análisis son posibles gracias a los métodos de correlación y regresión.

TABLA 10-4 Cálculo de r con la fórmula 10-2

x (Chocolate)	y (Nobel)	z_x	z_y	$z_x \cdot z_y$
5	6	0.239046	−0.052164	−0.012470
6	9	1.434274	0.730297	1.047446
4	3	−0.956183	−0.834625	0.798054
4	2	−0.956183	−1.095445	1.047446
5	11	0.239046	1.251937	0.299270
				$\Sigma(z_x \cdot z_y) = 3.179746$

Si se usa $\Sigma(z_x \cdot z_y) = 3.179746$ de la tabla 10-4, el valor de r se calcula usando la fórmula 10-2, como se muestra a continuación.

$$r = \frac{\Sigma\,(z_x \cdot z_y)}{n-1} = \frac{3.179746}{4} = 0.795$$

SU TURNO Resuelva el ejercicio 13 "Internet y premios Nobel".

¿Existe una correlación lineal?

Sabemos por los tres ejemplos anteriores que el valor del coeficiente de correlación lineal es $r = 0.795$ para los cinco pares de datos de muestra en la tabla 10-2, pero si usamos los 23 pares de datos muestrales en la tabla 10-1 del problema del capítulo, obtenemos los siguientes resultados de Statdisk, que incluyen $r = 0.801$ (redondeado).

Statdisk

Correlation Results:
Correlation coeff, r: 0.8006078
Critical r: ±0.4132467
P-value (two-tailed): 0.000

Ahora procedemos a interpretar el significado de $r = 0.801$ a partir de los 23 pares de datos de chocolate/Nobel, y nuestro objetivo en esta sección es decidir si parece haber una correlación lineal entre el consumo de chocolate y el número de premios Nobel. Usando los criterios dados en el recuadro de elementos clave anterior, podemos basar nuestra interpretación en un valor P o un valor crítico de la tabla A-6. Consulte los criterios de "Interpretación del coeficiente de correlación lineal r" que se indican en el recuadro de elementos clave anterior.

EJEMPLO 4 ¿Existe una correlación lineal?

Con base en el valor de $r = 0.801$ para los 23 pares de datos en la tabla 10-1 y usando un nivel de significancia de 0.05, ¿hay suficiente evidencia para respaldar la afirmación de que existe una correlación lineal entre el consumo de chocolate y el número de premios Nobel?

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS El primer requisito de una muestra aleatoria simple de datos cuantitativos es cuestionable. Los datos son cuantitativos, pero al examinar los datos originales, no parece que sean pares seleccionados aleatoriamente. Debido a que este requisito no se cumple, los resultados obtenidos pueden no ser válidos. El segundo requisito de un diagrama de dispersión que muestre un patrón en línea recta se satisface. Vea el diagrama de dispersión de la figura 10-1 en la página 469. El diagrama de dispersión de la figura 10-1 también muestra que se cumple el tercer requisito de que no haya valores atípicos. ✓

Podemos basar nuestra conclusión sobre la correlación, ya sea en el valor P obtenido con la tecnología, o en el valor crítico que se encuentra en la tabla A-6. (Consulte los criterios de "Interpretación del coeficiente de correlación lineal r" que se indican en el recuadro de elementos clave anterior).

- ***Uso del valor P de la tecnología para interpretar r:*** Utilice el valor P y el nivel de significancia α de la siguiente manera:

  Valor $P \leq \alpha$: Respalda la afirmación de una correlación lineal.

  Valor $P > \alpha$: No respalda la afirmación de una correlación lineal.

  La pantalla de Statdisk anterior muestra que el valor P es 0.000 cuando se redondea. Debido a que ese valor P es menor o igual al nivel de significancia de 0.05, concluimos que *hay evidencia suficiente para respaldar la conclusión de que, para los países, existe una correlación lineal entre su consumo de chocolate y el número de premios Nobel obtenidos*.

- ***Uso de la tabla A-6 para interpretar r:*** Considere los valores críticos de la tabla A-6 como positivos y como negativos, y trace una gráfica similar a la figura 10-3. Para los 23 pares de datos de la tabla 10-1, la tabla 4-6 arroja un valor crítico que está entre $r = 0.396$ y $r = 0.444$; la tecnología genera un valor crítico de $r = 0.413$. Ahora podemos comparar el valor calculado de $r = 0.801$ con los valores críticos de $r = \pm 0.413$, como se muestra en la figura 10-3.

  Correlación Si el coeficiente de correlación lineal calculado r se encuentra en la región de la cola izquierda o derecha, más allá del valor crítico para esa cola, concluya que hay evidencia suficiente para respaldar la afirmación de una correlación lineal.

  Sin correlación Si el coeficiente de correlación lineal calculado se encuentra entre los dos valores críticos, concluya que no hay evidencia suficiente para respaldar la afirmación de una correlación lineal.

Debido a que la figura 10-3 muestra que el valor calculado de $r = 0.801$ está más allá del valor crítico superior, concluimos que *hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre el consumo de chocolate y el número de premios Nobel* para diferentes países.

INTERPRETACIÓN

Aunque hemos encontrado una correlación lineal, sería absurdo pensar que comer más chocolate ayudaría a ganar un Premio Nobel. Vea el siguiente análisis en "*Interpretación de r como causalidad: ¡No lo haga!*" Además, el requisito de una muestra aleatoria simple no se cumple, por lo que la conclusión de una correlación lineal es cuestionable.

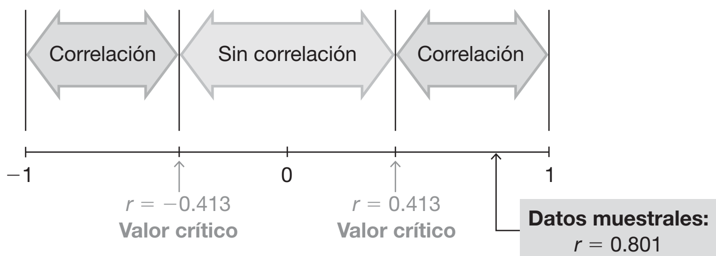


FIGURA 10-3 Valores críticos *r* y el valor calculado *r*

SU TURNO Haga el ejercicio 15 "Pizza y el metro".

Interpretación de *r*: variación explicada

Si concluimos que existe una correlación lineal entre x y y, podemos encontrar una ecuación lineal que exprese y en términos de x, y esa ecuación puede usarse para predecir valores de y para valores dados de x. En la sección 10-2 describiremos un procedimiento para encontrar tales ecuaciones y mostraremos cómo predecir los valores de y cuando se le dan valores de x. Pero un valor predicho de y no será necesariamente el resultado exacto que se produce

Las evaluaciones de los profesores se correlacionan con las calificaciones

Con frecuencia, las evaluaciones que hacen los estudiantes de los profesores se utilizan para medir la eficacia de la enseñanza. Muchos estudios revelan una correlación de altas calificaciones de los estudiantes con evaluaciones positivas de los profesores. Un estudio realizado en la Universidad Duke incluyó evaluaciones de los estudiantes, recabadas antes y después de la entrega de las calificaciones finales. El estudio reveló que "las expectativas de las calificaciones o las calificaciones recibidas causaron un cambio en la forma en que los estudiantes percibían a los maestros y la calidad de su enseñanza". Se señaló que las evaluaciones de los estudiantes "aumentan los incentivos de los profesores para manipular sus políticas de calificación con la intención de mejorar sus evaluaciones". Se concluyó que "la consecuencia final de este tipo de manipulaciones es la degradación de la calidad de la educación en Estados Unidos". (Vea "Teacher Course Evaluations and Student Grades: An Academic Tango", de Valen Johnson, *Chance*, vol. 15, núm. 3).

porque, además de x, hay otros factores que afectan a y, como la variación aleatoria y otras características no incluidas en el estudio. En la sección 10-3, presentaremos un análisis razonado y más detalles sobre el siguiente principio:

El valor de r^2 es la proporción de la variación en y que se explica por la relación lineal entre x y y.

EJEMPLO 5 Variación explicada

Con base en los 23 pares de datos de chocolate/Nobel de la tabla 10-1 en el problema del capítulo, obtenemos $r = 0.801$. ¿Qué proporción de la variación en el número de premios Nobel se puede explicar por la variación en el consumo de chocolate?

SOLUCIÓN

Con $r = 0.801$ obtenemos $r^2 = 0.642$.

INTERPRETACIÓN

Concluimos que 0.642 (o aproximadamente 64%) de la variación en el número de premios Nobel se puede explicar por la relación lineal entre el consumo de chocolate y el número de premios Nobel. Esto implica que alrededor del 36% de la variación en el número de premios Nobel no puede explicarse por las tasas de consumo de chocolate.

Interpretación de r como causalidad: ¡No lo haga!

En el ejemplo 4 llegamos a la conclusión de que hay evidencia suficiente para respaldar la afirmación de una correlación lineal entre el consumo de chocolate y el número de premios Nobel para diferentes países. No debemos llegar a ninguna conclusión que incluya una declaración sobre una relación de causa-efecto entre las dos variables. No deberíamos concluir que ganar premios Nobel hace que las personas consuman más chocolate, ni que consumir más chocolate es causa de un aumento en los premios Nobel. Vea el primero de los siguientes errores comunes y recuerde lo siguiente:

¡La correlación no implica causalidad!

Se observa en el problema del capítulo que debemos usar el sentido común al interpretar los resultados. Claramente, sería absurdo pensar que comer más chocolate ayudaría a ganar un Premio Nobel.

Errores comunes relacionados con la correlación

A continuación se describen tres de los errores más comunes que se producen en la interpretación de resultados que involucran correlación:

1. *Suponer que la correlación implica causalidad*. Un ejemplo clásico incluye datos pareados que consisten en la población de cigüeñas en Copenhague y el número de nacimientos humanos. Durante varios años, los datos sugirieron una correlación lineal. *Boletín:* Las cigüeñas en realidad no causan nacimientos y los nacimientos no causan cigüeñas. Ambas variables se vieron afectadas por otra variable que acechaba en el fondo. (Una **variable interventora** es aquella que afecta las variables que se estudian pero que no se incluye en el estudio). Aquí, una población humana en aumento resultó en más nacimientos y una mayor construcción de techos de paja atrajo a las cigüeñas.
2. *Usar datos basados en promedios*. Los promedios suprimen la variación individual y pueden inflar el coeficiente de correlación. Un estudio produjo un coeficiente de correlación lineal de 0.4 para datos pareados que relacionan ingresos y educación

entre individuos, pero el coeficiente de correlación lineal se convirtió en 0.7 cuando se usaron promedios regionales.

3. *Ignorar la posibilidad de una relación no lineal*. Si no existe una correlación lineal, puede haber alguna otra correlación que no sea lineal, como en la figura 10-2(d).

PARTE 2 Prueba de hipótesis formal

Hipótesis Al realizar una prueba de hipótesis formal para determinar si existe una correlación lineal significativa entre dos variables, considere las siguientes hipótesis nula y alternativa que utilizan ρ para representar el coeficiente de correlación lineal de la población:

Hipótesis nula H_0: $\rho = 0$ (sin correlación)

Hipótesis alternativa H_1: $\rho \neq 0$ (correlación)

Dato estadístico de prueba Se pueden usar los mismos métodos de la parte 1 con el dato estadístico de prueba r, o se puede encontrar el dato estadístico de prueba t usando lo siguiente:

Dato estadístico de prueba $$t = \frac{r}{\sqrt{\dfrac{1 - r^2}{n - 2}}} \quad \text{(con } n - 2 \text{ grados de libertad)}$$

Si se usa el dato estadístico de prueba t anterior, es posible encontrar los valores P y los valores críticos utilizando tecnología, o la tabla A-3 como se describió en los capítulos anteriores. Vea el siguiente ejemplo.

EJEMPLO 6 Prueba de hipótesis usando el valor *P* de la prueba *t*

Utilice los datos pareados de chocolate/Nobel de la tabla 10-1 de la página 469 para llevar a cabo una prueba de hipótesis formal de la afirmación de que existe una correlación lineal entre las dos variables. Use un nivel de significancia de 0.05 con el método del valor P para pruebas de hipótesis.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Los requisitos se abordaron en el ejemplo 4.
Afirmar que existe una correlación lineal es declarar que el coeficiente de correlación lineal poblacional ρ es diferente de 0. Por lo tanto, tenemos las siguientes hipótesis:

$$H_0: \rho = 0 \text{ (No hay correlación lineal)}$$
$$H_1: \rho \neq 0 \text{ (Hay una correlación lineal)}$$

El coeficiente de correlación lineal es $r = 0.801$ (obtenido con tecnología) y $n = 23$ (porque hay 23 pares de datos muestrales), por lo que el dato estadístico de prueba es

$$t = \frac{r}{\sqrt{\dfrac{1 - r^2}{n - 2}}} = \frac{0.801}{\sqrt{\dfrac{1 - 0.801^2}{23 - 2}}} = 6.131$$

Las tecnologías operan con más precisión para obtener un dato estadístico de prueba más preciso, $t = 6.123$. Con $n - 2 = 21$ grados de libertad, la tabla A-3 muestra que el dato estadístico de prueba $t = 6.123$ genera un valor P que es menor que 0.01. Las tecnologías muestran que el valor P es 0.000 al ser redondeado. Debido a que el valor P de 0.000 es menor que el nivel de significancia de 0.05, rechazamos H_0. ("Si el valor P es bajo, la nula debe irse". El valor P de 0.000 es bajo).

INTERPRETACIÓN

Concluimos que, para los países, existe evidencia suficiente para respaldar la afirmación de una correlación lineal entre el consumo de chocolate y los premios Nobel.

SU TURNO Resuelva el ejercicio 17 "Estadísticas del CSI".

Lectura de las manos

Algunas personas piensan que la longitud de la línea de la vida en la palma de las manos puede utilizarse para predecir la longevidad. En una carta publicada en el *Journal of the American Medical Association*, los autores M. E. Wilson y L. E. Mather refutaron esta creencia con un estudio de cadáveres. Se registraron las edades de los sujetos al morir, junto con las longitudes de la línea de la vida de sus palmas. Los autores concluyeron que no existe una correlación significativa entre la edad al morir y la longitud de la línea de la vida. La quiromancia pierde: ¡manos abajo!

Pruebas de una cola Por lo general, los ejemplos y ejercicios en esta sección involucran pruebas de dos colas, pero se pueden tener pruebas de una cola con afirmaciones de una correlación lineal positiva o negativa. En tales casos, las hipótesis serán como sigue:

Afirmación de correlación negativa (prueba de cola izquierda)	Afirmación de correlación positiva (prueba de cola derecha)
H_0: $\rho = 0$	H_0: $\rho = 0$
H_1: $\rho < 0$	H_1: $\rho > 0$

Para estas pruebas de una cola, es posible usar el método del valor *P* como en capítulos anteriores.

Justificación de los métodos de esta sección Hemos presentado las fórmulas 10-1 y 10-2 para calcular *r* y hemos ilustrado su uso. Tales fórmulas se presentan a continuación junto con algunas otras fórmulas "equivalentes", en el sentido de que todas producen los mismos valores.

FÓRMULA 10-1
$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

FÓRMULA 10-2
$$r = \frac{\Sigma(z_x z_y)}{n - 1}$$

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n - 1)s_x s_y} \qquad r = \frac{\Sigma\left[\dfrac{(x - \bar{x})}{s_x}\dfrac{(y - \bar{y})}{s_y}\right]}{n - 1}$$

$$r = \frac{s_{xy}}{\sqrt{s_{xx}}\sqrt{s_{yy}}}$$

Utilizaremos la fórmula 10-2 como una ayuda para comprender el razonamiento que subyace al desarrollo del coeficiente de correlación lineal. Debido a que la fórmula 10-2 utiliza puntuaciones z, el valor de $\Sigma(z_x z_y)$ no depende de la escala que se usa para los valores x y y. La figura 10-1 de la página 469 muestra el diagrama de dispersión de los datos de chocolate/Nobel presentados en la tabla 10-1, y la figura 10-4 en la página siguiente muestra el diagrama de dispersión de las puntuaciones z para los mismos datos muestrales. Compare la figura 10-1 con la figura 10-4 y observe que son esencialmente los mismos diagramas de dispersión con diferentes escalas. Las líneas grises en la figura 10-4 forman los mismos ejes de coordenadas que todos hemos llegado a conocer y amar en cursos de matemáticas anteriores. Esas líneas dividen la figura 10-4 en cuatro cuadrantes.

Si los puntos del diagrama de dispersión se aproximan a una línea ascendente (como en la figura), los valores individuales del producto $z_x \cdot z_y$ tienden a ser positivos (porque la mayoría de los puntos se encuentran en el primer y tercer cuadrantes, donde los valores de z_x y z_y son ambos positivos o ambos negativos), entonces $\Sigma(z_x z_y)$ tiende a ser positiva. 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 z_x y z_y son opuestos en signo, por lo que $\Sigma(z_x z_y)$ tiende a ser negativa. Los puntos que siguen un patrón no lineal tienden a dispersarse entre los cuatro cuadrantes, por lo que el valor de $\Sigma(z_x z_y)$ tiende a ser cercano a 0.

Por lo tanto, podemos usar $\Sigma(z_x z_y)$ como una medida de cómo se configuran los puntos entre los cuatro cuadrantes. Una gran suma positiva sugiere que los puntos están predominantemente en el primer y tercer cuadrantes (que corresponden a una correlación lineal positiva), una gran suma negativa sugiere que los puntos están predominantemente en el segundo

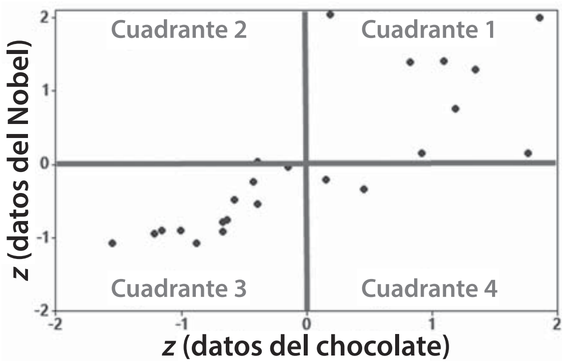


FIGURA 10-4 Diagrama de dispersión de las puntuaciones *z* para los datos de chocolate/Nobel en la tabla 10-1

y el cuarto cuadrantes (lo que corresponde a una correlación lineal negativa), y una suma cercana a 0 sugiere que los puntos están dispersos entre los cuatro cuadrantes (sin correlación lineal). Dividimos $\Sigma(z_x z_y)$ por $n - 1$ para obtener un promedio en lugar de un estadístico que se hace más grande simplemente porque hay más valores de datos. (Las razones para dividir por, $n - 1$ en lugar de n son esencialmente las mismas razones que se relacionan con la desviación estándar). El resultado final es la fórmula 10-2, que puede manipularse algebraicamente en cualquiera de las otras expresiones para obtener r.

CENTRO DE TECNOLOGÍA

Correlación

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Correlation and Regression** en el menú desplegable.
3. Ingrese el nivel de significancia deseado y seleccione las columnas a evaluar.
4. Haga clic en **Evaluate**.
5. Haga clic en **Scatterplot** para obtener un diagrama de dispersión con la línea de regresión incluida.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y elija **Correlation** en el submenú.
3. Seleccione las columnas que se evaluarán en *Variables*.
4. Seleccione el *Método* **Pearson correlation** y marque la casilla de **Display P-values**.
5. Haga clic en **OK**.

Diagrama de dispersión

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression–Fitted Line Plot** en el menú desplegable.
3. Seleccione las columnas deseadas para la variable *y* y la variable *x*.
4. Seleccione **Linear** en el *Tipo de modelo de regresión* y haga clic en **OK**.

SUGERENCIA: Otro procedimiento es hacer clic en **Assistant** en el menú superior, luego seleccionar **Regression** y **Simple Regression**, y finalmente completar el cuadro de diálogo para obtener los resultados.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable, luego elija **Simple Linear** en el submenú.
3. Seleccione las columnas que se usarán para la variable *x* y la variable *y*.
4. Haga clic en **Compute!**
5. Haga clic en la flecha de la parte inferior de la ventana de resultados para ver el diagrama de dispersión.

CENTRO DE TECNOLOGÍA *continuación*

Correlación

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **LinRegTTest** en el menú y presione ENTER.
3. Ingrese los nombres de las listas para las variables *x* y *y*. Introduzca **1** para *Freq* y para β & ρ seleccione $\neq$ **0** para probar la hipótesis nula de no correlación.
4. Seleccione **Calculate** y presione ENTER.

Diagrama de dispersión

1. Abra el menú **STAT PLOTS** presionando 2ND, Y=.
2. Presione ENTER para acceder a la pantalla de configuración Plot 1 mostrada:
   a. Seleccione **ON** y presione ENTER.
   b. Seleccione la primera opción de gráfica (diagrama de dispersión), después presione ENTER.
   c. Ingrese el nombre de las listas que contienen datos para las variables *x* y *y*.
3. Presione ZOOM y luego 9 (ZoomStat) para generar el diagrama de dispersión.

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Modeling Data**.
2. Seleccione **Linear Regression** en el menú desplegable.
3. Ingrese el rango de celdas que contiene los datos de la *variable Y/dependiente* y los datos de la *variable X/Explicativa*. Marque la casilla **Cuantitativa** bajo de la variable *X/Explicativa*. Si la primera fila de datos incluye una etiqueta, marque la casilla de **Variable labels**.
4. Haga clic en la pestaña **Outputs** y asegúrese de que las casillas de **Correlations** y **Analysis of Variance** estén marcadas.
5. Haga clic en **OK** y se mostrarán los resultados del coeficiente de correlación lineal, el valor *P*, el diagrama de dispersión y la prueba de hipótesis. El coeficiente lineal *r* se encuentra en la *matriz de correlación* y el valor *P* se encuentra en la tabla del *análisis de varianza* como *Pr* > *F*.

Excel

1. Haga clic en **Insert Function** f_x seleccione la categoría **Statistical** y seleccione la función **CORREL**. Haga clic en **OK**.
2. Para *Array1* ingrese el rango de datos para la variable independiente *x*. Para *Array2* ingrese el rango de datos para la variable dependiente *y*.
3. Haga clic en **OK** para el coeficiente de correlación lineal *r*.

Diagrama de dispersión (Excel)

1. Seleccione el rango de datos.
2. Haga clic en la pestaña **Insert** de la cinta de opciones.
3. En la sección de *gráficos* del menú superior, seleccione el tipo **Scatter**.
4. Haga clic con el botón derecho sobre la gráfica para personalizarla.

10-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Notación Se seleccionan al azar veinte diferentes estudiantes de estadística. Se mide la temperatura corporal (°C) y la circunferencia de la cabeza (cm) de cada uno de ellos.

a. Para esta muestra de datos pareados, ¿qué representa *r* y qué representa ρ?

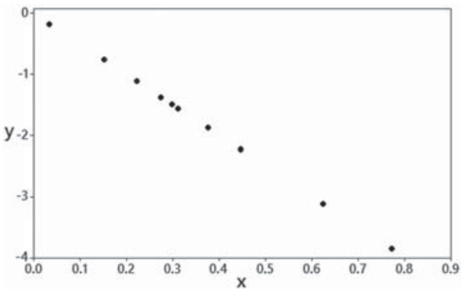
b. Sin hacer ninguna investigación o cálculo, estime el valor de *r*.

c. ¿Cambia *r* si las temperaturas corporales se convierten a grados Fahrenheit?

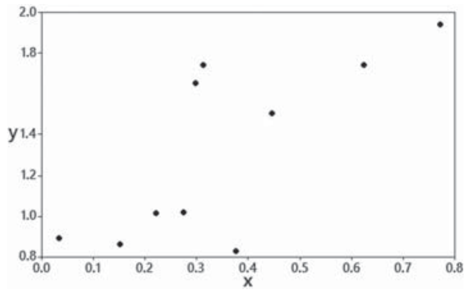
2. Interpretación de *r* Para las mismas dos variables descritas en el ejercicio 1, si encontramos que $r = 0$, ¿eso indica que no hay asociación entre las dos variables?

3. Calentamiento global Si descubrimos que existe una correlación lineal entre la concentración de dióxido de carbono (CO_2) en nuestra atmósfera y la temperatura media global, ¿eso indica que los cambios en el CO_2 provocan cambios en la temperatura global media? ¿Por qué sí o por qué no?

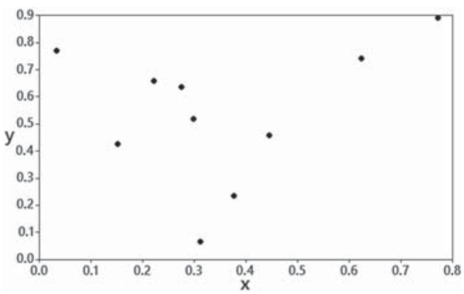
4. Diagramas de dispersión Relacione los siguientes valores de *r* con los cinco diagramas de dispersión que se muestran a continuación: 0.268, 0.992, −1, 0.746 y 1.



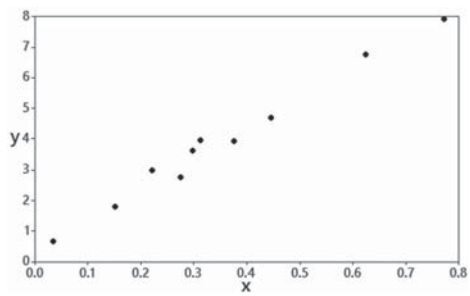
(a)



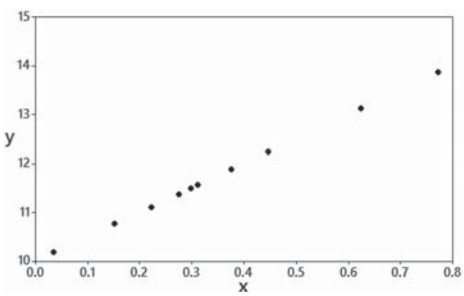
(b)



(c)



(d)



(e)

Interpretación de *r*. ***En los ejercicios 5 a 8, use un nivel de significancia de $\alpha = 0.05$ y refiérase a las pantallas adjuntas.***

5. Peso del oso y tamaño del pecho Se anestesiaron cincuenta y cuatro osos salvajes, y después se midieron sus pesos y tamaños de pecho, los cuales se listan en el conjunto de datos 9 "Mediciones de osos" en el apéndice B; los resultados se muestran en la pantalla de Statdisk adjunta. ¿Hay suficiente evidencia para respaldar la afirmación de que existe una correlación lineal entre el peso de los osos y el tamaño de sus pechos? Al medir un oso anestesiado, ¿es más fácil medir el tamaño del pecho que el peso? Si es así, ¿parece que se puede usar un tamaño de pecho medido para predecir el peso?

Correlation Results:
Correlation coeff, r: 0.963141
Critical r: ±0.2680855
P-value (two-tailed): 0.000

6. Tamaño del casino e ingresos El *New York Times* publicó los tamaños (en pies cuadrados) y los ingresos (en dólares) de siete casinos diferentes en Atlantic City. ¿Existe evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre el tamaño y los ingresos? ¿Los resultados sugieren que un casino puede aumentar sus ingresos al expandir su tamaño?

```
Correlation between Size and Revenue is:
0.44456896
```

7. Basura El conjunto de datos 31 "Peso de la basura" en el apéndice B incluye los pesos de basura desechados en una semana de 62 hogares diferentes. Los pesos pareados de papel y vidrio se usaron para obtener los resultados en XLSTAT que se muestran aquí. ¿Existe evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre los pesos del papel y el vidrio desechados?

XLSTAT

Correlation matrix (Pearson):

Variables	Paper	Glass
Paper	1	0.1174
Glass	0.1174	1

8. Asesinos cereales Se registraron las cantidades de azúcar (gramos de azúcar por gramo de cereal) y las calorías (por gramo de cereal) para una muestra de 16 cereales diferentes. A continuación se muestran los resultados de la calculadora TI-83/84 Plus. ¿Hay suficiente evidencia para respaldar la afirmación de que existe una correlación lineal entre el azúcar y las calorías en un gramo de cereal? Explique.

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
        LinRegTTest
 y=a+bx
 β≠0 and ρ≠0
↑df=14
 a=-188.3423729
 b=5.789830508
 s=11.17312583
 r²=.5858430396
 r=.7654038409
```

¡Explore! ***Los ejercicios 9 y 10 proporcionan dos conjuntos de datos de "Graphs in Statistical Analysis", de F. J. Anscombe,* American Statistician*, vol. 27. Para cada ejercicio:***

a. Elabore un diagrama de dispersión.

b. Encuentre el valor del coeficiente de correlación lineal *r*, luego determine si hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre las dos variables.

c. Identifique la característica de los datos que se perderían si el inciso (b) se resolviera sin elaborar el diagrama de dispersión.

9.

x	10	8	13	9	11	14	6	4	12	7	5
y	9.14	8.14	8.74	8.77	9.26	8.10	6.13	3.10	9.13	7.26	4.74

10.

x	10	8	13	9	11	14	6	4	12	7	5
y	7.46	6.77	12.74	7.11	7.81	8.84	6.08	5.39	8.15	6.42	5.73

11. Valor atípico Consulte el diagrama de dispersión generado por Minitab.

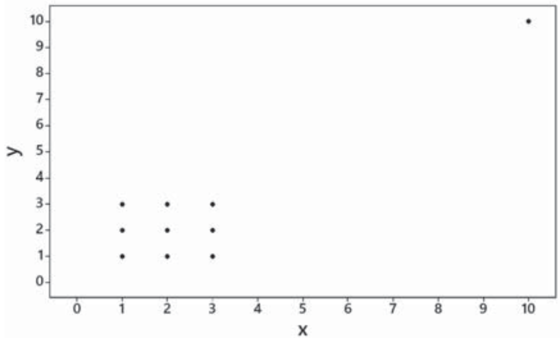
a. Examine el patrón de los 10 puntos y determine subjetivamente si parece haber una correlación entre *x* y *y*.

b. Después de identificar los 10 pares de coordenadas correspondientes a los 10 puntos. Encuentre el valor del coeficiente de correlación *r* y determine si existe una correlación lineal.

continúa

c. Ahora elimine el punto con coordenadas (10, 10) y repita los incisos (a) y (b).

d. ¿Qué se puede concluir sobre el posible efecto de un solo par de valores?



12. Aglomeraciones Consulte el siguiente diagrama de dispersión generado por Minitab. Los cuatro puntos en la esquina inferior izquierda son medidas de mujeres, y los cuatro puntos en la esquina superior derecha son de hombres.

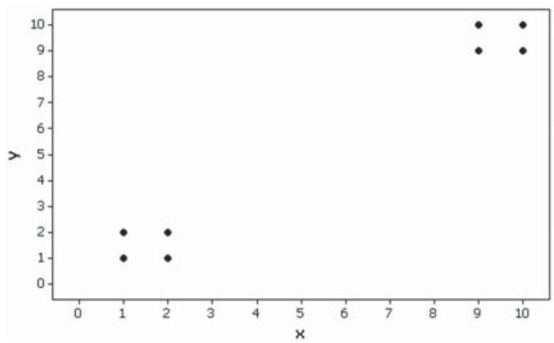
a. Examine el patrón de los cuatro puntos en la esquina inferior izquierda (de mujeres) solamente, y subjetivamente determine si parece haber una correlación entre *x* y *y* para las mujeres.

b. Examine el patrón de los cuatro puntos en la esquina superior derecha (de hombres) solamente, y subjetivamente determine si parece haber una correlación entre *x* y *y* para los hombres.

c. Encuentre el coeficiente de correlación lineal usando sólo los cuatro puntos en la esquina inferior izquierda (de mujeres). ¿Los cuatro puntos en la esquina superior derecha (de hombres) tienen el mismo coeficiente de correlación lineal?

d. Encuentre el valor del coeficiente de correlación lineal usando los ocho puntos. ¿Qué sugiere ese valor sobre la relación entre *x* y *y*?

e. Con base en los resultados anteriores. ¿Qué concluye usted? ¿Deberían considerarse conjuntamente los datos de las mujeres y los datos de los hombres, o parecen representar dos poblaciones diferentes y distintas que deberían analizarse por separado?



Prueba de una correlación lineal. ***En los ejercicios 13 a 28, elabore un diagrama de dispersión y encuentre el valor del coeficiente de correlación lineal r. Encuentre también el valor P o los valores críticos de r a partir de la tabla A-6. Use un nivel de significancia de α = 0.05. Determine si hay suficiente evidencia para respaldar una afirmación de una correlación lineal entre las dos variables. (Guarde su trabajo porque los mismos conjuntos de datos se usarán en los ejercicios de la sección 10-2).***

13. Internet y premios Nobel A continuación se listan las cantidades de usuarios de Internet por cada 100 personas y el número de premios Nobel por cada 10 millones de personas (del conjunto de datos 16 "Premios Nobel y chocolate" del apéndice B) para diferentes países. ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre los usuarios de Internet y los premios Nobel?

Usuarios de Internet	79.5	79.6	56.8	67.6	77.9	38.3
Premios Nobel	5.5	9.0	3.3	1.7	10.8	0.1

14. Old Faithful A continuación se listan los tiempos de duración (en segundos) y los intervalos de tiempo (en minutos) hasta la próxima erupción, de eventos eruptivos seleccionados al azar en el géiser Old Faithful del Parque Nacional Yellowstone. ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre los tiempos de duración de la erupción y el intervalo de tiempo posterior?

Duración	242	255	227	251	262	207	140
Intervalo posterior	91	81	91	92	102	94	91

15. Pizza y el metro La "conexión pizza" es el principio de que el precio de una rebanada de pizza en la ciudad de Nueva York es siempre aproximadamente igual a la tarifa del metro. Use los datos listados a continuación para determinar si existe una correlación lineal significativa entre el costo de una rebanada de pizza y la tarifa del metro.

Año	1960	1973	1986	1995	2002	2003	2009	2013	2015
Costo de la pizza	0.15	0.35	1.00	1.25	1.75	2.00	2.25	2.30	2.75
Tarifa del metro	0.15	0.35	1.00	1.35	1.50	2.00	2.25	2.50	2.75
CPI	30.2	48.3	112.3	162.2	191.9	197.8	214.5	233.0	237.2

16. IPC y el metro Use los datos de IPC/metro del ejercicio anterior para determinar si existe una correlación lineal significativa entre el IPC (índice de precios al consumidor) y la tarifa del metro.

17. Estadísticas del CSI En ocasiones la policía mide las huellas de los zapatos en las escenas del crimen para tener pistas acerca de los delincuentes. A continuación se listan longitudes de las huellas de los zapatos, longitudes de los pies y estaturas de hombres (del conjunto de datos 2 "Pies y estatura" en el apéndice B). ¿Hay evidencia suficiente para concluir que existe una correlación lineal entre la longitud de la huella del zapato y la estatura de los hombres? Con base en estos resultados, ¿parece que la policía puede usar la longitud de los pies para estimar la estatura de un hombre?

Huella del zapato (cm)	29.7	29.7	31.4	31.8	27.6
Longitud del pie (cm)	25.7	25.4	27.9	26.7	25.1
Estatura (cm)	175.3	177.8	185.4	175.3	172.7

18. Estadísticas del CSI Use los datos pareados de longitud del pie y estatura del ejercicio anterior. ¿Hay evidencia suficiente para concluir que existe una correlación lineal entre la longitud del pie y la estatura de los hombres? Con base en estos resultados ¿parece que la policía puede usar la longitud de los pies para estimar la estatura de un hombre?

19. Limones y accidentes automovilísticos A continuación se listan datos anuales de varios años. Los datos son pesos (en toneladas métricas) de los limones importados desde México y las tasas de mortalidad por accidentes automovilísticos en Estados Unidos por cada 100,000 habitantes [según datos de "The Trouble with QSAR (or How I Learned to Stop Worrying and Embrace Fallacy", de Stephen Johnson, *Journal of Chemical Information and Modeling*, vol. 48, núm. 1). ¿Existe evidencia suficiente para concluir que existe una correlación lineal entre el peso de las importaciones de limón de México y las tasas de mortalidad en automóvil en Estados Unidos? ¿Los resultados sugieren que los limones importados causan muertes en automóvil?

Importaciones de limón	230	265	358	480	530
Tasa de mortalidad en choques	15.9	15.7	15.4	15.3	14.9

20. Calificaciones de mpg revisadas A continuación se listan las calificaciones combinadas de ahorro de combustible en la ciudad y en carretera (en mi/gal) para diferentes automóviles. Las calificaciones antiguas se basan en pruebas utilizadas antes de 2008 y las calificaciones nuevas se basan en pruebas que entraron en vigencia en 2008. ¿Existe evidencia suficiente para concluir que existe una correlación lineal entre las calificaciones antiguas y las nuevas? ¿Qué sugieren los datos sobre las calificaciones antiguas?

Antiguas	16	27	17	33	28	24	18	22	20	29	21
Nuevas	15	24	15	29	25	22	16	20	18	26	19

21. Oscar A continuación se listan las edades de los ganadores del Oscar relacionadas de acuerdo con los años en que obtuvieron los premios (del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B). ¿Hay evidencia suficiente para concluir que existe una correlación lineal entre las edades de las mejores actrices y los mejores actores? ¿Deberíamos esperar que haya una correlación?

Mejor actriz	28	30	29	61	32	33	45	29	62	22	44	54
Mejor actor	43	37	38	45	50	48	60	50	39	55	44	33

22. Grillos y temperatura Una aplicación clásica de la correlación implica la asociación entre la temperatura y el número de veces que un grillo chirría en un minuto. A continuación se listan el número de chirridos en 1 minuto y las temperaturas correspondientes en °F (según datos de *The Song of Insects*, de George W. Pierce, Harvard University Press). ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre el número de chirridos en 1 min y la temperatura?

Chirridos en 1 minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (°F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

23. Pesaje de focas con una cámara A continuación se listan los anchos de cabeza (en cm) de focas medidos a partir de fotografías y los pesos (en kg) de las focas (según "Mass Estimation of Weddell Seals Using Techniques of Photogrammetry", de R. Garrott de la Montana State University). El objetivo del estudio fue establecer si el peso de las focas podría determinarse a partir de fotografías aéreas. ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre los anchos de las cabezas medidos en fotografías y su peso?

Ancho de cabeza	7.2	7.4	9.8	9.4	8.8	8.4
Peso	116	154	245	202	200	191

24. Manatíes A continuación se listan las cantidades de embarcaciones recreativas registradas en Florida (decenas de miles) y el número de muertes de manatíes por encuentros con embarcaciones en Florida durante cada uno de los últimos años. Los valores provienen del conjunto de datos 10 "Muertes de manatíes" en el apéndice B. ¿Existe evidencia suficiente para concluir que existe una correlación lineal entre el número de embarcaciones recreativas registradas y el número de muertes de manatíes por encuentros con embarcaciones?

Embarcaciones recreativas	99	99	97	95	90	90	87	90	90
Muertes de manatíes	92	73	90	97	83	88	81	73	68

25. Propinas A continuación se listan montos de cuentas de restaurante durante la cena y montos de las propinas que se dieron. Los datos fueron recolectados por estudiantes del autor. ¿Existe evidencia suficiente para concluir que existe una correlación lineal entre los montos de las cuentas y los tamaños de las propinas? Si todos los comensales tuvieran que dar una propina con el mismo porcentaje, ¿cuál debería ser el valor de r?

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

26. POTUS Los medios de comunicación discuten periódicamente el tema de las estaturas de los candidatos presidenciales ganadores y las estaturas de sus principales oponentes. A continuación se listan esas estaturas (en cm), para varias elecciones presidenciales recientes (del conjunto de datos 15 "Presidentes" en el apéndice B). ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre las estaturas de los candidatos presidenciales ganadores y las estaturas de sus principales oponentes? ¿Debería haber tal correlación?

Presidente	178	182	188	175	179	183	192	182	177	185	188	188	183	188
Oponente	180	180	182	173	178	182	180	180	183	177	173	188	185	175

27. Deportes En la tabla siguiente se listan los diámetros (cm), las circunferencias (cm), y los volúmenes (cm^3) de las pelotas utilizadas en diferentes deportes. ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre los diámetros y las circunferencias? ¿El diagrama de dispersión confirma una asociación *lineal*?

	Béisbol	Básquetbol	Golf	Fútbol	Tenis	Ping pong	Voleibol	Softbol
Diámetro	7.4	23.9	4.3	21.8	7.0	4.0	20.9	9.7
Circunferencia	23.2	75.1	13.5	68.5	22.0	12.6	65.7	30.5
Volumen	212.2	7148.1	41.6	5424.6	179.6	33.5	4780.1	477.9

28. Deportes Repita el ejercicio anterior usando los diámetros y los volúmenes.

Conjuntos de datos del apéndice B. ***En los ejercicios 29 a 34, use los datos del apéndice B para elaborar un diagrama de dispersión, encuentre el valor del coeficiente de correlación lineal r, y determine el valor P o los valores críticos de r a partir de la tabla A-6 y usando un nivel de significancia de $\alpha = 0.05$. Determine si hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre las dos variables. (Guarde su trabajo porque los mismos conjuntos de datos se usarán en los ejercicios de la sección 10-2).***

29. Internet y premios Nobel Use todos los datos pareados de Internet/Nobel listados en el conjunto de datos 16 "Premios Nobel y chocolate" del apéndice B.

30. Old Faithful Utilice todos los tiempos pareados de duración/intervalo posterior listados en el conjunto de datos 23.

31. Estadísticas del CSI Utilice todas las longitudes de la huella del zapato y las estaturas de los 19 hombres del conjunto de datos 2 "Pie y estatura" en el apéndice B.

32. Estadísticas del CSI Utilice todas las longitudes del pie y las estaturas de los 19 hombres del conjunto de datos 2 "Pie y estatura" en el apéndice B.

33. Conteo de palabras para hombres y mujeres Consulte el conjunto de datos 24 "Las palabras cuentan" en el apéndice B y use los conteos de palabras de hombres y mujeres en una relación de pareja que se listan en las dos primeras columnas del conjunto de datos.

34. Terremotos Consulte el conjunto de datos 21 "Terremotos" en el apéndice B y utilice las profundidades y las magnitudes de los terremotos. ¿Parece que las profundidades de los terremotos están asociadas con sus magnitudes?

10-1 Más allá de lo básico

35. Datos transformados Además de probar una correlación lineal entre, x y y, a menudo podemos usar transformaciones de datos para explorar otras relaciones. Por ejemplo, podríamos reemplazar cada valor x por x^2 y usar los métodos de esta sección para determinar si existe una correlación lineal entre y y x^2. Dados los datos pareados de la tabla adjunta, elabore el diagrama de dispersión y luego pruebe una correlación lineal entre y y cada uno de los siguientes valores, ¿cuál de ellos da como resultado el mayor valor de r?

a. x **b.** x^2 **c.** $\log x$ **d.** $\sqrt{x}$ **e.** $1/x$

x	2	3	20	50	95
y	0.3	0.5	1.3	1.7	2.0

36. Determinación de los valores críticos de *r* En la tabla A-6 se listan los valores críticos de r para valores seleccionados de n y α. De manera más general, los valores críticos de r se pueden encontrar mediante la fórmula

$$r = \frac{t}{\sqrt{t^2 + n - 2}}$$

donde el valor t se encuentra en la tabla de valores críticos t (tabla A-3) suponiendo un caso de dos colas con $n - 2$ grados de libertad. Use la fórmula de r dada aquí y en la tabla A-3 (con $n - 2$ grados de libertad) para encontrar los valores críticos de r correspondientes a H_1: $\rho \neq 0$, $\alpha = 0.02$, y $n = 27$.

10-2 Regresión

Concepto clave En esta sección se presentan métodos para encontrar la ecuación de la línea recta que mejor se ajusta a los puntos en un diagrama de dispersión de datos muestrales pareados. Esa línea recta con el mejor ajuste se denomina *línea de regresión*, y su ecuación se llama *ecuación de regresión*. La ecuación de regresión para hacer predicciones para el valor de una de las variables, dado algún valor específico de la otra variable. En la parte 2 de esta sección, analizamos el cambio marginal, los puntos influyentes y los diagramas residuales como herramientas para analizar los resultados de correlación y regresión.

PARTE 1 Conceptos básicos de regresión

En algunos casos, dos variables se relacionan de forma *determinística*, lo que significa que dado un valor para una variable, el valor de la otra se determina exactamente sin ningún error, como en la ecuación $y = 2.54x$ para convertir una distancia r de pulgadas a centímetros. Dichas ecuaciones se consideran en los cursos de álgebra, pero los cursos de estadística se centran en modelos *probabilísticos*, que son ecuaciones con una variable que no está completamente determinada por la otra variable. Por ejemplo, la estatura de un niño no puede determinarse completamente por la estatura del padre y de la madre. Sir Francis Galton (1822-1911) estudió el fenómeno de la herencia y demostró que cuando las parejas altas o bajas tienen hijos, las estaturas de los niños tienden a *regresar*, o volver a la estatura media más típica para las personas del mismo sexo. Continuamos utilizando la terminología "regresión", aunque nuestros datos no involucren los mismos fenómenos de la estatura estudiados por Galton.

DEFINICIONES

Dada una colección de datos muestrales pareados, la **línea de regresión** (o *línea de mejor ajuste*, o *línea de mínimos cuadrados*) es la línea recta que "mejor" se ajusta al diagrama de dispersión de los datos. (El criterio específico para la línea recta de "mejor ajuste" es la propiedad de los "mínimos cuadrados" que se describirá posteriormente).

La **ecuación de regresión**

$$\hat{y} = b_0 + b_1x$$

describe algebraicamente la línea de regresión. La ecuación de regresión expresa una relación entre x (llamada variable explicativa, **variable predictora**, o **variable independiente**) y $\hat{y}$ (llamada **variable de respuesta** o **variable dependiente**).

La definición anterior muestra que en estadística, la ecuación típica de una recta $y = mx + b$ se expresa en la forma $\hat{y} = b_0 + b_1x$, donde b_0 es la intersección y y b_1 es la pendiente. Los valores de b_1 y b_0 se pueden encontrar fácilmente utilizando cualquiera de los muchos programas de software y calculadoras diseñados para proporcionar esos valores, como se ilustra en el ejemplo 1. Los valores de b_1 y b_0 también se pueden determinar mediante cálculos manuales, como se muestra en el ejemplo 2.

ELEMENTOS CLAVE

Determinación de la ecuación de la línea de regresión

Objetivo

Encontrar la ecuación de una línea de regresión.

Notación para la ecuación de una línea de regresión

	Estadístico muestral	Parámetro poblacional
intersección *y* de la ecuación de regresión	b_0	β_0
Pendiente de la ecuación de regresión	b_1	β_1
Ecuación de la línea de regresión	$\hat{y} = b_0 + b_1x$	$y = \beta_0 + \beta_1 x$

Requisitos

1. La muestra de datos pareados (x, y) es una muestra *aleatoria* de datos cuantitativos.
2. El examen visual del diagrama de dispersión muestra que los puntos se aproximan a un patrón en línea recta.*
3. Los valores atípicos pueden tener un fuerte efecto en la ecuación de regresión; por lo tanto, elimine los valores atípicos si se sabe que son errores. Considere los efectos de los valores atípicos que no sean errores conocidos.*

**Nota:* Los requisitos 2 y 3 anteriores son intentos simplificados para verificar los siguientes requisitos formales para el análisis de regresión:

- Para cada valor fijo de x, los valores correspondientes de y tienen una distribución normal.
- Para los diferentes valores fijos de x, las distribuciones de los correspondientes valores de y tienen la misma desviación estándar. (Esto se infringe si parte del diagrama de dispersión muestra puntos muy cercanos a la línea de regresión, mientras que otra parte del diagrama presenta puntos mucho más alejados de la línea de regresión. Consulte el análisis de gráficas residuales en la parte 2 de esta sección).
- Para los diferentes valores fijos de x, las distribuciones de los valores de y correspondientes tienen medias que se encuentran en la misma línea recta.

Los métodos de esta sección no se ven seriamente afectados si las desviaciones de las distribuciones normales y las desviaciones estándar iguales no son demasiado extremas.

Fórmulas para determinar la pendiente b_1 y la intersección *y* b_0 en la ecuación de regresión $\hat{y} = b_0 + b_1x$

FÓRMULA 10-3 **Pendiente:** $b_1 = r\dfrac{s_y}{s_x}$

FÓRMULA 10-4 **intersección *y*:** $b_0 = \bar{y} - b_1\bar{x}$

donde r es el coeficiente de correlación lineal, s_y es la desviación estándar de los valores y, y s_x es la desviación estándar de los valores x.

La pendiente b_1 y la intersección *y* b_0 también se pueden encontrar usando las siguientes fórmulas que son útiles para cálculos manuales o para escribir programas de computadora:

$$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2} \qquad b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Redondeo de la pendiente b_1 y la intersección *y* b_0

Redondee b_1 y b_0 a tres dígitos significativos. Es difícil proporcionar una regla universal simple para redondear los valores de b_1 y b_0, pero esta regla funcionará para la mayoría de las situaciones en este libro. (Dependiendo de cómo redondee, las respuestas del libro a los ejemplos y ejercicios pueden diferir ligeramente de sus respuestas).

EJEMPLO 1 Uso de tecnología para encontrar la ecuación de regresión

Consulte los datos muestrales que se presentan en la tabla 10-1 del problema del capítulo en la página 469. Use la tecnología para encontrar la ecuación de la línea de regresión en la que la variable explicativa (o variable x) es el consumo de chocolate y la variable de respuesta (o variable y) es la tasa de premios Nobel correspondiente.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Se supone que los datos son una muestra aleatoria simple. (2) La figura 10-1 es un diagrama de dispersión que muestra un patrón de puntos, el cual se aproxima a un patrón en línea recta. (3) No hay valores atípicos. Los requisitos se satisfacen. ✓

Tecnología Se recomienda el uso de tecnología para encontrar la ecuación de una línea de regresión. A continuación se muestran los resultados de diferentes tecnologías. Minitab y XLSTAT proporcionan la ecuación real; las otras tecnologías listan los valores de la intersección con el eje y y la pendiente. Todas estas tecnologías muestran que la ecuación de regresión puede expresarse como $\hat{y} = -3.37 + 2.49x$, donde $\hat{y}$ es la tasa de premios Nobel predicha y x es la cantidad de chocolate consumido.

Statdisk

Regression Results:
Y= b0 + b1x:
Y Intercept, b0: −3.366668
Slope, b1: 2.493134

Excel (XLSTAT)

Equation of the model:

Nobel = -3.36667+2.49313*Choc

Minitab

Regression Equation

Nobel = −3.37 + 2.493 Choc

TI-83/84 Plus

NORMAL FLOAT AUTO REAL RADIAN MP

LinRegTTest
y=a+bx
β≠0 and ρ≠0
t=6.123022389
p=4.4774676E-6
df=21
a=-3.366667586
b=2.493133741
↓s=6.262664763

SPSS

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-3.367	2.700		-1.247	.226
	Choc	2.493	.407	.801	6.123	.000

JMP

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-3.366668	2.700151	-1.25	0.2262
Choc	2.4931337	0.407174	6.12	<.0001*

StatCrunch

Simple linear regression results:
Dependent Variable: Nobel
Independent Variable: Choc
Nobel = -3.3666676 + 2.4931337 Choc

Se debe tener en cuenta que la ecuación de regresión es una *estimación* de la ecuación de regresión verdadera para la población de datos pareados. Esta estimación se basa en un conjunto particular de datos muestrales, pero otra muestra extraída de la misma población probablemente conduzca a una ecuación ligeramente diferente.

SU TURNO Resuelva el ejercicio 13 “Internet y premios Nobel”.

EJEMPLO 2 Uso de los cálculos manuales para encontrar la ecuación de regresión

Consulte los datos muestrales que se presentan en la tabla 10-1 del problema del capítulo en la página 469. Use las fórmulas 10-3 y 10-4 para encontrar la ecuación de la línea de regresión en la que la variable explicativa (o la variable x) es el consumo de chocolate y la variable de respuesta (o variable y) es el número correspondiente de premios Nobel.

continúa

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Los requisitos se verificaron en el ejemplo 1.

Iniciamos con la determinación de la pendiente b_1 usando la fórmula 10-3 de la siguiente manera (con dígitos adicionales incluidos para una mayor precisión). Recuerde: r es el coeficiente de correlación lineal, s_y es la desviación estándar de los valores muestrales de y, y s_x es la desviación estándar de los valores muestrales de x.

$$b_1 = r\frac{s_y}{s_x} = 0.800608 \cdot \frac{10.211601}{3.279201} = 2.493135$$

Después de encontrar la pendiente b_1, ahora podemos usar la fórmula 10-4 para determinar la intersección y como sigue:

$$b_0 = \bar{y} - b_1\bar{x} = 11.104348 - (2.493135)(5.804348) = -3.366675$$

Después del redondeo, la pendiente es $b_1 = 2.49$ y la intersección con el eje y es $b_0 = -3.37$. Ahora podemos expresar la ecuación de regresión como $\hat{y} = -3.37 + 2.49x$, donde $\hat{y}$ es la tasa de premios Nobel predicha y x es la cantidad chocolate consumido.

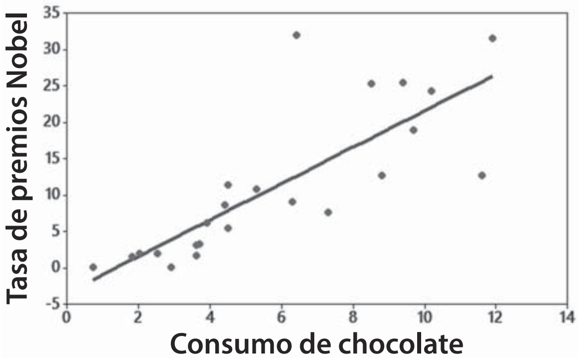
SU TURNO Resuelva el ejercicio 13 "Internet y premios Nobel".

EJEMPLO 3 Gráfica de la línea de regresión

Grafique la ecuación de regresión $\hat{y} = -3.37 + 2.49x$ (encontrada en los ejemplos 1 y 2) sobre el diagrama de dispersión de los datos de chocolate/Nobel de la tabla 10-1 y examine la gráfica para determinar subjetivamente qué tan bien se ajusta la línea de regresión a los datos.

SOLUCIÓN

A continuación se muestra la representación en Minitab del diagrama de dispersión incluyendo la gráfica de la línea de regresión. Podemos ver que la línea de regresión se ajusta bien a los puntos, pero éstos no se encuentran muy cerca de la línea.



Hacer predicciones

Con frecuencia, las ecuaciones de regresión son útiles para *predecir* el valor de una variable, dado algún valor específico de la otra variable. Al hacer predicciones, es necesario tener en cuenta lo siguiente:

1. **Modelo malo:** Si la ecuación de regresión no parece ser útil para hacer predicciones, *no* use la ecuación de regresión para realizarlas. Para los malos modelos, el mejor valor predicho de una variable es simplemente su media muestral.
2. **Buen modelo:** Utilice la ecuación de regresión para realizar predicciones sólo si la gráfica de la línea de regresión sobre el diagrama de dispersión confirma que la línea de regresión se ajusta razonablemente bien a los puntos.

3. **Correlación:** Use la ecuación de regresión para realizar predicciones sólo si el coeficiente de correlación lineal r indica que existe una correlación lineal entre las dos variables (como se describe en la sección 10-1).
4. **Alcance:** Utilice la línea de regresión para realizar predicciones sólo si los datos no van más allá del alcance de los datos muestrales disponibles. (La predicción fuera del alcance de los datos muestrales disponibles se llama *extrapolación*, y podría dar lugar a predicciones erróneas).

En la figura 10-5 se resume una estrategia para predecir valores de una variable y cuando se tiene algún valor dado de x. La figura 10-5 muestra que si la ecuación de regresión es un buen modelo, entonces sustituimos el valor de x en la ecuación de regresión para encontrar el valor predicho de y. Sin embargo, si la ecuación de regresión no es un buen modelo, el mejor valor predicho de y es simplemente $\hat{y}$, la media de los valores y. Recuerde, esta estrategia se aplica a patrones *lineales* de puntos en un diagrama de dispersión. Si el diagrama de dispersión muestra un patrón que no es lineal (no es una línea recta), se aplican otros métodos.

Estrategia para predecir valores de *y*

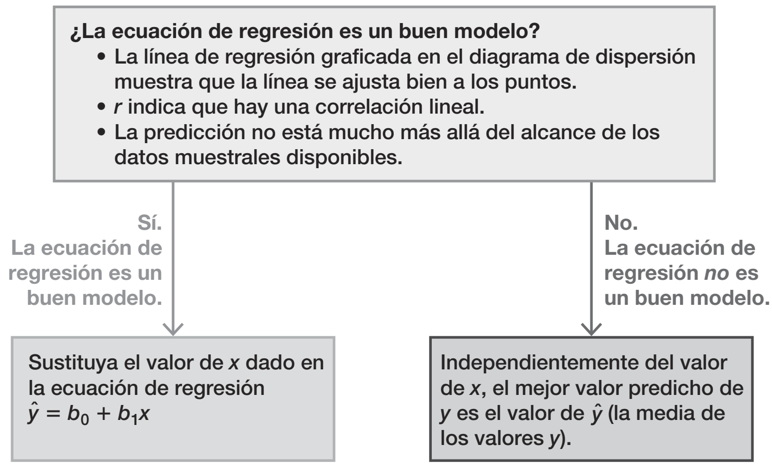


FIGURA 10-5 Estrategia recomendada para predecir valores de *y*

EJEMPLO 4 Realización de predicciones

a. Use los datos de chocolate/Nobel de la tabla 10-1 en la página 469 para predecir la tasa de premios Nobel para un país con un consumo de chocolate de 10 kg *per cápita*.

b. Prediga la puntuación de IQ para un adulto que tiene exactamente 175 cm de altura.

SOLUCIÓN

a. **Buen modelo: Utilice la ecuación de regresión para las predicciones.** La línea de regresión se ajusta bien a los puntos, como se muestra en el ejemplo 3. Además, hay una relación lineal entre el consumo de chocolate y la tasa del premios Nobel, como se muestra en la sección 10-1. Debido a que la ecuación de regresión $\hat{y} = -3.37 + 2.49x$ es un buen modelo, sustituya $x = 10$ en la ecuación de regresión para obtener una tasa de premios Nobel de 21.5 premios Nobel por cada 10 millones de personas.

b. **Modelo malo: Use $\hat{y}$ para las predicciones.** Sabiendo que no existe una correlación entre la estatura y la puntuación de IQ, sabemos que una ecuación de regresión no es un buen modelo, por lo que el mejor valor predicho para la puntuación de IQ es la media, que es 100.

continúa

Posposición de la muerte

Varios estudios tratan sobre la capacidad de las personas para retrasar su muerte hasta después de un suceso importante. Por ejemplo, el sociólogo David Phillips analizó las tasas de mortalidad de hombres judíos que murieron cerca de la Pascua judía, y descubrió que la tasa de mortalidad disminuía drásticamente una semana antes de esa festividad, pero que aumentaba la semana posterior a ella. Otros investigadores de los pacientes con cáncer concluyeron que "no hay un patrón que respalde el concepto de que la muerte toma vacaciones". (Vea "Holidays, Birthdays, and Postponement of Cancer Death", de Young y Hade, *Journal of the American Medical Association*, vol. 292, núm. 24). Con base en los registros de 1.3 millones de muertes, este estudio no encontró una relación entre el momento de la muerte y la Navidad, el Día de Acción de Gracias o el cumpleaños del individuo. David Phillips rebatió estos resultados y dijo que el estudio se enfocó en pacientes con cáncer, quienes tienen menos probabilidades de presentar efectos psicosomáticos.

INTERPRETACIÓN

Note que, en el inciso (a), los datos pareados resultan en un buen modelo de regresión, por lo que la tasa de premios Nobel predicha se encuentra al sustituir el valor de $x = 10$ en la ecuación de regresión. Sin embargo, en el inciso (b) no existe una correlación entre la estatura y el IQ, por lo que la mejor puntuación de IQ predicha es la media $\overline{y} = 100$.

Punto clave: Use la ecuación de regresión para las predicciones sólo si se trata de un buen modelo. Si la ecuación de regresión no es un buen modelo, use el valor predicho de $\overline{y}$.

SU TURNO Resuelva el ejercicio 5 “Citas rápidas”.

PARTE 2 Más allá de los fundamentos de la regresión

En la parte 2, consideramos el concepto de cambio marginal, que es útil para interpretar una ecuación de regresión; después abordamos los efectos de los valores atípicos y puntos especiales llamados *puntos influyentes*. También consideramos las gráficas residuales.

Interpretación de la ecuación de regresión: Cambio marginal

Podemos usar la ecuación de regresión para ver el efecto en una variable cuando la otra variable cambia en una cantidad específica.

DEFINICIÓN

Al trabajar con dos variables relacionadas por una ecuación de regresión, el **cambio marginal** de una variable es la cantidad que cambia cuando la otra se modifica en exactamente una unidad. La pendiente b_1 en la ecuación de regresión representa el cambio marginal en y que ocurre cuando x cambia en una unidad.

Consideremos los 23 pares de datos de chocolate/Nobel incluidos en la tabla 10-1. Esos 23 pares de datos dan como resultado la siguiente ecuación de regresión: $\hat{y} = -3.37 + 2.49x$. La pendiente de 2.49 nos dice que si aumentamos x (el consumo de chocolate) en 1 (kg *per cápita*), la tasa de premios Nobel predicha aumentará en 2.49 (por cada 10 millones de personas). Es decir, por cada aumento adicional de 1 kg *per cápita* en el consumo de chocolate, esperamos que la tasa de premios Nobel aumente en 2.49 por cada 10 millones de personas.

Valores atípicos y puntos influyentes

Un análisis de correlación/regresión de datos bivariados (pareados) debe incluir una investigación de *valores atípicos* y *puntos influyentes*, definidos de la siguiente manera.

DEFINICIONES

En un diagrama de dispersión, un **valor atípico** es un punto que cae lejos de los demás puntos de datos.

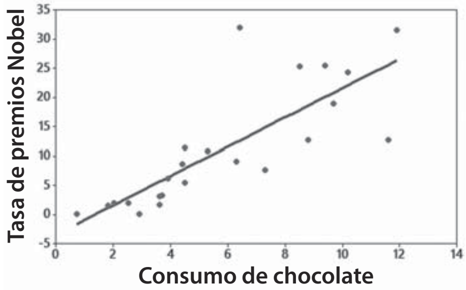
Los datos muestrales pareados pueden incluir uno o más **puntos influyentes**, que son puntos que afectan fuertemente la gráfica de la línea de regresión.

Para determinar si un punto es un valor atípico, examine el diagrama de dispersión para ver si el punto está lejos de los demás. Aquí se muestra cómo determinar si un punto es influyente: Primero, grafique la línea de regresión resultante de los datos con el punto incluido, luego grafique la línea de regresión resultante de los datos con el punto excluido. Si la línea de regresión cambia en una forma considerable, el punto es influyente.

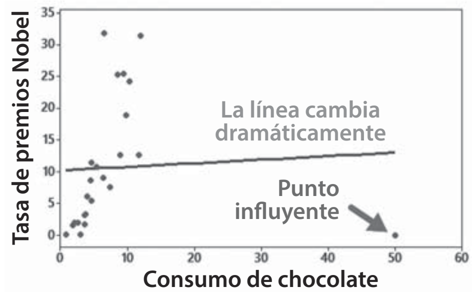
EJEMPLO 5 Punto influyente

Considere los 23 pares de datos de chocolate/ Nobel de la tabla 10-1 en el problema del capítulo. El diagrama de dispersión que se muestra enseguida a la izquierda muestra la línea de regresión. Si incluimos un par adicional de datos, $x = 50$ y $y = 0$, obtenemos la línea de regresión que se muestra a la derecha. El punto adicional (50, 0) es un punto influyente porque la gráfica de la línea de regresión cambia considerablemente, como lo muestra la línea de regresión ubicada en la porción derecha. Compare las dos gráficas para ver claramente que la inclusión de este par de valores tiene un efecto muy dramático en la línea de regresión, por lo que ese punto adicional se considera un punto influyente. El punto adicional también es un valor atípico porque está alejado de los demás puntos.

Datos originales de chocolate/Nobel en la tabla 10-1



Datos de chocolate/Nobel con un punto adicional: (50, 0)



Error de pronóstico de 1°F = mil millones de dólares

El pronóstico de las temperaturas parece una ciencia inexacta; sin embargo, muchas compañías están trabajando con ahínco para obtener estimaciones más precisas. El reportero de *USA Today*, Del Jones, escribió que "el costo anual de la electricidad podría disminuir por lo menos $1000 millones si se mejorara la exactitud de las predicciones del tiempo en 1 grado Fahrenheit". Al referirse a las autoridades de Tennessee Valley (TVA), afirma que los pronósticos han fallado un promedio de 2.35 grados, lo cual es bastante representativo de los pronósticos que se hacen en todo Estados Unidos. Si se mejorara el error de 2.35 a 1.35 grados, la TVA ahorraría hasta $100,000 diarios. El pronóstico de temperaturas se utiliza para determinar la distribución de la energía proveniente de generadores, plantas nucleares, hidroeléctricas, de carbón, de gas natural y eólicas. Las técnicas de pronóstico estadístico se encuentran en proceso de refinamiento, de manera que permitan ahorrar dinero y recursos naturales.

Residuos y propiedad de los mínimos cuadrados

Se ha establecido que la ecuación de regresión representa la línea recta que "mejor" se ajusta a los datos. El criterio para determinar la línea que es mejor que todas las demás se basa en las distancias verticales entre los puntos de datos originales y la línea de regresión. Tales distancias se denominan *residuos*.

DEFINICIÓN

Para un par de valores muestrales x y y, el **residuo** es la diferencia entre el valor *muestral observado* de y y el valor de y que se *predice* usando la ecuación de regresión. Es decir,

$$\text{Residuo} = y \text{ observada} - y \text{ predicha} = y - \hat{y}$$

Hasta ahora, esta definición no ha ganado ningún premio por simplicidad, pero usted puede entender fácilmente los residuos consultando la figura 10-6 en la página siguiente, que corresponde a los datos muestrales pareados que se muestran al margen. En la figura 10-6, los residuos están representados por líneas discontinuas y los datos pareados se grafican como puntos en gris oscuro.

x	8	12	20	24
y	4	24	8	32

Considere el punto muestral con coordenadas (8, 4). Si sustituimos $x = 8$ en la ecuación de regresión $\hat{y} = 1 + x$, obtenemos un valor predicho de $\hat{y} = 9$. Pero para $x = 8$, el valor muestral realmente observado es $y = 4$. La diferencia $y - \hat{y} = 4 - 9 = -5$ es un residuo.

x	8	12	20	24
y	4	24	8	32

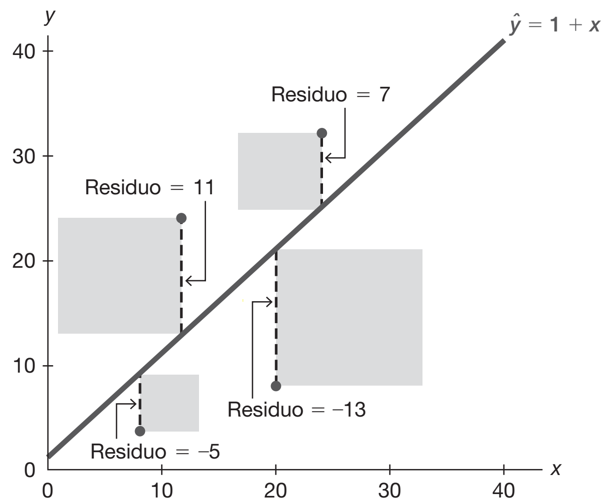


FIGURA 10-6 Residuos y cuadrados de los residuos

La ecuación de regresión representa la línea que "mejor" se ajusta los puntos, de acuerdo con la siguiente propiedad de mínimos cuadrados.

> **DEFINICIÓN**
>
> Una línea recta satisface la **propiedad de mínimos cuadrados** si la suma de los cuadrados de los residuos es la suma más pequeña posible.

En la figura 10-6 se observa que los residuos son -5, 11, -13 y 7, por lo que la suma de sus cuadrados es

$$(-5)^2 + 11^2 + (-13)^2 + 7^2 = 364$$

Podemos visualizar la propiedad de mínimos cuadrados haciendo referencia a la figura 10-6, donde los cuadrados de los residuos se representan mediante las áreas cuadradas sombreadas. La suma de estas áreas es 364, que es la suma más pequeña posible. Use cualquier otra línea recta, y los cuadrados sombreados se combinarán para producir un área más grande que el área combinada de 364.

Afortunadamente, no necesitamos tratar directamente con la propiedad de mínimos cuadrados cuando queremos encontrar la ecuación de la línea de regresión. Se ha utilizado el cálculo para elaborar la propiedad de mínimos cuadrados en las fórmulas 10-3 y 10-4. Debido a que las deducciones de estas fórmulas requieren cálculo, el proceso de derivación no se incluye en el presente texto.

Gráficas residuales

En esta sección y en la anterior, listamos los requisitos simplificados para el análisis efectivo de los resultados de correlación y regresión. Notamos que siempre debemos comenzar con un diagrama de dispersión, y debemos verificar que el patrón de puntos sea aproximadamente un patrón en línea recta. También deberíamos considerar los valores atípicos. Una *gráfica residual* puede ser otra herramienta útil para analizar los resultados de correlación y regresión, así como para verificar los requisitos necesarios para hacer inferencias sobre la correlación y la regresión.

> **DEFINICIÓN**
>
> Una **gráfica residual** es un diagrama de dispersión de los valores (x, y) después de remplazar cada uno de los valores de la coordenada y por el valor del residuo $y - \hat{y}$ (donde $\hat{y}$ expresa el valor predicho de y). Es decir, una gráfica residual es una gráfica de los puntos (x, $y - \hat{y}$).

Para elaborar una gráfica residual, trace una línea de referencia horizontal a través del valor residual de 0, luego trace los valores pareados $(x, y - \hat{y})$. Debido a que la elaboración manual de gráficas residuales puede ser tediosa, se recomienda el uso de la tecnología. Al analizar una gráfica residual, busque un patrón en la forma en que se configuran los puntos, y use los siguientes criterios:

- La gráfica residual no debe tener ningún patrón obvio (ni siquiera un patrón en línea recta). (Esta ausencia de un patrón confirma que un diagrama de dispersión de los datos muestrales tiene un patrón en línea recta en vez de otro patrón).
- La gráfica residual no debe ser mucho más ancha (o más delgada) al recorrerla de izquierda a derecha. (Esto confirma el requisito de que para los diferentes valores fijos de x, las distribuciones de los valores y correspondientes tienen la misma desviación estándar).

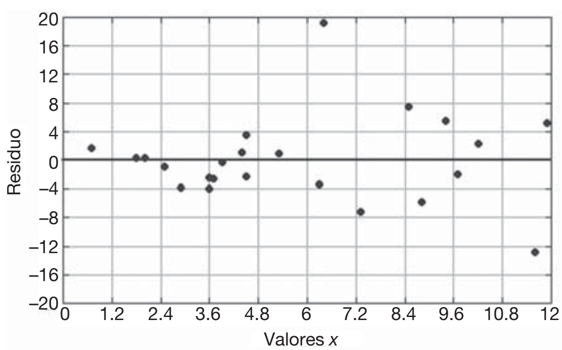
EJEMPLO 6 Gráfica residual

Los datos de chocolate/Nobel de la tabla 10-1 se usan para obtener la gráfica residual generada por Statdisk. Cuando el primer valor muestral x de 4.5 se sustituye en la ecuación de regresión $\hat{y} = -3.37 + 2.49x$ (determinada en los ejemplos 1 y 2), obtenemos el valor predicho de $\hat{y} = 7.84$. Para el primer valor de x de 4.5, el valor real correspondiente de y es 5.5, por lo que el valor del residuo es

$$y \text{ observada} - y \text{ predicha} = y - \hat{y} = 5.5 - 7.84 = -2.34$$

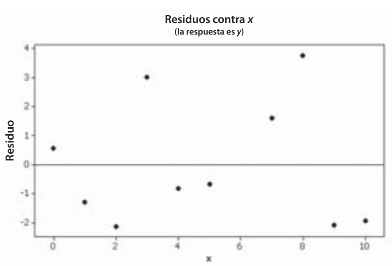
A partir del valor x de 4.5 y el residuo de -2.34, obtenemos las coordenadas del punto $(4.5, -2.34)$, que es uno de los puntos de la gráfica residual que se muestra a continuación.

Statdisk

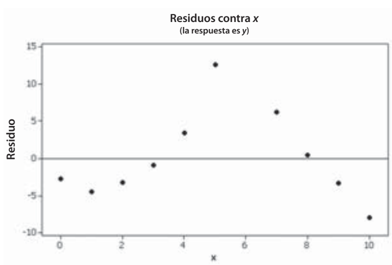


Vea las tres gráficas residuales siguientes. La gráfica de la izquierda sugiere que la ecuación de regresión es un buen modelo. La gráfica de en medio muestra un patrón distinto, lo que sugiere que los datos muestrales no siguen un patrón en línea recta como es requerido. La gráfica de la derecha se vuelve más ancha de izquierda a derecha, lo que sugiere que se viola el requisito de las desviaciones estándar iguales.

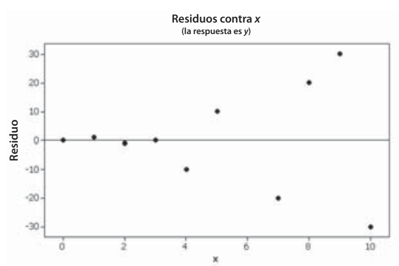
Gráfica residual que sugiere que la ecuación de regresión es un buen modelo



Gráfica residual con un patrón obvio, lo que sugiere que la ecuación de regresión no es un buen modelo



Gráfica residual que se hace más ancha, lo que sugiere que la ecuación de regresión no es un buen modelo



CENTRO DE TECNOLOGÍA

Regresión

Acceda a complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Correlation and Regression** en el menú desplegable.
3. Ingrese el nivel de significancia deseado y seleccione las columnas a evaluar.
4. Haga clic en **Evaluate**.
5. Haga clic en **Scatterplot** para obtener un diagrama de dispersión con la línea de regresión.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable y elija **Regression-Fit Regression Model** en el submenú.
3. En las *respuestas*, seleccione la columna que contiene los valores dependientes *y*. En *predictores continuos*, seleccione la columna que contiene los valores de independientes *x*.
4. Haga clic en **OK**. La ecuación de regresión se incluye en los resultados.

Gráfica de dispersión

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression–Fitted Line Plot** en el menú desplegable.
3. Seleccione las columnas deseadas para la variable *y* y la variable *x*.
4. Seleccione **Linear** para el *tipo de modelo de regresión* y haga clic en **OK**.

SUGERENCIA: Otro procedimiento es hacer clic en **Assistant** del menú superior, luego seleccionar **Regression** y **Simple Regression**. Complete el cuadro de diálogo para obtener los resultados, incluyendo la ecuación de regresión.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable, luego elija **Simple Linear** en el submenú.
3. Seleccione las columnas que se usarán para la variable *x* y la variable *y*.
4. ¡Haga clic en **Compute!**
5. Haga clic en la flecha de la parte inferior de la ventana de resultados para ver el diagrama de dispersión con la línea de regresión.

Calculadora TI-83/84 Plus

1. Presione STAT, luego seleccione **TESTS** en el menú superior.
2. Seleccione **LinRegTTest** en el menú y presione ENTER.
3. Ingrese los nombres de las listas para las variables *x* y *y*. Ingrese **1** para *Freq* y para β & ρ seleccione $\neq$ **0** con el fin de probar la hipótesis nula de no correlación.
4. Seleccione **Calculate** y presione ENTER para ver los resultados, que incluyen la intersección en *y* (a) y la pendiente (b) de la ecuación de regresión.

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Modeling Data.**
2. Seleccione **Linear Regression** en el menú desplegable.
3. Ingrese el rango de celdas que contiene los datos de la *variable dependiente y* y los datos de la *variable explicativa x*. Marque la casilla **Quantitative** bajo la *variable explicativa x*. Si la primera fila de datos incluye una etiqueta, marque la casilla **Variable Labels**.
4. Haga clic en la pestaña **Outputs** y asegúrese de marcar **Correlations**, **Analysis of Variance** y **Prediction and Residuals**.
5. Haga clic en **OK**, y la ecuación de la línea de regresión se mostrará en los resultados.

Excel (complemento de análisis de datos)

1. Haga clic en la pestaña **Data** en la cinta de opciones y luego haga clic en la pestaña **Data Analysis**.
2. Seleccione **Regression** en las *herramientas de análisis* y haga clic en **OK**.
3. Para el *rango de entrada y*, ingrese el rango de datos para la variable dependiente *y*. Para el *rango de entrada x* introduzca el rango de datos para la variable independiente *x*.
4. Marque la casilla **Labels** si la primera fila contiene una etiqueta.
5. Marque las casillas de **Line Fit Plots** y **Residual Plot**, y haga clic en **OK** para ver los resultados. En la tabla de *coeficientes*, la pendiente se etiqueta *X Variable 1* y la intersección en *y* se etiqueta *Intercept*.

SUGERENCIA: La gráfica que se muestra incluirá un diagrama de dispersión de los puntos muestrales originales junto con los puntos que predeciría la ecuación de regresión. Puede obtener la línea de regresión conectando los puntos "*y* predichos".

10-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Notación Se seleccionaron al azar diferentes hoteles ubicados en Las Vegas Boulevard ("The Strip"), y sus calificaciones y precios se obtuvieron en Travelocity. Usando tecnología, donde x representa las calificaciones y y representa el precio, se encontró que la ecuación de regresión tiene una pendiente de 130 y una intersección en y de -368.

a. ¿Cuál es la ecuación de la línea de regresión?

b. ¿Qué representa el símbolo $\hat{y}$?

2. Notación ¿Cuál es la diferencia entre la ecuación de regresión $\hat{y} = b_0 + b_1x$ y la ecuación de regresión $y = \beta_0 + \beta_1x$?

3. Línea de mejor ajuste

a. ¿Qué es un residuo?

b. ¿En qué sentido la línea de regresión es la línea recta que "mejor" se ajusta a los puntos en un diagrama de dispersión?

4. Correlación y pendiente ¿Cuál es la relación entre el coeficiente de correlación lineal r y la pendiente b_1 de una línea de regresión?

Realización de predicciones. ***En los ejercicios 5 a 8, sea la variable de predicción x la primera variable dada. Use los datos dados para encontrar la ecuación de regresión y el mejor valor predicho de la variable de respuesta. Asegúrese de seguir el procedimiento de predicción resumido en la figura 10-5 de la página 493. Use un nivel de significancia de 0.05.***

5. Citas rápidas Para 50 citas rápidas seleccionadas al azar, se registran las calificaciones de atractivo de las mujeres otorgadas por sus parejas masculinas (x) junto con las calificaciones de atractivo de los hombres dadas por sus parejas femeninas (y); las calificaciones provienen del conjunto de datos 18 "Citas rápidas" en el apéndice B. Los 50 pares de calificaciones producen $\bar{x} = 6.5$, $\bar{y} = 5.9$, $r = -0.277$, valor $P = 0.051$ y $\hat{y} = 8.18 - 0.345x$. Encuentre el mejor valor predicho (calificación de atractivo del hombre dada por la mujer) para una cita en la que la calificación de atractivo de la mujer dada por el hombre es $x = 8$.

6. Mediciones de osos Se midieron los anchos de cabeza (pulg) y los pesos (lb) de 20 osos seleccionados al azar (del conjunto de datos 9 "Mediciones de osos" en el apéndice B). Los 20 pares de mediciones producen $\bar{x} = 6.9$ pulg, $\bar{y} = 214.3$ lb, $r = 0.879$, valor $P = 0.000$ y $\hat{y} = -212 + 61.9x$. Encuentre el mejor valor predicho de $\hat{y}$ (peso) dado un oso con un ancho de cabeza de 6.5 pulg.

7. Estatura y peso Se miden las estaturas (cm) y los pesos (kg) de 100 varones adultos seleccionados al azar (del conjunto de datos 1 "Datos corporales" en el apéndice B). Las 100 medidas pareadas producen $\bar{x} = 173.79$ cm, $\bar{y} = 85.93$ kg, $r = 0.418$, valor $P = 0.000$ y $\hat{y} = -106 + 1.10x$. Encuentre el mejor valor predicho de $\hat{y}$ (peso) dado un hombre adulto que mide 180 cm de estatura.

8. Los mejores actores y actrices de reparto En 30 ceremonias recientes de entrega de premios de la Academia, se registraron las edades de los mejores actores de reparto (x) y las edades de las mejores actrices de reparto (y). Las 30 edades pareadas producen $\bar{x} = 52.1$ años, $\bar{y} = 37.3$ años, $r = 0.076$, valor $P = 0.691$ y $\hat{y} = 34.4 + 0.0547x$. Encuentre el mejor valor predicho de $\hat{y}$ (edad de la mejor actriz de reparto) en 1982, cuando la edad del mejor actor de reparto (x) fue de 46 años.

Determinación de la ecuación de la línea de regresión. ***En los ejercicios 9 y 10, use los datos dados para encontrar la ecuación de la línea de regresión. Examine el diagrama de dispersión e identifique una característica de los datos que ignora la línea de regresión.***

9.

x	10	8	13	9	11	14	6	4	12	7	5
y	9.14	8.14	8.74	8.77	9.26	8.10	6.13	3.10	9.13	7.26	4.74

10.

x	10	8	13	9	11	14	6	4	12	7	5
y	7.46	6.77	12.74	7.11	7.81	8.84	6.08	5.39	8.15	6.42	5.73

11. Efectos de un valor atípico Consulte el diagrama de dispersión generado por Minitab que se proporciona en el ejercicio 11 de la sección 10-1, en la página 485.

a. Con base en los pares de valores para los 10 puntos, encuentre la ecuación de la línea de regresión.

b. Después de eliminar el punto con coordenadas (10, 10), use los pares de valores para los 9 puntos restantes y encuentre la ecuación de la línea de regresión.

c. Compare los resultados de los incisos (a) y (b).

12. Efectos de las aglomeraciones Refiérase al diagrama de dispersión generado por Minitab que se proporciona en el ejercicio 12 de la sección 10-1, en la página 485.

a. Con base en los pares de valores para los 8 puntos, encuentre la ecuación de la línea de regresión.

b. Usando sólo los pares de valores para los 4 puntos en la esquina inferior izquierda, determine la ecuación de la línea de regresión.

c. Usando sólo los pares de valores para los 4 puntos en la esquina superior derecha, determine la ecuación de la línea de regresión.

d. Compare los resultados de los incisos (a), (b) y (c).

Regresión y predicciones. ***Los ejercicios 13 a 28 usan los mismos conjuntos de datos que los ejercicios 13 a 28 en la sección 10-1. En cada caso, encuentre la ecuación de regresión, considerando que la primera variable es la variable de predicción (x). Encuentre el valor predicho indicado mediante el procedimiento de predicción resumido en la figura 10-5 de la página 493.***

13. Internet y premios Nobel Encuentre la mejor tasa de premios Nobel predicha para Japón, que tiene 79.1 usuarios de Internet por cada 100 personas. ¿Cómo se compara con la tasa real de premios Nobel de Japón, que es de 1.5 por cada 10 millones de personas?

Usuarios de Internet	79.5	79.6	56.8	67.6	77.9	38.3
Premios Nobel	5.5	9.0	3.3	1.7	10.8	0.1

14. Old Faithful Con base en los tiempos de duración y los intervalos de tiempo posteriores indicados, encuentre el mejor tiempo predicho para el "intervalo posterior" a una erupción con duración de 253 segundos. ¿Cómo se compara con una erupción real que duró 253 segundos y tuvo un intervalo de tiempo posterior de 83 minutos?

Duración	242	255	227	251	262	207	140
Intervalo posterior	91	81	91	92	102	94	91

15. Pizza y el metro Use los costos de la pizza y las tarifas del metro para encontrar la mejor tarifa predicha del metro, dado que el costo de una porción de pizza es de $3.00. ¿Es probable que se implemente la mejor tarifa predicha del metro?

Año	1960	1973	1986	1995	2002	2003	2009	2013	2015
Costo de la pizza	0.15	0.35	1.00	1.25	1.75	2.00	2.25	2.30	2.75
Tarifa del metro	0.15	0.35	1.00	1.35	1.50	2.00	2.25	2.50	2.75
IPC	30.2	48.3	112.3	162.2	191.9	197.8	214.5	233.0	237.2

16. IPC y el metro Use los datos de IPC/tarifas del metro del ejercicio anterior y encuentre la mejor tarifa predicha del metro para un momento en que el IPC llega a 500. ¿Qué es erróneo en esta predicción?

17. Estadísticas del CSI Utilice las longitudes de la huella del zapato y las estaturas para encontrar la mejor estatura pronosticada para un hombre cuya huella de zapato tiene una longitud de 31.3 cm. ¿El resultado sería útil para los investigadores policiales de la escena del crimen al tratar de describir al hombre?

Huella del zapato (cm)	29.7	29.7	31.4	31.8	27.6
Longitud del pie (cm)	25.7	25.4	27.9	26.7	25.1
Estatura (cm)	175.3	177.8	185.4	175.3	172.7

18. Estadísticas del CSI Usa las longitudes del pie y las estaturas para encontrar la mejor estatura predicha para un hombre que tiene una longitud de pie de 28 cm. ¿El resultado sería útil para los investigadores policiales de la escena del crimen al tratar de describir al hombre?

19. Limones y accidentes automovilísticos Utilizando los datos listados de limón/choques, encuentre la mejor tasa de mortalidad por accidente predicha para un año en el que hay importaciones por 500 toneladas métricas de limón. ¿Esta predicción tiene algún valor?

Importaciones de limón	230	265	358	480	530
Tasa de mortalidad por accidente	15.9	15.7	15.4	15.3	14.9

20. Clasificaciones de mpg revisadas Con base en las clasificaciones de mpg viejas/nuevas, encuentre la mejor calificación de mpg predicha para un auto con una clasificación anterior de 30 mpg. ¿Hay algo que sugiera que la predicción es bastante buena?

Antiguas	16	27	17	33	28	24	18	22	20	29	21
Nuevas	15	24	15	29	25	22	16	20	18	26	19

21. Oscar Con base en las edades de actrices/actores listadas, encuentre la mejor edad predicha para el mejor actor dado que la edad de la mejor actriz es de 54 años. ¿Es el resultado razonablemente cercano a la edad real de 33 años del mejor actor (Eddie Redmayne), que ocurrió en 2015, cuando la mejor actriz fue Julianne Moore, que tenía 54 años de edad?

Mejor actriz	28	30	29	61	32	33	45	29	62	22	44	54
Mejor actor	43	37	38	45	50	48	60	50	39	55	44	33

22. Grillos y temperatura Encuentre la mejor temperatura predicha en un momento en que un grillo chirría 3000 veces en 1 minuto. ¿Qué es erróneo en esta temperatura prevista?

Chirridos en 1 minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (°F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

23. Pesaje de focas con una cámara Con base en los datos de ancho/peso listados, encuentre el mejor peso predicho de una foca si el ancho de cabeza medido en una fotografía es de 2 cm. ¿Puede la predicción ser correcta? Si no es así, ¿qué es erróneo?

Ancho de cabeza	7.2	7.4	9.8	9.4	8.8	8.4
Peso	116	154	245	202	200	191

24. Manatíes Use los datos de botes/manatíes. En un año no incluido en los datos siguientes, hubo 970,000 embarcaciones recreativas registradas en Florida. Encuentre la mejor cantidad predicha de muertes de manatíes como resultado de encuentros con embarcaciones. ¿El resultado es razonablemente cercano a 79, que fue el número real de muertes de manatíes?

Embarcaciones recreativas	99	99	97	95	90	90	87	90	90
Muertes de manatíes	92	73	90	97	83	88	81	73	68

25. Propinas Con base en los datos de cuenta/propina, encuentre el mejor monto predicho para una cuenta de $100 durante la cena. ¿Qué regla para dejar propinas sugiere la ecuación de regresión?

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

26. POTUS Utilizando las estaturas del presidente/oponente, encuentre la mejor estatura predicha para un oponente de un presidente que mide 190 cm de altura. ¿Parece que las alturas de los oponentes pueden predecirse a partir las estaturas de los presidentes?

Presidente	178	182	188	175	179	183	192	182	177	185	188	188	183	188
Oponente	180	180	182	173	178	182	180	180	183	177	173	188	185	175

27. Deportes Usando los datos de diámetro/circunferencia, encuentre la mejor circunferencia predicha para una canica con un diámetro de 1.50 cm. ¿Cómo se compara el resultado con la circunferencia real de 4.7 cm?

	Béisbol	Básquetbol	Golf	Fútbol	Tenis	Ping pong	Voleibol	Softbol
Diámetro	7.4	23.9	4.3	21.8	7.0	4.0	20.9	9.7
Circunferencia	23.2	75.1	13.5	68.5	22.0	12.6	65.7	30.5
Volumen	212.2	7148.1	41.6	5424.6	179.6	33.5	4780.1	477.9

28. Deportes Utilizando los datos de diámetro/volumen del ejercicio anterior, encuentre el mejor volumen predicho de una canica con un diámetro de 1.50 cm. ¿Cómo se compara el resultado con el volumen real de 1.8 cm^3?

Grandes conjuntos de datos. ***Los ejercicios 29 a 32 usan los mismos conjuntos de datos del apéndice B que los ejercicios 29 a 32 de la sección 10-1. En cada caso, encuentre la ecuación de regresión, considerando que la primera variable es la variable predictora (x). Encuentre los valores predichos indicados mediante el procedimiento de predicción resumido en la figura 10-5 de la página 493.***

29. Internet y premios Nobel Repita el ejercicio 13 usando todos los datos pareados de Internet/Nobel listados en el conjunto de datos 16 "Premios Nobel y chocolate" del apéndice B.

30. Old Faithful Repita el ejercicio 14 usando todos los tiempos pareados de duración/intervalo posterior listados en el conjunto de datos 23.

31. Estadísticas del CSI Repita el ejercicio 17 usando las longitudes de la huella del zapato y las estaturas de los 19 hombres del conjunto de datos 2 "Pie y estatura".

32. Estadísticas del CSI Repita el ejercicio 18 usando las longitudes de los pies y las estaturas de los 19 hombres del conjunto de datos 2 "Pie y estatura".

33. Conteos de palabras de hombres y mujeres Consulte conjunto de datos 24 "Las palabras cuentan" en el apéndice B y use los conteos de palabras medidos de hombres (x) y mujeres (y) en relaciones de pareja listados en las dos primeras columnas del conjunto de datos 24. Encuentre la mejor predicción para el número de palabras pronunciadas por una mujer, dado que su pareja masculina pronuncia 16,000 palabras en un día.

34. Terremotos Consulte el conjunto de datos 21 "Terremotos" en el apéndice B y utilice las profundidades (x) y las magnitudes (y) de los terremotos. Encuentre la mejor magnitud predicha de un terremoto con una profundidad de 5 km.

10-2 Más allá de lo básico

35. Propiedad de mínimos cuadrados De acuerdo con la propiedad de mínimos cuadrados, la línea de regresión minimiza la suma de los cuadrados de los residuos. Consulte los datos de la tabla 10-1 en la página 469.

a. Encuentre la suma de los cuadrados de los residuos.

b. Demuestre que la ecuación de regresión $\hat{y} = -3 + 2.5x$ resulta en una suma mayor de los cuadrados de los residuos.

10-3 Intervalos de predicción y variación

Concepto clave En la sección 10-2 presentamos un método para usar una ecuación de regresión con el fin de encontrar un valor predicho de y, pero sería grandioso contar con una forma de determinar la *precisión* de tales predicciones. En esta sección, presentamos el *intervalo de predicción*, que es una estimación de intervalo para un valor predicho de y. Vea las siguientes definiciones que distinguen entre un *intervalo de confianza* y un *intervalo de predicción*.

DEFINICIONES

Un **intervalo de predicción** es un rango de valores utilizados para estimar una *variable* (como un valor predicho de y en una ecuación de regresión).

Un **intervalo de confianza** es un rango de valores utilizados para estimar un *parámetro* poblacional (como ρ, μ o σ).

En el ejemplo 4(a) de la sección anterior, mostramos que cuando se usan los 23 pares de consumos de chocolate y tasas de premios Nobel, la ecuación de regresión es $\hat{y} = -3.37 + 2.49x$. Dado un país con un consumo de chocolate de $x = 10$ kg *per cápita*, la mejor tasa de premios Nobel es de 21.5 galardonados por cada 10 millones de personas (que se encuentra al sustituir $x = 10$ en la ecuación de regresión). Para $x = 10$, la "mejor" predicción de la tasa de premios Nobel es 21.5, pero no tenemos conocimiento de la precisión de tal estimación, por lo que necesitamos una estimación de intervalo. Se puede encontrar una estimación del intervalo de predicción para un valor predicho $\hat{y}$ usando los componentes del siguiente recuadro de elementos clave. Dada la naturaleza de los cálculos, se recomienda el uso de la tecnología.

ELEMENTOS CLAVE

Intervalos de predicción

Objetivo

Encontrar un intervalo de predicción, que es una estimación del intervalo para un valor predicho de y.

Requisito

Para cada valor fijo de x, los valores muestrales correspondientes de y se distribuyen normalmente alrededor de la línea de regresión, y esas distribuciones normales tienen la misma varianza.

continúa

Fórmulas para crear un intervalo de predicción

Dado un valor fijo y conocido x_0, el intervalo de predicción para un valor y individual es

$$\hat{y} - E < y < \hat{y} + E$$

donde el margen de error 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 es un valor dado de x, $t_{\alpha/2}$ tiene $n - 2$ grados de libertad, y s_e es el **error estándar de la estimación** encontrado a partir de la fórmula 10-5 o la fórmula 10-6. (El error estándar de la estimación s_e es una medida de la variación de los residuos, que son las diferencias entre los valores muestrales observados y y los valores predichos $\hat{y}$ que se determinan a partir de la ecuación de regresión).

FÓRMULA 10-5 $$s_e = \sqrt{\frac{\Sigma(y - \hat{y})^2}{n - 2}}$$

FÓRMULA 10-6 $$s_e = \sqrt{\frac{\Sigma y^2 - b_0\Sigma y - b_1\Sigma xy}{n - 2}}$$

(Esta es una forma equivalente de la fórmula 10-5 que puede usarse para cálculos manuales o para escribir programas de computadora).

EJEMPLO 1 Chocolate y premios Nobel: Determinación de un intervalo de predicción

Para los datos pareados de chocolate/Nobel en la tabla 10-1 del problema del capítulo, encontramos suficiente evidencia para respaldar la afirmación de una correlación lineal entre esas dos variables, y la ecuación de regresión es $\hat{y} = -3.37 + 2.49x$.

a. Si un país tiene un consumo de chocolate dado de $x = 10$ kg *per cápita*, encuentre el mejor valor predicho de la tasa de premios Nobel.

b. Use un consumo de chocolate de $x = 10$ kg *per cápita* para elaborar un intervalo de predicción del 95% para la tasa de premios Nobel.

SOLUCIÓN

a. Sustituya $x = 10$ en la ecuación de regresión $\hat{y} = -3.37 + 2.49x$ para obtener un valor predicho de $\hat{y} = 21.5$ premios Nobel por cada 10 millones de personas.

b. Las pantallas adjuntas de StatCrunch y Minitab proporcionan el intervalo de predicción del 95%, que es $7.8 < y < 35.3$ con valores redondeados.

StatCrunch

Predicted values:

X value	Pred. Y	s.e.(Pred. y)	95% C.I. for mean	95% P.I. for new
10	21.56467	2.1502909	(17.092895, 26.036445)	(7.7944348, 35.334905)

Minitab

```
Prediction for Nobel
Regression Equation
Nobel = -3.37 + 2.493 Chocolate

Variable   Setting
Chocolate       10

    Fit   SE Fit        95% CI               95% PI
21.5647  2.15029  (17.0929, 26.0364)  (7.79443, 35.3349)
```

El mismo intervalo de predicción del 95% podría calcularse manualmente utilizando los siguientes componentes:

$x_0 = 10$ (dado)

$s_e = 6.262665$ (proporcionado por muchas tecnologías, incluyendo Statdisk, Minitab, Excel, StatCrunch y las calculadoras TI-83/84 Plus)

$\hat{y} = 21.5$ (valor predicho de y encontrado al sustituir $x = 10$ en la ecuación de regresión)

$t_{\alpha/2} = 2.080$ (de la tabla A-3 con gl = 21 y un área de 0.05 en dos colas)

$n = 23$, $\bar{x} = 5.804348$, $\Sigma x = 133.5$, $\Sigma x^2 = 1011.45$

INTERPRETACIÓN

El intervalo de predicción del 95% es $7.8 < y < 35.3$. Esto significa que si seleccionamos un país con una tasa de consumo de chocolate de 10 kg *per cápita* ($x = 10$), tenemos un 95% de confianza en que los límites de 7.8 y 35.3 contienen la tasa de premios Nobel. Este es un amplio rango de valores. El intervalo de predicción sería mucho más estrecho y nuestra tasa estimada de premios Nobel sería mucho mejor si tuviésemos un conjunto mucho mayor de datos muestrales en vez de usar sólo los 23 pares de valores listados en la tabla 10-1.

SU TURNO Resuelva el ejercicio 13 “Embarcaciones”.

Variación explicable e inexplicable

Suponga que tenemos una muestra de datos pareados que tienen las siguientes propiedades mostradas en la figura 10-7:

- Hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre x y y.
- La ecuación de la recta de regresión es $\hat{y} = 3 + 2x$.
- La media de los valores de y está dada por $\bar{y} = 9$.
- Uno de los pares de datos muestrales es $x = 5$ y $y = 19$.
- El punto (5, 13) es uno de los puntos en la línea de regresión, porque al sustituir $x = 5$ en la ecuación de regresión $\hat{y} = 3 + 2x$ se obtiene $\hat{y} = 13$.

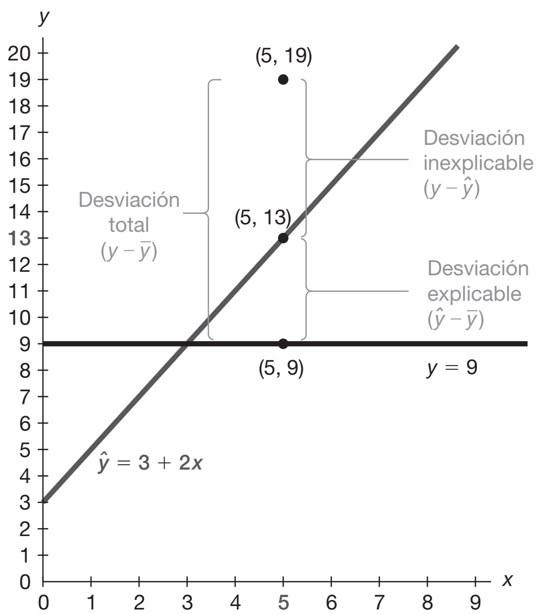


FIGURA 10-7 Desviaciones total, explicable e inexplicable

La figura 10-7 muestra que el punto (5, 13) se encuentra sobre la línea de regresión, pero el punto (5, 19) del conjunto de datos original no cae en la línea de regresión. Si ignoramos completamente los conceptos de correlación y regresión —y queremos predecir un valor de y dado un valor de x y una colección de datos pareados (x, y)— nuestra mejor estimación sería la media $\bar{y} = 9$. Pero en este caso hay una correlación lineal entre x y y, por lo que una mejor manera de predecir el valor de y cuando $x = 5$ es sustituir $x = 5$ en la ecuación de regresión para obtener $\hat{y} = 13$. Podemos explicar la discrepancia entre $\bar{y} = 9$ y $\hat{y} = 13$ al observar que hay una relación lineal mejor descrita por la línea de regresión. En consecuencia, cuando $x = 5$, el valor predicho de y es 13, no el valor medio de 9. Para $x = 5$, el valor predicho de y es 13, pero el valor muestral observado de y es en realidad 19. La discrepancia entre $\hat{y} = 13$ y $y = 19$ no puede explicarse por la línea de regresión, y se llama una *desviación residual* o *inexplicable*, que se puede expresar en el formato general de $y - \hat{y}$.

Como en la sección 3-2, donde definimos la desviación estándar, nuevamente consideramos que una *desviación* es una diferencia entre un valor y la media. (En este caso, la media es $\bar{y} = 9$). Examine cuidadosamente la figura 10-7 y note las siguientes desviaciones específicas de $\bar{y} = 9$:

Desviación total (desde $\bar{y} = 9$) del punto (5, 19) $= y - \bar{y} = 19 - 9 = 10$

Desviación explicable (desde $\bar{y} = 9$) del punto (5, 19) $= \hat{y} - \bar{y} = 13 - 9 = 4$

Desviación inexplicable (desde $\bar{y} = 9$) del punto (5, 19) $= y - \hat{y} = 19 - 13 = 6$

Estas desviaciones de la media se generalizan y se definen formalmente de la siguiente manera.

DEFINICIONES

Supongamos que tenemos una colección de datos pareados que contiene el punto muestral (x, y), que $\hat{y}$ es el valor predicho de y (obtenido mediante el uso de la ecuación de regresión) y que la media de los valores muestrales de y es $\bar{y}$.

La **desviación total** de (x, y) es la distancia vertical $y - \bar{y}$, que es la distancia entre el punto (x, y) y la línea horizontal que pasa a través de la media muestral $\bar{y}$.

La **desviación explicable** es la distancia vertical $\hat{y} - \bar{y}$, que es la distancia entre el valor predicho de y y la línea horizontal que pasa a través de la media muestral $\bar{y}$.

La **desviación inexplicable** es la distancia vertical $y - \hat{y}$, que es la distancia vertical entre el punto (x, y) y la línea de regresión. (La distancia $y - \hat{y}$ también se denomina *residuo*, tal como se definió en la sección 10-2).

En la figura 10-7 podemos ver la siguiente relación para un punto individual (x, y):

$$\begin{array}{ccccc} \text{(desviación total)} & = & \text{(desviación explicable)} & + & \text{(desviación inexplicable)} \\ (y - \bar{y}) & = & (\hat{y} - \bar{y}) & + & (y - \hat{y}) \end{array}$$

La expresión anterior involucra desviaciones de la media, y se aplica a cualquier punto particular (x, y). Si sumamos los cuadrados de las desviaciones usando todos los puntos (x, y), obtenemos las cantidades de *variación*. La misma relación se aplica a las sumas de cuadrados que se muestran en la fórmula 10-7, aunque la expresión anterior no es algebraicamente equivalente a la fórmula 10-7. En la fórmula 10-7, la variación total es la suma de los cuadrados de los valores de la desviación total, la variación explicable es la suma de los cuadrados de los valores de desviación explicable, y la variación inexplicable es la suma de los cuadrados de los valores de desviación inexplicable.

FÓRMULA 10-7

$$\begin{array}{ccccc}(\text{variación total}) & = & (\text{variación explicable}) & + & (\text{variación inexplicable}) \\ \Sigma(y-\bar{y})^2 & = & \Sigma(\hat{y}-\bar{y})^2 & + & \Sigma(y-\hat{y})^2\end{array}$$

Coeficiente de determinación

En la sección 10-1 vimos que el coeficiente de correlación lineal r puede usarse para encontrar la proporción de la variación total en y que puede explicarse por la correlación lineal. Esta declaración se hizo en la sección 10-1:

El valor de r^2 es la proporción de la variación en y que se explica por la relación lineal entre x y y.

Esta declaración sobre la variación explicable se formaliza con la siguiente definición.

DEFINICIÓN

El **coeficiente de determinación** es la proporción de la variación en y que se explica por la línea de regresión. Se calcula como

$$r^2 = \frac{\text{variación explicable}}{\text{variación total}}$$

Podemos calcular r^2 usando la definición que acabamos de dar con la fórmula 10-7, o simplemente podemos elevar al cuadrado el coeficiente de correlación lineal r. Se recomienda usar el cuadrado de r.

EJEMPLO 2 Datos de chocolate/Nobel: Obtención de un coeficiente de determinación

Si usamos los 23 pares de datos de chocolate/Nobel en la tabla 10-1 del problema del capítulo, encontramos que el coeficiente de correlación lineal es $r = 0.801$. Obtenga el coeficiente de determinación. Además, determine el porcentaje de la variación total en y (tasa de premios Nobel) que puede explicarse por la correlación lineal entre el consumo de chocolate y la tasa de premios Nobel.

SOLUCIÓN

Con $r = 0.801$ el coeficiente de determinación es $r^2 = 0.642$.

INTERPRETACIÓN

Dado que r^2 es la proporción de la variación total que puede explicarse, concluimos que 64.2% de la variación total en la tasa de premios Nobel puede explicarse por el consumo de chocolate, y el otro 35.8% no puede explicarse por dicho consumo. El 35.8% podría explicarse por algunos otros factores y/o por variación aleatoria. Pero el sentido común sugiere que es un poco inocente pensar seriamente que la tasa de premios Nobel de un país se ve afectada por la cantidad de chocolate consumido.

SU TURNO Resuelva el ejercicio 5 “Grillos y temperatura”.

CENTRO DE TECNOLOGÍA

Intervalos de predicción

Acceda a complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Statdisk proporciona la intersección y la pendiente de la ecuación de regresión, el error estándar de la estimación (etiquetado como "Standard Error") y el coeficiente de determinación. Estos resultados son útiles para encontrar un intervalo de predicción, pero no se proporciona el intervalo de predicción en sí.

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Correlation and Regression** en el menú desplegable.
3. Ingrese el nivel de significancia deseado y seleccione las dos columnas a evaluar.
4. Haga clic en **Evaluate**.

Minitab

1. Complete el procedimiento de regresión en Minitab de la sección 10-2 para obtener la ecuación de regresión. Minitab usará automáticamente esta ecuación en el siguiente procedimiento.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Regresión** en el menú desplegable y elija **Regression–Predict** en el submenú.
4. Seleccione **Enter individual values** en el menú desplegable.
5. Ingrese los valores deseados para la variable *x*.
6. Haga clic en el botón **Options** y cambie el nivel de confianza al valor deseado.
7. Haga clic en **OK** dos veces.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable, luego elija **Simple Linear** en el submenú.
3. Seleccione las columnas que se utilizarán para la variable *x* y la variable *y*.
4. Para la *predicción* de *y*, ingrese los valores de *x* deseados y el nivel de significancia.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

Los resultados de la calculadora TI-83/84 Plus incluyen la intersección (a) y la pendiente de la ecuación de regresión (b), el error estándar de estimación (s) y el coeficiente de determinación (r^2). Estos resultados son útiles para encontrar un intervalo de predicción, pero no se proporciona el intervalo de predicción en sí.

1. Presione STAT y luego seleccione **TESTS** en el menú superior.
2. Seleccione **LinRegTTest** en el menú y presione ENTER.
3. Ingrese los nombres de las listas para las variables *x* y *y*. Ingrese **1** para *Freq* y para β & ρ seleccione **≠ 0** con el fin de probar la hipótesis nula de no correlación.
4. Seleccione **Calculate** y presione ENTER para ver los resultados.

Excel

Complemento XLSTAT

1. Ingrese los datos muestrales en columnas de la hoja de trabajo.
2. Ingrese los valores deseados de *x* que se usarán para el intervalo de predicción en una columna.
3. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Modeling Data**.
4. Seleccione **Linear Regression** en el menú desplegable.
5. Ingrese el rango de celdas que contiene los datos de la *variable dependiente y* y los datos de la *variable explicativa x*. Marque la casilla **Quantitative** bajo la *variable explicativa x*. Si la primera fila de datos incluye una etiqueta, marque la casilla **Variable labels**.
6. Haga clic en la pestaña **Options** e ingrese el intervalo de confianza deseado, por ejemplo **95**.
7. Haga clic en la pestaña **Prediction**.
8. Marque la casilla **Prediction** y en el cuadro *cuantitativo* ingrese el rango de celdas que contiene el(los) valor(es) deseado(s) de *x* del paso 2. La primera celda del rango debe contener un valor, no una etiqueta.
9. Haga clic en **OK**. El(los) intervalo(s) de predicción se encuentran en la tabla de *predicciones para las nuevas observaciones*.

Excel (complemento de análisis de datos)

Excel proporciona la intersección y la pendiente de la ecuación de regresión, el error estándar de la estimación s_e (etiquetado como "Standard Error") y el coeficiente de determinación (etiquetado como "R Square"). Estos resultados son útiles para encontrar un intervalo de predicción, pero no se proporciona el intervalo de predicción en sí.

1. Haga clic en la pestaña **Data** en la cinta de opciones y luego haga clic en la pestaña **Data Analysis**.
2. Seleccione **Regression** en *herramientas de análisis* y haga clic en **OK**.
3. Para el *rango de entrada y*, ingrese el rango de datos para la variable dependiente *y*. Para el *rango de entrada x* ingrese el rango de datos para la variable independiente *x*.
4. Marque la casilla **Labels** si la primera fila contiene una etiqueta.
5. Haga clic en **OK** para desplegar los resultados.

10-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Notación s_e Con base en el conjunto de datos 1 "Datos corporales" del apéndice B, considere que la variable predictora x representa las estaturas de los hombres y que la variable de respuesta y representa los pesos de los hombres. La muestra de 153 estaturas y pesos resulta en $s_e = 16.27555$ cm. Describa con sus propias palabras qué representa ese valor de s_e.

2. Intervalo de predicción Usando las estaturas y los pesos descritos en el ejercicio 1, se considera una estatura de 180 cm para encontrar que el peso pronosticado es de 91.3 kg, y el intervalo de predicción del 95% es (59.0 kg, 123.6 kg). Escriba una declaración que interprete ese intervalo de predicción. ¿Cuál es la principal ventaja de utilizar un intervalo de predicción en lugar de simplemente usar el peso predicho de 91.3 kg? ¿Por qué se usa la terminología del *intervalo de predicción* en lugar del *intervalo de confianza*?

3. Coeficiente de determinación Usando las estaturas y los pesos descritos en el ejercicio 1, el coeficiente de correlación lineal r es 0.394. Encuentre el valor del coeficiente de determinación. ¿Qué información práctica proporciona el coeficiente de determinación?

4. Error estándar de la estimación Se obtiene una muestra aleatoria de 118 diferentes estudiantes femeninas de estadística y sus pesos se miden en kilogramos y en libras. Usando los 118 pesos pareados (peso en kg, peso en lb), ¿cuál es el valor de s_e? Para una estudiante de estadística que pesa 100 lb, el peso pronosticado en kilogramos es 45.4 kg. ¿Cuál es el intervalo de predicción del 95%?

Interpretación del coeficiente de determinación. ***En los ejercicios 5 a 8, use el valor del coeficiente de correlación lineal r para encontrar el coeficiente de determinación y el porcentaje de la variación total que puede explicarse por la relación lineal entre las dos variables.***

5. Grillos y temperatura $r = 0.874$ (x = número de chirridos del grillo en un minuto, y = temperatura en °F).

6. Pizza y el metro $r = 0.992$ (x = costo de una rebanada de pizza, y = tarifa del metro en la ciudad de Nueva York).

7. Peso/cintura $r = 0.885$ (x = peso del hombre, y = tamaño de la cintura del hombre).

8. Osos $r = 0.783$ (x = ancho de la cabeza del oso, y = peso del oso).

Interpretación de una pantalla de computadora. ***En los ejercicios 9 a 12, consulte la pantalla obtenida utilizando los datos pareados que consisten en las embarcaciones registradas en Florida (decenas de miles) y el número de muertes de manatíes por encuentros con embarcaciones en Florida durante los años recientes (del conjunto de datos 10 en el apéndice B). Junto con los datos muestrales de embarcaciones/manatíes, se introdujo a StatCrunch el valor de 85 (decenas de miles) embarcaciones que se utilizarán para predecir las muertes de manatíes.***

StatCrunch

Manatees = −49.048987 + 1.4062442 Boats
Sample size: 24
R (correlation coefficient) = 0.85014394
Estimate of error standard deviation: 9.6605284

Predicted values:

X value	Pred. Y	s.e.(Pred. y)	95% C.I. for mean	95% P.I. for new
85	70.481772	1.9724935	(66.391071, 74.572473)	(50.033706, 90.929839)

9. Prueba de correlación Utilice la información proporcionada en la pantalla para determinar el valor del coeficiente de correlación lineal. ¿Existe evidencia suficiente para respaldar la afirmación de una correlación lineal entre el número de embarcaciones registradas y el número de muertes de manatíes por encuentros con embarcaciones?

10. Identificación de la variación total ¿Qué porcentaje de la variación total en las muertes de manatíes se puede explicar por la correlación lineal entre las embarcaciones registradas y las muertes de manatíes?

11. Predicción de muertes de manatíes Considere $x = 85$ (para 850,000 embarcaciones registradas). ¿Cuál es el valor único de la mejor cantidad predicha de muertes de manatíes como resultado de encuentros con embarcaciones?

12. Determinación de un intervalo de predicción Durante un año con 850,000 ($x = 85$) embarcaciones registradas en Florida, identifique una estimación del intervalo de predicción del 95% para el número de muertes de manatíes resultantes de encuentros con embarcaciones. Escriba una declaración que interprete el intervalo.

Determinación de un intervalo de predicción. ***En los ejercicios 13 a 16, use los datos pareados consistentes en las embarcaciones registradas en Florida (decenas de miles) y las muertes de manatíes por encuentros con embarcaciones que se listan en el conjunto de datos 10 "Muertes de manatíes" del apéndice B. Considere que x representa el número de embarcaciones registradas y que y es el número correspondiente de muertes de manatíes. Use la cantidad dada de barcos registrados y el nivel de confianza dado para elaborar una estimación del intervalo de predicción para las muertes de manatíes.***

13. Embarcaciones Use $x = 85$ (para 850,000 embarcaciones registradas) con un nivel de confianza del 99%.

14. Embarcaciones Use $x = 98$ (para 980,000 embarcaciones registradas) con un nivel de confianza del 95%.

15. Embarcaciones Use $x = 96$ (para 960,000 embarcaciones registradas) con un nivel de confianza del 95%.

16. Embarcaciones Use $x = 87$ (para 870,000 embarcaciones registradas) con un nivel de confianza del 99%.

Variación e intervalos de predicción. ***En los ejercicios 17 a 20, encuentre (a) la variación explicable, (b) la variación inexplicable y (c) el intervalo de predicción indicado. En cada caso, hay suficiente evidencia para respaldar la afirmación de una correlación lineal, por lo que es razonable usar la ecuación de regresión al realizar predicciones.***

17. Altitud y temperatura A continuación se listan las altitudes (miles de pies) y las temperaturas al aire libre (°F) registradas por el autor durante el vuelo 1053 de Delta desde Nueva Orleans hasta Atlanta. Para el intervalo de predicción, use un nivel de confianza del 95% con la altitud de 6327 pies (o 6.327 mil pies).

Altitud (miles de pies)	3	10	14	22	28	31	33
Temperatura (°F)	57	37	24	−5	−30	−41	−54

18. Tribunales de la ciudad A continuación se listan las cantidades de ingresos a los tribunales y los salarios pagados a los jueces de varias ciudades (según datos del *Poughkeepsie Journal*). Las cantidades se dan en miles de dólares y todas las ciudades están en el condado de Dutchess, Nueva York. Para el intervalo de predicción, use un nivel de confianza del 99% con un ingreso al tribunal de $800,000.

Ingresos al tribunal	65	404	1567	1131	272	252	111	154	32
Salario del juez	30	44	92	56	46	61	25	26	18

19. Grillos y temperatura La siguiente tabla lista el número de chirridos de un grillo en un minuto y la temperatura en °F. Para el intervalo de predicción, use 1000 chirridos en un minuto y un nivel de confianza del 90%.

Chirridos en 1 minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (°F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

20. Pesaje de focas con una cámara La siguiente tabla lista los anchos de cabeza (cm) de focas medidos en fotografías y los pesos (kg) de las focas (de acuerdo con "Mass Estimation of Weddell Seals Using Techniques of Photogrammetry", de R. Garrot de la Montana State University). Para el intervalo de predicción, use un nivel de confianza del 99% con un ancho de cabeza de 9.0 cm.

Ancho de cabeza	7.2	7.4	9.8	9.4	8.8	8.4
Peso	116	154	245	202	200	191

10-3 Más allá de lo básico

21. Intervalo de confianza para el valor medio predicho En el ejemplo 1 de esta sección se ilustró el procedimiento para encontrar un intervalo de predicción para un valor *individual* de y. Cuando se utiliza un valor específico x_0 para predecir la *media* de todos los valores de y, el intervalo de confianza es el siguiente:

$$\hat{y} - E < \bar{y} < \hat{y} + E$$

donde

$$E = t_{\alpha/2} \cdot s_e \sqrt{\frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

El valor crítico $t_{\alpha/2}$ se encuentra con $n - 2$ grados de libertad. Con base en los 23 pares de datos de chocolate/Nobel en la tabla 10-1 de la página 469, correspondientes al problema del capítulo, encuentre una estimación del intervalo de confianza del 95% para la tasa media de premios Nobel dado que el consumo de chocolate es de 10 kg *per cápita*.

10-4 Regresión múltiple

Concepto clave Hasta ahora en este capítulo hemos analizado la correlación lineal entre dos variables, pero en esta sección se presentan métodos para analizar una relación lineal con *más de dos variables*. Nos centramos en los siguientes dos elementos clave: (1) determinación de la ecuación de regresión múltiple y (2) uso del valor de R^2 ajustado y el valor P como medidas de qué tan bien se ajusta la ecuación de regresión múltiple a los datos muestrales. Debido a que los cálculos requeridos son tan complicados, su realización manual no resulta práctica y constituye una amenaza para la salud mental, por lo que esta sección enfatiza el uso y la interpretación de los resultados obtenidos con tecnología.

PARTE 1 Conceptos básicos de una ecuación de regresión múltiple

Como en las secciones anteriores de este capítulo, consideraremos solamente relaciones *lineales*. La siguiente ecuación de *regresión múltiple* describe relaciones lineales que involucran más de dos variables.

DEFINICIÓN

Una **ecuación de regresión múltiple** expresa una relación lineal entre una variable de respuesta y y dos o más variables predictoras ($x_1, x_2, \ldots, x_k$). La forma general de una ecuación de regresión múltiple obtenida a partir de datos muestrales es

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$$

El siguiente recuadro de elementos clave incluye los componentes principales de esta sección. En la notación, observe que los coeficientes $b_0, b_1, b_2, \ldots, b_k$ son estadísticos muestrales usados para estimar los parámetros poblacionales correspondientes $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$. Además, tenga en cuenta que la ecuación de regresión múltiple es una extensión natural del formato $\hat{y} = b_0 + b_1x_1$ utilizado en la sección 10-2 para ecuaciones de regresión con una sola variable independiente x_1. En la sección 10-2, habría sido razonable preguntar por qué no usábamos el formato más común y familiar de $y = mx + b$, y ahora podemos ver que el uso de $\hat{y} = b_0 + b_1x_1$ nos permite ampliar fácilmente ese formato con el fin de incluir variables predictoras adicionales.

ELEMENTOS CLAVE

Determinación de una ecuación de regresión múltiple

Objetivo

Utilizar datos muestrales relacionados de tres o más variables para encontrar una ecuación de regresión múltiple, que sea útil en la predicción de valores de la variable de respuesta y.

Notación

$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$ (ecuación de regresión múltiple encontrada a partir de datos *muestrales*)
$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \cdots + \beta_kx_k$ (ecuación de regresión múltiple para la *población* de datos)
$\hat{y} =$ valor predicho de y (calculado usando la ecuación de regresión múltiple)
$k =$ número de variables *predictoras* (también llamadas *variables independientes* o *variables x*)
$n =$ tamaño de muestra (número de valores para cualquiera de las variables)

Requisitos

Para cualquier conjunto específico de valores x, la ecuación de regresión se asocia con un error aleatorio, a menudo expresado por ε. Suponemos que dichos errores se distribuyen normalmente con una media de 0 y una desviación estándar de σ, y que los errores aleatorios son independientes.

Procedimiento para encontrar una ecuación de regresión múltiple

Los cálculos manuales no son prácticos, por lo que se debe usar tecnología. (Consulte las instrucciones del "Centro de tecnología" al final de esta sección).

En 1886, Francis Galton fue uno de los primeros en estudiar la genética utilizando los métodos de regresión que estamos considerando. Escribió el artículo "Regresión hacia la mediocridad en la estatura hereditaria", afirmando que las alturas de la descendencia retroceden o vuelven a la normalidad. Aunque continuamos usando el término "regresión", las aplicaciones actuales se extienden mucho más allá de aquellas que involucran estaturas.

EJEMPLO 1 Predicción del peso

El conjunto de datos 1 "Datos corporales" en el apéndice B incluye las estaturas (cm), las circunferencias de cintura (cm) y los pesos (kg) de una muestra de 153 hombres. Encuentre la ecuación de regresión múltiple en la cual la variable de respuesta (y) es el peso de un hombre y las variables predictoras son la altura (x_1) y la circunferencia de la cintura (x_2).

SOLUCIÓN

Mediante el uso de Statdisk y los datos muestrales del conjunto de datos 1, obtenemos los resultados que se muestran en la pantalla de la parte superior de la página siguiente. Los coeficientes b_0, b_1 y b_2 se usan en la ecuación de regresión múltiple:

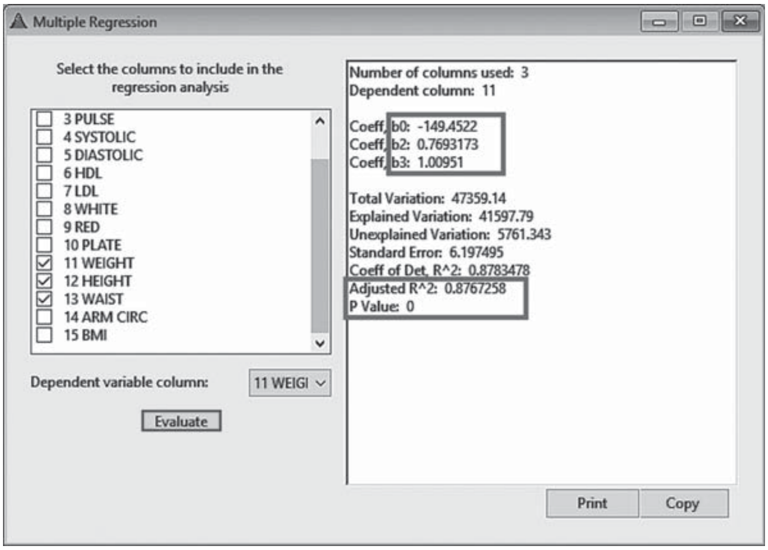
$$\hat{y} = -149 + 0.769x1 + 1.01x_2$$

o

$$\text{Peso} = -149 + 0.769\ \text{Altura} + 1.01\ \text{Cintura}$$

La ventaja obvia del segundo de estos formatos es que resulta más sencillo hacer un seguimiento de los papeles que desempeñan las variables.

Statdisk



SU TURNO Resuelva el ejercicio 13 "Predicción de nicotina en cigarrillos"

Si una ecuación de regresión múltiple se ajusta bien a los datos muestrales, puede usarse para realizar predicciones. Por ejemplo, si determinamos que la ecuación de regresión múltiple en el ejemplo 1 es adecuada para hacer predicciones, podemos usar la estatura y la circunferencia de la cintura de un hombre para predecir su peso. Pero, ¿cómo determinamos si la ecuación de regresión múltiple se ajusta bien a los datos muestrales? Dos herramientas muy útiles son el valor de R^2 ajustado y el valor P.

R^2 y R^2 ajustado

R^2 expresa el coeficiente de determinación múltiple, 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 tendría un valor cercano a 1. Un ajuste muy pobre daría como resultado un valor de R^2 cercano a 0. El valor de $R^2 = 0.878$ ("Coeff of Det, R^2") en la pantalla de Statdisk para el ejemplo 1 indica que el 87.8% de la variación en los pesos de los hombres se puede explicar por sus estaturas y circunferencias de cintura. Sin embargo, el coeficiente de determinación múltiple R^2 tiene un defecto grave: a medida que se incluyen más variables, R^2 aumenta. (R^2 puede conservarse igual, pero por lo general aumenta). Un valor mayor de R^2 se obtiene simplemente al incluir todas las variables disponibles, pero la mejor ecuación de regresión múltiple no necesariamente utiliza todas las variables disponibles. Debido a ese defecto, es mejor emplear el *coeficiente de determinación ajustado*, que es R^2 ajustado para el número de variables y el tamaño de la muestra.

DEFINICIÓN

El **coeficiente de determinación ajustado** es el coeficiente de determinación múltiple R^2 modificado para tener en cuenta el número de variables y el tamaño de la muestra. Se calcula utilizando la fórmula 10-8.

FÓRMULA 10-8

$$R^2 \text{ ajustado} = 1 - \frac{(n-1)}{[n-(k+1)]}(1-R^2)$$

donde

n = tamaño de la muestra
k = número de variables predictoras (x)

La pantalla de Statdisk anterior muestra el coeficiente de determinación ajustado como "Adjusted R^2" = 0.877 (redondeado). Si utilizamos la fórmula 10-8 con $R^2 = 0.8783478$, $n = 153$ y $k = 2$, obtenemos el R^2 ajustado = 0.877 (redondeado). Al comparar esta ecuación de regresión múltiple con otras, es mejor usar el R^2 ajustado de 0.877. Al considerar el R^2 ajustado de 0.877 por sí mismo, vemos que es bastante alto (cercano a 1), lo que sugiere que la ecuación de regresión se ajusta bien a los datos muestrales.

Valor *P*

El valor *P* es una medida de la significancia general de la ecuación de regresión múltiple. El valor *P* mostrado de 0 (redondeado) es pequeño, lo que indica que la ecuación de regresión múltiple tiene una buena significancia general y es utilizable para realizar predicciones. Podemos predecir el peso de los hombres según su estatura y la circunferencia de su cintura. Al igual que el R^2 ajustado, este valor *P* es una buena medida de qué tan bien se ajusta la ecuación a los datos muestrales. El valor *P* resulta de una prueba de hipótesis nula de que $\beta_1 = \beta_2 = 0$. El rechazo de $\beta_1 = \beta_2 = 0$ implica que al menos β_1 o al menos β_2 no es 0, lo que indica que esta ecuación de regresión es efectiva para predecir pesos de los hombres. Un análisis completo de los resultados podría incluir otros elementos importantes, como la significancia de los coeficientes individuales, pero deseamos mantener las cosas simples (¡!) al limitar el análisis a los tres componentes clave-ecuación de regresión múltiple, R^2 ajustado y valor *P*.

Determinación de la mejor ecuación de regresión múltiple

Al tratar de encontrar la mejor ecuación de regresión múltiple, no necesariamente debemos incluir todas las variables de predicción disponibles. La determinación de la mejor ecuación de regresión múltiple requiere un uso intensivo de juicio y sentido común, y no existe un procedimiento exacto y automático que pueda usarse para lograrlo. *La determinación de la mejor ecuación de regresión múltiple suele ser bastante difícil y está fuera del alcance de esta sección*, pero las siguientes directrices pueden ser útiles.

Directrices para determinar la mejor ecuación de regresión múltiple

1. ***Use consideraciones prácticas y de sentido común para incluir o excluir variables.*** Por ejemplo, cuando tratamos de encontrar una buena ecuación de regresión múltiple para predecir la estatura de una hija, debemos excluir la estatura del médico que recibió a la hija, porque obviamente esa altura resulta irrelevante.
2. ***Considere el valor P***. Seleccione una ecuación que tenga significancia general, según lo determina un valor *P* bajo desplegado en la pantalla de resultados de una tecnología.
3. ***Considere ecuaciones con valores altos del R^2 ajustado e intente incluir sólo algunas variables***. En vez de incluir casi todas las variables disponibles, intente incluir relativamente pocas variables de predicción (x). Utilice las siguientes pautas:
   - Seleccione una ecuación que tenga un valor del R^2 ajustado con esta propiedad: Si se incluye una variable de predicción adicional, el valor del R^2 ajustado no aumenta mucho.
   - Para un número particular de variables predictoras (x), seleccione la ecuación con el valor más grande del R^2 ajustado.
   - Al excluir variables predictoras (x) que no tienen mucho efecto en la variable de respuesta (y), podría ser útil encontrar el coeficiente de correlación lineal r para cada par de variables bajo consideración. Si dos valores de predicción tienen un coeficiente de correlación lineal muy alto (lo que se conoce como *multicolinealidad*), no hay necesidad de incluirlos a ambos, y se recomienda excluir la variable con el valor más bajo del R^2 ajustado.

El siguiente ejemplo ilustra que el sentido común y el *pensamiento crítico* son herramientas esenciales para el uso efectivo de los métodos estadísticos.

EJEMPLO 2 Predicción de la estatura con base en una huella de zapato

El conjunto de datos 2 "Pie y estatura" en el Apéndice B incluye la edad, la longitud del pie, la longitud de la huella del zapato, el tamaño del zapato y la estatura de 40 sujetos diferentes. Con base en estos datos muestrales, encuentre la mejor ecuación de regresión para predecir la estatura. ¿La "mejor" ecuación de regresión es una *buena* ecuación para predecir la estatura?

SOLUCIÓN

Utilizando la variable de respuesta de la estatura y las posibles variables predictoras de la edad, la longitud del pie, la longitud de la huella del zapato y el tamaño del zapato, hay 15 combinaciones posibles de variables predictoras. La tabla 10-5 incluye los resultados principales de cinco de esas combinaciones. La aplicación ciega e irreflexiva de los métodos de regresión sugeriría que la mejor ecuación de regresión usa las cuatro variables predictoras, porque esa combinación arroja el valor de R^2 ajustado más alto de 0.7585. Sin embargo, dado el objetivo de usar evidencias para estimar la estatura de un sospechoso, usamos el *pensamiento crítico* de la siguiente manera:

1. Elimine la variable de la edad, porque los delincuentes rara vez dejan evidencia que identifique su edad.
2. Elimine la variable del tamaño del zapato, porque en realidad es una forma redondeada de la longitud del pie.
3. Para las variables restantes de la longitud del pie y la longitud de la huella del zapato, use sólo la longitud del pie porque su valor de R^2 ajustado de 0.7014 es mayor que el de 0.6520 para la longitud del zapato, y no es mucho menor que el valor de 0.7484 para la longitud del pie y la longitud de la impresión del zapato juntos. En este caso, es mejor usar una variable de predicción que utilizar dos.
4. Aunque parece que el uso de la variable individual de la longitud del pie es lo más conveniente, también notamos que los delincuentes usualmente usan zapatos, por lo que es más probable que se encuentren longitudes de las huellas de los zapatos que de los pies.

TABLA 10-5 Seleccione los resultados clave del conjunto de datos 2 "Pie y estatura" en el apéndice B

Variables predictoras	R^2 ajustado	Valor *P*	
Edad	0.1772	0.004	← **No es lo mejor:** El R^2 ajustado es mucho menor que 0.7014 para la longitud del pie.
Longitud del pie	**0.7014**	**0.000**	← **Lo mejor:** R^2 ajustado alto y el valor *P* más bajo.
Longitud de la huella del zapato	0.6520	0.000	← **No es lo mejor:** R^2 ajustado es menor que 0.7014 para la longitud del pie.
Longitud del pie/longitud de la huella del zapato	0.7484	0.000	← **No es lo mejor:** El valor ajustado de R^2 no es mucho mayor que 0.7014 para la variable única de longitud del pie.
Edad/Longitud del pie/Longitud de la huella del zapato/tamaño del zapato	0.7585	0.000	← **No es lo mejor:** Hay otros casos que usan menos variables con R^2 ajustados que no son mucho más pequeños.

INTERPRETACIÓN

El uso ciego de los métodos de regresión sugiere que al estimar la estatura de un sujeto, deberíamos usar todos los datos disponibles, incluyendo las cuatro variables predictoras de edad, longitud del pie, longitud del zapato y tamaño del zapato, pero otras consideraciones prácticas sugieren que es mejor usar solamente la variable de predicción de la longitud del

continúa

pie. Entonces, la mejor ecuación de regresión parece ser la siguiente: Estatura = 64.1 + 4.29 (Longitud del pie). Sin embargo, dado que los delincuentes suelen usar zapatos, lo mejor es usar la variable predictora única de la longitud de la huella del zapato, por lo que la mejor ecuación de regresión práctica parece ser: Estatura = 80.9 + 3.22 (Longitud de la huella del zapato). El valor *P* de 0.000 sugiere que la ecuación de regresión arroja un buen modelo para estimar la estatura.

Debido a que los resultados de este ejemplo se basan en datos muestrales de sólo 40 sujetos, las estimaciones de estaturas no serán muy precisas. Como suele ser el caso, se pueden obtener mejores resultados utilizando muestras más grandes.

SU TURNO Resuelva el ejercicio 13 “Predicción de nicotina en cigarrillos”.

Pruebas de coeficientes de regresión Las directrices anteriores para encontrar la mejor ecuación de regresión múltiple se basan en el R^2 ajustado y el valor *P*, pero también podríamos realizar pruebas de hipótesis individuales basadas en los valores de los coeficientes de regresión. Considere el coeficiente de regresión de β_1. Una prueba de la hipótesis nula $\beta_1 = 0$ puede decirnos si la variable predictora correspondiente debe incluirse en la ecuación de regresión. El rechazo de $\beta_1 = 0$ sugiere que β_1 tiene un valor distinto de cero y, por lo tanto, es útil para predecir el valor de la variable de respuesta. Los procedimientos para tales pruebas se describen en el ejercicio 17.

Predicciones con regresión múltiple

Cuando analizamos la regresión en la sección 10-2, listamos (en la página 492) cuatro puntos a considerar cuando se usan ecuaciones de regresión para realizar predicciones. Estos mismos puntos se deben considerar cuando se emplean ecuaciones de regresión múltiple.

PARTE 2 Variables ficticias y regresión logística

Hasta ahora en este capítulo, todas las variables han representado datos continuos, pero muchas situaciones involucran una variable con sólo dos posibles valores cualitativos (como masculino/femenino, muerto/vivo o curado/no curado). Para obtener ecuaciones de regresión que incluyan tales variables, de alguna manera debemos asignar números a las dos categorías diferentes. Un procedimiento común es representar los dos valores posibles por 0 y 1, donde 0 representa un “fracaso” y 1 representa un “éxito”. En cuanto a los resultados de una enfermedad, a menudo se usa 1 para representar el evento de la enfermedad o la muerte, y 0 para representar la no ocurrencia del evento.

DEFINICIÓN

Una **variable ficticia** es una variable que tiene sólo los valores de 0 y 1, utilizados para representar las dos categorías diferentes de una variable cualitativa.

Una variable ficticia se llama en ocasiones una *variable dicotómica*. La palabra “ficticia” se usa porque la variable en realidad no tiene ningún valor cuantitativo, pero lo usamos como un sustituto para representar las diferentes categorías de la variable cualitativa.

Variable ficticia como variable predictora

Los procedimientos de análisis de regresión difieren drásticamente, dependiendo de si la variable ficticia es una variable predictora (*x*) o la variable de respuesta (*y*). Si incluimos una variable ficticia como otra variable *predictora* (*x*), podemos usar los mismos métodos de la parte 1 de esta sección, como se ilustra en el ejemplo 3.

EJEMPLO 3 Uso de una variable ficticia como variable predictora

La tabla 10-6 está adaptada del conjunto de datos 5 "Estaturas familiares" en el apéndice B y se muestra en un formato más conveniente para este ejemplo. Use la variable ficticia del sexo (codificada como 0 = mujer, 1 = hombre). Dado que un padre mide 69 pulgadas de alto y una madre mide 63 pulgadas de alto, encuentre la ecuación de regresión múltiple y úsela para predecir la estatura de (a) una hija y (b) un hijo.

TABLA 10-6 Estaturas (en pulgadas) de padres, madres y sus descendientes

Estatura del padre	Estatura de la madre	Estatura del descendiente	Sexo del descendiente (1 = hombre)
66.5	62.5	70.0	1
70.0	64.0	68.0	1
67.0	65.0	69.7	1
68.7	70.5	71.0	1
69.5	66.0	71.0	1
70.0	65.0	73.0	1
69.0	66.0	70.0	1
68.5	67.0	73.0	1
65.5	60.0	68.0	1
69.5	66.5	70.5	1
70.5	63.0	64.5	0
71.0	65.0	62.0	0
70.5	62.0	60.0	0
66.0	66.0	67.0	0
68.0	61.0	63.5	0
68.0	63.0	63.0	0
71.0	62.0	64.5	0
65.5	63.0	63.5	0
64.0	60.0	60.0	0
71.0	63.0	63.5	0

SOLUCIÓN

Mediante el uso de los métodos de regresión múltiple de la parte 1 de esta sección y el software de computadora, obtenemos la siguiente ecuación de regresión:

$$\text{Estatura del descendiente} = 36.5 - 0.0336\,(\text{Estatura del padre}) + 0.461\,(\text{Estatura de la madre}) + 6.14\,(\text{Sexo})$$

donde el valor de la variable ficticia de sexo es 0 para una hija o 1 para un hijo.

a. Para encontrar la estatura predicha de una *hija*; sustituimos 0 por la variable sexo, y también sustituimos 69 pulgadas por la estatura del padre y 63 pulgadas por la estatura de la madre. El resultado es una estatura predicha de 63.2 pulgadas para una hija.

b. Para encontrar la estatura pronosticada de un *hijo*, sustituimos 1 por la variable sexo, y también sustituimos 69 pulgadas por la estatura del padre y 63 pulgadas por la estatura de la madre. El resultado es una estatura pronosticada de 69.4 pulgadas para un hijo.

El coeficiente de 6.14 en la ecuación de regresión muestra que cuando se da la estatura de un padre y la estatura de una madre, un hijo tendrá una estatura pronosticada que es 6.14 pulgadas mayor que la estatura de una hija.

Congelar al pateador

Una estrategia común en el fútbol americano consiste en que, justo en el momento en que un pateador está a punto de intentar un gol de campo, el entrenador del equipo opuesto pide un tiempo fuera para "congelar" al pateador. La teoría sostiene que el pateador tiene tiempo para pensar, sentirse nervioso y perder la confianza. Sin embargo, ¿realmente funciona esta práctica? En el artículo "The Cold-Foot Effect", publicado por la revista *Chance*, Scott M. Berry presentó su análisis estadístico de los resultados de dos temporadas de la NFL. Utilizó un modelo de regresión logística con variables como el viento, las nubes, la lluvia, la temperatura, la presión de realizar la patada y si hubo o no una petición de tiempo fuera antes de esta acción. El autor escribió: "La conclusión a partir del modelo es que congelar al pateador sí funciona: parece que congelar al pateador reduce la probabilidad de una patada exitosa".

Regresión logística En el ejemplo 3, podríamos usar los mismos métodos de la parte 1 de esta sección porque la variable ficticia del sexo es una variable *predictora*. Sin embargo, si la variable ficticia es la variable de respuesta (y), no podemos utilizar los métodos de la parte 1 de esta sección, y debemos usar un método diferente conocido como regresión logística. Esta sección no incluye procedimientos detallados para usar la regresión logística, pero muchos libros están dedicados a ese tema. El ejemplo 4 ilustra brevemente el método de la regresión logística.

EJEMPLO 4 Regresión logística

Considere que un conjunto de datos muestrales consiste en las estaturas (cm) y las circunferencias del brazo (cm) de las mujeres y los hombres que se listan en el conjunto de datos 1 "Datos corporales" del apéndice B. La variable de *respuesta y* representa el género ($0 =$ mujer, $1 =$ hombre). Utilizando los valores de género de y y la lista combinada de las estaturas correspondientes y las circunferencias del brazo, se podría utilizar la regresión logística para obtener el siguiente modelo:

$$\ln\left(\frac{p}{1-p}\right) = -40.6 + 0.242(\text{EST}) + 0.000129(\text{CircBrazo})$$

En la expresión anterior, p es la probabilidad de un hombre, entonces $p = 1$ indica que el sujeto es definitivamente un hombre y $p = 0$ indica que el sujeto definitivamente no es un hombre (entonces el sujeto es una mujer). [Para resolver p, sustituya los valores de estatura y circunferencia del brazo para obtener un valor v, luego $p = e^v(1 + e^v)$]. Vea los siguientes dos conjuntos de resultados.

- Si utilizamos el modelo anterior y sustituimos para una altura de 183 cm (o 72.0 pulg) y una circunferencia del brazo de 33 cm (o 13.0 pulg), podemos resolver para obtener $p = 0.976$, lo que indica que esa persona tiene un 97.6% de probabilidad de ser un hombre.
- En contraste, una persona más pequeña con una altura de 150 cm (o 59.1 pulg) y una circunferencia del brazo de 20 cm (o 7.9 pulg) tiene una probabilidad de $p = 0.0134$, lo que indica que una persona tan pequeña es muy poco probable que sea un hombre.

CENTRO DE TECNOLOGÍA

Regresión múltiple

Acceda a complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Multiple Regression** en el menú desplegable.
3. Seleccione las columnas que se incluirán en el análisis de regresión. Para la columna de la *variable dependiente*, seleccione la columna que se utilizará para la variable *y*.
4. Haga clic en **Evaluate**.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable y elija **Regression–Fit Regression Model** en el submenú.
3. En *respuestas*, seleccione la columna que contiene los valores dependientes *y*. En *predictores continuos*, seleccione las columnas que contienen las variables que desea incluir como variables *x*.
4. Haga clic en **OK**. La ecuación de regresión se incluye en los resultados.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable, luego elija **Linear Multiple** en el submenú.
3. Seleccione las columnas que se usarán para la variable *x* y la columna que se usará para la variable *y*.
4. Haga clic en **Compute!**

CENTRO DE TECNOLOGÍA *continuación*

Regresión múltiple

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

Requiere el programa *A2MULREG* (disponible en www.pearsonenespañol.com/triola)

1. Los datos se deben ingresar como columnas en la *Matriz D*, donde la primera columna contiene los valores de la variable dependiente *y*: *Introducción manual de los datos*: Presione 2ND luego x^{-1} para ir al menú *MATRIX*, seleccione **EDIT** en el menú superior, elija **[D]** y presione ENTER. Ingrese el número de filas y columnas necesarias, presione ENTER y proceda a ingresar los valores muestrales.

   Uso de listas existentes: Las listas se pueden combinar y almacenar en la *Matriz D*. Presione 2ND después x^{-1} para ir al menú *MATRIX*, seleccione **MATH** en el menú superior y elija el elemento **List → matr**. Ingrese los nombres de lista (la primera lista debe contener valores para la variable dependiente *y*), seguidos del nombre de la matriz **[D]**, todos separados por ,.

   Importante: El nombre de la matriz se debe ingresar presionando 2ND luego x^{-1}, seleccionando **[D]**, y presionando ENTER. El siguiente es un resumen de los comandos utilizados para crear una matriz a partir de tres listas (L1, L2, L3): **Lista → matr(L1, L2, L3, [D])**.
2. Presione PRGM, seleccione *A2MULREG*, presione ENTER tres veces, seleccione **MULT REGRESSION** y presione ENTER.
3. Ingrese el número de variables independientes *x*, luego introduzca el número de columna de cada variable independiente *x*. Presione ENTER después de cada entrada.
4. Se mostrarán los resultados, incluido el valor *P* y el R^2 ajustado. Presione ENTER para ver resultados adicionales, incluyendo los valores utilizados en la ecuación de regresión múltiple.
5. Presione ENTER para seleccionar la opción **QUIT**.

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Modeling Data**.
2. Seleccione **Linear Regression** en el menú desplegable.
3. Ingrese el rango de celdas que contiene los datos de la *variable dependiente y* y los datos de la *variable explicativa x* (columnas múltiples). Marque la casilla **Quantitative** en la *variable explicativa x*. Si la primera fila de datos incluye una etiqueta, marque la casilla **Variable labels**.
4. Haga clic en la pestaña **Outputs** y asegúrese de que **Correlations** y el **Analysis of Variance** estén marcados.
5. Haga clic en **OK**, y la ecuación de la línea de regresión múltiple se mostrará en los resultados.

Excel (Complemento de análisis de datos)

1. Haga clic en la pestaña **Data** en la cinta de opciones y luego haga clic en la pestaña **Data Analysis**. Seleccione **Regression** en las *herramientas de análisis*.
2. Para el *rango de entrada y*, ingrese el rango de datos para la variable dependiente *y*. Para el *rango de entrada x*, ingrese el rango de datos para las variables independientes *x*. Los datos de las variables *x* deben estar ubicados en columnas adyacentes.
3. Marque la casilla **Labels** si la primera fila contiene una etiqueta.
4. Haga clic en **OK** para mostrar los resultados.

10-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Terminología Utilizando las longitudes (pulg), los tamaños del pecho (pulg) y los pesos (lb) de los osos del conjunto de datos 9 “Mediciones de osos” en el apéndice B, obtenemos la siguiente ecuación de regresión: Peso = −274 + 0.426 Longitud + 12.1 Tamaño del pecho. Identifique las variables de respuesta y predictoras.

2. Mejor ecuación de regresión múltiple Para la ecuación de regresión dada en el ejercicio 1, el valor *P* es 0.000 y el valor de R^2 ajustado es 0.925. Si tuviésemos que incluir una variable predictiva adicional del tamaño del cuello (pulg), el valor *P* se convertiría en 0.000 y R^2 ajustado sería 0.933. Dado que el valor de R^2 ajustado de 0.933 es mayor que 0.925, ¿es mejor usar la ecuación de regresión con las tres variables predictoras de longitud, tamaño del pecho y tamaño del cuello? Explique.

3. Coeficiente de determinación ajustado Para el ejercicio 2, ¿por qué es mejor usar valores de R^2 ajustado en vez de simplemente utilizar valores de R^2?

4. Interpretación de R^2 Para la ecuación de regresión múltiple dada en el ejercicio 1, obtenemos $R^2 = 0.928$. ¿Qué nos dice ese valor?

Interpretación de una pantalla de computadora. ***En los ejercicios 5 a 8, queremos considerar la correlación entre las estaturas de padres y madres y las estaturas de sus hijos. Consulte la pantalla de StatCrunch y responda las preguntas o identifique los elementos indicados. La pantalla se basa en el conjunto de datos 5 "Estaturas familiares" del apéndice B.***

Parameter estimates:

Parameter	Estimate	Std. Err.	Alternative	DF	T-Stat	P-value
Intercept	17.966577	6.4779134	≠ 0	131	2.7735131	0.0064
Father	0.50354896	0.067077219	≠ 0	131	7.507004	<0.0001
Mother	0.27714316	0.078318967	≠ 0	131	3.5386467	0.0006

Analysis of variance table for multiple regression model:

Source	DF	SS	MS	F-stat	P-value
Model	2	320.94662	160.47331	37.637221	<0.0001
Error	131	558.54293	4.2636865		
Total	133	879.48955			

Summary of fit:
Root MSE: 2.0648696
R-squared: 0.3649
R-squared (adjusted): 0.3552

5. Estatura de un hijo Identifique la ecuación de regresión múltiple que expresa la estatura de un hijo en términos de la estatura de su padre y su madre.

6. Estatura de un hijo Identifique lo siguiente:

a. El valor *P* correspondiente a la significancia general de la ecuación de regresión múltiple.

b. El valor del coeficiente de determinación múltiple R^2

c. El valor ajustado de R^2

7. Estatura de un hijo ¿Debería usarse la ecuación de regresión múltiple para predecir la estatura de un hijo según la estatura de su padre y su madre? ¿Por qué sí o por qué no?

8. Estatura de un hijo Un hijo nacerá de un padre que mide 70 pulgadas de alto y una madre que mide 60 pulgadas de alto. Use la ecuación de regresión múltiple para predecir la estatura del hijo. ¿Es probable que el resultado sea un buen valor predicho? ¿Por qué sí o por qué no?

Consumo de combustible en la ciudad: Determinación de la mejor ecuación de regresión múltiple. ***En los ejercicios 9 a 12, consulte la tabla adjunta, que se obtuvo usando los datos de 21 automóviles listados en el conjunto de datos 20 "Mediciones de automóviles" del apéndice B. La variable de respuesta (y) es CIUDAD (consumo de combustible en mi/gal). Las variables de predicción (x) son PL (peso en libras), CIL (cilindrada en litros) y CCC (consumo de combustible en carretera en mi/gal).***

Variables predictoras (x)	Valor *P*	R^2	R^2 ajustado	Ecuación de regresión
PL/CIL/CCC	0.000	0.943	0.933	CIUDAD = 6.86 − 0.00128 PL − 0.257 CIL + 0.652 CCC
PL/CIL	0.000	0.748	0.720	CIUDAD = 38.0 − 0.00395 PL − 1.29 CIL
PL/CCC	0.000	0.942	0.935	CIUDAD = 6.69 − 0.00159 PL + 0.670 CCC
DISP/CCC	0.000	0.935	0.928	CIUDAD = 1.87 − 0.625 CIL + 0.706 CCC
PL	0.000	0.712	0.697	CIUDAD = 41.8 − 0.00607 PL
CIL	0.000	0.659	0.641	CIUDAD = 29.0 − 2.98 CIL
CCC	0.000	0.924	0.920	CIUDAD = −3.15 + 0.819 CCC

9. Si sólo se usa una variable predictora (x) para predecir el consumo de combustible en la ciudad, ¿qué variable individual es la mejor? ¿Por qué?

10. Si se van a usar exactamente dos variables predictoras (x) para predecir el consumo de combustible en la ciudad, ¿cuáles dos variables se deben elegir? ¿Por qué?

11. ¿Qué ecuación de regresión es la mejor para predecir el consumo de combustible en la ciudad? ¿Por qué?

12. Un Honda Civic pesa 2740 lb, tiene una cilindrada de 1.8 L, y su consumo de combustible en carretera es de 36 mi/gal. ¿Cuál es el mejor valor predicho del consumo de combustible en la ciudad? ¿Es probable que ese valor predicho constituya una buena estimación? ¿Es probable que ese valor predicho sea muy preciso?

Conjuntos de datos del apéndice B. ***En los ejercicios 13 a 16, consulte el conjunto de datos indicado en el apéndice B y use tecnología para obtener los resultados.***

13. Predicción de nicotina en cigarrillos Consulte el conjunto de datos 13 “Contenido del cigarrillo” en el apéndice B y use las cantidades de alquitrán, nicotina y CO para cigarrillos que tienen 100 mm de largo, con filtro, sin mentol y que no son light (el último conjunto de mediciones). Encuentre la mejor ecuación de regresión para predecir la cantidad de nicotina en un cigarrillo. ¿Por qué es la mejor? ¿Es la mejor ecuación de regresión una buena ecuación para predecir el contenido de nicotina? ¿Por qué sí o por qué no?

14. Predicción de nicotina en cigarrillos Repita el ejercicio anterior utilizando los datos muestrales de los cigarrillos mentolados que se listan en el conjunto de datos 13 “Contenido del cigarrillo” en el apéndice B.

15. Predicción de la puntuación de IQ Consulte el conjunto de datos 8 “IQ y tamaño del cerebro” en el apéndice B y encuentre la mejor ecuación de regresión con la puntuación de IQ como la variable de respuesta (y). Use las variables de predicción del volumen cerebral y/o el peso corporal. ¿Por qué esta ecuación es la mejor? Con base en estos resultados, ¿podemos predecir el puntaje de IQ de alguien si conocemos el volumen de su cerebro y su peso corporal? De acuerdo con estos resultados ¿Parece que las personas con cerebros más grandes tienen puntuaciones de IQ más altas?

16. Puntuación completa de IQ Consulte el conjunto de datos 7 “IQ y plomo” en el apéndice B y encuentre la mejor ecuación de regresión con IQ COMPLETA (la puntuación de IQ completa) como variable de respuesta (y). Use las variables de predicción IQ VERB (puntuación de IQ verbal) y IQ REND (puntuación de IQ de rendimiento). ¿Por qué esta ecuación es la mejor? Con base en estos resultados, ¿podemos predecir la puntuación completa de IQ de alguien si conocemos sus puntuaciones de IQ verbal e IQ de rendimiento? ¿Es probable que esa predicción sea muy precisa?

10-4 Más allá de lo básico

17. Pruebas de hipótesis sobre coeficientes de regresión Si el coeficiente β_1 tiene un valor distinto de cero, entonces es útil para predecir el valor de la variable de respuesta. Si $\beta_1 = 0$, no es útil para predecir el valor de la variable de respuesta y puede eliminarse de la ecuación de regresión. Para probar la afirmación de que $\beta_1 = 0$ usa el dato estadístico de prueba $t = (b_1 - 0)/s_{b_1}$. Los valores críticos o valores P se pueden encontrar usando la distribución t con $n - (k + 1)$ grados de libertad, donde k es el número de variables predictoras (x) y n es el número de observaciones en la muestra. Por lo general, el error estándar s_{b_1} es proporcionado por el software. Por ejemplo, vea la pantalla de StatCrunch adjunta para el ejemplo 1, que muestra que $s_{b_1} = 0.071141412$ (se encuentra en la columna con el encabezado “Std. Err.” y la fila correspondiente a la primera variable de predicción de la estatura). Utilice los datos muestrales en el conjunto de datos 1 “Datos corporales” y la pantalla de StatCrunch para probar la afirmación de que $\beta_1 = 0$. También pruebe la afirmación de que $\beta_2 = 0$. ¿Qué implican los resultados sobre la ecuación de regresión?

Parameter estimates:

Parameter	Estimate	Std. Err.	Alternative	DF	T-Stat	P-value
Intercept	-149.45217	12.523494	≠ 0	150	-11.933743	<0.0001
Height	0.76931731	0.071141412	≠ 0	150	10.813917	<0.0001
Waist	1.0095102	0.033812346	≠ 0	150	29.856261	<0.0001

18. Intervalos de confianza para coeficientes de regresión Un intervalo de confianza para el coeficiente de regresión β_1 se expresa como

$$b_1 - E < \beta_1 < b_1 + E$$

donde

$$E = t_{\alpha/2} s_{b_1}$$

continúa

La puntuación crítica t se encuentra usando $n - (k + 1)$ grados de libertad, donde k, n y s_{b_1} se describen en el ejercicio 17. Usando los datos muestrales del ejemplo 1, $n = 153$ y $k = 2$, entonces gl $= 150$ y las puntuaciones t críticas son ± 1.976 para un nivel de confianza del 95%. Use los datos muestrales para el ejemplo 1, la pantalla de Statdisk del ejemplo 1 en la página 513 y de StatCrunch en el ejercicio 17 para elaborar estimaciones del intervalo de confianza del 95% para β_1 (el coeficiente de la variable que representa la estatura) y β_2 (el coeficiente de la variable que representa la circunferencia de la cintura). ¿El intervalo de confianza incluye 0, lo que sugiere que la variable debe eliminarse de la ecuación de regresión?

19. Variable ficticia Consulte el conjunto de datos 9 "Mediciones de osos" en el apéndice B y use el sexo, la edad y el peso de los osos. Para el sexo, considere que 0 representa una hembra y 1 representa un macho. La variable de respuesta (y) representa el peso; use la variable de la edad y la variable ficticia del sexo para encontrar la ecuación de regresión múltiple. Utilice la ecuación para hallar el peso predicho de un oso con las características que se detallan a continuación. ¿El sexo parece tener mucho efecto sobre el peso de un oso?

a. Oso femenino de 20 años de edad

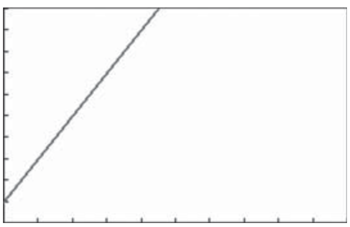
b. Oso masculino de 20 años de edad

10-5 Regresión no lineal

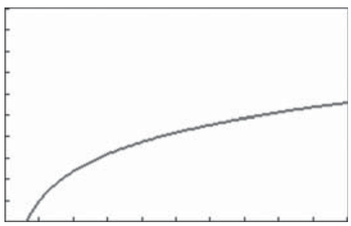
Concepto clave Las secciones anteriores de este capítulo tratan sólo con relaciones *lineales*, pero no todas en el mundo son lineales. Esta sección es una breve introducción a los métodos para encontrar algunas funciones *no lineales* que se ajustan a los datos muestrales. Nos enfocamos en el uso de la tecnología porque los cálculos requeridos son bastante complejos.

A continuación se muestran cinco modelos genéricos básicos considerados en esta sección. Cada uno de los modelos se proporciona con una fórmula genérica junto con un ejemplo de una función específica y su gráfica.

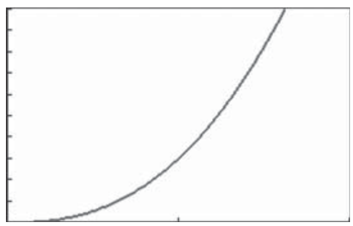
Lineal: $y = a + bx$
Ejemplo: $y = 1 + 2x$



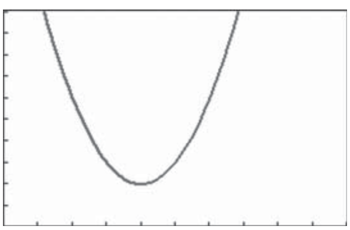
Logarítmica: $y = a + b \ln x$
Ejemplo: $y = 1 + 2 \ln x$



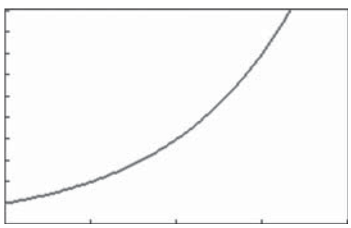
De potencia: $y = ax^b$
Ejemplo: $y = 3x^{2.5}$



Cuadrática: $y = ax^2 + bx + c$
Ejemplo: $y = x^2 - 8x + 18$



Exponencial: $y = ab^x$
Ejemplo: $y = 2^x$



Las siguientes tres reglas son básicas para identificar un buen modelo matemático:

1. ***Busque un patrón en la gráfica.*** Elabore una gráfica, compárela con las que se muestran aquí e identifique el modelo más parecido.
2. ***Compare los valores de R^2.*** Para cada modelo considerado, utilice la tecnología para encontrar el valor del coeficiente de determinación R^2. Elija las funciones que den como resultado valores más grandes de R^2, porque dichos valores corresponden a funciones que se ajustan mejor a los datos muestrales observados.

- No le dé mucha importancia a las pequeñas diferencias, como la diferencia entre $R^2 = 0.984$ y $R^2 = 0.989$.
- A diferencia de la sección 10-4, no necesitamos usar valores del R^2 ajustado. Debido a que todos los ejemplos de esta sección implican una sola variable de predicción, tiene sentido comparar los valores de R^2.
- Además de R^2, otra medida utilizada para evaluar la calidad de un modelo es la suma de los cuadrados de los residuos. Vea el ejercicio 18 "Criterio de la suma de los cuadrados".

3. ***Piense.*** Use el sentido común. No utilice un modelo que conduzca a valores predichos que no son realistas. Emplee el modelo para calcular valores futuros, valores pasados y valores de datos faltantes, y después determine si los resultados son realistas y tienen sentido. No se aleje demasiado del alcance de los datos muestrales disponibles.

EJEMPLO 1 Determinación del mejor modelo de población

La tabla 10-7 lista la población de Estados Unidos para diferentes intervalos de 20 años. Encuentre un modelo matemático para el tamaño de la población, luego pronostique el tamaño de la población de Estados Unidos en el año 2040.

TABLA 10-7 Población (en millones) de Estados Unidos

Año	1800	1820	1840	1860	1880	1900	1920	1940	1960	1980	2000
Año codificado	1	2	3	4	5	6	7	8	9	10	11
Población	5	10	17	31	50	76	106	132	179	227	281

SOLUCIÓN

Primero, "codificamos" los valores de los años usando 1, 2, 3, ..., en lugar de 1800, 1820, 1840, ... El motivo de esta codificación es usar valores de x que sean mucho más pequeños y con menos probabilidad de causar dificultades de cómputo.

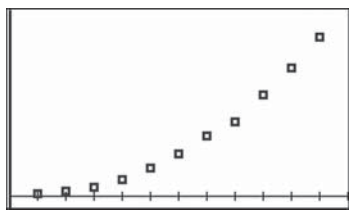
1. *Busque un patrón en la gráfica*. Examine el patrón de valores de datos en la pantalla TI-83/84 Plus (que se muestra en el margen), y compare ese patrón con los modelos genéricos que presentados anteriormente en esta sección. El patrón de esos puntos claramente no es una línea recta, por lo que descartamos un modelo lineal. Los buenos candidatos para el modelo parecen ser las funciones cuadrática, exponencial y de potencia.
2. *Encuentre y compare los valores de R^2*. La pantalla TI-83/84 para el modelo cuadrático se muestra en el margen. Para el modelo cuadrático, $R^2 = 0.9992$ (redondeado), que es bastante alto. La tabla 10-8 incluye este resultado junto con los valores para otros dos modelos potenciales. Al comparar los valores del coeficiente R^2, parece que el modelo cuadrático es el mejor porque tiene el valor más alto de 0.9992. 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 con $x = 1$ representando 1800, $x = 2$ representando 1820, etcétera) y la población y (en millones).

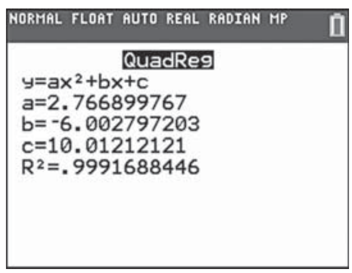
   Con base en su valor de $R^2 = 0.9992$, el modelo cuadrático parece ser el mejor, pero los otros valores de R^2 también son bastante altos. Nuestro conocimiento general del crecimiento de la población podría sugerir que el modelo exponencial es el más apropiado. (Con una tasa de natalidad constante y sin factores limitantes, la población crecerá exponencialmente).

TABLA 10-8 Modelos para los datos de población

Modelo	R^2	Ecuación
Cuadrático	**0.9992**	$y = 2.77x^2 - 6.00x + 10.01$
Exponencial	0.9631	$y = 5.24(1.48^x)$
De potencia	0.9764	$y = 3.35x^{1.77}$

TI-83/84 Plus





Atajo para un ensayo clínico

¿Qué haría usted si estuviera sometiendo a prueba un tratamiento y, antes de concluir su estudio, se diera cuenta de que es claramente eficaz? Debe acortar el estudio e informar a todos los participantes acerca de la eficacia del tratamiento. Esto sucedió cuando se sometió a prueba la hidroxiurea como tratamiento para la anemia falciforme. El estudio estaba programado para durar cerca de 40 meses, pero la eficacia del tratamiento se hizo evidente y el estudio se detuvo después de 36 meses (Vea "Trial Halted as Sickle Cell Treatment Proves Itself", de Charles Marwick, *Journal of the American Medical Association*, vol. 273, núm. 8).

Para predecir la población de EE.UU. en el año 2040, primero tenga en cuenta que el año 2040 está codificado como $x = 13$ (consulte la tabla 10-7). Al sustituir $x = 13$ en el modelo cuadrático de $y = 2.77x^2 - 6.00x + 10.01$ se obtiene $y = 400$, lo que indica que la población de EE.UU. se estima en 400 millones para el año 2040.

3. *Piense*. El resultado previsto de 400 millones en 2040 parece razonable. (Al momento de escribir estas líneas, las últimas cifras de la Oficina del Censo de EE.UU. utilizaban métodos mucho más sofisticados para proyectar que la población de EE.UU. en 2040 será de 380 millones). Sin embargo, existe un riesgo considerable al hacer estimaciones para tiempos que van más allá del alcance de los datos disponibles. Por ejemplo, el modelo cuadrático sugiere que en 1492, la población de EE.UU. era de 671 millones, un resultado que los estadísticos considerarían *ridículo*. El modelo cuadrático parece ser bueno para los datos disponibles (1800-2000), pero hay otros modelos que podrían ser mejores si es necesario hacer estimaciones futuras de la población.

SU TURNO Resuelva el ejercicio 5 “Dejar caer la pelota”.

EJEMPLO 2 Interpretación de R^2

En el ejemplo 1, obtuvimos el valor de $R^2 = 0.9992$ para el modelo cuadrático. Interprete ese valor en lo que se refiere a la variable de predictora del año y la variable de respuesta del tamaño de la población.

SOLUCIÓN

En el contexto de los datos de año/población de la tabla 10-7, el valor de $R^2 = 0.9992$ puede interpretarse de la siguiente manera: el 99.92% de la variación en el tamaño de la población puede explicarse por la ecuación de regresión cuadrática (dada en el ejemplo 1) que relaciona el año y el tamaño de la población.

SU TURNO Resuelva el ejercicio 3 “Interpretación de R^2”.

En “Modeling the U.S. Population” (*AMATYC Review*, vol. 20, núm. 2), Sheldon Gordon establece el siguiente punto importante que se aplica a todos los usos de métodos estadísticos:

“La mejor opción (de un modelo) depende del conjunto de datos que se analizan y requiere un ejercicio de juicio, no sólo cálculos”.

CENTRO DE TECNOLOGÍA

Regresión no lineal

Acceda a complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Statdisk puede encontrar el modelo cuadrático usando la función de regresión múltiple.
El siguiente procedimiento presenta datos de la tabla 10-7.

1. Ingrese los datos poblacionales de la tabla 10-7 en la columna 1 del Editor de muestras.
2. Ingrese los valores del año codificado correspondiente (1, 2, 3 ..., 11) en la columna 2.
3. Ingrese los cuadrados de los valores del año codificado (1, 4, 9, ..., 121) en la columna 3.
4. Haga clic en **Analysis** en el menú superior.
5. Seleccione **Multiple Regression** en el menú desplegable.
6. Seleccione las columnas 1, 2, 3 y elija la columna 1 como la variable dependiente.
7. Haga clic en **Evaluate**. Statdisk proporciona los coeficientes para la ecuación de regresión y el valor de R^2.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable y elija **Fitted Line Plot** en el submenú.
3. Seleccione las columnas que se utilizarán para la *variable de respuesta y* y la *variable predictora x*.
4. Elija el tipo deseado de modelo de regresión *lineal*, *cuadrático* o *cúbico*.
5. Haga clic en **OK**.

StatCrunch

StatCrunch puede encontrar el modelo para una función cuadrática (polinomio de orden 2).

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Regression** en el menú desplegable, luego elija **Polynomial** en el submenú.
3. Seleccione las columnas que se utilizarán para la variable *x* y la variable *y*.
4. Para una función cuadrática, seleccione **2** en *Poly. order*.
5. Haga clic en **Compute!**

CENTRO DE TECNOLOGÍA *continuación*

Regresión no lineal

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

1. Encienda la característica de *diagnóstico estadístico* presionando el botón MODE, desplazándose hasta **Stat Diagnostics**, resaltando **ON** y presionando ENTER.
2. Presione STAT, luego seleccione CALC en el menú superior.
3. Seleccione el modelo deseado de la lista de opciones disponibles, luego presione ENTER.
4. Ingrese los nombres de la lista de datos deseados para las variables *x* y *y* (para las calculadoras TI-83, ingrese los nombres de la lista separados por ,).
5. Seleccione **Calculate** y presione ENTER.

SUGERENCIA: Para las calculadoras TI-83 Plus, active el diagnóstico estadístico presionando 2ND 0 *para el menú de Catálogo. Desplácese hasta* ***DiagnosticON*** *y presione* ENTER *dos veces.*

Excel

El complemento XLSTAT no se puede usar para crear modelos de regresión no lineales, por lo que se debe utilizar Excel.

1. Seleccione el rango de celdas que contiene los datos pareados.
2. Haga clic en la pestaña **Insert** en la cinta de opciones y seleccione **Scatter** en la sección de *gráficos*.
3. Haga clic derecho en cualquier punto de datos en el diagrama de dispersión y seleccione **Add Trendline...**
4. Seleccione el modelo deseado y marque **Display Equation on chart** y **Display R-squared value on chart**.

10-5 Habilidades y conceptos básicos

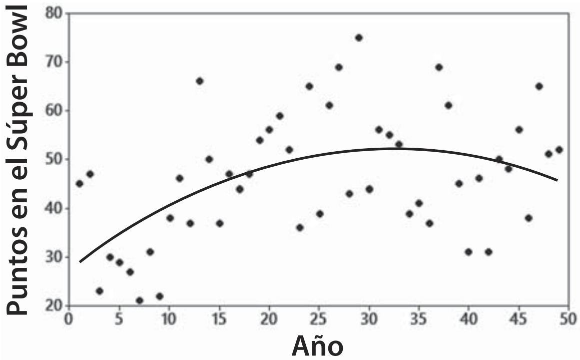
Conocimiento estadístico y pensamiento crítico

1. Identificación de un modelo y R^2 Se recogen diferentes muestras, y cada muestra se compone de las puntuaciones de IQ de 25 estudiantes de estadística. Considere que x representa la desviación estándar de las 25 puntuaciones de IQ en una muestra, y que y representa la varianza de los 25 puntajes de IQ en una muestra. ¿Qué fórmula describe mejor la relación entre x y y? ¿Cuál de los cinco modelos describe esta relación? ¿Cuál debería ser el valor de R^2?

2. Súper Bowl y R^2 Considere que x representa años codificados como 1, 2, 3, ..., para los años que comienzan en 1980, y y representa el número de puntos anotados en cada Súper Bowl desde 1980. Usando los datos de 1980 hasta el último Súper Bowl en el momento de escribir esto, obtenemos los siguientes valores de R^2 para los diferentes modelos: lineal: 0.147; cuadrático: 0.255; logarítmico: 0.176; exponencial: 0.175; de potencia: 0.203. Con base en estos resultados, ¿qué modelo es el mejor? ¿Es ese mejor modelo un buen modelo? ¿Qué sugieren los resultados para predecir el número de puntos anotados en un futuro juego del Súper Bowl?

3. Interpretación de R^2 En el ejercicio 2, el modelo cuadrático da como resultado $R^2 = 0.255$. Identifique el porcentaje de variación en los puntos del Súper Bowl que puede ser explicado por el modelo cuadrático que relaciona la variable del año y la variable de los puntos anotados. (*Sugerencia:* Vea el ejemplo 2). ¿Qué sugiere el resultado sobre la utilidad del modelo cuadrático?

4. Interpretación de una gráfica La gráfica adjunta presenta la cantidad de puntos anotados en cada Súper Bowl hasta el último de ellos en el momento de escribir estas líneas. La gráfica de la ecuación cuadrática que mejor se ajusta a los datos también se muestra en negro. ¿Qué característica de la gráfica justifica el valor de $R^2 = 0.255$ para el modelo cuadrático?



Determinación del mejor modelo. ***En los ejercicios 5 a 16, elabore un diagrama de dispersión e identifique el modelo matemático que mejor se ajusta a los datos dados. Suponga que el modelo se debe usar sólo para el alcance de los datos dados, y considere solamente los modelos lineales, cuadráticos, logarítmicos, exponenciales y de potencia.***

5. Dejar caer la pelota La tabla lista la distancia d (en metros) por encima del suelo para un objeto soltado en un vacío de baja gravedad desde una altura de 300 m. El tiempo t (seg) es el tiempo después de soltar el objeto.

t (segundos)	1	2	3	4	5
d (metros)	295.1	280.5	256.1	222.0	178.1

6. Tierra abonada La compañía Tierra Abonada Dirt Guy en Durham, CT, vende (lo adivinó) tierra abonada. Se vende por “yarda”, que en realidad es una yarda cúbica, y la variable x es la longitud (en yardas) de cada lado de un cubo de tierra abonada.

x (yd)	1	2	3	4	6
Costo (dólares)	25	200	675	1600	5400

7. Ganancias de CD La tabla lista el valor y (en dólares) de $1000 depositados en un certificado de depósito en el Banco de Nueva York (basado en tasas actualmente vigentes).

Año	1	2	3	4	5
Valor	1012.20	1024.55	1037.05	1049.70	1062.51

8. Intensidad del sonido La tabla lista intensidades del sonido como múltiplos de un sonido de referencia básico. Se usa una escala similar a la escala de decibeles para medir la intensidad del sonido.

Intensidad de sonido	316	500	750	2000	5000
Valor en la escala	25.0	27.0	28.75	33.0	37.0

9. Cultivo bacteriano En un experimento cuidadosamente controlado, las bacterias pueden crecer durante una semana. La cantidad de bacterias se registra al final de cada día con los siguientes resultados: 20, 40, 80, 160, 320, 640, 1280.

10. Muertes por choques de vehículos automotores A continuación se lista el número de muertes en Estados Unidos resultantes de accidentes automovilísticos. Utilice el mejor modelo para encontrar el número proyectado de tales muertes para el año 2025.

Año	1975	1980	1985	1990	1995	2000	2005	2010
Muertes	44,525	51,091	43,825	44,599	41,817	41,945	43,443	32,708

11. Escala de Richter La tabla lista diferentes cantidades (toneladas métricas) del explosivo TNT y el valor correspondiente medido en la escala de Richter resultante de las explosiones del TNT.

TNT	2	10	15	50	100	500
Escala de Richter	3.4	3.9	4.0	4.4	4.6	5.0

12. Ley de Benford De acuerdo con la ley de Benford, una variedad de conjuntos de datos diferentes incluye números con los primeros dígitos que se listan en la siguiente tabla, los cuales ocurren con las proporciones indicadas.

Primer dígito	1	2	3	4	5	6	7	8	9
Proporción	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

13. Mercado de valores En la parte superior de la página siguiente, ordenados por filas, se encuentran los valores anuales más altos del promedio industrial Dow Jones para cada año que comienza en 1990. Encuentre el mejor modelo y luego pronostique el valor para el año 2014 (el último año listado). ¿El valor predicho está cerca del valor real de 18,054?

3000	3169	3413	3794	3978	5216	6561	8259	9374	11,568
11,401	11,350	10,635	10,454	10,855	10,941	12,464	14,198	13,338	10,606
11,625	12,929	13,589	16,577	18,054					

14. Números de manchas solares A continuación se listan, ordenados por fila, los números anuales de manchas solares que comienzan en 1980. ¿Es el mejor modelo un buen modelo? Examine cuidadosamente el diagrama de dispersión e identifique el patrón de los puntos. ¿Cuál de los modelos se ajusta a ese patrón?

154.6	140.5	115.9	66.6	45.9	17.9	13.4	29.2	100.2	157.6
142.6	145.7	94.3	54.6	29.9	17.5	8.6	21.5	64.3	93.3
119.6	123.3	123.3	65.9	40.4	29.8	15.2	7.5	2.9	3.1
16.5	55.7	57.6	64.7	79.3					

15. Dióxido de carbono A continuación se listan las cantidades medias de concentraciones de dióxido de carbono (partes por millón) en nuestra atmósfera para cada década, comenzando en la década de 1880. Encuentre el mejor modelo y luego pronostique el valor para la década de 2090 a 2099. Comente el resultado.

292 294 297 300 304 307 309 314 320 331 345 360 377

16. Calentamiento global A continuación se listan las temperaturas medias anuales (°C) en la Tierra para cada década, comenzando en la década de 1880. Encuentre el mejor modelo y luego pronostique el valor para la década de 2090 a 2099. Comente el resultado.

13.819	13.692	13.741	13.788	13.906	14.016	14.052
13.983	13.938	14.014	14.264	14.396	14.636	

10-5 Más allá de lo básico

17. Ley de Moore En 1965, el cofundador de Intel, Gordon Moore, inició lo que desde entonces se conoce como la ley de Moore: el número de transistores por pulgada cuadrada en circuitos integrados se duplicará aproximadamente cada 18 meses. En la tabla siguiente, la primera fila lista diferentes años y la segunda fila enumera la cantidad de transistores (en miles) para diferentes años.

1971	1974	1978	1982	1985	1989	1993	1997	2000	2002	2003	2007	2011
2.3	5	29	120	275	1180	3100	7500	42,000	220,000	410,000	789,000	2,600,000

a. Asumiendo que la ley de Moore es correcta y que los transistores por pulgada cuadrada se duplican cada 18 meses, ¿qué modelo matemático describe mejor esta ley: lineal, cuadrático, logarítmico, exponencial, de potencia? ¿Qué función específica describe la ley de Moore?

b. ¿Qué modelo matemático se ajusta mejor a los datos muestrales listados?

c. Compare los resultados de los incisos (a) y (b). ¿La ley de Moore parece funcionar razonablemente bien?

18. Criterio de la suma de cuadrados Además del valor de R^2, otra medida utilizada para evaluar la calidad de un modelo es la *suma de los cuadrados de los residuos*. Recuerde de la sección 10-2 que un residuo es la diferencia entre un valor observado de y y el valor de y predicho a partir del modelo, que se expresa como $\hat{y}$. Los mejores modelos tienen sumas de cuadrados más pequeñas. Consulte los datos de la tabla 10-7 en la página 523.

a. Encuentre $\Sigma(y - \hat{y})^2$, la suma de los cuadrados de los residuos resultantes del modelo lineal.

b. Encuentre la suma de los cuadrados de los residuos resultantes del modelo cuadrático.

c. Verifique que, de acuerdo con el criterio de la suma de cuadrados, el modelo cuadrático es mejor que el modelo lineal.

Examen rápido del capítulo

Los siguientes ejercicios se basan en los datos muestrales dados que consisten en el número de estudiantes inscritos (en miles) y el número de robos en grandes universidades seleccionadas al azar en un año reciente (según datos del* New York Times*).

Inscripción (miles)	53	28	27	36	42
Robos	86	57	32	131	157

1. Conclusión El coeficiente de correlación lineal r es 0.499, el valor P es 0.393, y los valores críticos para un nivel de significancia de 0.05 son ± 0.878. ¿Qué se puede concluir de esto?

2. Variables conmutadas ¿Qué cambia si las dos variables de inscripción y robos se conmutan: el valor de $r = 0.499$, el valor P de 0.393, los valores críticos de ± 0.878?

3. Cambio en la escala El ejercicio 1 indicó que r es 0.499. ¿Cambia ese valor si se usan los valores de inscripción reales de 53,000, 28,000, 27,000, 36,000 y 42,000 en lugar de 53, 28, 27, 36 y 42?

4. Valores de *r* Si usted calcula un valor del coeficiente de correlación lineal de 1.500, ¿qué debería concluir?

5. Predicciones Los datos muestrales dan como resultado un coeficiente de correlación lineal de $r = 0.499$ y la ecuación de regresión $\hat{y} = 3.83 + 2.39x$. ¿Cuál es el mejor número predicho de robos, dada una inscripción de 50 (miles), y cómo se encontró?

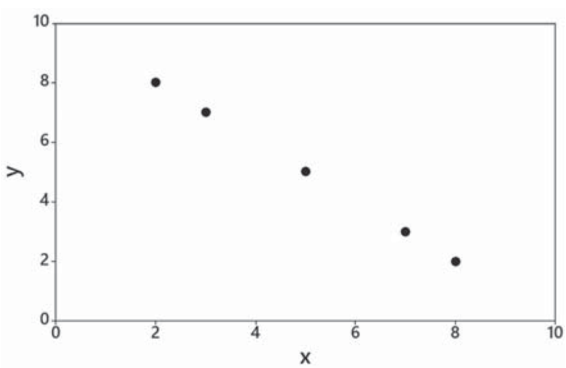
6. Predicciones Repita el ejercicio anterior, suponiendo que el coeficiente de correlación lineal es $r = 0.997$.

7. Variación explicable Dado que el coeficiente de correlación lineal r es 0.499, ¿cuál es la proporción de la variación en el número de robos que se explica por la relación lineal entre la inscripción y el número de robos?

8. Correlación lineal y relaciones Verdadero o falso: si no hay una correlación lineal entre la inscripción y el número de robos, entonces esas dos variables no están relacionadas de ninguna manera.

9. Causalidad Verdadero o falso: si los datos muestrales nos llevan a concluir que hay evidencia suficiente para respaldar la afirmación de una correlación lineal entre la inscripción y el número de robos, también podríamos concluir que un número más alto de inscripciones causa un aumento en el número de robos.

10. Interpretación de un diagrama de dispersión Si los datos muestrales resultaran en el diagrama de dispersión que se muestra aquí, ¿cuál es el valor del coeficiente de correlación lineal r?



Ejercicios de repaso

1. Alquitrán y nicotina en el cigarrillo La tabla siguiente lista las cantidades medidas (mg) de alquitrán, monóxido de carbono (CO) y nicotina en cigarrillos extra-largos de diferentes marcas (del conjunto de datos 13 "Contenido del cigarrillo" en el apéndice B).

a. ¿Hay suficiente evidencia para respaldar una afirmación de una correlación lineal entre el alquitrán y la nicotina?

b. ¿Qué porcentaje de la variación en la nicotina se puede explicar por la correlación lineal entre la nicotina y el alquitrán?

c. Considere que y representa la cantidad de nicotina y que x representa la cantidad de alquitrán, identifique la ecuación de regresión.

d. El cigarrillo de la marca Raleigh no está incluido en la tabla, y contiene 23 mg de alquitrán. ¿Cuál es la mejor cantidad predicha de nicotina? ¿Cómo se compara la cantidad predicha con la cantidad real de 1.3 mg de nicotina?

Alquitrán	25	27	20	24	20	20	21	24
CO	18	16	16	16	16	16	14	17
Nicotina	1.5	1.7	1.1	1.6	1.1	1.0	1.2	1.4

2. Nicotina y monóxido de carbono en cigarrillos Consulte la tabla de datos proporcionada en el ejercicio 1 y use las cantidades de nicotina y monóxido de carbono (CO).

a. Elabore un diagrama de dispersión usando la nicotina para la escala x, o eje horizontal. ¿Qué sugiere el diagrama de dispersión sobre una correlación lineal entre las cantidades de nicotina y de monóxido de carbono?

b. Encuentre el valor del coeficiente de correlación lineal y determine si hay suficiente evidencia que respalde una afirmación de una correlación lineal entre las cantidades de nicotina y de monóxido de carbono.

c. Considere que y representa la cantidad de monóxido de carbono y que x representa la cantidad de nicotina, encuentre la ecuación de regresión.

d. El cigarrillo extra-largo de Raleigh no está incluido en la tabla, y tiene 1.3 mg de nicotina. ¿Cuál es la mejor cantidad predicha de monóxido de carbono? ¿Cómo se compara la cantidad predicha con la cantidad real de 1.5 mg de monóxido de carbono?

3. Tiempo y movimiento En un experimento de física en Doane College, un balón de fútbol fue lanzado hacia arriba desde la cama de un camión en movimiento. La siguiente tabla muestra el tiempo (seg) que ha transcurrido desde el lanzamiento y la altura (m) del balón de fútbol. ¿Qué se puede concluir sobre la relación entre el tiempo y la altura? ¿Qué error terrible podría cometerse fácilmente si el análisis se realizase sin un diagrama de dispersión?

Tiempo (seg)	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8
Altura (m)	0.0	1.7	3.1	3.9	4.5	4.7	4.6	4.1	3.3	2.1

4. Regresión múltiple con cigarrillos Use los datos muestrales dados en el ejercicio de repaso 1 "Alquitrán y nicotina en el cigarrillo".

a. Encuentre la ecuación de regresión múltiple con la variable de respuesta (y) de la cantidad de nicotina y las variables predictoras (x) de las cantidades de alquitrán y monóxido de carbono.

b. Identifique el valor del coeficiente de determinación múltiple R^2, el R^2 ajustado y el valor P que representan el significado general de la ecuación de regresión múltiple.

c. Use un nivel de significancia de 0.05 y determine si la ecuación de regresión puede usarse para predecir la cantidad de nicotina dadas las cantidades de alquitrán y monóxido de carbono.

d. El cigarrillo extra-largo de Raleigh no está incluido en la tabla, y tiene 23 mg de alquitrán y 15 mg de monóxido de carbono. ¿Cuál es la mejor cantidad predicha de nicotina? ¿Cómo se compara la cantidad predicha con la cantidad real de 1.3 mg de nicotina?

Ejercicios de repaso acumulado

Acciones y manchas solares. ***A continuación se listan los valores máximos anuales del índice industrial Dow Jones (DJIA, por sus siglas en inglés) y el promedio anual de manchas solares durante ocho años recientes. Use los datos para los ejercicios 1 a 5. Un número de manchas solares es una medida de la cantidad de manchas o grupos de ellas en la superficie del sol. El DJIA es un índice comúnmente utilizado y consiste en una media ponderada que se calcula a partir de diferentes valores de acciones.***

DJIA	14,198	13,338	10,606	11,625	12,929	13,589	16,577	18,054
Número de manchas solares	7.5	2.9	3.1	16.5	55.7	57.6	64.7	79.3

1. Análisis de datos Use sólo los números de manchas solares para lo siguiente.

a. Encuentre la media, la mediana, el rango, la desviación estándar y la varianza.

b. ¿Los números de manchas solares son datos categóricos o datos cuantitativos?

c. ¿Cuál es el nivel de medición de los datos? (nominal, ordinal, de intervalo, de razón)

2. Correlación Use un nivel de significancia de 0.05 para probar una correlación lineal entre los valores del DJIA y los números de manchas solares. ¿Es el resultado que usted esperaba? ¿Debería alguien considerar invertir en acciones con base en el número de manchas solares?

3. Puntuaciones *z* Usando sólo los números de manchas solares, identifique el número más alto y conviértalo en una puntuación z. En el contexto de estos datos muestrales, ¿ese valor más alto es "significativamente alto"? ¿Por qué sí o por qué no?

4. Prueba de hipótesis El número medio de manchas solares durante los últimos tres siglos es de 49.7. Use un nivel de significancia de 0.05 para probar la afirmación de que los ocho números de manchas solares listados son de una población con una media igual a 49.7.

5. Intervalo de confianza Elabore una estimación del intervalo de confianza del 95% para el número medio de manchas solares. Escriba una breve declaración que interprete el intervalo de confianza.

6. Teléfonos celulares y conducción En la ciudad natal del autor, Madison, CT, hubo 2733 paradas de tráfico policial en un año reciente, y el 7% de ellas fueron atribuibles al uso indebido de teléfonos celulares. Use un nivel de significancia de 0.05 para probar la afirmación de que la muestra proviene de una población en la que menos del 10% de las paradas de tráfico policial son atribuibles al uso indebido del teléfono celular.

7. Edades de los cinéfilos La siguiente tabla muestra la distribución de las edades de los cinéfilos (según datos de la Asociación de Películas de América). Use los datos para estimar la media, la desviación estándar y la varianza de las edades de los cinéfilos. *Sugerencia:* Para la categoría abierta de "60 años o más", suponga que la categoría es en realidad de 60 a 80 años.

Años	2–11	12–17	18–24	25–39	40–49	50–59	60 años o más
Porcentaje	7	15	19	19	15	11	14

8. Edades de los cinéfilos Con base en los datos del ejercicio de repaso acumulado 7, suponga que las edades de los cinéfilos se distribuyen normalmente con una media de 35 años y una desviación estándar de 20 años.

a. ¿Cuál es el porcentaje de cinéfilos menores de 30 años?

b. Encuentra P_{25}, que es el percentil 25.

c. Encuentre la probabilidad de que una muestra aleatoria simple de 25 cinéfilos tenga una edad media inferior a 30 años.

d. Encuentre la probabilidad de que para una muestra aleatoria simple de 25 cinéfilos, cada uno de ellos tenga menos de 30 años. Para una película y un horario particulares, ¿por qué no sería inusual tener 25 cinéfilos menores de 30 años?

Proyecto de tecnología

Citas rápidas El conjunto de datos 18 "Citas rápidas" en el apéndice B incluye datos de 199 citas. Debido al gran tamaño de este conjunto de datos, se encuentran disponibles en www.pearsonenespañol.com/triola. Descargue el conjunto de datos y proceda a investigar las correlaciones entre pares de variables usando los datos de la quinta, séptima y novena columnas, todas basadas en respuestas de mujeres. Use las medidas "me gusta" de las mujeres como la variable *y* en cada caso.

1. ¿Existe una correlación entre las medidas "me gusta" y las medidas de atractivo?

2. ¿Existe una correlación entre las medidas "me gusta" y las medidas de atributos?

3. ¿Existe una correlación entre las medidas de atractivo y las medidas de atributos?

4. Si se cubrió la sección 10-5 (regresión múltiple), investigue la correlación y la regresión usando las medidas "me gusta" como la variable *y* y las medidas de atractivo y de atributos como las otras dos variables *x*.

5. Repita lo anterior usando la sexta, octava y décima columnas, que se basan en las respuestas de los hombres.

6. De acuerdo con los resultados, ¿qué concluye usted? Escriba un breve informe e incluya los resultados de computadora adecuados.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿El medicamento para el dolor Duragesic es eficaz para reducir el dolor?

A continuación se listan las medidas de la intensidad del dolor antes y después de usar el medicamento Duragesic (fentanilo) (según los datos de Janssen Pharmaceutical Products, L.P.). Los datos se listan ordenados por fila, y las medidas correspondientes son del mismo sujeto antes y después del tratamiento. Por ejemplo, el primer sujeto tenía una medida de 1.2 antes del tratamiento y una medida de 0.4 después del tratamiento. Cada par de mediciones es de un sujeto, y la intensidad del dolor se midió usando la puntuación analógica visual estándar. Una puntuación más alta corresponde a una mayor intensidad del dolor.

Intensidad del dolor antes del tratamiento con Duragesic

1.2	1.3	1.5	1.6	8.0	3.4	3.5	2.8	2.6	2.2
3.0	7.1	2.3	2.1	3.4	6.4	5.0	4.2	2.8	3.9
5.2	6.9	6.9	5.0	5.5	6.0	5.5	8.6	9.4	10.0
7.6									

Intensidad del dolor después del tratamiento con Duragesic

0.4	1.4	1.8	2.9	6.0	1.4	0.7	3.9	0.9	1.8
0.9	9.3	8.0	6.8	2.3	0.4	0.7	1.2	4.5	2.0
1.6	2.0	2.0	6.8	6.6	4.1	4.6	2.9	5.4	4.8
4.1									

Análisis de los resultados

1. Correlación Utilice los datos proporcionados para elaborar un diagrama de dispersión, luego use los métodos de la sección 10-1 para probar una correlación lineal entre la intensidad del dolor antes y después del tratamiento. Si parece que hay una correlación lineal, ¿es posible concluir que el tratamiento con el medicamento es efectivo?

2. Regresión Use los datos dados para encontrar la ecuación de la línea de regresión. Considere que la variable de respuesta (*y*) es la intensidad del dolor después del tratamiento. ¿Cuál sería la ecuación de la línea de regresión para un tratamiento que no tiene absolutamente ningún efecto?

3. Dos muestras independientes Los métodos de la sección 9-2 se pueden usar para evaluar la afirmación de que dos poblaciones tienen la misma media. Identifique la afirmación específica de que el tratamiento es efectivo, luego use los métodos de la sección 9-2 para probar esa afirmación. Los métodos de la sección 9-2 se basan en el requisito de que las muestras sean independientes. ¿Son independientes en este caso?

4. Pares relacionados Los métodos de la sección 9-3 se pueden usar para evaluar una afirmación sobre datos relacionados. Identifique la afirmación específica de que el tratamiento es efectivo, luego use los métodos de la sección 9-3 para probar esa afirmación.

5. ¿El mejor método? ¿Cuál de los resultados anteriores es el mejor para determinar si el tratamiento farmacológico es efectivo para reducir el dolor? Con base en los resultados anteriores, ¿el medicamento parece ser efectivo?

Actividades de cooperación en grupo

1. Actividad en clase Mida la estatura y la longitud de la huella del zapato de cada alumno de la clase. Pruebe una correlación lineal e identifique la ecuación de la línea de regresión. Mida la longitud de la huella del zapato del profesor y utilícela para estimar su estatura. ¿Qué tan cerca está la estatura estimada de la estatura real?

2. Actividad fuera de clase Cada estudiante debe estimar la cantidad de pasos que caminaría entre la puerta del salón de clases y la puerta utilizada para salir del edificio. Después de registrar todas las estimaciones, cada alumno debe contar el número de pasos mientras camina desde la puerta del salón de clases hasta la puerta utilizada para salir del edificio. Una vez recogidas todas las estimaciones y conteos reales, explore la correlación y la regresión utilizando las herramientas presentadas en este capítulo.

3. Actividad en clase Divídanse en grupos de 8 a 12 personas. Mida la estatura y la altura del ombligo de cada miembro del grupo. ¿Hay una correlación entre la estatura y la altura del ombligo? Si es así, encuentre la ecuación de regresión con la estatura expresada en términos de la altura del ombligo. De acuerdo con una teoría, la razón promedio de la estatura entre la altura del ombligo de una persona es la proporción dorada: $(1 = \sqrt{5})/2 \approx 1.6$. ¿Esta teoría parece ser razonablemente precisa?

4. Actividad en clase Divídanse en grupos de 8 a 12 personas. Mida la estatura y el alcance de los brazos de cada miembro del grupo. Para el alcance de los brazos el sujeto debe pararse con los brazos extendidos, como las alas de un avión. Usando los datos muestrales pareados, ¿existe una correlación entre la estatura y el alcance de los brazos? Si es así, encuentre la ecuación de regresión con la estatura expresada en términos del alcance de los brazos. ¿Se puede usar el alcance de los brazos como un buen predictor de la estatura?

5. Actividad en clase Divídanse en grupos de 8 a 12 personas. Use una cuerda y una regla para medir la circunferencia de la cabeza y la longitud del antebrazo de cada miembro del grupo. ¿Existe una relación entre estas dos variables? Si es así, ¿cuál es?

6. Actividad en clase Use una regla como dispositivo para medir el tiempo de reacción. Una persona debe suspender la regla sosteniéndola en la parte superior mientras el sujeto coloca su pulgar e índice en el borde inferior, listo para atrapar la regla cuando sea soltada. Registre la distancia que la regla cae antes de ser atrapada. Convierta esa distancia en el tiempo (segundos) que tardó el sujeto en reaccionar y atrapar la regla. (Si la distancia se mide en pulgadas, use $t = \sqrt{d/192}$. Si la distancia se mide en centímetros, use $t = \sqrt{d/487.68}$). Pruebe a cada sujeto una vez con la mano derecha y una vez con la mano izquierda, y registre los datos pareados. Pruebe la existencia de una correlación. Encuentre la ecuación de la línea de regresión. ¿La ecuación de la línea de regresión sugiere que la mano dominante tiene un tiempo de reacción menor?

7. Actividad en clase Divídanse en grupos de 8 a 12 personas. Registre el pulso de cada miembro del grupo mientras él o ella está sentado. Luego registre el pulso de cada miembro del grupo mientras él o ella esté de pie. ¿Existe una relación entre el pulso sentado y parado? Si es así, ¿cuál es?

8. Actividad en clase Divídanse en grupos de tres o cuatro personas. El apéndice B incluye muchos conjuntos de datos aún no incluidos en ejemplos o ejercicios de este capítulo. Busque en el apéndice B un par de variables de interés, luego investigue la correlación y la regresión. Indique sus conclusiones y trate de identificar aplicaciones prácticas.

9. Actividad fuera de clase Divídanse en grupos de tres o cuatro personas. Investigue la relación entre dos variables mediante la recolección de sus propios datos muestrales pareados y el uso de los métodos de este capítulo para determinar si existe una correlación lineal significativa. También identifique la ecuación de regresión y describa un procedimiento para predecir valores de una de las variables cuando se den valores de la otra variable. Temas sugeridos:

- ¿Existe una relación entre el sabor y el costo de las diferentes marcas de galletas con chispas de chocolate (o bebidas de cola)? El sabor se puede medir en una escala numérica, por ejemplo de 1 a 10.
- ¿Existe una relación entre los salarios de los jugadores profesionales de béisbol (o básquetbol, o fútbol) y sus logros en la temporada?
- ¿Existe una relación entre los promedios de calificaciones de los estudiantes y el tiempo que miran televisión? Si es así, ¿cuál es?

11

BONDAD DE AJUSTE Y TABLAS DE CONTINGENCIA

PROBLEMA DEL CAPÍTULO

Ciberseguridad: detección de intromisiones en sistemas informáticos

De acuerdo con la ley de Benford, muchos conjuntos de datos formados con el primer dígito a la izquierda de todos y cada uno de los números que lo integran, tiene la propiedad de seguir la distribución descrita por las dos filas superiores la tabla 11-1 de la página siguiente. Los conjuntos de datos con valores que tienen dígitos iniciales que se ajustan a la ley de Benford incluyen los números de seguidores en Twitter, tamaños de población, cantidades en declaraciones de impuestos, longitudes de ríos y montos de cheques. En el artículo del *New York Times* "Following Benford's Law, or Looking Out for No. 1", Malcolm Browne escribió que "las agencias de impuestos sobre la renta de varias naciones y varios estados, incluyendo California, están usando software de detección (para identificar intromisiones en sistemas informáticos) con base en la Ley de Benford, al igual que una veintena de grandes empresas y negocios contables".

Ahora parece que la ley de Benford puede ser útil para detectar ataques a sistemas informáticos mediante el análisis de los tiempos entre llegadas, es decir, los tiempos transcurridos entre llegadas consecutivas del tráfico en Internet. La idea básica es detectar anomalías en los tiempos entre llegadas del flujo de tráfico en Internet analizando los dígitos iniciales correspondientes y determinar si la distribución de los primeros dígitos es una desviación significativa de la distribución que sigue la ley de Benford. (Vea "Benford's Law Behavior of Internet Traffic", de Arshadi y Jahangir, *Journal of Network and Computer Applications*, vol. 40, núm. 2014). Las principales ventajas de este método son que es relativamente simple, no requiere cálculos difíciles, se puede hacer en tiempo real y los hackers no podrían configurar su *malware* para evitar la detección.

En las dos filas inferiores de la tabla 11-1, listamos los primeros dígitos de los tiempos entre llegadas del flujo de tráfico en Internet. Una de las dos filas inferiores representa el tráfico normal en Internet y la otra proviene del tráfico en Internet con una intromisión de un pirata informático. La sección 11-1 presentará métodos para identificar cuál de las dos filas inferiores indica que ha ocurrido una intrusión.

TABLA 11-1 Ley de Benford: distribución de los primeros dígitos

Primer dígito	1	2	3	4	5	6	7	8	9
Ley de Benford	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%
Muestra 1	76	62	29	33	19	27	28	21	22
Muestra 2	69	40	42	26	25	16	16	17	20

OBJETIVOS DEL CAPÍTULO

Los capítulos 7 y 8 presentaron métodos importantes de estadística inferencial, que incluyen intervalos de confianza para estimar parámetros poblacionales (capítulo 7) y métodos para probar hipótesis o afirmaciones (capítulo 8). En los capítulos 9 y 10, consideramos inferencias que involucran dos poblaciones y la correlación/regresión con datos pareados. En este capítulo utilizamos métodos estadísticos para analizar datos categóricos (o cualitativos, o de atributos) que se pueden separar en diferentes categorías. Los métodos de este capítulo usan la misma distribución χ^2 (ji-cuadrada) que se presentó en la sección 7-3 y nuevamente en la sección 8-4. Consulte la sección 7-3 o la sección 8-4 para realizar un repaso rápido de las propiedades de la distribución χ^2. Los objetivos del capítulo son:

11-1 Bondad de ajuste

- Usar conteos de frecuencia de datos categóricos divididos en diferentes categorías y determinar si los datos se ajustan a alguna distribución afirmada.

11-2 Tablas de contingencia

- Utilizar datos categóricos resumidos como frecuencias en una tabla bidireccional, con al menos dos filas y dos columnas, para realizar una prueba formal de independencia entre la variable de fila y la variable de columna.
- Ser capaz de realizar una prueba formal de una afirmación de que diferentes poblaciones tienen las mismas proporciones de algunas características.

11-1 Bondad de ajuste

Concepto clave Por "bondad de ajuste" queremos decir que los datos muestrales, que consisten en conteos de frecuencia observados y se disponen en una sola fila o columna (llamada *tabla de frecuencias unidireccional*) concuerdan con alguna distribución particular (por ejemplo normal o uniforme) en consideración. Usaremos una prueba de hipótesis para la afirmación de que los conteos de frecuencia observados concuerdan con la distribución afirmada.

DEFINICIÓN

Una **prueba de bondad de ajuste** se usa para probar la hipótesis de que una distribución de frecuencia observada se ajusta a (o concuerda con) alguna distribución afirmada.

ELEMENTOS CLAVE

Prueba de bondad de ajuste

Objetivo

Realizar una prueba de bondad de ajuste, que es una prueba de hipótesis para determinar si una sola fila (o columna) de conteos de frecuencias concuerda con alguna distribución específica (como uniforme o normal).

Notación

O representa la *frecuencia observada* de un resultado, que se encuentra a partir de datos muestrales.
E representa la *frecuencia esperada* de un resultado, que se encuentra al suponer que la distribución es como se afirma.
k representa la cantidad de *diferentes categorías* o celdas.
n representa el *número total de ensayos* (o el total de valores muestrales observados).
p representa la *probabilidad* de que un valor muestral caiga dentro de una categoría particular.

Requisitos

1. Los datos han sido seleccionados 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 de al menos 5. (La frecuencia esperada para una categoría es la frecuencia que ocurriría si los datos realmente tuvieran la distribución que se afirma, una frecuencia esperada menor a 5 puede llevar a una inferencia incorrecta en la prueba de hipótesis ji-cuadrada. No se requiere que la frecuencia *observada* para cada categoría sea al menos de 5).

Hipótesis nula y alternativa

H_0: Los conteos de frecuencias concuerdan con la distribución afirmada.
H_1: Los conteos de frecuencias no concuerdan con la distribución afirmada.

Dato estadístico de prueba para pruebas de bondad de ajuste

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Valores *P*: Por lo general, los valores P se obtienen utilizando el método discutido en la sección 8-1, también es posible encontrar un rango de valores P a partir de la tabla A-4.

Valores críticos:

1. Los valores críticos se encuentran en la tabla A-4 usando $k - 1$ grados de libertad, donde k es el número de categorías.
2. Las pruebas de hipótesis de bondad de ajuste siempre son *de cola derecha*.

Determinación de frecuencias esperadas

La realización de una prueba de bondad de ajuste requiere que identifiquemos las frecuencias *observadas* que O expresa, luego encontramos las frecuencias *esperadas* (expresadas por E) con la distribución afirmada. Hay dos métodos para determinar las frecuencias esperadas E:

- **Si las frecuencias esperadas son todas iguales: Calcule $E = n/k$.**
- **Si las frecuencias esperadas no son todas iguales: Calcule $E = np$ para cada categoría individual.**

Aunque las dos fórmulas anteriores para E podrían ser muy buenas, es mejor usar un método informal simplemente preguntando, "¿Cómo se pueden dividir las frecuencias observadas entre las diferentes categorías para que haya una coincidencia perfecta con la distribución afirmada?" Además, tenga en cuenta que las frecuencias *observadas* son números enteros porque representan conteos reales, pero las frecuencias *esperadas* no necesitan ser números enteros.

Ejemplos:

a. Igualmente probables Un solo dado se lanza 45 veces con los siguientes resultados. Suponiendo que el dado es legal y que todos los resultados son igualmente probables, encuentre la frecuencia esperada E para cada celda vacía.

Resultado	1	2	3	4	5	6
Frecuencia observada *O*	13	6	12	9	3	2
Frecuencia esperada *E*						

Con $n = 45$ resultados y $k = 6$ categorías, la frecuencia esperada para cada celda es la misma: $E = n/k = 45/6 = 7.5$. Si el dado es legal y los resultados son todos igualmente probables, esperamos que cada resultado se produzca alrededor de 7.5 veces.

b. No son igualmente probables Utilizando los mismos resultados del inciso (a), suponga que afirmamos que, en lugar de ser legal, el dado se carga para que el resultado de 1 ocurra 50% de las veces y cada uno de los otros cinco resultados ocurran 10% de las veces. Las probabilidades se listan en la segunda fila de la siguiente tabla. Usando $n = 45$ y las probabilidades listadas a continuación, encontramos que para la primera celda, $E = np = (45)(0.5) = 22.5$. Cada una de las otras cinco celdas tendrá el valor esperado de $E = np = (45)(0.1) = 4.5$.

Resultado	1	2	3	4	5	6
Probabilidad	0.5	0.1	0.1	0.1	0.1	0.1
Frecuencia observada *O*	13	6	12	9	3	2
Frecuencia esperada *E*	22.5	4.5	4.5	4.5	4.5	4.5

Medición de la discrepancia con la distribución afirmada

Sabemos que las frecuencias muestrales generalmente difieren un poco de los valores que teóricamente esperamos, por lo que consideramos la pregunta clave:

¿Son *significativas* las diferencias entre las frecuencias reales *observadas O* y las frecuencias teóricamente *esperadas E*?

Para medir la discrepancia entre los valores O y E, usamos el dato estadístico de prueba dado en el recuadro de elementos clave anterior. (Más adelante explicaremos cómo se desarrolló este dato estadístico de prueba, pero tiene a las diferencias de $O - E$ como el componente clave).

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

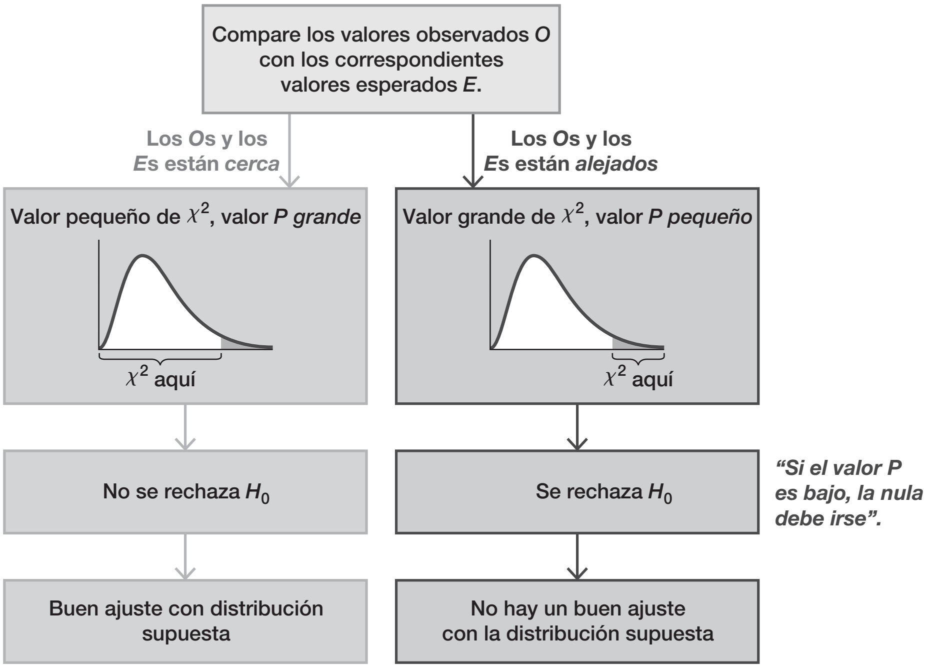


FIGURA 11-1 Relaciones entre el dato estadístico de prueba χ^2, el valor P y la bondad de ajuste

El dato estadístico de prueba χ^2 se basa en las diferencias entre los valores observados y los esperados. Si los valores observados y esperados son *cercanos*, el dato estadístico de prueba χ^2 será pequeño y el valor P será grande. Si las frecuencias observadas y esperadas están *muy separadas*, el dato estadístico de prueba χ^2 será grande y el valor P será pequeño. La figura 11-1 resume esta relación. Las pruebas de hipótesis de esta sección siempre son de cola derecha, porque el valor crítico y la región crítica se encuentran en el extremo derecho de la distribución. Si en algún momento se siente confundido, recuerde este truco mnemotécnico:

"Si el valor P es bajo, la nula debe irse".
(Si el valor P es pequeño, rechace la hipótesis nula de que la distribución es como se afirma).

En cifras

3.141592653: Los primeros 10 dígitos de π, que es un número con una cantidad infinita de dígitos sin ningún patrón de repetición.

EJEMPLO 1 Últimos dígitos de pesos

Se obtiene una muestra aleatoria de 100 pesos californianos, y los últimos dígitos de esos pesos se resumen en la tabla 11-2 (según datos del Departamento de Salud Pública de California). Al obtener los pesos de los sujetos, es extremadamente importante medir realmente su peso en lugar de pedirles que los reporten. Al analizar los *últimos dígitos* de los pesos, los investigadores pueden verificar que se obtuvieron a través de mediciones reales en lugar de reportes propios. Cuando las personas reportan pesos, tienden a redondear hacia abajo y con frecuencia redondean *muy* hacia abajo, por lo que un peso de 197 lb se puede redondear y reportar como un peso más deseable de 170 lb. Los pesos reportados tienden a tener muchos últimos dígitos 0 o 5. Por el contrario, si las personas realmente se pesan, los datos tienden a tener dígitos finales uniformemente distribuidos, donde 0, 1, 2, . . ., 9 ocurren aproximadamente con las mismas frecuencias. Podríamos examinar subjetivamente las frecuencias en la tabla 11-2 para ver que los dígitos de 0 y 5 parecen ocurrir mucho más a menudo que los otros dígitos, pero procederemos con una prueba de hipótesis formal para reforzar esa conclusión subjetiva.

Pruebe la afirmación de que la muestra proviene de una población de pesos en la que los últimos dígitos *no* ocurren con la misma frecuencia. Con base en los resultados, ¿qué podemos concluir sobre el procedimiento utilizado para obtener los pesos?

TABLA 11-2
Últimos dígitos de pesos

Último dígito	Frecuencia
0	46
1	1
2	2
3	3
4	3
5	30
6	4
7	0
8	8
9	3

continúa

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los datos provienen de sujetos seleccionados al azar. (2) Los datos consisten en conteos de frecuencias, como se muestra en la tabla 11-2. (3) Con 100 valores de muestra y 10 categorías que se consideran igualmente probables, cada frecuencia esperada es 10, por lo que cada frecuencia esperada satisface el requisito de ser un valor de al menos 5. Se cumplen todos los requisitos.

La afirmación de que los dígitos no ocurren con la misma frecuencia es equivalente a la afirmación de que las frecuencias relativas o las probabilidades de las 10 celdas ($p_0, p_1, \ldots, p_9$) no son todas iguales. (Esto es equivalente a probar la afirmación de que la distribución de dígitos no es una distribución uniforme).

Paso 1: La afirmación 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 demás.

Paso 2: Si la afirmación original es falsa, entonces todas las probabilidades son iguales; es decir, $p_0 = p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9$.

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 demás.

Paso 4: No se especificó ningún nivel de significancia, por lo que seleccionamos la opción común de $\alpha = 0.05$.

Paso 5: Debido a que estamos probando una afirmación sobre la distribución de los últimos dígitos que es una distribución uniforme (donde todos los dígitos tienen la misma probabilidad). Usamos la prueba de "bondad de ajuste" descrita en esta sección. La distribución χ^2 se usa con el dato estadístico de prueba dado en el recuadro de elementos clave anterior.

XLSTAT

Chi-square (Observed value)	212.8000
Chi-square (Critical value)	16.9190
DF	9
p-value	< 0.0001
alpha	0.05

Paso 6: Las frecuencias observadas O se listan en la tabla 11-2. Cada frecuencia esperada correspondiente E es igual a 10 (porque los 100 dígitos se distribuirán uniformemente entre las 10 categorías). El complemento XLSTAT de Excel se usa para obtener los resultados que se muestran en la pantalla adjunta, y la tabla 11-3 muestra el cálculo manual del dato estadístico de prueba χ^2. El dato estadístico de prueba es $\chi^2 = 212.800$. El valor crítico es $\chi^2 = 16.919$ (encontrado en la tabla A-4 con $\alpha = 0.05$ en la cola derecha y $k - 1 = 9$ grados de libertad). El valor P es menor que 0.0001. El dato estadístico de prueba y el valor crítico se muestran en la figura 11-2.

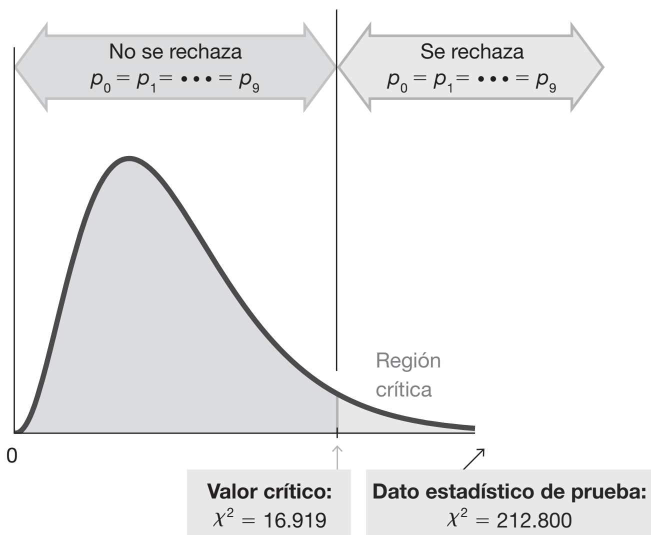


FIGURA 11-2 **Prueba de $p_0 = p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9$**

Paso 7: Si usamos el método del valor *P* para pruebas de hipótesis, vemos que el valor *P* es pequeño (menos de 0.0001), por lo que rechazamos la hipótesis nula. Si usamos el método del valor crítico para probar hipótesis, la figura 11-2 muestra que el dato estadístico de prueba se encuentra en la región crítica, por lo que hay suficiente evidencia para rechazar la hipótesis nula.

Paso 8: Hay suficiente evidencia para respaldar la afirmación de que los últimos dígitos no ocurren con la misma frecuencia relativa.

TABLA 11-3 Cálculo del dato estadístico de prueba χ^2 para los últimos dígitos de los pesos

Último dígito	Frecuencia observada *O*	Frecuencia esperada *E*	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
0	46	10	36	1296	129.6
1	1	10	−9	81	8.1
2	2	10	−8	64	6.4
3	3	10	−7	49	4.9
4	3	10	−7	49	4.9
5	30	10	20	400	40.0
6	4	10	−6	36	3.6
7	0	10	−10	100	10.0
8	8	10	−2	4	0.4
9	3	10	−7	49	4.9
					$\chi^2 = \sum\frac{(O-E)^2}{E} = 212.8$

INTERPRETACIÓN

Esta prueba de bondad de ajuste sugiere que los últimos dígitos no tienen un buen ajuste con la distribución uniforme declarada de frecuencias igualmente probables. En vez de pesar realmente a los sujetos, parece que estos reportaron sus pesos. De hecho, los pesos son del Estudio de Entrevistas de la Salud de California (CHIS, por sus siglas en inglés), y el título de ese estudio indica que los sujetos fueron entrevistados, no medidos. Debido a que los pesos se reportan, la confiabilidad de los datos es muy cuestionable.

SU TURNO Resuelva el ejercicio 20 "Últimos dígitos de pesos".

El ejemplo 1 implica una situación en la que las frecuencias esperadas *E* para las diferentes categorías son todas iguales. Los métodos de esta sección también se pueden usar cuando las frecuencias esperadas son diferentes, como en el ejemplo 2.

EJEMPLO 2 Ley de Benford: Detección de intromisiones informáticas

El problema del capítulo presentó la *ley de Benford*, según la cual una variedad de conjuntos de datos diferentes incluyen números cuyos *primeros* dígitos siguen la distribución mostrada en las primeras dos filas de la tabla 11-4. La fila inferior lista las frecuencias de los dígitos iniciales de los tiempos entre llegadas para el tráfico en Internet (con base en el problema del capítulo). ¿Las frecuencias en la fila inferior se ajustan a la distribución descrita por la Ley de Benford?

TABLA 11-4 Dígitos iniciales de los tiempos entre llegadas para el tráfico en Internet

Dígito inicial	1	2	3	4	5	6	7	8	9
Ley de Benford: Distribución de los dígitos iniciales	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%
Muestra 2 de dígitos iniciales	69	40	42	26	25	16	16	17	20

continúa

¿Qué asientos de automóvil son los más seguros?

Muchas personas consideran que el lugar más seguro de un automóvil es el asiento trasero. ¿Será verdad? Investigadores de la Universidad de Buffalo analizaron más de 60,000 accidentes automovilísticos fatales, y encontraron que el asiento trasero intermedio es el lugar más seguro en un automóvil. Los autores descubrieron que, al sentarse en ese lugar, los pasajeros tienen 86% más probabilidades de sobrevivir que los pasajeros sentados en los asientos delanteros, y 25% más probabilidades de sobrevivir que los pasajeros que se ubican en el asiento trasero, junto a las ventanas. Un análisis del uso del cinturón de seguridad reveló que, cuando los pasajeros no lo utilizan en el asiento trasero, tienen tres veces más probabilidades de morir en un accidente que los que sí utilizan el cinturón. Las personas preocupadas por la seguridad deben sentarse en la parte media del asiento trasero y utilizar siempre el cinturón de seguridad.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los datos muestrales se seleccionan aleatoriamente de una población mayor. (2) Los datos muestrales consisten en conteos de frecuencias. (3) Cada frecuencia esperada es de al menos 5. La frecuencia más baja esperada es $271 \cdot 0.46 = 12.466$. Todos los requisitos se satisfacen.

Paso 1: La afirmación original es que los dígitos iniciales se ajustan a la distribución dada como la ley de Benford. Utilizando subíndices correspondientes a los dígitos iniciales, podemos expresar esta afirmación como $p_1 = 0.301$ y $p_2 = 0.176$ y $p_3 = 0.125$ y . . . y $p_9 = 0.046$.

Paso 2: Si la afirmación original es falsa, al menos una de las proporciones no tiene el valor que se afirma.

Paso 3: La hipótesis nula debe contener la condición de igualdad, por lo que tenemos

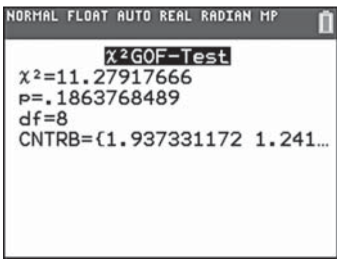
$$H_0: p_1 = 0.301 \text{ y } p_2 = 0.176 \text{ y } p_3 = 0.125 \text{ y . . . y } p_9 = 0.046.$$

H_1: Al menos una de las proporciones no es igual al valor que se afirma.

Paso 4: El nivel de significancia no se especifica, por lo que usamos la opción común de $\alpha = 0.05$.

Paso 5: Dado que estamos probando una afirmación de que la distribución de los dígitos iniciales se ajusta a la distribución dada por la ley de Benford, utilizamos la prueba de bondad de ajuste que se describe en esta sección. La distribución χ^2 se usa con el dato estadístico de prueba dado en el cuadro de elementos clave anterior.

TI-84 Plus C



Paso 6: La tabla 11-5 muestra los cálculos de los componentes del dato estadístico de prueba χ^2 para los dígitos iniciales de 1 y 2. Si incluimos los nueve dígitos iniciales, obtenemos el dato estadístico de prueba de $\chi^2 = 11.2792$, como se muestra en la pantalla adjunta de la calculadora TI-84 Plus C. El valor crítico es $\chi^2 = 15.507$ (que se encuentra en la tabla A-4, con $\alpha = 0.05$ en la cola derecha y $k - 1 = 8$ grados de libertad). La pantalla de la calculadora TI-84 Plus C muestra el valor del dato estadístico de prueba, así como el valor P de 0.186. (La fila inferior completa de la pantalla se puede ver al desplazarse hacia la derecha. CNTRB es una forma abreviada de "contribución", y los valores son las contribuciones individuales al valor total del dato estadístico de prueba χ^2).

TABLA 11-5 Cálculo del dato estadístico de prueba χ^2 para los dígitos iniciales de la tabla 11-4

Dígito inicial	Frecuencia observada O	Frecuencia esperada $E = np$	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
1	69	$271 \cdot 0.301 = 81.5710$	−12.5710	158.0300	1.9373
2	40	$271 \cdot 0.176 = 47.6960$	−7.6960	59.2284	1.2418

Paso 7: El valor P de 0.186 es mayor que el nivel de significancia de 0.05, por lo que no hay suficiente evidencia para rechazar la hipótesis nula. (Además, el dato estadístico de prueba $\chi^2 = 11.2792$ no cae en la región crítica limitada por el valor crítico de 15.507, por lo que no hay evidencia suficiente para rechazar la hipótesis nula).

Paso 8: No hay evidencia suficiente para justificar el rechazo de la afirmación de que los 271 dígitos iniciales se ajustan a la distribución dada por la ley de Benford.

INTERPRETACIÓN

La muestra de dígitos iniciales no proporciona suficiente evidencia para concluir que la distribución de la ley de Benford no se está siguiendo. No hay evidencia suficiente para respaldar la conclusión de que los dígitos iniciales son de tiempos entre llegadas que no corresponden al tráfico normal, por lo que no hay pruebas suficientes para concluir que se haya producido una intrusión en Internet.

SU TURNO Resuelva el ejercicio 21 "Detección de fraude".

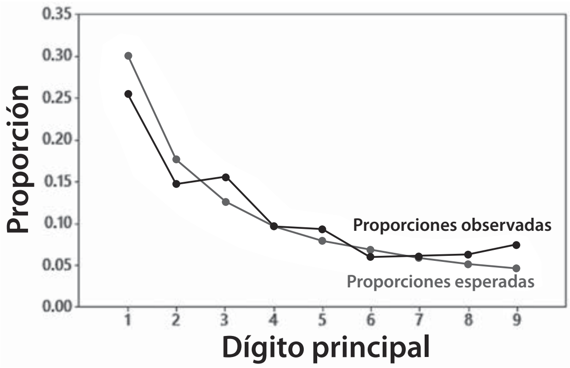


FIGURA 11-3 Tiempos entre llegadas: Proporciones observadas y proporciones esperadas con la Ley de Benford

En la figura 11-3, usamos una línea gris para graficar las proporciones esperadas, según la ley de Benford (como en la tabla 11-4) junto con una línea negra para las proporciones observadas de la tabla 11-4. La figura 11-3 nos permite visualizar la "bondad de ajuste" entre la distribución dada por la ley de Benford y las frecuencias que se observaron. En la figura 11-3, las líneas gris y negra concuerdan razonablemente bien, por lo que parece que los datos observados se ajustan razonablemente bien a los valores esperados.

Justificación para los estadísticos de prueba Los ejemplos 1 y 2 muestran que el dato estadístico de prueba χ^2 es una medida de la discrepancia entre las frecuencias observadas y esperadas. La simple suma de las diferencias $O - E$ entre los valores observados y los esperados no nos dice nada, porque esa suma siempre es 0. La elevación al cuadrado de los valores $O - E$ nos da un mejor estadístico. (Las razones para elevar al cuadrado los valores $O - E$ son esencialmente las mismas que para elevar al cuadrado los valores $x - \bar{x}$ en la fórmula para la desviación estándar). El valor de $\Sigma(O - E)^2$ mide sólo la magnitud de las diferencias, pero necesitamos encontrar la magnitud de las diferencias en relación con lo que se esperaba. Necesitamos un tipo de promedio en lugar de un total acumulado. Esta magnitud relativa se encuentra a través de la división por las frecuencias esperadas, como en el dato estadístico de prueba $\Sigma(O - E)^2/E$.

La distribución teórica de $\Sigma(\mathrm{O} - E)^2/E$ es una distribución discreta porque el número de valores posibles es finito. La distribución se puede aproximar mediante una distribución ji cuadrada, que es continua. Esta aproximación generalmente se considera aceptable, siempre que todos los valores esperados E sean al menos 5. (Hay formas de eludir el problema de una frecuencia esperada menor que 5, como combinar algunas categorías para que todas las frecuencias esperadas sean de al menos 5. Además, existen diferentes procedimientos que se pueden usar cuando no todas las frecuencias esperadas son al menos 5).

El número de grados de libertad refleja el hecho de que podemos asignar frecuencias libremente a $k - 1$ categorías antes de determinar la frecuencia para cada una de ellas. (Aunque decimos que podemos asignar frecuencias "libremente" a $k - 1$ categorías, no podemos tener frecuencias negativas, ni frecuencias tan grandes que su suma exceda el total de las frecuencias observadas para todas las categorías combinadas).

¿Los datos de Mendel se falsificaron?

Debido a que algunos de los datos de los famosos experimentos genéticos de Mendel parecen demasiado perfectos para ser verdaderos, el especialista en estadística R. A. Fisher concluyó que los datos probablemente se falsificaron; utilizó una distribución ji cuadrada para demostrar que, cuando un dato estadístico de prueba se localiza a la extrema izquierda y da como resultado un valor *P* muy cercano a 1, los datos muestrales se ajustan a la distribución establecida casi de manera perfecta, lo cual es evidencia de que los datos muestrales no fueron seleccionados al azar. Se ha sugerido que el jardinero de Mendel sabía cuáles valores eran los esperados según la teoría mendeliana y que ajustó los resultados para que coincidieran con ésta.

Ira Pilgrim escribió en *The Journal of Heredity* que este uso de la distribución ji cuadrada no es adecuado; señaló que no se trata de la bondad de ajuste con una distribución en particular, sino más bien de determinar si los datos provienen de una muestra realmente aleatoria. Pilgrim utilizó la fórmula de probabilidad binomial para calcular las probabilidades de los resultados obtenidos en los experimentos de Mendel. Con base en tales resultados, Pilgrim concluyó que "no hay razón alguna para cuestionar la honestidad de Mendel". Al parecer, sus resultados no son demasiado buenos para ser verdaderos y pudieron haberse obtenido de un proceso realmente aleatorio.

CENTRO DE TECNOLOGÍA

Prueba de bondad de ajuste

Acceda a los complementos de software, videos y conjuntos de datos en **pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Goodness-of-Fit** en el menú desplegable.
3. Seleccione **Equal Expected Frecuencies** o **Unequal Expected Frecuencies**.
4. Ingrese el nivel de significancia deseado y seleccione la columna que contiene las frecuencias observadas. Para *frecuencias esperadas desiguales* también indique si los datos están en el formato de conteos o proporciones y seleccione la columna que contiene los datos esperados.
5. Haga clic en **Evaluate**

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Tables** en el menú desplegable y elija **Chi-Square Goodness-of-Fit Test** en el submenú.
3. Haga clic en **Observed Counts** y seleccione la columna que contiene las frecuencias observadas.
4. En *Test*, seleccione **Equal Proportions** si las frecuencias esperadas son todas iguales. Para frecuencias o proporciones esperadas desiguales, seleccione **Proportions specified by historical counts** y seleccione la columna que contiene las frecuencias o proporciones esperadas).
5. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Goodness-of-fit** en el menú desplegable, luego elija **Chi-Square Test** en el submenú.
3. Seleccione la columna con las frecuencias observadas.
4. Seleccione la columna que contiene las frecuencias esperadas si éstas no son todas iguales. De lo contrario, haga clic en **All cells in equal proportion**.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

Las calculadoras TI-83/84 requieren frecuencias esperadas. Las proporciones esperadas no se pueden usar.

1. Ingrese los valores observados en una lista (*L1*) y las frecuencias esperadas en una lista distinta (*L2*).
2. Presione STAT, luego seleccione **TESTS** en el menú superior.
3. Seleccione χ^2 **GOF-Test** en el menú y presione ENTER
4. Ingrese los nombres de las listas para las frecuencias observadas y esperadas. Para *df* ingrese los grados de libertad, que es 1 menos que el número de categorías.
5. Seleccione **Calculate** y presione ENTER.

Sugerencia: Las calculadoras TI-83 requieren el programa **X2GOF**, que está disponible en pearsonenespañol.com/triola.

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Parametric Tests**.
2. Seleccione **Multinomial goodness of fit test** en el menú desplegable.
3. En el cuadro de *frecuencias*, ingrese el rango de celdas que contienen las frecuencias observadas. En el cuadro de *frecuencias esperadas*, ingrese el rango de celdas que contienen las frecuencias esperadas. Si usa proporciones esperadas, marque la casilla de **Proportions** bajo el *formato de datos*.
4. Marque la casilla de **Chi-square test**. Si el rango de datos incluye una etiqueta de datos, también marque la casilla de **Columna labels**.
5. Ingrese un nivel de significancia y haga clic en **OK**. El dato estadístico de prueba se etiqueta como *Chi-square* (*Observed Value*).

Excel

1. Haga clic en **Insert function** f_x, seleccione la categoría **Statistical** y elija la función **CHISQ.TEST**.
2. Para *el rango real*, ingrese el rango de celdas para las frecuencias observadas. Para el *rango esperado*, ingrese las frecuencias esperadas.
3. Haga clic en **OK** para el valor *P*.

11-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Ciberseguridad La siguiente tabla muestra los primeros dígitos de 317 tiempos entre llegadas del tráfico en Internet para una computadora, junto con las frecuencias de los dígitos iniciales que se esperan con la ley de Benford (con base en la tabla 11-1 del problema del capítulo).

a. Identifique la notación utilizada para los valores observados y esperados.

b. Identifique los valores observados y esperados para el dígito inicial de 2.

c. Utilice los resultados del inciso (b) para encontrar la contribución al dato estadístico de prueba χ^2 de la categoría que representa el dígito inicial 2.

Dígito inicial	1	2	3	4	5	6	7	8	9
Ley de Benford	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%
Dígitos iniciales de los tiempos entre llegadas del tráfico	76	62	29	33	19	27	28	21	22

2. Ciberseguridad Al usar los datos del ejercicio 1 para probar la bondad de ajuste con la distribución descrita por la ley de Benford, identifique las hipótesis nula y alternativa.

3. Ciberseguridad Los resultados de Statdisk que se muestran al margen se obtienen a partir de los datos proporcionados en el ejercicio 1. ¿Qué se debe concluir al probar la afirmación de que los dígitos iniciales tienen una distribución que se ajusta bien a la ley de Benford?

Test Statistic, X^2:	20.9222
Critical X^2:	15.5073
P-Value:	0.0074

4. Ciberseguridad ¿Qué sugieren los resultados de los ejercicios anteriores acerca de la posibilidad de que la computadora haya sido pirateada? ¿Hay alguna acción correctiva que se deba tomar?

En los ejercicios 5 a 20, realice la prueba de hipótesis y proporcione el dato estadístico de prueba, el valor P y/o el valor crítico, y establezca la conclusión.

5. Prueba de una máquina tragamonedas El autor compró una máquina tragamonedas (Bally Modelo 809) y la probó jugando 1197 veces. Hay 10 categorías de resultados, incluyendo no ganar nada, ganar el premio mayor, ganar con tres campanas, etcétera. Al probar la afirmación de que los resultados observados concuerdan con las frecuencias esperadas, el autor obtuvo un dato estadístico de prueba de $\chi^2 = 8.185$. Use un nivel de significancia de 0.05 para probar la afirmación de que los resultados reales concuerdan con las frecuencias esperadas. ¿Parece que la máquina tragamonedas funciona como se esperaba?

6. Neumático desinflado y clase perdida Una historia clásica involucra a cuatro estudiantes que compartían el auto y que faltaron a un examen; su excusa fue un neumático desinflado. En la reposición del examen, el profesor les pidió a los estudiantes que identificaran el neumático en particular que se desinfló. Si realmente no tuvieran un neumático desinflado, ¿podrían identificar el mismo neumático? El autor le pidió a otros 41 estudiantes que identificaran el neumático que seleccionarían. Los resultados se listan en la siguiente tabla (excepto por un estudiante que seleccionó el neumático de repuesto). Use un nivel de significancia de 0.05 para probar la afirmación del autor de que los resultados se ajustan a una distribución uniforme. ¿Qué sugiere el resultado sobre la probabilidad de que cuatro estudiantes identifiquen el mismo neumático cuando realmente nunca tuvieron el accidente?

Neumático	Delantero izquierdo	Delantero derecho	Trasero izquierdo	Trasero derecho
Número de selecciones	11	15	8	6

7. Dado cargado El autor perforó un agujero en un dado y lo llenó con un peso de plomo, luego procedió a lanzarlo 200 veces. A continuación se listan las frecuencias observadas para los resultados de 1, 2, 3, 4, 5 y 6, respectivamente: 27, 31, 42, 40, 28 y 32. Use un nivel de significancia de 0.05 para probar la afirmación de que los resultados no son igualmente probables. ¿Parece que el dado cargado se comporta de forma diferente a un dado legal?

8. ¿Sesgo en ensayos clínicos? Los investigadores investigaron el tema de la raza y la igualdad de acceso a los ensayos clínicos. La siguiente tabla muestra la distribución de la población y el número de participantes en ensayos clínicos que involucran cáncer de pulmón (según datos de “Participation in Cancer Clinical Trials”, de Murthy, Krumholz y Gross, *Journal of the American Medical Association*, vol. 291, núm. 22). Use un nivel de significancia de 0.01 para evaluar la afirmación de que la distribución de los participantes en ensayos clínicos se ajusta bien a la distribución de la población. ¿Hay una raza/grupo étnico que parezca estar muy poco representado?

Raza/etnicidad	Blanco no hispano	Hispano	Negro	Asiático/Islas del Pacífico	Indio americano/ nativo de Alaska
Distribución de la población	75.6%	9.1%	10.8%	3.8%	0.7%
Cantidad en ensayos clínicos de cáncer de pulmón	3855	60	316	54	12

9. Genética mendeliana Se realizan experimentos con híbridos de dos tipos de chícharos. Si los descendientes siguen la teoría de la herencia de Mendel, las semillas que se producen son amarillas lisas, verdes lisas, amarillas corrugadas y verdes corrugadas, y deben ocurrir en una proporción de 9:3:3:1, respectivamente. Se diseñó un experimento para probar la teoría de Mendel con el resultado de que las semillas de los descendientes estuvieron formadas por 307 lisas amarillas, 77 verdes lisas, 98 amarillas corrugadas y 18 verdes corrugadas. Use un nivel de significancia de 0.05 para probar la afirmación de que los resultados contradicen la teoría de Mendel.

10. ¿Los impactos de bombas en la Segunda Guerra Mundial se ajustan a una distribución de Poisson? Al analizar los impactos de las bombas V-1 buzz en la Segunda Guerra Mundial, el sur de Londres se subdividió en regiones, cada una con un área de 0.25 km^2. A continuación se muestra una tabla de frecuencias reales de los impactos y las frecuencias esperadas con la distribución de Poisson. (La distribución de Poisson se describe en la sección 5-3). Use los valores listados y un nivel de significancia de 0.05 para probar la afirmación de que las frecuencias reales se ajustan a una distribución de Poisson. ¿El resultado demuestra que los datos se ajustan a la distribución de Poisson?

Número de impactos de bomba	0	1	2	3	4 o más
Número real de regiones	229	211	93	35	8
Número esperado de regiones (a partir de la distribución de Poisson)	227.5	211.4	97.9	30.5	8.7

11. Llamadas a la policía El departamento de policía en Madison, Connecticut, recibió el siguiente número de llamadas en los diferentes días de la semana durante un mes de febrero que tuvo 28 días: lunes (114), martes (152), miércoles (160), jueves (164), viernes (179), sábado (196), domingo (130). Use un nivel de significancia de 0.01 para probar la afirmación de que los diferentes días de la semana tienen las mismas frecuencias de llamadas a la policía. ¿Hay algo notable sobre las frecuencias observadas?

12. Llamadas a la policía Repita el ejercicio 11 usando las siguientes frecuencias observadas para llamadas a la policía recibidas durante el mes de marzo: lunes (208), martes (224), miércoles (246), jueves (173), viernes (210), sábado (236), domingo (154). ¿Cuál es un error fundamental con este análisis?

13. Derby de Kentucky La siguiente tabla lista la frecuencia de victorias para las diferentes posiciones de salida hasta la edición número 141 del Derby de Kentucky. La posición de salida 1 es la más cercana al carril interior, por lo que el caballo en esa posición recorre la distancia más corta. (Debido a que el número de caballos varía cada año, sólo se incluyen las primeras 10 posiciones). Use un nivel de significancia de 0.05 para probar la afirmación de que la probabilidad de ganar es la misma para las diferentes posiciones de salida. Con base en el resultado, ¿deberían los apostadores considerar la posición de salida de los caballos en el Derby de Kentucky?

Posición de salida	1	2	3	4	5	6	7	8	9	10
Victorias	19	14	11	15	15	7	8	12	5	11

14. Lotería Daily 4 de California El autor registró todos los dígitos seleccionados en la lotería Daily 4 de California durante los 60 días previos al momento de escribir este ejercicio. Las frecuencias de los dígitos del 0 al 9 fueron 21, 30, 31, 33, 19, 23, 21, 16, 24 y 22. Use un nivel de significancia de 0.05 para probar la afirmación de los comisionados de la lotería acerca de que los dígitos se seleccionan de un modo en que son igualmente probables.

15. Juegos de la Serie Mundial La siguiente tabla muestra las cantidades de juegos realizados en 105 Series Mundiales de las Grandes Ligas de Béisbol (MLB). Esta tabla también incluye las proporciones esperadas para el número de juegos en una Serie Mundial, suponiendo que, en cada serie, ambos equipos tienen casi la misma posibilidad de ganar. Use un nivel de significancia de 0.05 para probar la afirmación de que el número real de juegos se ajusta a la distribución indicada por las proporciones esperadas.

Juegos realizados	4	5	6	7
Número de Series Mundiales	21	23	23	38
Proporción esperada	2/16	4/16	5/16	5/16

16. Nacimientos de beisbolistas En su libro *Outliers*, el autor Malcolm Gladwell argumenta que hay más beisbolistas con fechas de nacimiento en los meses inmediatamente posteriores al 31 de julio, porque esa era la fecha de corte de edad para las ligas no escolares de béisbol. A continuación se presenta una muestra de conteos de frecuencia para los meses de nacimiento de los jugadores de Grandes Ligas nacidos en Estados Unidos, comenzando de enero: 387, 329, 366, 344, 336, 313, 313, 503, 421, 434, 398, 371. Utilizando un nivel de significancia de 0.05, ¿hay suficiente evidencia para respaldar el rechazo de la afirmación de que los jugadores de las Grandes Ligas nacidos en Estados Unidos nacen en los diferentes meses con la misma frecuencia? ¿Los valores muestrales parecen respaldar la afirmación de Gladwell?

Los ejercicios 17 a 20 se basan en conjuntos de datos incluidos en el apéndice B. Los conjuntos de datos completos pueden encontrarse en www.pearsonenespañol.com/triola.

17. Admisiones para alumbramiento El conjunto de datos 4 "Nacimientos" incluye los días de la semana en que las futuras madres ingresaron en un hospital para dar a luz. Un médico afirma que debido a que muchos nacimientos son inducidos o implican una cesárea, están programados para días que no sean el sábado o el domingo, por lo que los nacimientos no ocurren en los siete distintos días de la semana con la misma frecuencia. Use un nivel de significancia de 0.01 para probar esa afirmación.

18. Altas después del nacimiento El conjunto de datos 4 "Nacimientos" incluye los días de la semana en que los recién nacidos fueron dados de alta del hospital. Un administrador del hospital afirma que tales altas ocurren los siete distintos días de la semana con la misma frecuencia. Use un nivel de significancia de 0.01 para probar esa afirmación.

19. Dulces M&M Mars, Inc. afirma que sus dulces planos M&M se distribuyen con los siguientes porcentajes de color: 16% verde, 20% naranja, 14% amarillo, 24% azul, 13% rojo y 13% café. Consulte el conjunto de datos 27 "Pesos de M&Ms" en el apéndice B y utilice los datos muestrales para probar la afirmación de que la distribución de color es la que afirma Mars, Inc. Use un nivel de significancia de 0.05.

20. Últimos dígitos de los pesos El conjunto de datos 1 "Datos corporales" en el apéndice B incluye los pesos (en kg) de 300 sujetos. Use un nivel de significancia de 0.05 para probar la afirmación de que la muestra proviene de una población de pesos en la que los últimos dígitos *no* ocurren con la misma frecuencia. Cuando las personas *reportan* su peso en vez de medirlo, tienden a redondear de modo que los últimos dígitos no ocurren con la misma frecuencia. ¿Los resultados sugieren que los pesos fueron reportados?

Ley de Benford. ***De acuerdo con la ley de Benford, una variedad de conjuntos de datos diferentes incluyen números donde los primeros dígitos siguen la distribución mostrada en la tabla siguiente. En los ejercicios 21 a 24, pruebe la bondad de ajuste con la distribución descrita por la ley de Benford.***

Primer dígito	1	2	3	4	5	6	7	8	9
Ley de Benford: Distribución de los primeros dígitos	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%

21. Detección de fraude Cuando trabajaba para el fiscal de distrito de Brooklyn, el investigador Robert Burton analizó los primeros dígitos de los montos de 784 cheques emitidos por siete compañías sospechosas. Las frecuencias fueron 0, 15, 0, 76, 479, 183, 8, 23 y 0, correspondientes a los primeros dígitos 1, 2, 3, 4, 5, 6, 7, 8 y 9, respectivamente. Si las frecuencias observadas son sustancialmente diferentes de las frecuencias esperadas con la ley de Benford, los montos de los cheques parecen ser el resultado de un fraude. Use un nivel de significancia de 0.01 para probar la bondad de ajuste con la ley de Benford. ¿Parece que los cheques son el resultado de un fraude?

22. Montos de los cheques del autor El ejercicio 21 lista las frecuencias observadas de los dígitos iniciales en los montos de los cheques de siete compañías sospechosas. A continuación se dan las frecuencias observadas de los dígitos iniciales en los montos de los cheques más recientes emitidos por el autor al momento de escribir este ejercicio: 83, 58, 27, 21, 21, 21, 6, 4, 9. (Esas frecuencias observadas corresponden a los dígitos iniciales 1, 2, 3, 4, 5, 6, 7, 8 y 9, respectivamente). Utilizando un nivel de significancia de 0.01, pruebe la afirmación de que estos dígitos iniciales provienen de una población de dígitos iniciales que se ajustan a la ley de Benford. ¿La conclusión cambia si el nivel de significancia es 0.05?

23. ¿Falsedad en los impuestos? Las frecuencias de los dígitos iniciales de varias declaraciones de impuestos recibidas por la Oficina de Hacienda son 152, 89, 63, 48, 39, 40, 28, 25 y 27 (correspondientes a los dígitos iniciales 1, 2, 3, 4, 5, 6, 7, 8 y 9, respectivamente; con base en datos de Mark Nigrini, un proveedor de software para el análisis de datos de Benford). Utilizando un nivel de significancia de 0.05, pruebe la bondad de ajuste con la ley de Benford. ¿Parece que las declaraciones de impuestos son legítimas?

24. Archivos de computadora del autor El autor registró los dígitos iniciales de los tamaños de los archivos electrónicos para la edición actual de este libro. Los dígitos iniciales tienen frecuencias de 55, 25, 17, 24, 18, 12, 12, 3 y 4 (correspondientes a los dígitos iniciales 1, 2, 3, 4, 5, 6, 7, 8 y 9, respectivamente). Utilizando un nivel de significancia de 0.05, prueba la bondad de ajuste con la ley de Benford.

11-1 Más allá de lo básico

25. Prueba de bondad de ajuste con una distribución normal Consulte en el conjunto de datos 1 "Datos corporales" del apéndice B, las estaturas de las mujeres.

Estatura (cm)	Menos de 155.45	155.45 – 162.05	162.05 – 168.65	Más de 168.65
Frecuencia				

a. Ingrese las frecuencias observadas en la tabla anterior.

b. Suponiendo una distribución normal con media y desviación estándar dadas por la media y la desviación estándar muestrales, use los métodos del capítulo 6 para encontrar la probabilidad de una estatura seleccionada al azar que pertenezca a cada clase.

c. Con base en las probabilidades determinadas en el inciso (b). encuentre la frecuencia esperada para cada categoría.

d. Use un nivel de significancia de 0.01 para probar la afirmación de que las estaturas fueron seleccionadas al azar de una población distribuida normalmente. ¿La prueba de bondad de ajuste sugiere que los datos provienen de una población distribuida normalmente?

11-2 Tablas de contingencia

Concepto clave Ahora consideraremos métodos para analizar *tablas de contingencia* (o tablas de frecuencias bidireccionales), que incluyen conteos de frecuencias para datos categóricos dispuestos en una tabla con al menos dos filas y al menos dos columnas. En la parte 1 de esta sección presentamos un método para realizar una prueba a la hipótesis nula de que las variables de fila y columna son independientes entre sí. Esta prueba de independencia se usa ampliamente en aplicaciones del mundo real. En la parte 2, consideraremos tres variaciones del método básico presentado en la parte 1: (1) prueba de homogeneidad, (2) prueba exacta de Fisher y (3) prueba de McNemar para pares relacionados.

PARTE 1 Conceptos básicos de las pruebas de independencia

En esta sección utilizamos métodos estadísticos estándar para analizar los conteos de frecuencias en una tabla de contingencia (o tabla de frecuencias bidireccional).

DEFINICIÓN

Una **tabla de contingencia** (o **tabla de frecuencias bidireccional**) es una tabla que consiste en conteos de frecuencias de datos categóricos correspondientes a dos variables diferentes (una variable se usa para categorizar las filas y una segunda variable se usa para categorizar las columnas).

La palabra *contingente* tiene unos cuantos significados diferentes, uno de los cuales se refiere a la *dependencia* de algún otro factor. Usamos el término *tabla de contingencia* porque probamos la *independencia* entre las variables de fila y de columna. Primero definimos una *prueba de independencia* y después proporcionamos los elementos clave de la prueba en el siguiente cuadro de elementos clave.

DEFINICIÓN

En una **prueba de independencia**, probamos la hipótesis nula de que en una tabla de contingencia, las variables de fila y de columna son independientes. (Es decir, que no hay dependencia entre las variables de fila y de columna).

ELEMENTOS CLAVE

Tabla de contingencia

Objetivo

Realizar una prueba de hipótesis de independencia entre la variable de fila y la variable de columna en una tabla de contingencia.

Notación

O representa la *frecuencia observada* en una celda de una tabla de contingencia.
E representa la *frecuencia esperada* en una celda, la cual se encuentra suponiendo que las variables de fila y de columna son independientes.
r representa el número de filas en una tabla de contingencia (sin incluir las etiquetas o los totales de las filas).
c representa el número de columnas en una tabla de contingencia (sin incluir las etiquetas o los totales de las columnas).

Requisitos

1. Los datos muestrales se seleccionan al azar.
2. Los datos muestrales se representan como conteos de frecuencias en una tabla bidireccional.
3. Para cada celda en la tabla de contingencia, la frecuencia esperada E es al menos 5. (No es necesario que cada frecuencia observada sea al menos 5).

Hipótesis nula y alternativa

Las hipótesis nula y alternativa son como sigue:

H_0: Las variables de fila y de columna son independientes.
H_1: Las variables de fila y de columna son dependientes.

Dato estadístico de prueba para una prueba de independencia

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

donde O es la frecuencia observada en una celda y E es la frecuencia esperada en una celda que se encuentra al evaluar

$$E = \frac{(\text{total de fila})\,(\text{total de columna})}{(\text{gran total})}$$

Valores *P*

Por lo general, los valores P pueden ser obtenidos utilizando el método descrito en la sección 8-1, también es posible determinar un rango de valores P a partir de la tabla A-4.

continúa

Valores críticos:

1. Los valores críticos se encuentran en la tabla A-4 usando

$$\textbf{Grados de libertad} = (r - 1)(c - 1)$$

donde r es el número de filas y c es el número de columnas.

2. Las pruebas de independencia con una tabla de contingencia son siempre de *cola derecha*.

La distribución del dato estadístico de prueba χ^2 se puede aproximar por la distribución ji cuadrada, siempre que todas las celdas tengan frecuencias esperadas que sean al menos 5. El número de grados de libertad $(r - 1)(c - 1)$ refleja el hecho de que, dado que conocemos el total de las frecuencias en una tabla de contingencia, podemos asignar frecuencias libremente a sólo $r - 1$ filas y $c - 1$ columnas antes de determinar la frecuencia para cada celda. Sin embargo, no podemos tener frecuencias negativas o frecuencias tan grandes que alguna suma de fila (o columna) exceda el total de las frecuencias observadas para esa fila (o columna).

Frecuencias observadas y esperadas El dato estadístico de prueba nos permite medir el tamaño de la discrepancia entre las frecuencias realmente observadas y aquellas que teóricamente esperaríamos cuando las dos variables son independientes. Los valores grandes del dato estadístico de prueba χ^2 están en la región más a la derecha de la distribución ji cuadrada, y reflejan diferencias significativas entre las frecuencias observadas y esperadas. Como en la sección 11-1, si las frecuencias observadas y esperadas son cercanas entre sí, el dato estadístico de prueba χ^2 será pequeño y el valor P será grande. Si las frecuencias observadas y esperadas están muy separadas, el dato estadístico de prueba χ^2 será grande y el valor P será pequeño. Estas relaciones se resumen e ilustran en la figura 11-4.

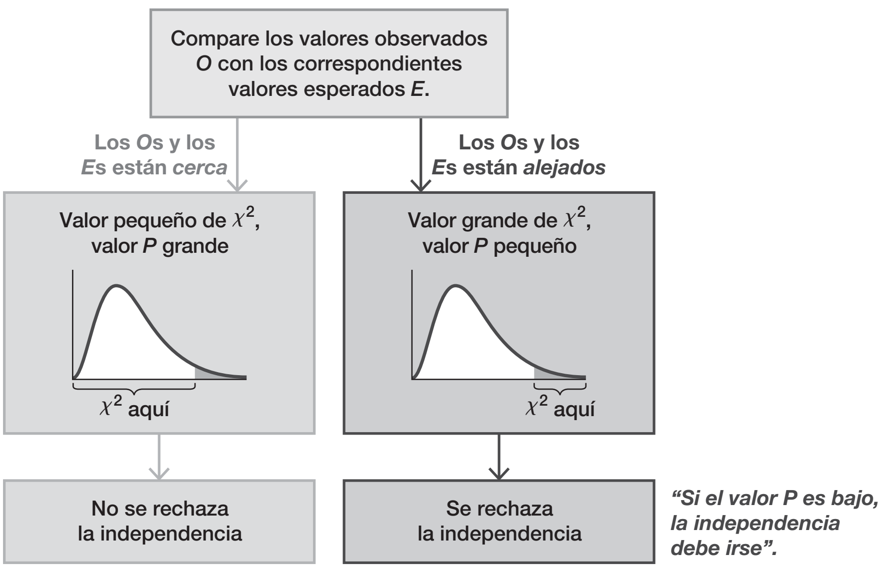


FIGURA 11-4 Relaciones entre componentes clave en una prueba de independencia

Determinación de los valores esperados *E*

Una frecuencia individual esperada *E* para una celda se puede encontrar simplemente al multiplicar el total de las frecuencias de fila por el total de las frecuencias de columna, para después dividir por el total de todas las frecuencias, como se muestra en el ejemplo 1.

$$E = \frac{\text{(total de fila) (total de columna)}}{\text{(gran total)}}$$

EJEMPLO 1 Determinación de la frecuencia esperada

La tabla 11-6 es una tabla de contingencia con cuatro filas y dos columnas. Las celdas de la tabla contienen conteos de frecuencias, que son los valores observados; por su parte, los valores esperados se muestran entre paréntesis. La variable de fila identifica el tratamiento utilizado para una fractura por esfuerzo en un hueso del pie y la variable de columna identifica el resultado como un éxito o un fracaso (según datos de "Surgery Unfounded for Tarsal Navicular Stress Fracture", de Bruce Jancin, *Internal Medicine News*, vol. 42, núm. 14). Consulte la tabla 11-6 y encuentre la frecuencia esperada para la celda en la primera fila y la primera columna, donde la frecuencia observada es 54.

SOLUCIÓN

TABLA 11-6 Tratamientos para fractura por esfuerzo en un hueso del pie

	Éxito	Fracaso
Cirugía	54 (E = 47.478)	12 (E = 18.522)
Yeso con carga de peso	41 (E = 66.182)	51 (E = 25.818)
Yeso sin carga de peso por 6 semanas	70 (E = 52.514)	3 (E = 20.486)
Yeso sin carga de peso por menos de 6 semanas	17 (E = 15.826)	5 (E = 6.174)

La primera celda se encuentra en la primera fila (con una frecuencia total de 66) y la primera columna (con una frecuencia total de 182). El "gran total" es la suma de todas las frecuencias en la tabla, que es 253. La frecuencia esperada de la primera celda es

$$E = \frac{\text{(total de fila) (total de columna)}}{\text{(gran total)}} = \frac{(66)\ (182)}{253} = 47.478$$

INTERPRETACIÓN

Sabemos que la primera celda tiene una *frecuencia observada* de $O = 54$ y una *frecuencia esperada* de $E = 47.478$. Podemos interpretar el valor esperado al afirmar que si suponemos que el éxito es independiente del tratamiento, entonces esperamos encontrar que 47.478 sujetos serían tratados con cirugía y que el tratamiento sería exitoso. Hay una discrepancia entre $O = 54$ y $E = 47.478$, y tales discrepancias son componentes clave del dato estadístico de prueba que es una medida colectiva de la discrepancia general entre las frecuencias observadas y las frecuencias esperadas con independencia entre las variables de fila y de columna.

SU TURNO Resuelva el inciso (a) del ejercicio 1 "Dominancia manual y uso del teléfono celular".

El ejemplo 2 ilustra el procedimiento para realizar una prueba de hipótesis de independencia entre las variables de fila y de columna en una tabla de contingencia.

Encuestas y psicólogos

Los resultados de las encuestas pueden verse dramáticamente afectados por la redacción de las preguntas. Una frase como "en los últimos años" es interpretada de manera diferente por las distintas personas. En los últimos años (en realidad, desde 1980), los investigadores de encuestas y los psicólogos han estado trabajando juntos para mejorar este tipo de estudios disminuyendo el sesgo y aumentando la precisión. En cierto caso, los psicólogos estudiaron el hallazgo de que entre el 10 el 15 por ciento de los encuestados dicen que votaron en las últimas elecciones cuando en realidad no lo hicieron. Experimentaron con teorías de la memoria defectuosa, el deseo de ser considerados responsables y la tendencia de quienes generalmente votan, a decir que votaron en las elecciones más recientes, incluso si no lo hicieron. Sólo se descubrió que la última teoría era parte del problema.

En cifras

$112,962: El salario anual que una madre que se queda en casa ganaría si le pagaran por su trabajo. 95: El número de horas trabajadas cada semana por la madre típica que se queda en casa.

EJEMPLO 2 ¿La elección del tratamiento para una fractura afecta el éxito?

Utilice los mismos datos muestrales del ejemplo 1 con un nivel de significancia de 0.05 para probar la afirmación de que el éxito del tratamiento es independiente del tipo de tratamiento. ¿Qué indica el resultado sobre la creciente tendencia a utilizar la cirugía?

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Con base en la descripción del estudio, trataremos los sujetos como aleatoriamente seleccionados y asignados al azar a los diferentes grupos de tratamiento. (2) Los resultados se expresan como conteos de frecuencias en la tabla 11-6. (3) Las frecuencias esperadas son todas de al menos 5. (La frecuencia más baja esperada es 6.174). Se cumplen los requisitos. ✓

Las hipótesis nula y alternativa son:

H_0: El éxito es independiente del tratamiento.

H_1: El éxito y el tratamiento son dependientes.

El nivel de significancia es $\alpha = 0.05$.

Debido a que los datos en la tabla 11-6 están en la forma de una tabla de contingencia, usamos la distribución χ^2 con el siguiente dato estadístico de prueba:

$$\chi^2 = \sum \frac{(O - E)^2}{E} = \frac{(54 - 47.478)^2}{47.478} + \cdots + \frac{(5 - 6.174)^2}{6.174}$$
$$= 58.393$$

XLSTAT

Chi-square (Observed value)	58.3933
Chi-square (Critical value)	7.8147
DF	3
p-value	< 0.0001
alpha	0.05

Valor *P* dado por la tecnología Si se usa una tecnología, los resultados suelen incluir el dato estadístico de prueba χ^2 y el valor P. Por ejemplo, vea la pantalla de XLSTAT adjunta que muestra el dato estadístico de prueba $\chi^2 = 58.393$ y el valor P es menor que 0.0001.

Valor *P* dado por la tabla A-4 Si se usa la tabla A-4 en lugar de la tecnología, primero se debe encontrar el número de grados de libertad: $(r - 1)(c - 1) = (4 - 1)(2 - 1) = 3$ grados de libertad. Dado que el dato estadístico de prueba $\chi^2 = 58.393$ excede el valor más alto (12.838) en la tabla A-4, para la fila correspondiente a 3 grados de libertad, sabemos que el valor $P < 0.005$.

Debido a que el valor P es menor que el nivel de significancia de 0.05, rechazamos la hipótesis nula de independencia entre el éxito y el tratamiento.

Valor crítico Si se usa el método del valor crítico para pruebas de hipótesis, el valor crítico de $\chi^2 = 7.815$ se encuentra en la tabla A-4, con $\alpha = 0.05$ en la cola derecha y el número de grados de libertad dados por $(r - 1)(c - 1) = (4 - 1)(2 - 1) = 3$. El dato estadístico de prueba y el valor crítico se muestran en la figura 11-5. Debido a que el dato estadístico de prueba cae en la región crítica, rechazamos la hipótesis nula de independencia entre el éxito y el tratamiento.

INTERPRETACIÓN

Parece que el éxito es dependiente del tratamiento. Aunque los resultados de esta prueba no nos dicen qué tratamiento es mejor, podemos ver en la tabla 11-6 que las tasas de éxito de 81.8%, 44.6%, 95.9% y 77.3% sugieren que el mejor tratamiento es usar un yeso sin peso durante 6 semanas. Estos resultados sugieren que el uso creciente de la cirugía es una estrategia de tratamiento que no está respaldada por la evidencia.

FIGURA 11-5 Prueba χ^2 de independencia

SU TURNO Resuelva el ejercicio 5 “Detector de mentiras”.

Justificación de las frecuencias esperadas *E* Para comprender mejor las frecuencias esperadas, considere que sólo conocemos los totales de filas y columnas en la tabla 11-6. Suponga además que las variables de fila y columna son independientes y que 1 de los 253 sujetos de estudio se selecciona al azar. La probabilidad de obtener a alguien de la primera celda de la tabla 11-6 se encuentra de la siguiente manera:

$$P(\text{cirugía}) = 66/253 \text{ y } P(\text{éxito}) = 182/253$$

Si las variables de fila y columna son independientes, como suponemos, es posible usar la regla de la multiplicación para eventos independientes (vea la sección 4-2) de la siguiente manera:

$$P(\text{tratamiento quirúrgico } y \text{ éxito}) = \frac{66}{253} \cdot \frac{182}{253} = 0.187661$$

Con una probabilidad de 0.187661 para la primera celda, esperamos que entre 253 sujetos, haya $253 \cdot 0.187661 = 47.478$ sujetos en la primera celda. Si generalizamos estos cálculos, obtenemos lo siguiente:

$$\text{Frecuencia esperada } E = (\text{gran total}) \cdot \frac{(\text{total de fila})}{(\text{gran total})} \cdot \frac{(\text{total de columna})}{(\text{gran total})}$$

Esta expresión se puede simplificar como

$$E = \frac{(\text{total de fila})\,(\text{total de columna})}{(\text{gran total})}$$

PARTE 2 Prueba de homogeneidad, prueba exacta de Fisher, prueba de McNemar para pares relacionados

Prueba de homogeneidad

En la parte 1 de esta sección, nos enfocamos en la prueba de *independencia* entre las variables de fila y de columna en una tabla de contingencia. En la parte 1, los datos muestrales provienen de una población, y los resultados muestrales individuales se categorizan con las

Alternativa a los ensayos clínicos

La reumatóloga Jennifer Frankovich diagnosticó a una paciente con lupus, pero notó una combinación específica de síntomas que habían provocado coágulos de sangre en el pasado. Sus colegas del Hospital para Niños Stanford Packard recomendaron no tratarla con medicamentos anticoagulantes, por lo que buscó algún estudio relacionado pero no pudo encontrar ninguno. Entonces recolectó los datos de todos los pacientes con lupus tratados en el hospital durante los últimos cinco años y usó estadísticas básicas para descubrir que su paciente tenía un mayor riesgo de desarrollar coágulos de sangre, por lo que procedió a tratarla con medicamentos anticoagulantes. Un ensayo clínico aleatorizado con grupos de tratamiento y placebo sería mejor, pero tales ensayos rara vez se realizan para complicaciones específicas.

Asientos más seguros en un jet comercial

Un estudio del escritor e investigador en aviación David Noland mostró que sentarse más atrás en un avión comercial aumentará sus posibilidades de sobrevivir en caso de un accidente. El estudio sugiere que la probabilidad de sobrevivir no es la misma para cada asiento, por lo que una prueba de bondad de ajuste llevaría al rechazo de la hipótesis nula de que cada asiento tiene la misma probabilidad de que un pasajero sobreviva. Se analizaron los registros de los 20 accidentes en jets comerciales ocurridos desde 1971. Se descubrió que quien se sienta en *business class* o primera clase, tiene 49% de probabilidad de sobrevivir a un accidente; quien se sienta junto al ala o delante de ella, tiene una probabilidad de 56% de sobrevivir; y si alguien se sienta detrás del ala, tiene un 69% de probabilidad de sobrevivir.

Al comentar sobre este estudio, David Noland declaró que no busca un asiento trasero cuando vuela. Afirma que debido a que la posibilidad de un accidente es tan pequeña, no se preocupa por dónde está sentado, pero prefiere un asiento junto a la ventana.

variables de fila y de columna. En una *prueba de homogeneidad con ji cuadrada*, tenemos muestras seleccionadas al azar de diferentes poblaciones, y queremos determinar si esas poblaciones tienen las mismas proporciones de alguna característica bajo consideración. (La palabra *homogénea* significa “tener la misma calidad” y, en este contexto, estamos probando para determinar si las proporciones son iguales). En la sección 9-1 se presentó un procedimiento para probar una afirmación sobre *dos* poblaciones con datos categóricos y dos posibles resultados, pero una prueba de homogeneidad con ji cuadrada nos permite usar dos o más poblaciones con resultados de varias categorías.

DEFINICIÓN

Una **prueba de homogeneidad con ji cuadrada** es una prueba de la afirmación de que *diferentes poblaciones* tienen las mismas proporciones de algunas características.

Muestreo de diferentes poblaciones En una prueba típica de independencia, como la descrita en la parte 1 de esta sección, los sujetos de muestra se seleccionan aleatoriamente de una población (como personas tratadas por fracturas de esfuerzo en un hueso del pie) y se observan los valores de dos variables diferentes (como el éxito/fracaso de las personas que reciben diferentes tratamientos). En una prueba típica de homogeneidad con ji cuadrada, los sujetos se seleccionan aleatoriamente de diferentes poblaciones y por separado.

Procedimiento Al realizar una prueba de homogeneidad, podemos usar los mismos requisitos, notación, dato estadístico de prueba, valor crítico y procedimientos dados en el cuadro de elementos clave de la parte 1 en la página 547 de esta sección; con una sola excepción: En vez de probar la hipótesis nula de independencia entre las variables de fila y de columna, probamos la hipótesis nula *de que las diferentes poblaciones tienen la misma proporción de alguna característica*.

EJEMPLO 3 Experimento de la billetera perdida

La tabla 11-7 lista los resultados de un experimento del *Reader’s Digest* en el que se perdieron intencionalmente 12 billeteras en cada una de 16 ciudades diferentes, incluidas Nueva York, Londres, Ámsterdam, etcétera. Use un nivel de significancia de 0.05 con los datos de la tabla 11-7 para probar la hipótesis nula de que las ciudades tienen la misma proporción de billeteras devueltas. El artículo del *Reader’s Digest* titulado “Most Honest Cities: The *Reader’s Digest* Lost Wallet Test” implica que la devolución de una billetera depende de la ciudad en la que se perdió. Pruebe la afirmación de que la proporción de billeteras devueltas no es la misma en las 16 diferentes ciudades.

TABLA 11-7 Experimento de la billetera perdida

Ciudad	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Billeteras devueltas	8	5	7	11	5	8	6	7	3	1	4	2	4	6	4	9
Billeteras no devueltas	4	7	5	1	7	4	6	5	9	11	8	10	8	6	8	3

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) De acuerdo con la descripción del estudio, trataremos a los sujetos como seleccionados aleatoriamente y asignados aleatoriamente a las diferentes ciudades. (2) Los resultados se expresan como conteos de frecuencias en la tabla 11-7. (3) Las frecuencias esperadas son todas de al menos 5. (Todos los valores esperados son 5.625 o 6.375). Se cumplen los requisitos. ✓

Las hipótesis nula y alternativa son:

H_0: La devolución de una billetera perdida es independiente de la ciudad en la que se perdió.

H_1: La devolución de una billetera perdida depende de la ciudad en la que se perdió.

La pantalla StatCrunch adjunta muestra el dato estadístico de prueba $\chi^2 = 35.388$ (redondeado) y el valor P de 0.002 (redondeado). Debido a que el valor P de 0.002 es menor que el nivel de significancia de 0.05, rechazamos la hipótesis nula de independencia entre las dos variables. ("Si el valor P es bajo, la nula debe irse").

StatCrunch

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	15	35.388235	0.0022

INTERPRETACIÓN

Rechazamos la hipótesis nula de independencia, por lo que parece que la proporción de billeteras devueltas depende de la ciudad en la que se perdieron. Hay suficiente evidencia para concluir que la proporción de billeteras devueltas no es la misma en las 16 ciudades.

Prueba exacta de Fisher

Los procedimientos para probar hipótesis con tablas de contingencia tienen el requisito de que cada celda debe tener una frecuencia esperada de al menos 5. Este requisito es necesario para que la distribución χ^2 sea una aproximación adecuada a la distribución exacta del dato estadístico de prueba χ^2. La *prueba exacta de Fisher* se usa a menudo para una tabla de contingencia de 2×2 con una o más frecuencias esperadas que están por debajo de 5. La prueba exacta de Fisher proporciona un valor *P exacto* y no requiere una técnica de aproximación. Debido a que los cálculos son bastante complejos, es una buena idea usar tecnología para abordar la prueba exacta de Fisher. Statdisk, Minitab, XLSTAT y StatCrunch tienen la capacidad de realizar la prueba exacta de Fisher.

EJEMPLO 4 ¿Bostezar hace que otros bostecen?

El programa "Cazadores de mitos" de Discovery Channel probó la teoría de que cuando alguien bosteza, es más probable que otros bostecen. Los resultados se resumen en la tabla 11-8. Los métodos de la parte 1 de esta sección no se deben usar porque una de las celdas tiene una frecuencia esperada de 4.480, lo que viola el requisito de que cada celda debe tener una frecuencia esperada E de al menos 5. Utilizando la prueba exacta de Fisher se obtiene un valor P de 0.513, por lo que no hay evidencia suficiente para respaldar el mito de que las personas expuestas a bostezos realmente bostezan más que las que no están expuestas a éstos. (Para probar la afirmación de que no hay diferencia, el valor P es 1.000, lo que indica que no hay una diferencia significativa entre los dos grupos).

TABLA 11-8 Experimento de la teoría de los bostezos

		¿Sujeto expuesto a bostezos?	
		Sí	No
¿El sujeto bostezó?	Sí	10	4
	No	24	12

Prueba de McNemar para pares relacionados

Los métodos de la parte 1 de esta sección se basan en datos independientes. Para tablas de 2×2 que constan de conteos de frecuencias resultantes de pares relacionados, los conteos de frecuencias dentro de cada par relacionado no son independientes y, para tales casos, podemos usar la prueba de McNemar para la hipótesis nula de que las frecuencias de las categorías discordantes (diferentes) ocurren en la misma proporción.

La tabla 11-9 muestra un formato general para resumir resultados de datos consistentes en conteos de frecuencias de pares relacionados. La tabla 11-9 se refiere a dos tratamientos diferentes (por ejemplo dos soluciones diferentes de gotas oculares) aplicados a dos partes diferentes de cada sujeto (por ejemplo el ojo izquierdo y el ojo derecho). Debemos tener cuidado al leer una tabla como la tabla 11-9. Si $a = 100$, entonces 100 sujetos se curaron con ambos tratamientos. Si $b = 50$ en la tabla 11-9, entonces cada uno de los 50 sujetos no se curaron con el tratamiento X pero sí se curaron con el tratamiento Y. El número total de sujetos es $a + b + c + d$, y cada uno de los sujetos arroja resultados de cada una de las dos partes de un par relacionado. Recuerde, las entradas en la tabla 11-9 son conteos de frecuencias de sujetos, no el número total de componentes individuales en los pares relacionados. Si 500 personas han tratado cada ojo con dos soluciones diferentes, el valor de $a + b + c + d$ es 500 (el número de sujetos), no 1000 (el número de ojos tratados).

TABLA 11-9 Tabla de 2 × 2 con conteos de frecuencias de pares relacionados

		Tratamiento *X*	
		Curado	No curado
Tratamiento *Y*	Curado	*a*	*b*
	No curado	*c*	*d*

La prueba de McNemar requiere que para una tabla como la tabla 11-9, las frecuencias sean tales que $b + c \geq 10$. La prueba es una prueba ji cuadrada de cola derecha, con el siguiente dato estadístico de prueba:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

Los valores P generalmente son proporcionados por el software, y los valores críticos se pueden encontrar en la tabla A-4 usando 1 grado de libertad. *Precaución:* Al aplicar la prueba de McNemar, tenga cuidado de usar sólo los dos conteos frecuencias de pares *discordantes* (diferentes), como la frecuencia b en la tabla 11-9 (con diferentes pares de curado/no curado) y la frecuencia c en la tabla 11-9 (con diferentes pares de no curado/curado).

EJEMPLO 5 ¿Los protectores de cadera son efectivos?

Se diseñó un ensayo aleatorizado y controlado para evaluar la efectividad de los protectores de cadera para la prevención de fracturas de cadera en los ancianos. Los residentes de las residencias de ancianos tenían protección en una cadera, pero no en la otra. Los resultados se resumen en la tabla 11-10 (basada en datos de "Efficacy of Hip Protector to Prevent Hip Fracture in Nursing Home Residents", de Kiel *et al.*, *Journal of the American Medical Association*, vol. 298, núm. 4). La prueba de McNemar puede usarse para probar la hipótesis nula de que las siguientes dos proporciones son iguales:

- La proporción de sujetos sin fractura en la cadera protegida y con fractura en la cadera desprotegida.
- La proporción de sujetos con fractura en la cadera protegida y sin fractura en la cadera desprotegida.

TABLA 11-10 Ensayo controlado aleatorizado de protectores de cadera

		Sin uso de protector de cadera	
		Sin fractura de cadera	Fractura de cadera
Uso de protector de cadera	Sin fractura de cadera	309	10
	Fractura de cadera	15	2

Usando los pares discordantes (diferentes) con el formato general de la tabla 11-9, tenemos $b = 10$ y $c = 15$, por lo que la estadística de prueba se calcula de la siguiente manera:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c} = \frac{(|10 - 15| - 1)^2}{10 + 15} = 0.640$$

Con un nivel de significancia de 0.05 y grados de libertad dados por gl = 1, nos referimos a la tabla A-4 para encontrar el valor crítico de $\chi^2 = 3.841$ para esta prueba de cola derecha. El dato estadístico de prueba de $\chi^2 = 0.640$ no excede el valor crítico de $\chi^2 = 3.841$, por lo que no se puede rechazar la hipótesis nula. (Además, el valor P es 0.424, que es mayor que 0.05, lo que indica que no podemos rechazar la hipótesis nula).La proporción de fracturas de cadera con el uso de protectores no es significativamente diferente de la proporción de fracturas de cadera sin el uso de protectores. Los protectores no parecen ser efectivos en la prevención de fracturas de cadera.

CENTRO DE TECNOLOGÍA

Tablas de contingencia

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Contingency Tables** en el menú desplegable.
3. Ingrese el nivel de significancia deseado y seleccione las columnas que se incluirán en el análisis.
4. Haga clic en **Evaluate**.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Tables** en el menú desplegable y seleccione **Chi-Square Test for Association**.
3. Seleccione **Summarized data in a two-way table** del cuadro desplegable.
4. Seleccione las columnas que contienen las frecuencias observadas.
5. Haga clic en **OK**.

SUGERENCIA: Las frecuencias observadas deben presentarse en columnas tal como aparecen en la tabla de contingencia.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Tables** en el menú desplegable, luego elija **Contingency—With Summary** en el submenú.
3. Seleccione las columnas de datos que se incluirán en el análisis.
4. Para las *etiquetas de fila*, seleccione la columna que contiene los nombres de fila.
5. Haga clic en **Compute!** El dato estadístico de prueba y el valor P se muestran en la parte inferior de los resultados.

SUGERENCIA: Debe ingresar los nombres de fila en la primera columna.

Calculadora TI-83/84 Plus

1. Ingrese los datos de contingencia como una matriz:

   Introducción manual de datos: Presione 2ND y luego x^{-1} si desea acceder al menú *MATRIX*, seleccione **EDIT** en el menú superior, elija una letra de matriz y presione ENTER. Ingrese el número de filas y columnas necesarias, presione ENTER y proceda a introducir los valores muestrales.

   Uso de listas existentes: Las listas se pueden combinar y almacenar en una matriz. Presione 2ND luego x^{-1} para ir al menú *MATRIX*, seleccione **MATH** en el menú superior y elija el elemento **List → matr**. Ingrese los nombres de lista (la primera lista debe contener valores para la variable dependiente *y*), seguidos por el nombre de la matriz, todos separados por comas. *Importante:* El nombre de la matriz se debe ingresar presionando 2ND, luego x^{-1}, seleccionando la letra de la matriz y presionando ENTER. El siguiente es un resumen de los comandos utilizados para crear una matriz a partir de tres listas (L1, L2, L3): **List → matr(L1, L2, L3,[D])**.
2. Presione STAT, luego seleccione **TESTS** en el menú superior.
3. Seleccione **χ^2-Test** en el menú y presione ENTER.
4. Para los datos *observados* ingrese la matriz creada en el paso 1 presionando 2ND y luego x^{-1} y seleccione la letra de la matriz. Los datos *esperados* muestran la matriz que se utilizará para almacenar automáticamente las frecuencias esperadas que se calculan.
5. Seleccione **Calculate** y presione ENTER.

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Correlation/Association tests**.
2. Seleccione **Tests on contingency tables** en el menú desplegable.
3. En el cuadro de *tabla de contingencia*, ingrese el rango de celdas que contiene los conteos de frecuencias de la tabla de contingencia. Si el rango incluye etiquetas de datos, marque la casilla de **Labels included**.
4. En el *formato de datos*, seleccione **Contingency table**.
5. Haga clic en la pestaña **Options**.
6. Marque la casilla de **Chi-square test** e ingrese un nivel de significancia.
7. Haga clic en **OK** para desplegar los resultados.

11-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

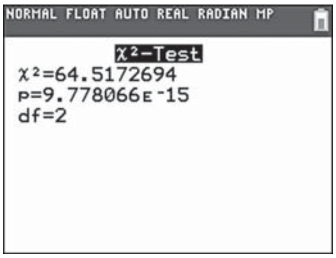
1. Predominio manual y uso del teléfono celular La tabla adjunta proviene de un estudio realizado con el objetivo declarado de abordar la seguridad del teléfono celular al comprender por qué utilizamos un oído en particular para el uso del teléfono celular. (Vea "Hemispheric Dominance and Cell Phone Use", de Seidman, Siegel, Shah y Bouyer, *JAMA Otolaryngology—Head & Neck Surgery*, vol. 139, núm. 5). El objetivo era determinar si la elección del oído está asociada con el predominio hemisférico cerebral auditivo o de lenguaje. Supongamos que queremos probar la afirmación de que el predominio manual y la preferencia de oído para usar el teléfono celular son independientes entre sí.

a. Use los datos en la tabla para encontrar el valor esperado para la celda que tiene una frecuencia observada de 3. Redondee el resultado a tres lugares decimales.

b. ¿Qué indica el valor esperado sobre los requisitos para la prueba de hipótesis?

Preferencia de oído para usar el teléfono celular

	Oído derecho	Oído izquierdo	Sin preferencia
Diestro	436	166	40
Zurdo	16	50	3



2. Hipótesis Consulte los datos proporcionados en el ejercicio 1 y suponga que todos los requisitos se satisfacen y queremos realizar una prueba de hipótesis de independencia utilizando los métodos de esta sección. Identifique las hipótesis nula y alternativa.

3. Prueba de hipótesis La pantalla adjunta de la calculadora TI-83/84 Plus es el resultado de la prueba de hipótesis descrita en el ejercicio 1. Suponga que se satisfacen todos los requisitos de la prueba de hipótesis. Identifique el dato estadístico de prueba y el valor *P* (expresado en forma estándar y redondeado a tres lugares decimales), y luego formule la conclusión sobre la hipótesis nula.

4. De cola derecha, de cola izquierda, de dos colas ¿La prueba de hipótesis descrita en el ejercicio 1 es de cola derecha, de cola izquierda o de dos colas? Explique su elección.

En los ejercicios 5 a 18, pruebe la afirmación dada.

5. Detector de mentiras La siguiente tabla incluye los resultados de experimentos con el polígrafo (detector de mentiras) realizados por los investigadores Charles R. Honts (Universidad Estatal de Boise) y Gordon H. Barland (Departamento de Poligrafía del Departamento de Defensa). En cada caso se supo si el sujeto mintió o no mintió, por lo que la tabla indica cuando la prueba del polígrafo fue correcta. Use un nivel de significancia de 0.05 para probar la afirmación de que la falsedad de un sujeto es independiente de la indicación de la prueba con el polígrafo. ¿Los resultados sugieren que los polígrafos son efectivos para distinguir entre verdades y mentiras?

	¿El sujeto realmente mintió?	
	No (no mintió)	Sí (mintió)
La prueba del polígrafo indicó que el sujeto mintió.	15	42
La prueba de polígrafo indicó que el sujeto no mintió.	32	9

6. ¿Férula o cirugía? Se diseñó un ensayo aleatorizado y controlado para comparar la efectividad de la férula contra la cirugía en el tratamiento del síndrome del túnel carpiano. Los resultados se dan en la siguiente tabla (con base en datos de "Splinting *vs*. Surgery in the Treatment of Carpal Tunnel Syndrome", de Gerritsen *et al*., *Journal of the American Medical Association*, vol. 288, núm. 10). Los resultados se basan en evaluaciones realizadas un año después del tratamiento. Utilizando un nivel de significancia de 0.01, pruebe la afirmación de que el éxito es independiente del tipo de tratamiento. ¿Qué sugieren los resultados sobre el tratamiento del síndrome del túnel carpiano?

	Tratamiento exitoso	Tratamiento sin éxito
Tratamiento con férula	60	23
Tratamiento con cirugía	67	6

7. Mensajes de texto y bebida En un estudio realizado con estudiantes de preparatoria de al menos 16 años de edad, los investigadores obtuvieron los resultados de una encuesta que se resumen en la tabla adjunta (basada en datos de "Texting While Driving and Other Risky Motor Vehicle Behaviors Among U.S. High School Students", de O'Malley, Shults y Eaton, *Pediatrics*, vol. 131, núm. 6). Use un nivel de significancia de 0.05 para probar la afirmación de independencia entre el envío de mensajes de texto al conducir y la conducción cuando se ha bebido alcohol. ¿Son esos dos comportamientos riesgosos independientes el uno del otro?

	¿Condujo cuando había bebido alcohol?	
	Sí	**No**
Envió mensajes de texto mientras conducía	731	3054
No envió mensajes de texto mientras conducía	156	4564

8. Mensajes de texto y uso de cinturones de seguridad En un estudio realizado con estudiantes de preparatoria de al menos 16 años de edad, los investigadores obtuvieron los resultados de encuesta resumidos en la tabla adjunta (según datos de "Texting While Driving and Other Risky Motor Vehicle Behaviors Among U.S. High School Students", de O'Malley, Shults y Eaton, *Pediatrics*, vol. 131, núm. 6). Use un nivel de significancia de 0.05 para probar la afirmación de independencia entre enviar mensajes de texto mientras se conduce y el uso irregular del cinturón de seguridad. ¿Son esas dos conductas riesgosas independientes entre sí?

	¿Uso irregular del cinturón de seguridad?	
	Sí	**No**
Envió mensajes de texto mientras conducía	1737	2048
No envió mensajes de texto mientras conducía	1945	2775

9. ¿Cuatro monedas de ¢0.25 son iguales que $1? En un estudio del "efecto de la denominación", a 43 estudiantes universitarios se les dio un dólar en forma de cuatro monedas de ¢0.25, mientras que a otros 46 estudiantes universitarios se les dio un dólar en forma de un billete de 1 dólar. A todos los estudiantes se les dieron dos opciones: (1) guardar el dinero; (2) gastar el dinero en goma de mascar. Los resultados se dan en la tabla adjunta (basada en "The Denomination Effect", de Priya Raghubir y Joydeep Srivastava, *Journal of Consumer Research*, vol. 36). Use un nivel de significancia de 0.05 para probar que, si los estudiantes compraron goma de mascar o mantuvieron el dinero es independiente de si recibieron cuatro monedas de ¢0.25 o un billete de $1. ¿Hay un "efecto de denominación"?

	Compró goma de mascar	Guardó el dinero
Estudiantes que recibieron cuatro monedas de ¢0.25	27	16
Estudiantes que recibieron un billete de $1	12	34

10. Regla del tiempo extra en el fútbol americano La tabla adjunta lista los resultados de los juegos de fútbol en tiempo extra antes y después de cambiar la regla correspondiente en la National Football League en 2011. Use un nivel de significancia 0.05 para probar la afirmación de independencia entre ganar un juego en tiempo extra y jugar bajo la regla anterior o la nueva regla. ¿Qué sugieren los resultados sobre la efectividad del cambio de la reglas?

	Antes del cambio de regla	Después del cambio de regla
El ganador del lanzamiento de moneda previo al tiempo extra ganó el juego	252	24
El ganador del lanzamiento de moneda previo al tiempo extra perdió el juego	208	23

11. Desafíos en el tenis La siguiente tabla muestra los resultados, desde 2006, de las decisiones arbitrales impugnadas en el Abierto de Estados Unidos. Use un nivel de significancia de 0.05 para probar la afirmación de que el género del tenista es independiente de si la decisión es revocada o no. ¿Los jugadores de algún género parecen ser mejores en el desafío de decisiones?

	¿El desafío a la decisión fue exitoso?	
	Sí	**No**
Hombres	161	376
Mujeres	68	152

12. ¿La enfermera es una asesina en serie? Las enfermeras de guardia del Veteran's Affairs Medical Center en Northampton, Massachusetts, notaron un número inusualmente alto de muertes cuando otra enfermera, Kristen Gilbert, estaba trabajando. Esas mismas enfermeras más tarde notaron suministros faltantes de la droga epinefrina, que es una adrenalina sintética que estimula el corazón. Kristen Gilbert fue arrestada y acusada de cuatro cargos de asesinato y dos de intento de asesinato. Cuando se buscaba una acusación para un gran jurado, los fiscales proporcionaron una pieza clave de evidencia que consistió en la siguiente tabla. Use un nivel de significancia de 0.01 para probar la afirmación de la defensa de que las muertes en los distintos turnos son independientes de si Gilbert estaba trabajando o no. ¿Qué sugiere el resultado sobre la culpabilidad o la inocencia de Gilbert?

	Turnos con una muerte	Turnos sin una muerte
Gilbert estaba trabajando	40	217
Gilbert no estaba trabajando	34	1350

13. Estrategia de fútbol En el fútbol soccer, las faltas graves en el área penal dan como resultado un penalti que un jugador ejecuta y es defendido por un portero. La siguiente tabla resume los resultados de 286 penaltis ejecutados durante los partidos entre varios de los mejores equipos (según datos de "Action Bias Among Elite Soccer Goalkeepers: The Case of Penalty Kicks", de Bar-Eli *et al*., *Journal of Economic Psychology*, vol. 28, núm. 5). En la tabla, la dirección del salto indica en qué dirección se movió el portero, donde la dirección del penalti se toma desde la perspectiva del portero. Utilice un nivel de significancia de 0.05 para probar la afirmación de que la dirección del penalti es independiente de la dirección en la que se mueve el portero. ¿Los resultados respaldan la teoría de que, debido a que los penaltis son tan rápidos, los porteros no tienen tiempo de reaccionar, por lo que las direcciones de sus movimientos son independientes de las direcciones de los penaltis?

	Salto de portero		
	Izquierda	Centro	Derecha
Penalti a la izquierda	54	1	37
Penalti al centro	41	10	31
Penalti a la derecha	46	7	59

14. ¿El uso del cinturón de seguridad es independiente del consumo de cigarrillos? Un estudio realizado con usuarios y no usuarios del cinturón de seguridad arrojó los datos muestrales seleccionados aleatoriamente que se resumen en la tabla dada (según datos de "What Kinds of People Do Not Use Seat Belts?", de Helsing y Comstock, *American Journal of Public Health*, vol. 67, núm. 11). Pruebe la afirmación de que la cantidad de cigarrillos consumidos es independiente del uso del cinturón de seguridad. Una teoría factible es que las personas que fuman más están menos preocupadas por su salud y seguridad y, por lo tanto, están menos inclinadas a usar cinturones de seguridad. ¿Esta teoría está respaldada por los datos muestrales?

	Cantidad de cigarrillos fumados por día			
	0	1–14	15–34	35 y más
Usan cinturones de seguridad	175	20	42	6
No usan cinturones de seguridad	149	17	41	9

15. Ensayo clínico con equinácea En un ensayo clínico de la eficacia de la equinácea para prevenir resfriados, se obtuvieron los resultados que se muestran en la tabla siguiente (según datos de "An Evaluation of Echinacea Angustifolia in Experimental Rhinovirus Infections", de Turner *et al*., *New England Journal of Medicine*, vol. 353, núm. 4). Use un nivel de significancia de 0.05 para probar si un resfriado es independiente del grupo de tratamiento. ¿Qué sugieren los resultados sobre la eficacia de la equinácea para prevenir los resfriados?

	Grupo de tratamiento		
	Placebo	Equinácea: Extracto al 20%	Equinácea: Extracto al 60%
Enfermó de resfriado	88	48	42
No enfermó de resfriado	15	4	10

16. Lesiones y color del casco de la motocicleta Se realizó un estudio de control de casos (o retrospectivo) para investigar la relación entre los colores de los cascos usados por los conductores de motocicletas y si resultaron heridos o murieron en un accidente. Los resultados se dan en la siguiente tabla (según datos de "Motorcycle Rider Conspicuity and Crash Related Injury: Case-Control Study", de Wells *et al*., *BMJ USA*, vol. 4). Pruebe la afirmación de que las lesiones son independientes del color del casco. ¿Deben los conductores de motocicletas elegir cascos con un color particular? En caso afirmativo, ¿qué color parece ser el mejor?

	Color del casco				
	Negro	Blanco	Amarillo/naranja	Rojo	Azul
Controles (sin lesiones)	491	377	31	170	55
Casos (heridos o muertos)	213	112	8	70	26

17. Rechazos en encuestas Un estudio de las personas que se negaron a responder preguntas de una encuesta proporcionó los datos muestrales seleccionados al azar que se presentan en la siguiente tabla (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). Con un nivel de significancia de 0.01, pruebe la afirmación de que la cooperación del sujeto (respuesta o rechazo) es independiente de la categoría de edad. ¿Algún grupo de edad en particular parece ser particularmente no cooperativo?

	Edad					
	18–21	22–29	30–39	40–49	50–59	60 y más
Respondió	73	255	245	136	138	202
Rechazó	11	20	33	16	27	49

18. Nacimientos de beisbolistas En su libro *Outliers*, el autor Malcolm Gladwell argumenta que más jugadores de béisbol nacidos en Estados Unidos tienen fechas de nacimiento en los meses inmediatamente posteriores al 31 de julio porque esa era la fecha del corte de edad para las ligas de béisbol no escolares. La siguiente tabla lista los meses de nacimiento de una muestra de jugadores de béisbol nacidos en Estados Unidos y beisbolistas nacidos en el extranjero. Con un nivel de significancia de 0.05, ¿hay suficiente evidencia para respaldar el rechazo de la afirmación de que los meses de nacimiento de los jugadores de béisbol son independientes de si nacieron en Estados Unidos? ¿Los datos parecen respaldar la afirmación de Gladwell?

	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sept.	Oct.	Nov.	Dic.
Nacido en EE.UU.	387	329	366	344	336	313	313	503	421	434	398	371
Nacido en el extranjero	101	82	85	82	94	83	59	91	70	100	103	82

19. Placas de automóvil California, Connecticut y Nueva York son estados con leyes que requieren que los automóviles tengan placas en las partes delantera y trasera. El autor seleccionó autos al azar en esos estados y los resultados se muestran en la tabla adjunta. Use un nivel de significancia de 0.05 para probar la afirmación de independencia entre el estado y si un auto tiene placas delanteras y traseras. ¿Parece que las leyes de las placas se siguen con las mismas tasas en los tres estados?

	California	Connecticut	Nueva York
Auto sólo con placa trasera	35	45	9
Auto con placas delantera y trasera	528	289	541

20. ¿La ventaja del local es independiente del deporte? Se recolectaron los datos del equipo ganador en juegos de diferentes deportes, con los resultados que se listan en la tabla de la parte superior de la página siguiente (según datos de "Predicting Professional Sports Game Outcomes from Intermediate Game Scores", de Copper, DeNeve y Mosteller, *Chance*, vol. 5, núm. 3-4). Use un nivel de significancia de 0.10 para probar la afirmación de que las victorias como local/visitante son independientes del deporte. Dado que entre los cuatro deportes incluidos aquí, el béisbol es el único en el que el equipo local puede modificar las dimensiones del campo para favorecer a sus propios jugadores, ¿parece que los equipos de béisbol son efectivos en el uso de esta ventaja?

continúa

	Básquetbol	Béisbol	Hockey sobre hielo	Fútbol americano
El equipo local gana	127	53	50	57
El equipo visitante gana	71	47	43	42

11-2 Más allá de lo básico

21. Pruebas equivalentes Una prueba χ^2 que involucre una tabla de 2×2 es equivalente a la prueba de la diferencia entre dos proporciones, como se describió en la sección 9-1. Use la afirmación y la tabla del ejercicio 9 "¿Cuatro monedas de ¢0.25 equivalen a $1?" verifique que los estadísticos de prueba χ^2 y z (que se encuentran para la prueba de igualdad de dos proporciones) estén relacionados de la siguiente manera: $z^2 = \chi^2$. También demuestre que los valores críticos tienen la misma relación.

22. Uso de la corrección de Yates para la continuidad La distribución ji cuadrada es continua, mientras que el dato estadístico de prueba utilizado en esta sección es discreto. Algunos estadísticos utilizan la *corrección de Yates para la continuidad* en celdas con una frecuencia esperada de menos de 10 o en todas las celdas de una tabla de contingencia con dos filas y dos columnas. Con la corrección de Yates, reemplazamos

$$\sum \frac{(O - E)^2}{E} \quad \text{por} \quad \sum \frac{(|O - E| - 0.5)^2}{E}$$

Dada la tabla de contingencia del ejercicio 9 "¿Cuatro monedas de ¢0.25 equivalen a $1?" encuentre el valor del dato estadístico de prueba χ^2 usando la corrección de Yates en todas las celdas. ¿Qué efecto tiene la corrección de Yates?

Examen rápido del capítulo

Los ejercicios 1 a 5 se refieren a los datos muestrales de la siguiente tabla, que resume los últimos dígitos de las estaturas (cm) de 300 sujetos seleccionados al azar (del conjunto de datos 1 "Datos corporales" en el apéndice B). Suponga que queremos usar un nivel de significancia de 0.05 para probar la afirmación de que los datos provienen de una población que tiene la propiedad de que los últimos dígitos son igualmente probables.

Último dígito	0	1	2	3	4	5	6	7	8	9
Frecuencia	30	35	24	25	35	36	37	27	27	24

1. ¿Cuáles son las hipótesis nula y alternativa que corresponden a la afirmación declarada?

2. Al probar la afirmación del ejercicio 1, ¿cuáles son las frecuencias observadas y esperadas para el último dígito 7?

3. ¿La prueba de hipótesis es de cola izquierda, de cola derecha o de dos colas?

4. Si se usa un nivel de significancia de 0.05 para probar la afirmación declarada, encuentre el número de grados de libertad.

5. Dado que el valor P para la prueba de hipótesis es 0.501. ¿Qué se puede concluir? ¿Parece que las estaturas se obtuvieron por medición o que los sujetos reportaron sus estaturas?

Las preguntas 6 a 10 se refieren a los datos muestrales en la siguiente tabla, que describe el destino de los pasajeros y la tripulación a bordo del* Titanic *cuando se hundió el 15 de abril de 1912. Supongamos que los datos son una muestra de una gran población y queremos usar un nivel de significancia de 0.05 para probar la afirmación de que la supervivencia es independiente de si la persona es un hombre, una mujer, un niño o una niña.

	Hombres	Mujeres	Niños	Niñas
Sobrevivió	332	318	29	27
Murió	1360	104	35	18

6. Identifique las hipótesis nula y alternativa correspondientes a la afirmación declarada.

7. ¿Qué distribución se usa para probar la afirmación declarada (normal, *t*, *F*, ji cuadrada, uniforme)?

8. ¿La prueba de hipótesis es de cola izquierda, de cola derecha o de dos colas?

9. Encuentre el número de grados de libertad.

10. Dado que el valor *P* para la prueba de hipótesis es 0.000 cuando se redondea a tres lugares decimales, ¿qué se puede concluir? ¿Qué indican los resultados sobre la regla de que las mujeres y los niños deben ser los primeros en ser salvados?

Ejercicios de repaso

1. Muertes en automóviles La siguiente tabla lista las muertes en automóviles los días de la semana de un año reciente (según datos del Instituto de Seguros para la Seguridad en las Carreteras). Use un nivel de significancia de 0.01 para probar la afirmación de que las muertes en auto ocurren en los diferentes días de la semana con la misma frecuencia. Proporcione una explicación de los resultados.

Día	Dom.	Lun.	Mar.	Mie.	Jue.	Vie.	Sáb.
Frecuencia	5304	4002	4082	4010	4268	5068	5985

2. Empastes dentales La siguiente tabla muestra los resultados de un estudio en el que algunos pacientes fueron tratados con empastes que contienen mercurio y otros con empastes que no lo contienen (según datos de “Neuropsychological and Renal Effects of Dental Amalgam in Children”, de Bellinger *et al*., *Journal of the American Medical Association*, vol. 295, núm. 15). Use un nivel de significancia de 0.05 para probar la independencia entre el tipo de empaste y la presencia de condiciones de salud adversas. ¿Los empastes que contienen mercurio parecen afectar las condiciones de salud?

	Empaste con mercurio	Empaste sin mercurio
Reporte de condición de salud adversa	135	145
Sin reporte de condición de salud adversa	132	122

3. *American Idol* Los concursantes en el programa de televisión *American Idol* competían para ganar un concurso de canto. En su momento, el sitio web WhatNotToSing.com listó el número real de eliminaciones para diferentes posiciones en el orden de presentación de los cantantes, y también mencionó el número esperado de eliminaciones. Los resultados se dan en la siguiente tabla. Use un nivel de significancia de 0.05 para probar la afirmación de que las eliminaciones reales concuerdan con los números esperados. ¿Parece haber respaldo para la afirmación de que los cantantes iniciales parecen estar en desventaja?

Orden de presentación	1	2	3	4	5	6	7–12
Eliminaciones reales	20	12	9	8	6	5	9
Eliminaciones esperadas	12.9	12.9	9.9	7.9	6.4	5.5	13.5

4. Ensayo clínico de Lipitor Lipitor es el nombre comercial del medicamento atorvastatina, que se usa para reducir el colesterol en pacientes. (Hasta que expiró su patente en 2011, era el medicamento más vendido en el mundo, con ventas anuales de 13 mil millones de dólares). Las reacciones adversas al medicamento se han estudiado en ensayos clínicos y la tabla siguiente resume los resultados de infecciones en pacientes de diferentes grupos de tratamiento (según datos de Parke-Davis). Use un nivel de significancia de 0.01 para probar la afirmación de que contraer una infección es independiente del tratamiento. ¿El tratamiento con atorvastatina (Lipitor) parece tener un efecto sobre las infecciones?

	Placebo	Atorvastatina 10 mg	Atorvastatina 40 mg	Atorvastatina 80 mg
Infección	27	89	8	7
Sin infección	243	774	71	87

5. Muertes relacionadas con el clima En un año reciente, el número de muertes relacionadas con el clima en Estados Unidos cada mes fue de 28, 17, 12, 24, 88, 61, 104, 32, 20, 13, 26, 25 (listados en orden iniciando con enero). Use un nivel de significancia de 0.01 para probar la afirmación de que las muertes relacionadas con el clima ocurren en los diferentes meses con la misma frecuencia. Proporcione una explicación para el resultado.

Ejercicios de repaso acumulado

1. Muertes relacionadas con el clima El ejercicio de repaso 5 involucró las muertes relacionadas con el clima en Estados Unidos. Entre las 450 muertes incluidas en ese ejercicio, 320 eran hombres. Use un nivel de significancia de 0.05 para probar la afirmación de que entre los fallecidos en relación con el clima, el porcentaje de hombres es igual a 50%. Proporcione una explicación de los resultados.

2. Chocolate y felicidad En una encuesta patrocinada por la compañía de chocolate Lindt, 1708 mujeres fueron encuestadas y 85% de ellas dijo que el chocolate las hacía más felices.

a. ¿Hay algo potencialmente erróneo en esta encuesta?

b. De las 1708 mujeres encuestadas, ¿cuál es la cantidad que dijo que el chocolate las hacía más felices?

3. Chocolate y felicidad Use los resultados del inciso (b) del ejercicio de repaso acumulado 2 para elaborar un cálculo del intervalo de confianza del 99% para el porcentaje de mujeres que dicen que el chocolate las hace más felices. Escriba un breve enunciado que interprete el resultado.

4. Chocolate y felicidad Use los resultados del inciso (b) del ejercicio de repaso acumulado 2 para probar la afirmación de que, cuando se les pregunta, más del 80% de las mujeres dicen que el chocolate las hace más felices. Use un nivel de significancia de 0.01.

5. Un gran billete o muchos billetes pequeños En un estudio del "efecto de la denominación", 150 mujeres en China recibieron un solo billete de 100 yuanes o un total de 100 yuanes en billetes más pequeños. El valor de 100 yuanes es de aproximadamente $15. A las mujeres se les dio la opción de gastar el dinero en artículos específicos o conservar el dinero. Los resultados se resumen en la siguiente tabla (con base en "The Denomination Effect", de Priya Raghubir y Joydeep Srivastava, *Journal of Consumer Research*, vol. 36). Use un nivel de significancia de 0.05 para probar la afirmación de que la forma en que se dan los 100 yuanes es independiente de si el dinero se gastó o no. ¿Qué sugiere el resultado sobre el efecto de la denominación?

	Gastó el dinero	Conservó el dinero
Mujeres que recibieron un solo billete de 100 yuanes	60	15
Mujeres que recibieron 100 yuanes en billetes más pequeños	68	7

6. Probabilidad Consulte los resultados de los 150 sujetos en el ejercicio de repaso acumulado 5.

a. Encuentre la probabilidad de que si 1 de los 150 sujetos se selecciona al azar, el resultado sea una mujer que gastó el dinero.

b. Encuentre la probabilidad de que si 1 de los 150 sujetos se selecciona al azar, el resultado sea una mujer que gastó el dinero o que recibió un solo billete de 100 yuanes.

c. Si dos mujeres diferentes son seleccionadas al azar, encuentre la probabilidad de que ambas hayan gastado el dinero.

7. Costos de reparación de automóviles A continuación se listan los costos de reparación (en dólares) de automóviles que chocaron a 6 mi/h en pruebas de choque completamente frontal y del mismo número de automóviles que chocaron a 6 mi/h en pruebas de choque completamente trasero (según datos del *Instituto de Seguros para la Seguridad en Carreteras*). Los autos son el Toyota Camry, Mazda 6, Volvo S40, Saturn Aura, Subaru Legacy, Hyundai Sonata y Honda Accord. ¿Hay suficiente evidencia para concluir que existe una correlación lineal entre los costos de reparación de los choques completamente frontales y traseros?

Frontal	936	978	2252	1032	3911	4312	3469
Trasero	1480	1202	802	3191	1122	739	2767

8. Alcance del agarre frontal y ergonomía Al diseñar automóviles y aviones, debemos considerar el alcance del agarre frontal de las mujeres. Las mujeres poseen alcances de agarre distribuidos normalmente con una media de 686 mm y una desviación estándar de 34 mm (según datos de una encuesta antropométrica de Gordon, Churchill, *et al*.).

a. Si el tablero de instrumentos de un automóvil se coloca de manera que pueda ser alcanzado por 95% de las mujeres. ¿Cuál es el alcance de agarre frontal más corto que se puede permitir en el tablero?

b. Si el tablero de instrumentos de un automóvil se coloca de manera que las mujeres puedan alcanzarlo con un alcance de agarre superior a 650 mm, ¿qué porcentaje de mujeres no pueden alcanzar el tablero? ¿Es ese porcentaje demasiado alto?

c. Encuentre la probabilidad de que 16 mujeres elegidas al azar tengan alcances de agarre frontal con una media mayor a 680 mm. ¿Este resultado tiene algún efecto en el diseño?

Proyecto de tecnología

Use Statdisk, Minitab, Excel, StatCrunch, una calculadora TI-83/84 Plus o cualquier otro paquete de software o calculadora capaz de generar dígitos aleatorios igualmente probables entre 0 y 9 inclusive. Genere 5000 dígitos y registre los resultados en la tabla adjunta. Use un nivel de significancia de 0.05 para probar la afirmación de que los dígitos de la muestra provienen de una población con una distribución uniforme (de modo que todos los dígitos son igualmente probables). ¿El generador de números aleatorios parece estar funcionando como debería?

Dígito	0	1	2	3	4	5	6	7	8	9
Frecuencia										

Statdisk: Seleccione **Data**, luego **Uniform Generator**.

Minitab: Seleccione **Calc**, **Random Data**, **Integer**.

Excel: Haga clic en **Insert function** f_x, luego seleccione la categoría **Math & Trig** y la función **RANDBETWEEN**. Haga clic y arrastre la celda hacia abajo por la columna para generar números aleatorios adicionales.

TI-83/84 Plus: Presione **MATH**, seleccione **PROB**, luego use la función **randInt** con el formato de **randInt** (***inferior***, ***superior***, ***n***).

StatCrunch: Seleccione **Data**, **Simulate**, **Discrete Uniform**.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿Se equivocó Allstate?

En cierta ocasión, la compañía de seguros Allstate emitió un comunicado de prensa en el que listaba los signos del zodíaco junto con las cantidades correspondientes de accidentes automovilísticos, como se muestra en la primera y última columnas de la siguiente tabla.

En el comunicado de prensa original, Allstate incluyó comentarios como uno que indica que los Virgo se preocupan mucho y son tímidos, y que estuvieron involucrados en 211,650 accidentes, lo que los convertía en los peores infractores. Allstate rápidamente emitió una disculpa y una retractación. En un comunicado de prensa, Allstate incluyó lo siguiente: "Los signos astrológicos no tienen absolutamente ningún papel en la forma en que basamos la cobertura y establecemos las tasas. La clasificación por astrología no sería actuarialmente sólida".

Signo del zodíaco	Fechas	Longitud (días)	Accidentes
Capricornio	Ene. 18–Feb. 15	29	128,005
Acuario	Feb. 16–Mar. 11	24	106,878
Piscis	Mar. 12–Abr. 16	36	172,030
Aries	Abr. 17–May. 13	27	112,402
Tauro	May. 14–Jun. 19	37	177,503
Géminis	Jun. 20–Jul. 20	31	136,904
Cáncer	Jul. 21–Ago. 9	20	101,539
Leo	Ago. 10–Sept. 15	37	179,657
Virgo	Sept. 16–Oct. 30	45	211,650
Libra	Oct. 31–Nov. 22	23	110,592
Escorpión	Nov. 23–Nov. 28	6	26,833
Ofiuco	Nov. 29–Dic. 17	19	83,234
Sagitario	Dic. 18–Ene. 17	31	154,477

Análisis de los resultados

El comunicado de prensa original de Allstate no incluía las longitudes (días) de los diferentes signos del zodíaco. La tabla anterior lista esas longitudes en la tercera columna. Una explicación razonable para las diferentes cantidades de accidentes es que deben ser proporcionales a las longitudes de los signos del zodíaco. Por ejemplo, las personas nacen bajo el signo de Capricornio en 29 de los 365 días del año, por lo que se espera que tengan 29/365 del número total de accidentes. Use los métodos de este capítulo para determinar si esto parece explicar los resultados en la tabla. Escriba un breve informe de sus descubrimientos.

Actividades en equipo

1. Actividad fuera de clase Divídanse en grupos de cuatro o cinco estudiantes. El problema del capítulo señaló que, de acuerdo con la ley de Benford, una variedad de conjuntos de datos diferentes incluye números donde los primeros dígitos siguen la distribución que se muestra en la siguiente tabla. Recolecte datos originales y utilice los métodos de la sección 11-1 para respaldar o refutar la afirmación de que los datos se ajustan razonablemente a la ley de Benford. A continuación se dan algunas sugerencias: (1) los dígitos iniciales de las contraseñas de teléfonos inteligentes; (2) los dígitos iniciales de los precios de acciones; (3) los dígitos iniciales de los números de amigos en Facebook; (4) los dígitos iniciales de la longitud de los ríos en el mundo.

Dígito principal	1	2	3	4	5	6	7	8	9
Ley de Benford	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%

2. Actividad fuera de clase Divídanse en grupos de cuatro o cinco estudiantes y recolecte resultados pasados de una lotería estatal. Tales resultados suelen estar disponibles en los sitios web de cada lotería estatal. Use los métodos de la sección 11-1 para probar la afirmación de que los números se seleccionan de modo que todos los posibles resultados son igualmente probables.

3. Actividad fuera de clase Divídanse en grupos de cuatro o cinco estudiantes. Cada miembro del grupo debe encuestar a por lo menos 15 estudiantes varones y 15 estudiantes mujeres en la misma universidad, y formularles dos preguntas: (1) ¿Qué partido político es el más favorecido por el sujeto? (2) Si el sujeto fuera a inventar una excusa de ausencia de un neumático desinflado, ¿qué neumático diría que se desinfló si el profesor le preguntara? (Consulte el ejercicio 6 en la sección 11-1). Pídale al sujeto que escriba las dos respuestas en una tarjeta índice, también registre el sexo del sujeto y si éste escribió con la mano derecha o con la izquierda. Use los métodos del presente capítulo para analizar los datos recopilados. Incluya las siguientes afirmaciones:

- Las cuatro opciones posibles para un neumático desinflado se seleccionan con la misma frecuencia.
- El neumático que se identifica como desinflado es independiente del género del sujeto.
- La elección del partido político es independiente del género del sujeto.
- La elección del partido político es independiente de si el sujeto es diestro o zurdo.
- El neumático que se identifica como desinflado es independiente de si el sujeto es diestro o zurdo.
- El género es independiente de si el sujeto es diestro o zurdo.
- La elección del partido político es independiente del neumático que se identifica como desinflado.

4. Actividad fuera de clase Divídanse en grupos de cuatro o cinco estudiantes. Cada miembro debe seleccionar aproximadamente a otros 15 estudiantes y primero pedirles que seleccionen "aleatoriamente" cuatro dígitos cada uno. Después de que los cuatro dígitos hayan sido registrados, pida a cada sujeto que escriba los últimos cuatro dígitos de su número de Seguridad Social (para mayor seguridad, escriba estos dígitos en cualquier orden). Tome los resultados de la muestra "aleatoria" de los dígitos individuales y mézclelos en una muestra grande, luego mezcle los dígitos individuales del número del Seguro Social en una segunda muestra grande. Usando el conjunto de muestra "aleatorio", pruebe la afirmación de que los estudiantes seleccionan dígitos aleatoriamente. Luego use los dígitos del seguro social para probar la afirmación de que provienen de una población de dígitos aleatorios. Compare los resultados. ¿Parece que los estudiantes pueden seleccionar dígitos al azar? ¿Es probable que seleccionen cualquier dígito con más frecuencia que otros? ¿Es probable que seleccionen cualquier dígito con menos frecuencia que otros? ¿Los últimos dígitos de los números del seguro social parecen estar seleccionados aleatoriamente?

5. Actividad en la clase Divídanse en grupos de tres o cuatro estudiantes. A cada grupo se le debe dar un dado junto con la instrucción de que debe probar su "legalidad". ¿El dado es legal o está cargado? Describa el análisis y los resultados.

6. Actividad fuera de clase Divídanse en grupos de dos o tres estudiantes. En ocasiones, el análisis de los últimos dígitos de los datos puede revelar si los valores resultaron de mediciones reales o si se trata de estimaciones reportadas. Encuentre las longitudes de los ríos en el mundo, luego analice los últimos dígitos para determinar si esas longitudes parecen ser mediciones reales o si parecen ser estimaciones reportadas. En lugar de longitudes de los ríos, podría usar otras variables, como las siguientes:

- Alturas de montañas
- Alturas de los edificios más elevados
- Longitudes de puentes
- Alturas de montañas rusas

12

ANÁLISIS DE VARIANZA

PROBLEMA DEL CAPÍTULO

¿La exposición al plomo afecta las puntuaciones de IQ de los niños?

Un importante estudio relacionado con la salud y el medio ambiente incluyó a niños que vivían a menos de 7 km (alrededor de 4 millas) de una gran fundidora de mineral en El Paso, Texas. En estas instalaciones se funde el mineral con la finalidad de separar los metales que lo componen. Debido a que la fundidora emitía contaminación con plomo, existía la preocupación de que estos niños de alguna manera se vieran afectados. El objetivo de este capítulo es investigar el posible efecto de la exposición al plomo en las puntuaciones del IQ de “desempeño” medidas mediante la escala de inteligencia de Wechsler. (Una puntuación de IQ completa es una combinación de una puntuación de IQ de desempeño y una puntuación de IQ verbal. La prueba de desempeño incluye componentes como el análisis de imágenes, de la disposición de imágenes y de patrones coincidentes).

Los datos del estudio se incluyen en el conjunto de datos 7 “IQ y plomo” del apéndice B. Según los niveles de plomo medidos en la sangre, los niños se dividieron en grupos con niveles bajo, medio y alto de plomo. (Consulte el conjunto de datos 7 para conocer los valores de corte específicos del nivel de plomo en la sangre).

TABLA 12-1 Puntuaciones del IQ de desempeño en niños

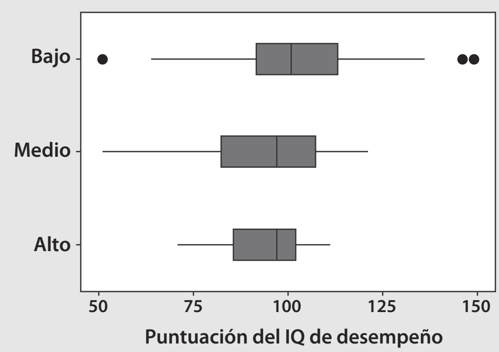
Nivel bajo de plomo en la sangre															
85	90	107	85	100	97	101	64	111	100	76	136	100	90	135	104
149	99	107	99	113	104	101	111	118	99	122	87	118	113	128	121
111	104	51	100	113	82	146	107	83	108	93	114	113	94	106	92
79	129	114	99	110	90	85	94	127	101	99	113	80	115	85	112
112	92	97	97	91	105	84	95	108	118	118	86	89	100		
Nivel medio de plomo en la sangre															
78	97	107	80	90	83	101	121	108	100	110	111	97	51	94	80
101	92	100	77	108	85										
Nivel alto de plomo en la sangre															
93	100	97	79	97	71	111	99	85	99	97	111	104	93	90	107
108	78	95	78	86											

Las puntuaciones del IQ de desempeño se incluyen en la tabla 12-1 (con base en datos de "Neuropsychological Dysfunction in Children with Chronic Low-Level Lead Absorption", de P. J. Landrigan, R. H. Whitworth, R. W. Baloh, N. W. Staehling, W. F. Barthel y B. F. Rosenblum, *Lancet*, vol. 1, artículo 7909).

Antes de saltar a la aplicación de un método estadístico particular, debemos explorar los datos. En la siguiente tabla se presentan los datos estadísticos muestrales. Vea también los diagramas de caja y bigotes de los tres conjuntos de puntuaciones del IQ de desempeño. Las comparaciones informales y subjetivas muestran que el grupo bajo tiene una media un poco mayor que las medias de los grupos medio y alto. Los diagramas de caja y bigotes se superponen, por lo que las diferencias no parecen ser grandes. Pero necesitamos más métodos formales que nos permitan reconocer cualquier diferencia significativa. Podríamos usar los métodos de la sección 9-2 para comparar las medias de las muestras recolectadas a partir de *dos* poblaciones diferentes, pero aquí tenemos que comparar las medias de muestras recolectadas a partir de *tres* poblaciones. Cuando tenemos muestras de tres o más poblaciones, podemos probar la igualdad de las medias poblacionales utilizando el método *análisis de varianza*, que se presentará en la sección 12-1. En la primera sección del capítulo, usaremos el análisis de varianza para probar la afirmación de que las tres muestras provienen de poblaciones con la misma media.

	Nivel bajo de plomo en la sangre	Nivel medio de plomo en la sangre	Nivel alto de plomo en la sangre
Tamaño de la muestra n	78	22	21
$\bar{x}$	102.7	94.1	94.2
s	16.8	15.5	11.4
Distribución	Aproximadamente normal	Aproximadamente normal	Aproximadamente normal
Valores atípicos	Potencial valor atípico bajo de 51 y posibles valores atípicos altos de 146 y 149, pero no están muy lejos de los demás valores de datos.	Ninguno	Ninguno

Diagramas de caja y bigotes del IQ de desempeño con Minitab



OBJETIVOS DEL CAPÍTULO

La sección 9-2 incluye métodos para probar la igualdad de medias de *dos* poblaciones independientes, pero este capítulo presenta un método para probar la igualdad de *tres o más* medias poblacionales. Los objetivos del capítulo son:

12-1 ANOVA de un factor

- Aplicar el método de análisis de varianza de un factor para realizar una prueba de hipótesis de igualdad de tres o más medias poblacionales. El enfoque de esta sección está en la interpretación de los resultados proporcionados por la tecnología.

12-2 ANOVA de dos factores

- Analizar datos muestrales provenientes de poblaciones separadas en categorías, utilizando dos características (o factores), como el sexo y el color de los ojos.
- Aplicar el método de análisis de varianza de dos factores a: (1) pruebas para una ***interacción*** entre dos factores, (2) pruebas para un efecto del factor de ***fila***, y (3) pruebas para un efecto del factor de ***columna***. El enfoque de esta sección está en la interpretación de los resultados proporcionados por la tecnología.

12-1 ANOVA de un factor

Concepto clave En esta sección presentamos el método del *análisis de varianza de un factor*, que se utiliza para realizar pruebas de hipótesis de que tres o más poblaciones tienen medias que son todas iguales, como en H_0: $\mu_1 = \mu_2 = \mu_3$. Debido a que los cálculos son muy complicados, hacemos énfasis en la interpretación de los resultados obtenidos mediante el uso de la tecnología.

Distribución *F*

Los métodos de análisis de varianza (ANOVA) de este capítulo requieren la distribución *F*, que se presentó en la sección 9-4. En esa sección notamos que la distribución *F* tiene las siguientes propiedades (vea la figura 12-1):

Hay una distribución *F* diferente para cada par diferente de grados de libertad en el numerador y en el denominador.

1. La distribución *F* no es simétrica. Está sesgada a la derecha.
2. Los valores de la distribución *F* no pueden ser negativos.
3. La forma exacta de la distribución *F* depende de los dos diferentes grados de libertad.

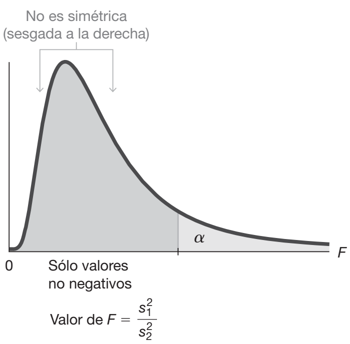


FIGURA 12-1 Distribución *F*

PARTE 1 Conceptos básicos del análisis de varianza de un factor

Cuando pruebe la igualdad de tres o más medias poblacionales, use el método del análisis de varianza de un factor.

DEFINICIÓN

El **análisis de varianza de un factor (ANOVA)** es un método para probar la igualdad de tres o más medias poblacionales mediante el análisis de varianzas muestrales. El análisis de varianza de un factor se utiliza con datos categorizados con *un* **factor** (o **tratamiento**), por lo que hay una característica que se usa para separar los datos muestrales en diferentes categorías.

El término *tratamiento* se usa porque las primeras aplicaciones del análisis de varianza involucraron experimentos agrícolas en los que diferentes parcelas de tierras agrícolas fueron tratadas con diferentes fertilizantes, tipos de semillas, insecticidas, etcétera. La tabla 12-1 usa el "tratamiento" (o factor) del nivel de plomo en la sangre. Ese factor tiene tres categorías: nivel de plomo en la sangre bajo, medio y alto (como se define en el conjunto de datos 7 del apéndice B).

ELEMENTOS CLAVE

Análisis de varianza de un factor para probar la igualdad de tres o más medias poblacionales

Objetivo

Usar muestras de tres o más poblaciones diferentes para probar la afirmación de que todas las poblaciones tienen la misma media.

Requisitos

1. Las poblaciones tienen distribuciones que son aproximadamente normales. (Este es un requisito poco estricto, porque el método funciona bien a menos que una población tenga una distribución muy alejada de la normal. Si una población tiene una distribución que está lejos de ser normal, use la prueba de Kruskal-Wallis descrita en la sección 13-5).
2. Las poblaciones tienen la misma varianza σ^2 (o la misma desviación estándar σ). Este es un requisito poco estricto, porque el método funciona bien a menos que las varianzas poblacionales difieran en grandes cantidades. El estadístico George E. P. Box mostró que mientras los tamaños de muestra sean iguales (o casi iguales), la varianza más grande puede ser hasta nueve veces la varianza más pequeña y los resultados del ANOVA continuarán siendo esencialmente confiables.
3. Las muestras son muestras aleatorias simples de datos cuantitativos.
4. Las muestras son independientes entre sí. (Las muestras no están relacionadas o pareadas de ninguna forma).
5. Las diferentes muestras son de poblaciones que se encuentran categorizadas de una sola manera.

Procedimiento para probar H_0: $\mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$

1. Use la tecnología para obtener resultados que incluyan el dato estadístico de prueba y el valor P.
2. Identifique el valor P en la pantalla. (La prueba de ANOVA es de cola derecha porque sólo los valores grandes del dato estadístico de prueba hacen que rechacemos la igualdad de las medias poblacionales).
3. Obtenga una conclusión con base en los siguientes criterios que usan el nivel de significancia α:
   - **Rechazar:** Si el valor $P \le \alpha$, rechace la hipótesis nula de la igualdad de medias y concluya que al menos una de las medias poblacionales es diferente de las demás.
   - **No rechazar:** Si el valor $P > \alpha$, no rechace la hipótesis nula de la igualdad de medias.

Debido a que los cálculos requeridos para el análisis de varianza de un factor son complicados, recomendamos utilizar la tecnología con la siguiente estrategia de estudio:

1. Entender que un valor *P* pequeño (como 0.05 o menos) conduce al rechazo de la hipótesis nula de la igualdad de medias. (“Si el valor *P* es bajo, la nula debe irse”). Con un valor *P* grande (por ejemplo mayor que 0.05), no se rechaza la hipótesis nula de la igualdad de medias.
2. Desarrolle una comprensión de los fundamentos subyacentes estudiando los ejemplos que se presentan en esta sección.

EJEMPLO 1 Plomo y puntuaciones del IQ de desempeño

Use las puntuaciones del IQ de desempeño listadas en la tabla 12-l y un nivel de significancia de $\alpha = 0.05$ para probar la afirmación de que las tres muestras provienen de poblaciones con medias que son todas iguales.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Con base en las tres muestras listadas en la tabla 12-1, las tres poblaciones parecen tener distribuciones que son aproximadamente normales, como lo indican las gráficas cuantilares normales. (2) Las tres muestras en la tabla 12-1 tienen desviaciones estándar que no son muy diferentes, por lo que las tres varianzas poblacionales parecen ser aproximadamente las mismas. (3) Con base en el diseño del estudio, podemos tratar las muestras como muestras aleatorias simples. (4) Las muestras son independientes entre sí; los puntajes del IQ de desempeño no se corresponden de ninguna manera. (5) Las tres muestras son de poblaciones categorizadas de acuerdo con el factor único del nivel de plomo (bajo, medio, alto). Los requisitos se satisfacen. ✓

La hipótesis nula y la hipótesis alternativa son:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1: Al menos una de las medias es diferente de las demás

El nivel de significancia es $\alpha = 0.05$.

Paso 1: Use la tecnología para obtener los resultados de ANOVA, por ejemplo una de las que se muestran en las pantallas adjuntas.

Statdisk

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Treatment:	2	2022.729906	1011.364953	4.071122	3.073087	0.01951
Error:	118	29314.046953	248.424127			
Total:	120	31336.77686				

Minitab

One-way ANOVA: Low, Medium, High

Source	DF	SS	MS	F	P
Factor	2	2023	1011	4.07	0.020
Error	118	29314	248		
Total	120	31337			

S = 15.76 R-Sq = 6.45% R-Sq(adj) = 4.87%

StatCrunch

ANOVA table

Source	df	SS	MS	F-Stat	P-value
Treatments	2	2022.7299	1011.3649	4.071122	0.0195
Error	118	29314.047	248.42413		
Total	120	31336.777			

Excel

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2022.729906	2	1011.364953	4.071122103	0.019510383	3.073090341
Within Groups	29314.04695	118	248.4241267			
Total	31336.77686	120				

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
      One-way ANOVA
 F=4.071122103
 p=.0195103826
 Factor
  df=2
  SS=2022.72991
  MS=1011.36495
 Error
↓ df=118
```

SPSS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2022.730	2	1011.365	4.071	.020
Within Groups	29314.047	118	248.424		
Total	31336.777	120			

JMP

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	2022.730	1011.36	4.0711
Error	118	29314.047	248.42	Prob > F
C. Total	120	31336.777		0.0195*

Paso 2: Además del dato estadístico de prueba $F = 4.0711$, todas las pantallas muestran que el valor P es 0.020 después de redondear.

Paso 3: Como el valor P de 0.020 es menor que el nivel de significancia de $\alpha = 0.05$, rechazamos la hipótesis nula de la igualdad de medias. (Si el valor P es bajo, la nula debe irse).

INTERPRETACIÓN

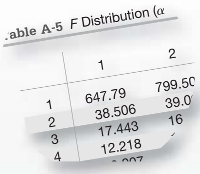
Existe suficiente evidencia para justificar el rechazo de la afirmación de que las tres muestras provienen de poblaciones con medias que son todas iguales. Usando las muestras de mediciones listadas en la tabla 12-1, concluimos que esos valores provienen de poblaciones que tienen medias que no son todas iguales. Sobre la base de esta prueba de ANOVA, no podemos concluir que una media particular sea diferente de las otras, pero podemos observar informalmente que la media muestral para el grupo con bajo nivel de plomo en la sangre es más alta que la media para los grupos con niveles de plomo en la sangre medio y alto. Parece que los mayores niveles de plomo en la sangre se asocian con puntuaciones menores del IQ de desempeño.

SU TURNO Resuelva el ejercicio 5 “Plomo y puntuaciones del IQ verbal”.

PRECAUCIÓN Cuando concluimos que hay suficiente evidencia para rechazar la afirmación de igualdad de las medias poblacionales, no podemos concluir por ANOVA que cualquier media particular es diferente de las demás. (Hay varios otros métodos que pueden usarse para identificar las medias específicas que son diferentes, y algunos de ellos se estudian en la parte 2 de esta sección).

¿Cómo se relaciona el valor *P* con el dato estadístico de prueba? Los valores más *grandes* del dato estadístico de prueba dan como resultado valores P más *pequeños*, por lo que la prueba ANOVA es de cola derecha. La figura 12-2 en la página siguiente muestra la relación entre el dato estadístico de prueba F y el valor P. Suponiendo que las poblaciones tienen la misma varianza σ^2 (como se requiere para la prueba), el estadístico de la prueba F es la relación de estas dos estimaciones de σ^2: (1) variación *entre* las muestras (basada en la variación entre medias muestrales); y (2) variación *dentro* de las muestras (basada en las varianzas muestrales).

¿Por qué 0.05?



En 1925, R. A. Fisher publicó un libro que presentaba el método del análisis de varianza y necesitaba una tabla de valores críticos basada en los grados de libertad del numerador y los grados de libertad del denominador, como en la tabla A-5 del apéndice A. Debido a que la tabla utiliza dos grados de libertad diferentes, se vuelve muy larga si se utilizan muchos valores críticos diferentes, por lo que Fisher incluyó una tabla usando solamente 0.05. En una edición posterior también incluyó el nivel de significancia de 0.01.

Stephen Stigler, un notable historiador de la estadística, escribió en la revista *Chance* que la elección de un nivel de significancia de 0.05 es un número redondo conveniente que resulta algo arbitrario. Aunque sea arbitrario, el valor de 0.05 cumple los siguientes objetivos importantes. (1) Un nivel de significancia de 0.05 da como resultado tamaños de muestra que son razonables y no demasiado grandes. (2) La elección de 0.05 es suficientemente grande como para proporcionar una posibilidad razonable de identificar efectos importantes (rechazando correctamente una hipótesis nula de no efecto cuando realmente hay un efecto). (3) La elección de 0.05 no es tan pequeña como para que nos obliga a perdernos efectos importantes (al cometer el error de no rechazar una hipótesis nula de no efecto cuando realmente lo hay).

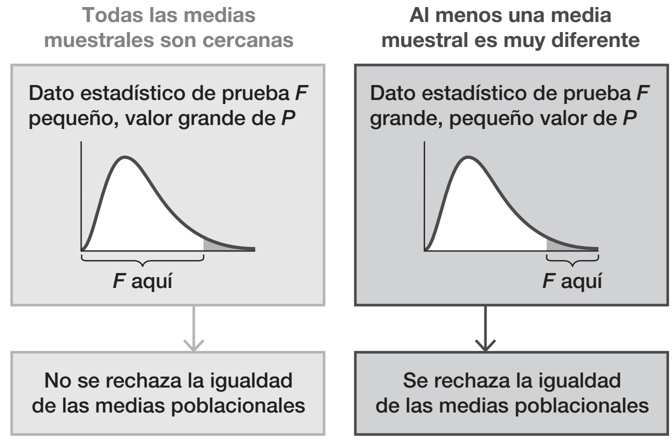


FIGURA 12-2 Relación entre el dato estadístico de prueba *F* y el valor *P*

Dato estadístico de prueba para ANOVA de un factor: $F = \dfrac{\text{varianza entre las muestras}}{\text{varianza dentro de las muestras}}$

El numerador del dato estadístico de prueba F mide la variación entre las medias muestrales. La estimación de la varianza en el denominador depende sólo de las varianzas muestrales y no se ve afectada por las diferencias entre las medias de las muestras. En consecuencia, las medias muestrales que tienen valores cercanos entre sí, resultan en un dato estadístico de prueba F pequeño y un valor P grande, por lo que concluimos que no hay diferencias significativas entre las medias muestrales. La medias muestrales que están muy separadas entre sí producen un dato estadístico de prueba F grande y un valor P pequeño, por lo que rechazamos la afirmación de la igualdad de medias.

¿Por qué no probar sólo dos muestras a la vez? Si deseamos probar la igualdad entre tres o más medias poblacionales. ¿Por qué necesitamos un nuevo procedimiento cuando podemos probar la igualdad de dos medias usando los métodos presentados en la sección 9-2? Por ejemplo, si queremos utilizar los datos muestrales de la tabla 12-1 para probar la afirmación de que las tres poblaciones tienen la misma media, ¿por qué no simplemente las emparejamos y probamos dos a la vez probando H_0: $\mu_1 = \mu_2$, H_0: $\mu_2 = \mu_3$, y H_0: $\mu_1 = \mu_3$? Para los datos de la tabla 12-1, el método de probar la igualdad de dos medias a la vez requiere tres pruebas de hipótesis diferentes. Si usamos un nivel de significancia de 0.05 para cada una de esas tres pruebas de hipótesis, el nivel de confianza general real podría ser tan bajo como 0.95^3 (o 0.857). En general, a medida que aumentamos el número de pruebas individuales de significancia, aumentamos el riesgo de encontrar una diferencia sólo por casualidad (en lugar de una diferencia real en las medias). El riesgo de un error tipo I —encontrar una diferencia en uno de los pares cuando en realidad no existe tal diferencia— es demasiado alto. El método del análisis de varianza nos ayuda a evitar ese inconveniente particular (rechazar una hipótesis nula verdadera) mediante el uso de *una prueba* para la igualdad de varias medias, en lugar de varias pruebas donde cada una compara dos medias a la vez.

PRECAUCIÓN Cuando se prueba la igualdad de tres o más poblaciones, use el análisis de varianza. (Usar múltiples pruebas de hipótesis con dos muestras a la vez podría afectar adversamente el nivel de significancia).

PARTE 2 Cálculos e identificación de medias que son diferentes

Cálculo del dato estadístico de prueba *F* con tamaños de muestra *n* iguales

La tabla 12-2 puede ser muy útil para comprender los métodos de ANOVA. En la tabla 12-2, compare el conjunto de datos A con el conjunto de datos B para ver que el conjunto de datos

A es igual que el conjunto de datos B con la siguiente excepción notable: los valores de Muestra 1 difieren cada uno en 10. Si todos los conjuntos de datos tienen el mismo tamaño de muestra (por ejemplo $n = 4$ para la tabla 12-2), los siguientes cálculos no son demasiado difíciles, como se muestra aquí.

TABLA 12-2 Efecto de una media en el dato estadístico de prueba F

Agregue 10 a los datos en la Muestra 1

	Conjunto de datos A			Conjunto de datos B		
	Muestra 1	Muestra 2	Muestra 3	Muestra 1	Muestra 2	Muestra 3
	7	6	4	17	6	4
	3	5	7	13	5	7
	6	5	6	16	5	6
	6	8	7	16	8	7
	$n_1 = 4$	$n_2 = 4$	$n_3 = 4$	$n_1 = 4$	$n_2 = 4$	$n_3 = 4$
	$\bar{x}_1 = 5.5$	$\bar{x}_2 = 6.0$	$\bar{x}_3 = 6.0$	$\bar{x}_1 = \mathbf{15.5}$	$\bar{x}_2 = 6.0$	$\bar{x}_3 = 6.0$
	$s_1^2 = 3.0$	$s_2^2 = 2.0$	$s_3^2 = 2.0$	$s_1^2 = 3.0$	$s_2^2 = 2.0$	$s_3^2 = 2.0$
	Conjunto de datos A			Conjunto de datos B		
Paso 1: Varianza *entre* las muestras	$ns_{\bar{x}}^2 = 4(0.0833) = 0.3332$			$ns_{\bar{x}}^2 = 4(30.0833) = 120.3332$		
Paso 2: Varianza *dentro* de las muestras	$s_p^2 = \frac{3.0 + 2.0 + 2.0}{3} = 2.3333$			$s_p^2 = \frac{3.0 + 2.0 + 2.0}{3} = 2.3333$		
Paso 3: Dato estadístico de prueba F	$F = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{0.3332}{2.3333} = \mathbf{0.1428}$			$F = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{120.3332}{2.3333} = \mathbf{51.5721}$		
Valor P	Valor P = **0.8688**			Valor P = **0.0000118**		

Paso 1: Encuentra la varianza entre muestras

Cálculo de la varianza *entre* muestras evaluando $ns_{\bar{x}}^2$ 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 que las medias muestrales son un conjunto ordinario de valores y calcule la varianza. (Del teorema del límite central, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ puede resolverse para σ para obtener $\sigma = \sqrt{n} \cdot \sigma_{\bar{x}}$, de modo que podamos estimar σ^2 con $ns_{\bar{x}}^2$). Por ejemplo, las medias muestrales para el conjunto de datos A en la tabla l2-2 son 5.5, 6.0 y 6.0, y estos tres valores tienen una varianza de $s_{\bar{x}}^2 = 0.0833$, por lo que

$$\text{varianza entre las muestras} = ns_{\bar{x}}^2 = 4(0.0833) = 0.3332$$

Paso 2: Determinación de la varianza dentro de las muestras

Estime la varianza *dentro* de las muestras calculando s_p^2, que es la varianza combinada obtenida al encontrar la media de las varianzas muestrales. Las varianzas muestrales en la tabla 12-2 son 3.0, 2.0 y 2.0, de modo que

$$\text{varianza dentro de las muestras} = s_p^2 = \frac{3.0 + 2.0 + 2.0}{3} = 2.3333$$

Paso 3: Cálculo del dato estadístico de prueba

Evalúe el dato estadístico de prueba F de la siguiente manera:

$$F = \frac{\text{varianza entre las muestras}}{\text{varianza dentro de las muestras}} = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{0.3332}{2.3333} = 0.1428$$

Determinación del valor crítico

El valor crítico F se encuentra suponiendo una prueba de cola derecha porque los valores grandes de F corresponden a diferencias significativas entre las medias. Con k muestras que tienen n valores, los números de grados de libertad son los siguientes.

Grados de libertad (usando k = número de muestras y n = tamaño de muestra)

$$\text{Grados de libertad del numerador} = k - 1$$

$$\text{Grados de libertad del denominador} = k(n - 1)$$

Para el conjunto de datos A en la tabla 12-2, $k = 3$ y $n = 4$, por lo que los grados de libertad son 2 para el numerador y $3(4 - 1) = 9$ para el denominador. Con $\alpha = 0.05$, 2 grados de libertad para el numerador, y 9 grados de libertad para el denominador, el valor crítico F de la tabla A-5 es 4.2565. Si tuviéramos que usar el método del valor crítico de la prueba de hipótesis con el conjunto de datos A en la tabla 12-2, veríamos que esta prueba de cola derecha tiene un dato estadístico de prueba $F = 0.1428$ y un valor crítico de $F = 4.2565$, por lo que el dato estadístico de prueba no está en la región crítica. Así que no podemos rechazar la hipótesis nula de la igualdad de medias.

Comprensión del efecto de una media sobre el dato estadístico de prueba *F* Para entender realmente cómo funciona el método del análisis de varianza, considere el conjunto de datos A y el conjunto de datos B en la tabla 12-2 y tenga en cuenta lo siguiente.

- Las tres muestras en el conjunto de datos A son idénticas a las tres muestras en el conjunto de datos B, a excepción de lo siguiente: cada valor en la muestra 1 del conjunto de datos B es 10 más que el valor correspondiente en el conjunto de datos A.
- Sumar 10 a cada valor de datos en la primera muestra del conjunto de datos A tiene un efecto significativo en el dato estadístico de prueba, donde F cambia de 0.1428 a 51.5721.
- Sumar 10 a cada valor de datos en la primera muestra del conjunto de datos A tiene un efecto notorio en el valor P, que cambia de 0.8688 (no significativo) a 0.0000118 (significativo).
- Las tres *medias* muestrales en el conjunto de datos A (5.5, 6.0, 6.0) están muy cerca, pero las medias muestrales en el conjunto de datos B (15.5, 6.0, 6.0) no están cerca.
- Las tres varianzas muestrales en el conjunto de datos A son idénticas a las del conjunto de datos B.
- La *varianza entre las muestras* en el conjunto de datos A es 0.3332, pero para el conjunto de datos B es 120.3332 (lo que indica que las medias muestrales en B están más separadas).
- La *varianza dentro de las muestras* es 2.3333 tanto en el conjunto de datos A como en el conjunto de datos B, porque la varianza *dentro* de una muestra no se ve afectada cuando sumamos una constante a cada valor muestral. *El cambio en el dato estadístico de prueba F y el valor P sólo se puede atribuir al cambio en* $\overline{x}_1$. Esto ilustra el punto clave que subyace al método del análisis de varianza de un factor:

  El dato estadístico de prueba F es muy sensible a las *medias* muestrales, aunque se obtiene a través de dos estimaciones diferentes de la *varianza* poblacional común.

Cálculos con tamaños de muestra desiguales

Si bien los cálculos para casos con tamaños de muestra iguales son algo razonables, se vuelven mucho más complicados cuando los tamaños de muestra no son todos iguales, pero se aplica el mismo razonamiento básico. En lugar de proporcionar las fórmulas desordenadas pertinentes que se requieren para casos con tamaños de muestra desiguales, suponemos prudente y convenientemente que la tecnología debe usarse para obtener el valor P para el análisis de varianza. No nos vemos complicados por cálculos complejos y podemos centrarnos en verificar los requisitos y en interpretar los resultados.

Calculamos un dato estadístico de prueba F que es la razón de dos estimaciones diferentes de la varianza poblacional común σ^2. Con tamaños de muestra desiguales, debemos usar medidas *ponderadas* que tengan en cuenta los tamaños de muestra. El dato estadístico de prueba es esencialmente el mismo que se dio anteriormente y su interpretación también es igual a la descrita previamente.

En cifras

$5816: Costo adicional en el que incurre un empleador privado cada año, atribuible a un empleado que fuma. Ese total incluye los costos de las pausas para fumar y los costos de atención médica debido a la gran cantidad de problemas de salud que sufren los fumadores.

Diseño de experimentos

Con el análisis de varianza de un factor (o de factor único), usamos un factor como base para dividir los datos en diferentes categorías. Si concluimos que las diferencias entre las medias son significativas, no podemos estar absolutamente seguros de que las diferencias puedan explicarse por el factor utilizado. Es posible que la variación de algún otro factor desconocido sea responsable. Una forma de reducir el efecto de los factores externos es diseñar el experimento de forma que éste sea **completamente aleatorio**, donde cada valor muestral tiene la misma probabilidad de pertenecer a los diferentes grupos de factores. Por ejemplo, es posible asignar sujetos a dos grupos de tratamiento diferentes y a un tercer grupo placebo a través de un proceso de selección aleatoria equivalente a escoger papelitos de un tazón. Otra forma de reducir el efecto de los factores externos es usar un diseño **rigurosamente controlado**, en el que los valores de las muestras se eligen cuidadosamente para que los demás factores no tengan variabilidad. En general, los buenos resultados requieren que el experimento sea cuidadosamente diseñado y ejecutado.

Identificación de las medias que son diferentes

Después de realizar una prueba de análisis de varianza, podemos concluir que hay suficiente evidencia para rechazar una afirmación de igualdad de medias poblacionales, pero no podemos concluir por ANOVA que cualquier media en *particular* sea diferente de las demás. Existen varios procedimientos formales e informales que se pueden usar para identificar las medias específicas que son diferentes. A continuación se presentan dos métodos *informales* para comparar las medias:

- Construir diagramas de caja y bigotes de las diferentes muestras y examinar cualquier superposición para ver si uno o más de los diagramas de caja son muy diferentes de los demás.
- Construir estimaciones del intervalo de confianza de las medias para cada una de las diferentes muestras, luego comparar esos intervalos de confianza para ver si uno o más de ellos no se superponen con los demás.

Existen varios procedimientos formales para identificar las medias que son diferentes. Algunas de las pruebas, llamadas **pruebas de rango**, nos permiten identificar subconjuntos de medias que no son significativamente diferentes entre sí. Otras pruebas, llamadas **pruebas de comparación múltiple**, usan pares de medias, pero hacen ajustes para superar el problema de tener un nivel de significancia que aumenta a medida que se incrementa el número de pruebas individuales. No hay consenso sobre qué prueba es mejor, pero algunas de las más comunes son la prueba de Duncan, prueba de Student-Newman-Keuls (o prueba de SNK), prueba de Tukey (o prueba de diferencia honestamente significativa de Tukey), prueba de Scheffé, prueba de Dunnett, prueba de diferencia mínima significativa, y prueba de Bonferroni. Consideremos la prueba de Bonferroni para ver un ejemplo de comparación múltiple. El procedimiento es el siguiente.

Prueba de comparación múltiple de Bonferroni

Paso 1: Haga una prueba t por separado para cada par de muestras, pero realice los ajustes descritos en los siguientes pasos.

continúa

Paso 2: Para una estimación de la varianza σ^2 que es común a todas las poblaciones involucradas, use el valor de MS(error), que utiliza todos los datos muestrales disponibles. El valor de MS(error) generalmente se proporciona con los resultados al realizar la prueba de análisis de varianza. Usando el valor de MS(error), calcule el valor del dato estadístico de prueba t, como se muestra a continuación. El dato estadístico de prueba particular calculado a continuación se basa en la elección de la Muestra 1 y la Muestra 2; cambie los subíndices y use otro par de muestras hasta que todos los diferentes pares posibles de muestras hayan sido probados.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\text{MS(error)} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Paso 3: Después de calcular el valor del dato estadístico de prueba t para un par particular de muestras, encuentre el valor crítico t o el valor P, pero realice el siguiente ajuste para que el nivel de significancia general no aumente.

Valor *P* Utilice el dato estadístico de prueba t con gl $= N - k$, donde N es el número total de valores de muestra y k es el número de muestras, y encuentre el valor P usando la tecnología o la tabla A-3, pero ajuste el valor P multiplicándolo por el número de diferentes emparejamientos posibles de dos muestras. (Por ejemplo, con tres muestras, hay tres emparejamientos posibles diferentes, así que ajuste el valor P multiplicándolo por 3).

Valor crítico Al encontrar el valor crítico, ajuste el nivel de significancia dividiéndolo entre el número de emparejamientos posibles diferentes de dos muestras. (Por ejemplo, con tres muestras, hay tres emparejamientos posibles diferentes, así que ajuste el nivel de significancia dividiéndolo por 3).

Tenga en cuenta que en el paso 3 del procedimiento anterior de Bonferroni se realiza una prueba individual con un nivel de significancia mucho más bajo, o bien se aumenta mucho el valor P. El rechazo de la igualdad de medias a ese respecto requiere diferencias mucho más alejadas. Este ajuste en el paso 3 compensa el hecho de que estamos haciendo varias pruebas en lugar de sólo una.

EJEMPLO 2 Prueba de Bonferroni

El ejemplo 1 en esta sección utilizó el análisis de varianza con los datos muestrales de la tabla 12-1. Concluimos que hay evidencia suficiente para justificar el rechazo de la afirmación de igualdad de medias. Use la prueba de Bonferroni con un nivel de significancia de 0.05 para identificar qué media es diferente de las demás.

SOLUCIÓN

La prueba de Bonferroni requiere una prueba t por separado para cada uno de los tres posibles pares de muestras posibles. Las hipótesis nulas que deben ser probadas son:

$$H_0\colon \mu_1 = \mu_2 \qquad H_0\colon \mu_1 = \mu_3 \qquad H_0\colon \mu_2 = \mu_3$$

Comenzamos con $H_0\colon \mu_1 = \mu_2$. Usando los datos muestrales dados en la tabla 12-1 y conservando algunos decimales adicionales para una mayor precisión en los cálculos, tenemos $n_1 = 78$ y $\bar{x}_1 = 102.705128$. Además, $n_2 = 22$ y $\bar{x}_2 = 94.136364$. A partir de los resultados de la tecnología mostrados en el ejemplo 1, también sabemos que MS(error) = 248.424127.

Ahora evaluamos el dato estadístico de prueba usando las media muestrales no redondeadas:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\text{MS(error)} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{102.705128 - 94.136364}{\sqrt{248.424127 \cdot \left(\frac{1}{78} + \frac{1}{22}\right)}} = 2.252$$

El número de grados de libertad es gl $= N - k = 121 - 3 = 118$. ($N = 121$ porque hay 121 valores muestrales diferentes en las tres muestras combinadas, y $k = 3$ porque hay tres muestras diferentes). Con un dato estadístico de prueba de $t = 2.252$ y con gl $= 118$, el valor P de dos colas es 0.026172, pero ajustamos este valor P multiplicándolo por 3 (el número de diferentes pares posibles de muestras) para obtener un valor P final de 0.078516 o 0.079 cuando se redondea. Debido a que este valor P no es pequeño (menos de 0.05) no podemos rechazar la hipótesis nula. Parece que las muestras 1 y 2 no tienen medias significativamente diferentes.

En lugar de continuar con pruebas de hipótesis separadas para los otros dos emparejamientos, consulte la pantalla de SPSS que muestra todos los resultados de las pruebas de Bonferroni. En estos resultados, los niveles bajos de plomo están representados por 1, los niveles medios están representados por 2, y los niveles altos están representados por 3. (La primera fila de resultados numéricos corresponde a los resultados encontrados aquí, vea el valor de 0.079, que se calculó previamente). La pantalla muestra que el emparejamiento de bajo/alto produce un valor P de 0.090, o no hay una diferencia significativa entre los promedios bajo y alto del nivel de sangre. Además, la pantalla SPSS muestra que el emparejamiento de medio/alto produce un valor P de 1.000, por lo que no hay una diferencia significativa entre los promedios de los niveles de plomo medio y alto.

Resultados de SPSS para la prueba Bonferroni

(I) Level	(J) Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	8.56876	3.80486	.079	-.6717	17.8092
	3.00	8.51465	3.87487	.090	-.8958	17.9251
2.00	1.00	-8.56876	3.80486	.079	-17.8092	.6717
	3.00	-.05411	4.80851	1.000	-11.7320	11.6238
3.00	1.00	-8.51465	3.87487	.090	-17.9251	.8958
	2.00	.05411	4.80851	1.000	-11.6238	11.7320

INTERPRETACIÓN

Aunque la prueba de análisis de varianza nos dice que al menos una de las medias es diferente de las demás, los resultados de las pruebas de Bonferroni no identifican ninguna media muestral particular que sea significativamente diferente. En el artículo original que analiza estos resultados, los autores afirman que "nuestros hallazgos indican que una absorción crónica de plomo particulado... puede resultar en un deterioro sutil pero estadísticamente significativo en las habilidades motoras cognitivas y perceptivas no verbales medidas por la escala de pruebas de inteligencia del desempeño de Wechsler". Esta declaración confirma los siguientes resultados: a partir del análisis de varianza sabemos que al menos una media es diferente de las otras, pero la prueba de Bonferroni no identificó una media particular como significativamente diferente [aunque las medias muestrales de 102.7 (bajo nivel de plomo en la sangre), 94.1 (nivel medio de plomo en la sangre) y 94.2 (alto nivel de plomo en la sangre) sugieren que los niveles medio y alto de plomo en la sangre parecen estar asociados con puntajes más bajos en el IQ de desempeño que el grupo con bajo nivel de plomo en la sangre.

SU TURNO Resuelva el ejercicio 18 "Prueba de Bonferroni".

En cifras

Desde el comienzo de la especie humana, han nacido aproximadamente 100 mil millones de seres humanos. Hay aproximadamente 7 mil millones de personas vivas ahora, por lo que aproximadamente 7% de todos los humanos todavía están vivos.

CENTRO DE TECNOLOGÍA

Análisis de varianza de un factor

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **One-Way Analysis of Variance** en el menú desplegable.
3. Ingrese el nivel de significancia deseado y seleccione al menos 3 columnas para incluir en el análisis.
4. Haga clic en **Evaluate**.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **ANOVA** en el menú desplegable y elija **One-Way** en el submenú.
3. Seleccione **Response data are in a separate column for each factor level**.
4. En el cuadro de *Respuesta*, seleccione las columnas que se incluirán en el análisis.
5. Haga clic en el botón **Options** y marque la casilla **Assume equal variances**.
6. Haga clic en **OK** dos veces.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **ANOVA** en el menú desplegable, luego elija **One Way** en el submenú.
3. Seleccione las columnas que se incluirán en el análisis.
4. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Presione **STAT** y luego seleccione **TESTS** en el menú superior.
2. Seleccione **ANOVA** en el menú y presione **ENTER**.
3. Ingrese los nombres de las listas que incluyen los datos que se usarán en el análisis. Separe los nombres de las listas con **,**, de modo que el comando aparezca en el formato **ANOVA(LI, L2, L3)**.
4. Presione **ENTER** y use los botones de flecha para desplazarse por los resultados.

Excel

Complemento XLSTAT

Requiere que todos los datos se apilen en una sola columna con el nombre de categoría correspondiente para cada valor de datos en una columna separada.

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Modeling Data**.
2. Seleccione **ANOVA** del menú desplegable.
3. Ingrese el rango de celdas que contienen los valores de datos de la *variable dependiente y*.
4. Seleccione el cuadro **Qualitative** e ingrese el rango de celdas que contienen los valores cualitativos (nombres de categoría) para la *variable explicativa x*.
5. Si la primera fila de datos incluye una etiqueta, marque la casilla **Variable labels**.
6. Haga clic en **OK**. La tabla de análisis de varianza incluye el dato estadístico de prueba *F* y el valor *P*.

Complemento de análisis de datos en Excel

1. Haga clic en la pestaña **Data** en la cinta de opciones y luego seleccione **Data Analysis** en el menú superior.
2. Seleccione **Anova: Single Factor** en las *herramientas de análisis* y haga clic en **OK**.
3. Ingrese el rango de datos deseado para **Input Range**.
4. En *agrupado por* seleccione **Columns** si los datos para cada categoría están contenidos en columnas separadas; seleccione **Rows** si los datos están organizados por filas.
5. Marque la casilla **Labels in First Row** si la primera celda contiene una etiqueta de categoría.
6. Haga clic en **OK** para ver los resultados, incluido el dato estadístico de prueba *F* y el valor *P*.

12-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

En los ejercicios 1 a 4, use los siguientes tiempos de demora en la llegada (minutos) para vuelos de American Airlines de Nueva York a Los Ángeles. Los valores negativos corresponden a vuelos que llegaron temprano. También se muestran los resultados de SPSS para el análisis de varianza. Suponga que planeamos usar un nivel de significancia de 0.05 para probar la afirmación de que los diferentes vuelos tienen el mismo tiempo medio de demora en la llegada.

Vuelo 1	–32	–25	–26	–6	5	–15	–17	–36
Vuelo 19	–5	–32	–13	–9	–19	49	–30	–23
Vuelo 21	–23	28	103	–19	–5	–46	13	–3

SPSS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2575.000	2	1287.500	1.334	.285
Within Groups	20271.500	21	965.310		
Total	22846.500	23			

1. ANOVA

a. ¿Qué característica de los datos anteriores indica que debemos usar el análisis de varianza *de un factor*?

b. Si el objetivo es probar la afirmación de que los tres vuelos tienen el mismo tiempo medio de demora en la llegada, ¿por qué el método se denomina análisis de *varianza*?

2. ¿Por qué no probar de dos en dos? Consulte los datos muestrales dados en el ejercicio 1. Si queremos probar la igualdad de las tres medias, ¿por qué no utilizamos tres pruebas de hipótesis separadas para $\mu_1 = \mu_2$, $\mu_1 = \mu_3$ y $\mu_2 = \mu_3$?

3. Dato estadístico de prueba ¿Cuál es el valor del dato estadístico de prueba? ¿Qué distribución se usa con el dato estadístico de prueba?

4. Valor *P* Si usamos un nivel de significancia de 0.05 en el análisis de varianza con los datos muestrales dados en el ejercicio l, ¿cuál es el valor *P*? ¿Qué deberíamos concluir? Si un pasajero aborrece las llegadas demoradas de los vuelos, ¿se puede ayudar a ese pasajero seleccionando uno de los vuelos?

En los ejercicios 5 a 16, use el análisis de varianza para la prueba indicada.

5. Plomo y puntuaciones de IQ verbal El ejemplo 1 usó puntuaciones de IQ del desempeño medidas para tres niveles diferentes de plomo en la sangre. Si usamos las mismas tres categorías de niveles de plomo en la sangre con puntuaciones medidas del IQ *verbal*, obtenemos la pantalla adjunta de Minitab. (Los datos se listan en el conjunto de datos 7 "IQ y plomo" del apéndice B). Utilizando un nivel de significancia de 0.05, evalúe la afirmación de que las tres categorías de nivel de plomo en la sangre tienen la misma puntuación media de IQ verbal. ¿La exposición al plomo parece tener un efecto en las puntuaciones de IQ verbal?

Minitab

Source	DF	SS	MS	F	P
LEAD	2	142	71	0.39	0.677
Error	118	21441	182		
Total	120	21584			

6. Plomo y puntuaciones de IQ completo El ejemplo 1 utilizó puntuaciones medidas de IQ del *desempeño* para tres niveles diferentes de plomo en la sangre. Si utilizamos las mismas tres categorías de niveles de plomo en la sangre con las puntuaciones del IQ *completo*, obtenemos la pantalla adjunta de Excel. (Los datos se listan en el conjunto de datos 7 "IQ y plomo" del apéndice B). Utilizando un nivel de significancia de 0.05, evalúe la afirmación de que las tres categorías de niveles de plomo en la sangre tienen la misma puntuación media del IQ total. ¿Parece que la exposición al plomo tiene un efecto en las puntuaciones del IQ completo?

Excel

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	938.3653	2	469.1827	2.303395	0.104395	3.07309
Within Groups	24035.63	118	203.6918			
Total	24974	120				

7. Tiempos de servicio de comida rápida en la cena El conjunto de datos 25 "Comida rápida" en el apéndice B lista los tiempos de servicio (en segundos) para cenas en McDonald's, Burger King y Wendy's. El uso de esos tiempos con una calculadora TI-83/84 Plus produce la siguiente pantalla.

Utilizando un nivel de significancia de 0.05, pruebe la afirmación de que las tres muestras provienen de poblaciones con la misma media. ¿Qué se puede concluir?

TI-83/84 Plus

```
NORMAL FLOAT AUTO REAL RADIAN MP
      One-way ANOVA
 F=5.596275825
 p=.0045462237
 Factor
  df=2
  SS=35696.44
  MS=17848.22
 Error
↓ df=147
```

8. Peso al nacer El conjunto de datos de 4 "Nacimientos" en el apéndice B lista el peso al nacer de bebés en el Albany Medical Center, Bellevue Hospital en la ciudad de Nueva York, Olean General Hospital y Strong Memorial Hospital en Rochester, Nueva York. Después de dividir los pesos al nacer según el hospital, obtenemos la pantalla de StatCrunch que se muestra a continuación. Use un nivel de significancia de 0.05 para probar la afirmación de que los diferentes hospitales tienen distintos promedios de peso al nacer. ¿Los pesos al nacer parecen ser diferentes en las áreas urbanas y rurales?

StatCrunch

ANOVA table

Source	DF	SS	MS	F-Stat	P-value
Columns	3	1701400	567133.33	1.1810493	0.3167
Error	396	1.90157e8	480194.44		
Total	399	1.918584e8			

9. Frecuencias de pulso y edad femeninas Usando los pulsos de mujeres en el conjunto de datos 1 "Datos corporales" en el apéndice B, después de que se particionan en los tres grupos de edad de 18 a 25, 26 a 40 y 41 a 80, obtenemos la siguiente pantalla de Statdisk. Utilizando un nivel de significancia de 0.05, pruebe la afirmación de que las mujeres de los tres grupos de edad tienen la misma frecuencia media de pulso. ¿Qué concluye usted?

Statdisk

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Treatment:	2	2280.049935	1140.024967	7.933788	3.058925	0.000539
Error:	144	20691.705167	143.692397			
Total:	146	22971.755102				

10. Frecuencias de pulso y edad masculinas Usando los pulsos de hombres del conjunto de datos 1 "Datos corporales" en el apéndice B, después de que se particionen en los tres grupos de edad de 18 a 25, 26 a 40 y 41 a 80, obtenemos la siguiente pantalla de SPSS. Utilizando un nivel de significancia de 0.05, pruebe la afirmación de que los hombres de los tres grupos de edad tienen la misma frecuencia media de pulso. ¿Qué concluye usted?

SPSS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	333.464	2	166.732	1.304	.275
Within Groups	19183.765	150	127.892		
Total	19517.229	152			

11. Tiempos en el triatlón Jeff Parent es un profesor de estadística que participa en triatlones. A continuación se listan los tiempos (en minutos y segundos) que registró mientras montaba la bicicleta en cinco etapas por cada milla de un circuito de 3 millas. Use un nivel de significancia de 0.05 para probar la afirmación de que se necesita el mismo tiempo para recorrer cada una de las millas. ¿Alguna de las millas parece tener una colina?

Milla 1	3:15	3:24	3:23	3:22	3:21
Milla 2	3:19	3:22	3:21	3:17	3:19
Milla 3	3:34	3:31	3:29	3:31	3:29

12. Arsénico en el arroz A continuación se listan las cantidades de arsénico en muestras de arroz integral de tres estados. Las cantidades están en microgramos de arsénico y todas las muestras tienen el mismo tamaño de porción. Los datos son de la Administración de Alimentos y Medicamentos. Use un nivel de significancia de 0.05 para probar la afirmación de que las tres muestras provienen de poblaciones con la misma media. ¿Las cantidades de arsénico parecen ser diferentes en los diferentes estados? Dado que las cantidades de arsénico en las muestras de Texas tienen la media más alta, ¿podemos concluir que el arroz integral de Texas representa el mayor problema de salud?

Arkansas	4.8	4.9	5.0	5.4	5.4	5.4	5.6	5.6	5.6	5.9	6.0	6.1
California	1.5	3.7	4.0	4.5	4.9	5.1	5.3	5.4	5.4	5.5	5.6	5.6
Texas	5.6	5.8	6.6	6.9	6.9	6.9	7.1	7.3	7.5	7.6	7.7	7.7

13. Demoras en la salida de vuelos A continuación se listan los tiempos de demora en la salida (minutos) para los vuelos de American Airlines de Nueva York a Los Ángeles. Los valores negativos corresponden a los vuelos que salieron temprano. Use un nivel de significancia de 0.05 para probar la afirmación de que los diferentes vuelos tienen el mismo tiempo medio de demora en su salida. ¿Qué característica notable de los datos se puede identificar al examinar visualmente los datos?

Vuelo 1	−2	−1	−2	2	−2	0	−2	−3
Vuelo 19	19	−4	−5	−1	−4	73	0	1
Vuelo 21	18	60	142	−1	−11	−1	47	13

14. Citas rápidas A continuación se listan las calificaciones de atributos de los hombres dadas por las mujeres que participaron en eventos de citas rápidas (del conjunto de datos 18 "Citas rápidas" en el apéndice B). Use un nivel de significancia de 0.05 para probar la afirmación de que las mujeres en los diferentes grupos de edad otorgan calificaciones de atributos con la misma media. ¿La edad parece ser un factor en las calificaciones de atributos que dan las mujeres?

Edad 20-22	38	42	30.0	39	47	43	33	31	32	28
Edad 23-26	39	31	36.0	35	41	45	36	23	36	20
Edad 27-29	36	42	35.5	27	37	34	22	47	36	32

En los ejercicios 15 y 16, use el conjunto de datos del apéndice B.

15. Galletas con chispas de chocolate Consulte el Conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B y utilice los recuentos de chispas de chocolate de los tres tipos diferentes de galletas Chips Ahoy. Use un nivel de significancia de 0.05 para probar la afirmación de que los tres tipos diferentes de galletas Chips Ahoy tienen la misma cantidad media de chispas de chocolate. ¿Las galletas Chips Ahoy bajas en grasa parecen tener menos chispas de chocolate?

16. Fumadores pasivos Consulte el conjunto de datos 12 "Fumadores pasivos y activos" en el apéndice B y utilice los niveles de cotinina sérica medidos (en mg/ml) de los tres grupos de sujetos (fumadores, no fumadores expuestos al humo de tabaco y no fumadores no expuestos al humo de tabaco). Cuando el cuerpo absorbe nicotina, se produce cotinina. Use un nivel de significancia de 0.05 para probar la afirmación de que las tres muestras provienen de poblaciones con la misma media. ¿Qué sugieren los resultados sobre los efectos del humo en los fumadores pasivos?

12-1 Más allá de lo básico

17. Prueba de Tukey En la página 577 se proporciona una pantalla de los resultados de la prueba de Bonferroni de la tabla 12-1 (que es parte del problema del capítulo). En la parte superior de la página siguiente se muestra la pantalla de los resultados de la prueba de Tukey generada por SPSS utilizando los mismos datos. Compare los resultados de las pruebas de Tukey con los de la prueba de Bonferroni.

continúa

SPSS

(I) Level	(J) Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	8.56876	3.80486	.067	-.4626	17.6002
	3.00	8.51465	3.87487	.076	-.6830	17.7123
2.00	1.00	-8.56876	3.80486	.067	-17.6002	.4626
	3.00	-.05411	4.80851	1.000	-11.4678	11.3596
3.00	1.00	-8.51465	3.87487	.076	-17.7123	.6830
	2.00	.05411	4.80851	1.000	-11.3596	11.4678

18. Prueba de Bonferroni A continuación se muestran los pesos (en kg) de álamos obtenidos de árboles plantados en una región rica y húmeda. Los árboles recibieron diferentes tratamientos identificados en la siguiente tabla. Los datos provienen de un estudio realizado por investigadores de la Universidad Estatal de Pensilvania y fueron proporcionados por Minitab, Inc. También se muestran los resultados parciales del uso de la prueba de Bonferroni con los datos muestrales.

Sin tratamiento	Fertilizante	Riego	Fertilizante y riego
1.21	0.94	0.07	0.85
0.57	0.87	0.66	1.78
0.56	0.46	0.10	1.47
0.13	0.58	0.82	2.25
1.30	1.03	0.94	1.64

a. Use un nivel de significancia de 0.05 para probar la afirmación de que los diferentes tratamientos resultan en el mismo peso promedio.

b. ¿Qué nos dicen los resultados de SPSS para la prueba de Bonferroni?

c. Utilice el procedimiento de prueba de Bonferroni con un nivel de significancia de 0.05 para evaluar una diferencia significativa entre el peso medio del grupo de tratamiento de riego y el grupo tratado con fertilizante y riego. Identifique el dato estadístico de prueba y el valor *P* o los valores críticos. ¿Qué indican los resultados?

Resultados de la prueba de Bonferroni con SPSS

(I) TREATMENT	(J) TREATMENT	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.02200	.26955	1.000	-.8329	.7889
	3.00	.23600	.26955	1.000	-.5749	1.0469
	4.00	-.84400*	.26955	.039	-1.6549	-.0331

12-2 ANOVA de dos factores

Concepto clave La sección 12-1 consideró datos particionados utilizando *un* factor, pero esta sección describe el método del *análisis de varianza de dos factores*, que se utiliza con datos divididos en categorías de acuerdo con *dos* factores. El método de esta sección requiere que primero probemos una *interacción* entre los dos factores; luego probamos un efecto del factor de fila y probamos un efecto a partir del factor de columna. La tabla 12-3 es un ejemplo de frecuencias de pulso (latidos por minuto) categorizadas con *dos* factores:

TABLA 12-3 Frecuencias de pulso con dos factores: intervalo de edad y género

	Mujeres										Hombres									
18-29	104	82	80	78	80	84	82	66	70	78	72	64	72	64	64	70	72	64	54	52
30-49	66	74	96	86	98	88	82	72	80	80	80	90	58	74	96	72	58	66	80	92
50-80	94	72	82	86	72	90	64	72	72	100	54	102	52	52	62	82	82	60	52	74

1. Rango de edad (años): Un factor es el intervalo de edad (18-29, 30-49, 50-80).
2. Género: el segundo factor es el género (mujer, hombre).

Las subcategorías en la tabla 12-3 se llaman *celdas*, por lo que esta tabla tiene seis celdas que contienen diez valores cada una.

Al analizar los datos muestrales en la tabla 12-3, ya hemos presentado el análisis de varianza de un factor, por lo que podría parecer razonable proceder simplemente con ANOVA de un factor para el rango de edad y otro ANOVA para el factor de género, pero ese método desperdicia información e ignora por completo una característica muy importante: el posible efecto de una *interacción* entre los dos factores.

DEFINICIÓN

Existe una **interacción** entre dos factores si el efecto de uno de ellos cambia para diferentes categorías del otro factor.

Como ejemplo de una *interacción* entre dos factores, considere parear los alimentos. La mantequilla de maní y la jalea interactúan bien, pero el ketchup y el helado interactúan de una manera que da como resultado un mal sabor, por lo que rara vez vemos a alguien comiendo helado cubierto con ketchup. En general, considere que un efecto de interacción es un efecto debido a la combinación de los dos factores.

Exploración de datos con medias y una gráfica de interacción

Exploremos los datos en la tabla 12-3 calculando la media para cada celda y construyendo una gráfica. Las medias de las celdas individuales se muestran en la tabla 12-4. Esas medias varían desde un mínimo de 64.8 hasta un máximo de 82.2, por lo que varían considerablemente. La figura l2-3 es una *gráfica de interacción*, que muestra las gráficas de esas medias. Podemos interpretar una gráfica de interacción de la siguiente manera:

- **Efecto de interacción:** Se sugiere un efecto de interacción cuando los segmentos de línea están lejos de ser paralelos.
- **Sin efecto de interacción:** Si los segmentos de línea son aproximadamente *paralelos*, como en la figura 12-3, parece que las diferentes categorías de una variable tienen el mismo efecto sobre las diferentes categorías de la otra variable, entonces no parece haber un efecto de interacción.

TABLA 12-4 Medias de celdas de la tabla 12-3

	Mujeres	Hombres
18-29	80.4	64.8
30-49	82.2	76.6
50-80	80.4	67.2

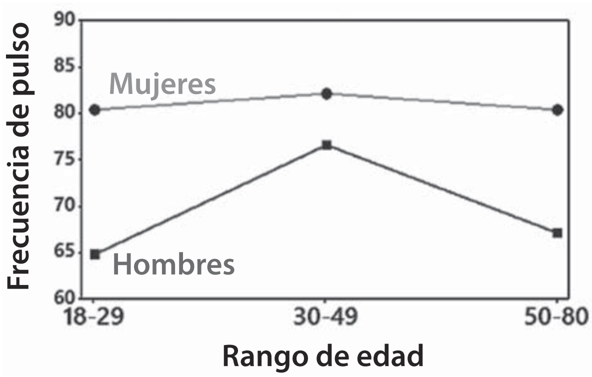


FIGURA 12-3 **Gráfica de interacción del rango de edad y el género: Medias de la tabla 12-4**

En lugar de basarnos únicamente en juicios subjetivos realizados al examinar las medias de la tabla l2-4 y la gráfica de interacción en la figura l2-3, realizaremos un procedimiento más objetivo, el análisis de varianza de dos factores. A continuación se presentan los requisitos y el procedimiento básico para el análisis de varianza (ANOVA) de dos factores. El procedimiento también se resume en la figura 12-4, que se presenta después del recuadro de elementos clave en la página siguiente.

ELEMENTOS CLAVE

Análisis de varianza de dos factores

Objetivo

Con los datos muestrales categorizados por dos factores (una variable de fila y una variable de columna), utilizar el análisis de varianza de dos factores para realizar las siguientes tres pruebas:

1. Probar un efecto de una interacción entre el factor de fila y el factor de columna.
2. Probar un efecto del factor de fila.
3. Probar un efecto del factor de columna.

Requisitos

1. **Normalidad** Para cada celda, los valores muestrales provienen de una población con una distribución que es aproximadamente normal. (Este procedimiento es robusto contra desviaciones razonables de las distribuciones normales).
2. **Variación** Las poblaciones tienen la misma varianza σ^2 (o desviación estándar σ). (Este procedimiento es robusto contra desviaciones razonables del requisito de varianzas iguales).
3. **Muestreo** Las muestras son muestras aleatorias simples de datos cuantitativos.
4. **Independencia** Las muestras son independientes entre sí. (Este procedimiento no se aplica a muestras que carecen de independencia).
5. **Bifactorial** Los valores muestrales se clasifican de dos maneras. (Esta es la base del nombre del método: análisis de varianza de dos factores).
6. **Diseño equilibrado** Todas las celdas tienen el mismo número de valores muestrales. (Esto se denomina *diseño equilibrado*. Esta sección no incluye métodos para un diseño que no sea equilibrado).

Procedimiento para ANOVA de dos factores (consulte la figura 12-4)

Paso 1: ***Efecto de interacción:*** En el análisis de varianza de dos factores, comience probando la hipótesis nula de que no hay interacción entre los dos factores. Use la tecnología para encontrar el valor P correspondiente al siguiente dato estadístico de prueba:

$$F = \frac{\text{MS(interacción)}}{\text{MS(error)}}$$

Conclusión:

- **Rechazar:** Si el valor P correspondiente al dato estadístico de prueba anterior es pequeño (por ejemplo menor o igual a 0.05), rechace la hipótesis nula de no interacción. Concluya que hay un efecto de interacción.
- **No rechazar:** Si el valor P es grande (por ejemplo mayor que 0.05), no se puede rechazar la hipótesis nula de que no hay interacción entre los dos factores. Concluya que no hay efecto de interacción.

Paso 2: ***Efectos de fila/columna:*** Si concluimos que hay un efecto de interacción, entonces debemos detenernos ahora; no deberíamos continuar con las dos pruebas adicionales. (Si hay una interacción entre factores, no deberíamos considerar los efectos de ninguno de los factores sin tener en cuenta los del otro).

Si concluimos que no hay efecto de interacción, entonces debemos proceder con las siguientes dos pruebas de hipótesis.

Factor de fila

Para el factor de fila, pruebe la hipótesis nula H_0: no hay efectos del factor de fila (es decir, los valores de fila son de poblaciones con la misma media). Encuentre el valor P correspondiente al dato estadístico de prueba $F = \text{MS(fila)}/\text{MS(error)}$.

Conclusión:

- **Rechazar:** Si el valor P correspondiente al dato estadístico de prueba es pequeño (por ejemplo menor o igual a 0.05), rechace la hipótesis nula de no efecto del factor de fila. Concluya que hay un efecto del factor de fila.
- **No rechazar:** Si el valor P es grande (por ejemplo mayor que 0.05), no se puede rechazar la hipótesis nula de que no hay efecto del factor de fila. Concluya que no hay efecto del factor de fila.

Factor de columna

Para el factor de columna, pruebe la hipótesis nula H_0: No hay efectos del factor de columna (es decir, los valores de columna provienen de poblaciones con la misma media). Encuentre el valor P correspondiente al dato estadístico de prueba F = MS(columna)/MS(error).

Conclusión:

- **Rechazar:** Si el valor P correspondiente al dato estadístico de prueba es pequeño (por ejemplo, menor o igual a 0.05), rechace la hipótesis nula de no efecto del factor de columna. Concluya que hay un efecto del factor de columna.
- **No rechazar:** Si el valor P es grande (por ejemplo, mayor que 0.05), no se puede rechazar la hipótesis nula de no efecto del factor de columna. Concluya que no hay efecto del factor de columna.

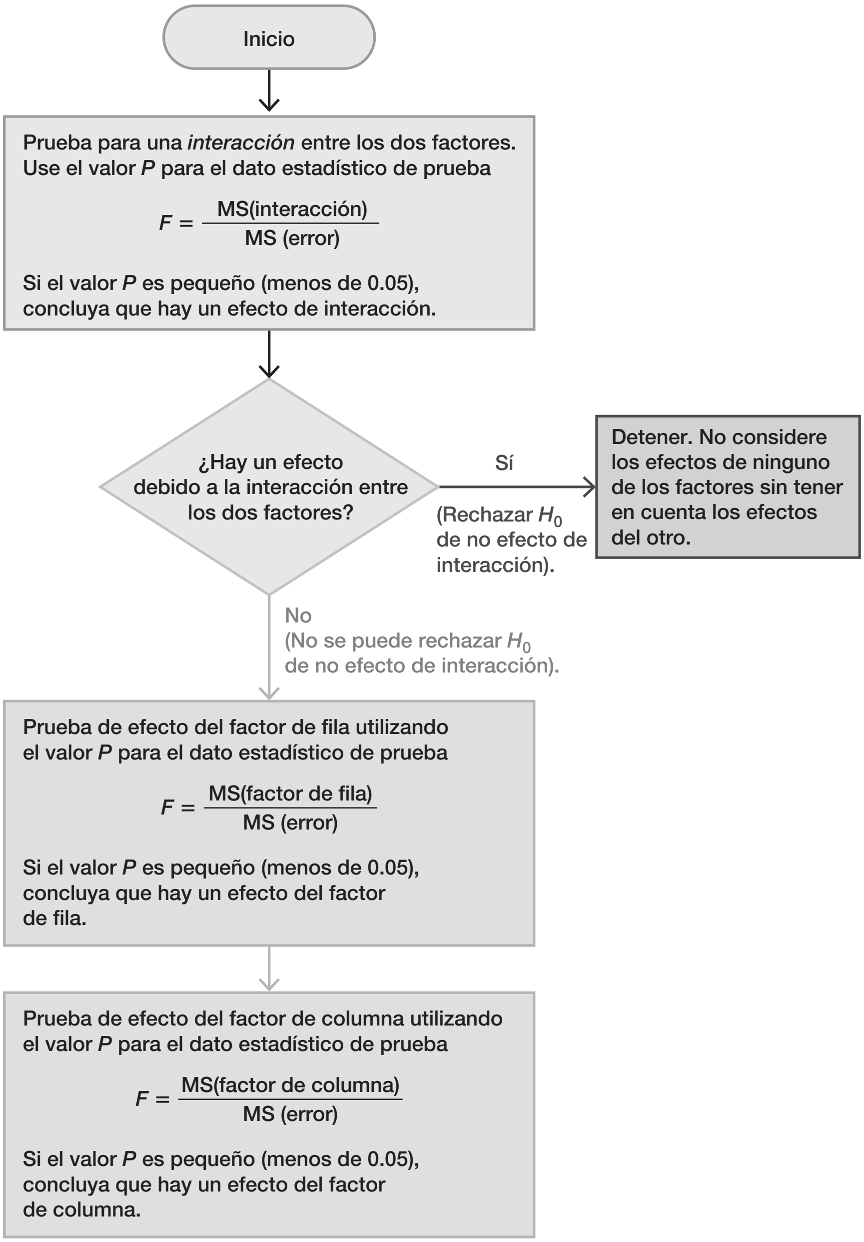


FIGURA 12-4 Procedimiento para el análisis de varianza de dos factores

EJEMPLO 1 Frecuencias de pulso

Dadas las frecuencias de pulso de la tabla 12-3 en la página 582 (del conjunto de datos 1 "Datos corporales" en el apéndice B), utilice el análisis de varianza de dos factores para probar un efecto de interacción, un efecto del factor de fila para el rango de edad, y un efecto del factor de columna para el género. Use un nivel de significancia de 0.05.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Para cada celda, los valores muestrales parecen provenir de una población con una distribución que es aproximadamente normal, como lo indican las gráficas cuantilares normales. (2) Las varianzas de las celdas (100.3, 51.7, 103.5, 183.2, 138.5, 293.5) difieren considerablemente, pero la prueba es robusta frente a desviaciones de las varianzas iguales. (3) Las muestras son muestras aleatorias simples de sujetos. (4) Las muestras son independientes entre sí; los sujetos no están pareados de ninguna manera. (5) Los valores muestrales se categorizan de dos maneras (rango de edad y sexo). (6) Todas las celdas tienen el mismo número (diez) de valores muestrales. Los requisitos se satisfacen.

Los cálculos son bastante complicados, por lo tanto se utiliza tecnología. A continuación se muestra la pantalla del análisis de varianza de dos factores en StatCrunch, para los datos de la tabla 12-3.

Statcrunch

ANOVA table

Source	DF	SS	MS	F-Stat	P-value
Age	2	526.93333	263.46667	1.8156202	0.1725
Gender	1	1972.2667	1972.2667	13.591424	0.0005
Interaction	2	272.53333	136.26667	0.93905054	0.3973
Error	54	7836	145.11111		
Total	59	10607.733			

Paso 1: Efecto de interacción: Comenzamos probando la hipótesis nula de que no hay interacción entre los dos factores. Al utilizar para los datos de la tabla 12-3, obtenemos los resultados que se muestran en la pantalla de StatCrunch anterior y podemos ver que el dato estadístico de prueba para la interacción es $F = 0.9391$. Este dato estadístico de prueba se puede calcular de la siguiente manera:

$$F = \frac{\text{MS(interacción)}}{\text{MS(error)}} = \frac{136.26667}{145.11111} = 0.9391$$

Interpretación: El valor P correspondiente se muestra en la pantalla StatCrunch como 0.3973, por lo que no se puede rechazar la hipótesis nula de no interacción entre los dos factores. No parece que las frecuencias del pulso se vean afectadas por una interacción entre el grupo de edad (18-29, 30-49, 50-80) y el género. No parece haber un efecto de interacción.

Paso 2: Efectos de fila/columna: Dado que no parece haber un efecto de interacción, procedemos a probar los efectos de los factores de fila y de columna. Las dos pruebas de hipótesis usan las siguientes hipótesis nulas:

H_0: No hay efectos del factor de fila (es decir, los valores de fila son de poblaciones con las mismas medias).

H_0: No hay efectos del factor de columna (es decir, los valores de columna provienen de poblaciones con las mismas medias).

Factor de fila: Para el factor de fila (rango de edad), nos referimos a la pantalla de resultados de StatCrunch anterior para ver que el dato estadístico de prueba para el factor de fila es $F = 1.8156$ (redondeado). Este dato estadístico de prueba se puede calcular de la siguiente manera:

$$F = \frac{\text{MS(rango de edad)}}{\text{MS(error)}} = \frac{263.46667}{145.11111} = 1.8156$$

Conclusión: El valor *P* correspondiente se muestra en la pantalla de StatCrunch como 0.1725. Debido a que ese valor *P* es mayor que el nivel de significancia de 0.05, no podemos rechazar la hipótesis nula de que no hay efectos del rango de edad. Es decir, las frecuencias del pulso no parecen verse afectadas por el rango de edad.

Factor de columna: Para el factor de columna (género), nos referimos a la pantalla de resultados StatCrunch anterior para ver que el dato estadístico de prueba para el factor de columna es $F = 13.5914$ (redondeado). Este dato estadístico de prueba se puede calcular de la siguiente manera:

$$F = \frac{\text{MS(género)}}{\text{MS(error)}} = \frac{1972.2667}{145.11111} = 13.5914$$

Conclusión: El valor *P* correspondiente se muestra en la pantalla de StatCrunch como 0.0005. Debido a que ese valor *P* es menor que el nivel de significancia de 0.05, rechazamos la hipótesis nula de que no hay efectos del género. Las tasas de pulso parecen estar afectadas por el género.

INTERPRETACIÓN

Con base en los datos muestrales de la tabla 12-3, concluimos que las frecuencias de pulso parecen estar afectadas por el género, pero no por el intervalo de edad y no por una interacción entre el grupo de edad y el género.

SU TURNO Resuelva el ejercicio 5 “Frecuencias de pulso”.

PRECAUCIÓN El análisis de varianza de dos factores no es un análisis de varianza de un factor hecho dos veces. Al realizar un análisis de varianza de este tipo, asegúrese de probar una *interacción* entre ambos factores.

CENTRO DE TECNOLOGÍA

Análisis de varianza de dos factores

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Two-Way Analysis of Variance** en el menú desplegable.
3. Ingrese el número de categorías para las variables de fila y las variables de columna.
4. Ingrese el número de valores en cada celda y haga clic en **Continue**.
5. En la tabla, ingrese o pegue los datos en la columna de *valor*.
6. Haga clic en **Evaluate**.

Minitab

1. Ingrese todos los valores muestrales en la columna C1.
2. Ingrese los números (o nombres) de fila correspondientes en la columna C2.
3. Ingrese los números (o nombres) de columna correspondientes en la columna C3.
4. Haga clic en **Stat** en el menú superior.
5. Seleccione **ANOVA** en el menú desplegable y seleccione **General Linear Model-Fit General Linear Model**.
6. Para *respuestas*, seleccione **C1** y elija **C2** y **C3** como *factores*.
7. Haga clic en el botón **Model**.
8. En *factores y covariables*, seleccione **C2** y **C3** y haga clic en el botón **Add**.
9. Haga clic en **OK** dos veces. Vea el *análisis de varianza* en los resultados.

SUGERENCIA: Use etiquetas descriptivas en lugar de C1, C2 y C3 para evitar confusiones.

StatCrunch

1. Ingrese todos los valores muestrales en una columna llamada “Respuestas”.
2. Ingrese los números (o nombres) de fila correspondientes en una segunda columna llamada “Factor de fila”.
3. Ingrese los números de columna correspondientes (o nombres) en una tercera columna llamada “Factor de columna”.
4. Haga clic en **Stat** en el menú superior.
5. Seleccione **ANOVA** en el menú desplegable, luego seleccione **Two Way** desde el submenú.
6. Seleccione las columnas que se utilizarán para las respuestas, factor de fila y factor de columna.
7. Haga clic en **Compute!**

continúa

CENTRO DE TECNOLOGÍA *continuación*

Análisis de varianza de dos factores

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Calculadora TI-83/84 Plus

Requiere el programa *A1ANOVA* (disponible en www.pearsonenespañol.com/triola**).**

1. El programa A1ANOVA requiere que primero creemos una Matriz [D] que contenga los datos muestrales: *Ingrese los datos manualmente*: Presione 2ND luego x^{-1} para ir al menú *MATRIX*, seleccione **EDIT** desde el menú superior, seleccione **[D]** y presione ENTER.

   Ingrese el número de filas y columnas necesarias, presione ENTER y proceda a ingresar los valores muestrales.

   Uso de una lista existente: Las listas se pueden combinar y almacenar en una matriz. Presione 2ND luego x^{-1} para ir al menú *MATRIX*, seleccione **MATH** en el menú superior y seleccione el elemento **List → matr**. Ingrese los nombres de la lista, seguidos por el nombre de la matriz **[D]**, todos separados por comas. *Importante:* El nombre de la matriz se debe ingresar presionando 2ND luego x^{-1}, seleccionando **[D]**, y presionando ENTER. El siguiente es un resumen de los comandos utilizados para crear una matriz a partir de tres listas (L1, L2, L3): **List → matr(L1, L2, L3,[D])**.
2. Presione **PRGRM**, luego seleccione **A1ANOVA** y presione ENTER dos veces.
3. Seleccione **RAN BLOCK DESIGN** y presione ENTER dos veces. Seleccione **Continue** y presione ENTER.
4. El programa trabajará con los datos en la Matriz **[D]** y mostrará los resultados. Los resultados no caben en una sola pantalla, entonces presione ENTER para ver los resultados restantes.

TIP: En los resultados, F(A) es el dato estadístico de prueba *F* para el factor de fila, F(B) es el dato estadístico de prueba *F* para el factor de columna y F(AB) es el dato estadístico de prueba *F* para el efecto de interacción.

Excel

Complemento XLSTAT

Requiere que todos los datos se apilen en una sola columna con los nombres de categoría correspondientes para cada valor de datos en dos columnas separadas y adyacentes. Los nombres de las filas deben estar en una de esas columnas y los nombres de las columnas deben estar en la otra columna.

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Modeling Data**.
2. Seleccione **ANOVA** del menú desplegable.
3. Ingrese el rango de celdas que contienen los valores muestrales en el cuadro de las *variables dependientes* Y (Y/*Dependent variables*).
4. Seleccione el cuadro **Qualitative** e ingrese el rango de celdas que contiene los nombres de fila y columna en el cuadro de las *variables independientes x*, (*X*/*Explanatory variables*), como B1:C30.
5. Si se incluye una etiqueta de variable en el rango de datos, marque la casilla **Variable labels**.
6. Haga clic en la pestaña **Options** y confirme que la casilla **Interactions**/**Level** esté marcada y configurada en **2**.
7. Haga clic en la pestaña **Output** y marque la casilla etiquetada **Type I**/**II**/**III SS**.
8. Haga clic en **OK**. Haga clic en **All** en la ventana **Factors and interactions** y haga clic en **OK** para ver los resultados. Busque los resultados clave bajo el título de “Type I Sum of Squares Analysis”. Los valores *P* se etiquetan como $Pr > F$.

Complemento de Excel de análisis de datos

Más de una entrada por celda

Para las tablas bidireccionales con más de una entrada por celda, las entradas de la misma celda se deben enumerar en una columna, no en una fila. Ingrese las etiquetas correspondientes al conjunto de datos en la columna A y la fila 1, como se muestra en el siguiente ejemplo:

	A	B	C	D
1		Bajo	Medio	Alto
2	Hombre	85	78	93
3	Hombre	90	107	97
⋮	⋮	⋮	⋮	⋮

1. Haga clic en la pestaña **Data** en la cinta de opciones y luego seleccione **Data Analysis** en el menú superior
2. Seleccione **Anova: Two-Factor With Replication** en *Analysis Tools* y haga clic en **OK**.
3. Ingrese el rango de datos deseado en **Input Range**.
4. En *Rows per sample*, ingrese el número de valores en cada celda.
5. Haga clic en **OK**.

12-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. ANOVA de dos factores A continuación se reproducen las frecuencias de pulso en la tabla 12-3 del ejemplo 1, con los datos impresos (en negro) utilizados para las frecuencias de pulso de mujeres de 30 a 49 años. ¿Qué característica de los datos sugiere que el método de análisis apropiado es el análisis de varianza de dos factores? Es decir, ¿qué es "bidireccional" sobre los datos ingresados en esta tabla?

	Mujeres										Hombres									
18-29	104	82	80	78	80	84	82	66	70	78	72	64	72	64	64	70	72	64	54	52
30-49	**46**	**54**	**76**	**66**	**78**	**68**	**62**	**52**	**60**	**60**	80	90	58	74	96	72	58	66	80	92
50-80	94	72	82	86	72	90	64	72	72	100	54	102	52	52	62	82	82	60	52	74

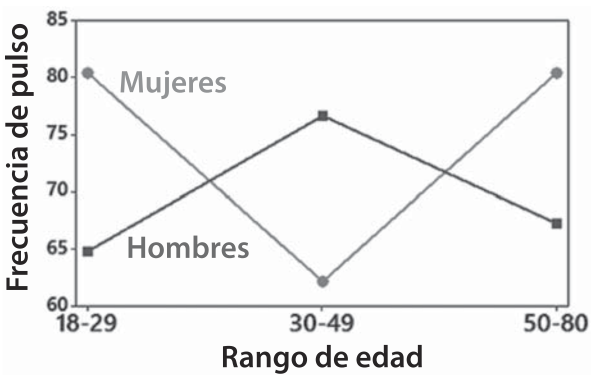
2. ANOVA de dos factores Si tenemos el objetivo de utilizar los datos descritos en el ejercicio 1 para (1) determinar si el grupo de edad tiene un efecto en las frecuencias de pulso y (2) determinar si el género tiene un efecto en las frecuencias de pulso, ¿Deberíamos usar el análisis de varianza de un factor para las dos pruebas individuales? ¿Por qué sí o por qué no?

3. Interacción

a. ¿Qué es una interacción entre dos factores?

b. En general, cuando se utiliza el análisis de varianza de dos factores, si encontramos que hay un efecto de interacción, ¿cómo afecta eso el procedimiento?

c. A continuación se muestra una gráfica de interacción construida a partir de los datos del ejercicio l. ¿Qué sugiere la gráfica?



4. Diseño equilibrado ¿La tabla dada en el ejercicio 1 constituye un *diseño equilibrado*? ¿Por qué sí o por qué no?

5. Frecuencias de pulso Si usamos los datos dados en el ejercicio 1 con un análisis de varianza de dos factores, obtenemos la pantalla adjunta. ¿Qué se puede concluir?

Statdisk

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Interaction:	2	2779.2	1389.6	9.58	3.1682	0.0003
Row Variable:	2	206.9333	103.4667	0.7130	3.1682	0.4947
Column Variable:	1	345.6	345.6	2.3816	4.0195	0.1286

6. Pesos Los pesos (kg) en la siguiente tabla provienen del conjunto de datos 1 “Datos corporales” en el apéndice B. También se muestran los resultados del análisis de varianza de dos factores. Use los resultados mostrados y un nivel de significancia de 0.05. ¿Qué se puede concluir?

	Mujeres						Hombres					
18-29	63.4	57.8	52.6	46.9	61.7	61.5	71.6	64.9	144.9	96.4	80.7	84.4
	77.2	50.4	97.0	76.1			63.9	79.0	99.4	64.1		
30-49	110.5	84.6	133.3	90.2	125.7	105.3	96.2	56.4	107.4	99.5	64.8	94.7
	115.5	75.3	92.8	57.7			74.2	112.8	72.6	91.4		
50-80	103.2	48.3	87.8	101.3	67.8	45.2	84.8	127.5	89.9	75.3	110.2	72.3
	79.8	60.1	68.5	43.3			77.2	86.5	71.3	73.1		

StatCrunch

ANOVA table

Source	DF	SS	MS	F-Stat	P-value
Age	2	3727.7583	1863.8792	4.3546612	0.0176
Gender	1	1013.526	1013.526	2.3679444	0.1297
Interaction	2	3137.611	1568.8055	3.6652678	0.0322
Error	54	23113.044	428.01933		
Total	59	30991.939			

7. Estaturas Las estaturas (cm) en la siguiente tabla provienen del conjunto de datos 1 “Datos corporales” del apéndice B. También se muestran los resultados del análisis de varianza de dos factores. Use los resultados mostrados y un nivel de significancia de 0.05. ¿Qué se puede concluir?

	Mujeres					Hombres				
18-29	161.2	170.2	162.9	155.5	168.0	172.8	178.7	183.1	175.9	161.8
	153.3	152.0	154.9	157.4	159.5	177.5	170.5	180.1	178.6	178.5
30-49	169.1	170.6	171.1	159.6	169.8	170.1	165.4	178.5	168.5	180.3
	169.5	156.5	164.0	164.8	155.6	178.2	174.4	174.6	162.8	174.4
50-80	146.7	160.9	163.3	176.1	163.1	181.9	166.6	171.7	170.0	169.1
	151.6	164.7	153.3	160.3	134.5	182.9	176.3	166.7	166.3	160.5

XLSTAT

Source	DF	Sum of squares	Mean squares	F	Pr > F
Age	2	222.0610	111.0305	2.0403	0.1399
Gender (1=M)	1	2365.0482	2365.0482	43.4607	< 0.0001
Age*Gender (1=M)	2	195.5803	97.7902	1.7970	0.1756

8. Experimento con panqueques A continuación se listan las calificaciones de panqueques otorgadas por expertos (con base en datos de Minitab). Los diferentes panqueques se hicieron con o sin un suplemento y con diferentes cantidades de suero de leche. Se muestran los resultados del análisis de varianza de dos factores. Use los resultados mostrados y un nivel de significancia de 0.05. ¿Qué concluye usted?

	Suero											
	0%			10%			20%			30%		
Sin suplemento	4.4	4.5	4.3	4.6	4.5	4.8	4.5	4.8	4.8	4.6	4.7	5.1
Suplemento	3.3	3.2	3.1	3.8	3.7	3.6	5.0	5.3	4.8	5.4	5.6	5.3

Minitab

Two-way ANOVA: Quality versus Supplement, Whey

```
Source       DF       SS       MS      F      P
Supplement    1   0.5104  0.51042  17.01  0.001
Whey          3   6.6912  2.23042  74.35  0.000
Interaction   3   3.7246  1.24153  41.38  0.000
Error        16   0.4800  0.03000
Total        23  11.4062
```

9. Tiempos en el Maratón A continuación se listan los tiempos registrados en la carrera de maratón de la ciudad de Nueva York (en segundos) para corredores elegidos al azar que completaron el maratón. ¿Los tiempos en la carrera se ven afectados por una interacción entre el género y el rango de edad? ¿Los tiempos de ejecución están afectados por el género? ¿Los tiempos de ejecución se ven afectados por el rango de edad? Use un nivel de significancia de 0.05.

Tiempos (en segundos) para corredores de la maratón de la ciudad de Nueva York

	Edad		
	21-29	**30-39**	**40 y más**
Hombres	13,615	14,677	14,528
	18,784	16,090	17,034
	14,256	14,086	14,935
	10,905	16,461	14,996
	12,077	20,808	22,146
Mujeres	16,401	15,357	17,260
	14,216	16,771	25,399
	15,402	15,036	18,647
	15,326	16,297	15,077
	12,047	17,636	25,898

10. Tabaquismo, género y temperatura corporal La siguiente tabla lista las temperaturas corporales obtenidas de sujetos seleccionados al azar (según el conjunto de datos 3 "Temperaturas corporales" en el apéndice B). Con un nivel de significancia de 0.05, pruebe una interacción entre el género y el tabaquismo, pruebe un efecto del género y un efecto del tabaquismo. ¿Qué concluye usted?

	Fuma				No fuma			
Hombres	98.8	97.6	98.0	98.5	98.4	97.8	98.0	97.0
Mujeres	98.0	98.5	98.3	98.7	97.7	98.0	98.2	99.1

12-2 Más allá de lo básico

11. Transformaciones de datos El ejemplo 1 ilustró el uso del ANOVA de dos factores para analizar los datos muestrales en la tabla 12-3 en la página 582. ¿Cómo se ven afectados los resultados en cada uno de los siguientes casos?

a. Se suma la misma constante a cada valor muestral.

b. Cada valor muestral se multiplica por la misma constante distinta de cero.

c. El formato de la tabla se transpone para que los factores de fila y de columna se intercambien.

d. Se cambia el primer valor muestral en la primera celda, de modo que se convierte en un valor atípico.

Examen rápido del capítulo

1. Pesos de bebidas de cola El conjunto de datos 26 "Pesos y volúmenes de bebidas de cola" en el apéndice B lista los pesos (lb) del contenido de latas de cuatro muestras diferentes de bebidas de cola: (1) Coca-Cola regular, (2) Coca-Cola Light, (3) Pepsi regular y (4) Pepsi Light. Los resultados del análisis de varianza se muestran al inicio de la página siguiente. ¿Cuál es la hipótesis nula para esta prueba de análisis de varianza? En función de los resultados que se muestran, ¿qué se debería concluir acerca de H_0? ¿Qué concluye usted sobre la igualdad de los pesos medios de las cuatro muestras?

continúa

Minitab

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	0.047979	0.015993	503.06	0.000
Error	140	0.004451	0.000032		
Total	143	0.052430			

2. Pesos de bebidas de cola Para las cuatro muestras descritas en el ejercicio 1, la muestra de Coca-Cola regular tiene un peso medio de 0.81682 lb, la muestra de Coca-Cola Ligth tiene un peso medio de 0.78479 lb, la muestra de Pepsi regular tiene un peso promedio de 0.82410 lb, y la muestra de Pepsi Ligth tiene un peso promedio de 0.78386 lb. Si usamos el análisis de varianza y llegamos a una conclusión para rechazar la igualdad de las cuatro medias muestrales, ¿podemos concluir que cualquiera de las muestras específicas tiene medias que son significativamente diferentes de las demás?

3. Pesos de bebidas de cola Para el análisis de la prueba de varianza descrito en el ejercicio 1, ¿esa prueba es de cola izquierda, de cola derecha o de dos colas?

4. Pesos de bebidas de cola Identifique el valor del dato estadístico de prueba en la pantalla incluida con el ejercicio 1. En general, ¿los estadísticos de prueba más grandes dan como resultado valores P más grandes, valores P más pequeños o valores P que no están relacionados con el valor del dato estadístico de prueba?

5. Peso de bebidas de Cola Los resultados mostrados en el ejercicio 1 son del análisis de varianza de un factor. ¿A qué se debe que esta prueba se caracterice como un análisis de varianza de un factor en lugar de un análisis de varianza de do factores?

6. ANOVA de un factor En general, ¿para qué sirve el análisis de varianza de un factor?

7. Uno contra dos ¿Cuál es la diferencia fundamental entre el análisis de varianza de un factor y el análisis de varianza de dos factores?

8. Estimación de longitud A continuación se presenta una pantalla de Minitab resultante del análisis de varianza de dos factores, con datos muestrales que constan de 18 diferentes estimaciones visuales de estudiantes sobre la longitud de un aula. Los valores se clasifican de acuerdo con el género y las especialidades (matemáticas, negocios, artes liberales). ¿Qué se puede concluir sobre una interacción entre el género y la especialidad?

Minitab

Source	DF	SS	MS	F	P
Sex	1	29.389	29.3889	0.78	0.395
Major	2	10.111	5.0556	0.13	0.876
Interaction	2	14.111	7.0556	0.19	0.832
Error	12	453.333	37.7778		
Total	17	506.944			

9. Estimación de longitud Usando los mismos resultados que se muestran en el ejercicio 8, ¿parece que las estimaciones de longitud se ven afectadas por el género del sujeto?

10. Estimación de longitud Usando los mismos resultados que se muestran en el ejercicio 8, ¿parece que las estimaciones de longitud se ven afectadas por la especialidad del sujeto?

Ejercicios de repaso

1. Citas rápidas El conjunto de datos 18 “Citas rápidas” en el apéndice B lista las calificaciones de atributos de las mujeres dadas por los hombres que participaron en eventos de citas rápidas, y algunos de esos valores se incluyen en la tabla al inicio de la página siguiente. El análisis de varianza se usa con los valores de esa tabla, y los resultados de StatCrunch se muestran en la siguiente página después de los datos. Use un nivel de significancia de 0.05 para probar la afirmación de que los varones en los diferentes grupos de edad otorgan calificaciones de atributos con la misma media. ¿La edad parece ser un factor en las calificaciones de atributos dadas por los hombres?

Edad 20-22	32	34	37	40.5	33	28	31	50	39	41
Edad 23-26	40	21	14	32	26	34	31	34	34	34
Edad 27-29	31	39	27	34	43	31	30	38	37	34

StatCrunch

ANOVA table

Source	DF	SS	MS	F-Stat	P-value
Columns	2	222.95	111.475	2.7346508	0.0829
Error	27	1100.625	40.763889		
Total	29	1323.575			

2. Legibilidad del autor El autor seleccionó páginas al azar de *El oso y el dragón* de Tom Clancy, *Harry Potter y la piedra del hechicero* de J. K. Rowling y *La guerra y la paz* de León Tolstoi. Las puntuaciones de facilidad de lectura para esas páginas se listan a continuación. ¿Los autores parecen tener el mismo nivel de legibilidad?

Clancy	58.2	73.4	73.1	64.4	72.7	89.2	43.9	76.3	76.4	78.9	69.4	72.9
Rowling	85.3	84.3	79.5	82.5	80.2	84.6	79.2	70.9	78.6	86.2	74.0	83.7
Tolstoi	69.4	64.2	71.4	71.6	68.5	51.9	72.2	74.4	52.8	58.4	65.4	73.6

3. Pruebas de colisiones de automóviles El conjunto de datos 19 "Pruebas de colisión de automóviles" en el apéndice B lista los resultados de pruebas de colisión realizadas con automóviles. El conjunto de datos incluye las cargas de la prueba de colisión (libras) en el fémur izquierdo y el fémur derecho. Cuando esas cargas se dividen en tres categorías por el tamaño del automóvil: pequeño, mediano y grande, el análisis de dos factores de los resultados de XLSTAT se muestra a continuación. (El factor de fila del fémur tiene los dos valores de fémur izquierdo y fémur derecho, y el factor de columna del tamaño tiene los tres valores de pequeño, mediano y grande). Use un nivel de significancia de 0.05 para aplicar el método del análisis de varianza de dos factores ¿Qué concluye usted?

XLSTAT

Source	DF	Sum of squares	Mean squares	F	Pr > F
Femur	1	166068.5952	166068.5952	1.3896	0.2462
Size	2	532911.8571	266455.9286	2.2296	0.1222
Femur*Size	2	410435.7619	205217.8810	1.7171	0.1940

4. Citas rápidas A continuación se listan las calificaciones de atributos de los hombres otorgadas por las mujeres que participaron en eventos de citas rápidas (del conjunto de datos 18 "Citas rápidas" en el apéndice B). Use un nivel de significancia de 0.05 para aplicar el método del análisis de varianza de dos factores. ¿Qué concluye usted?

		Hombres														
		20-23					**24-26**					**27-30**				
Mujeres	**20-23**	42	24	40	32	30	37	47	33	32	21	22	30	32	43	28
	24-26	34	31	25	36	30	36	41	36	33	48	34	32	27	43	35
	27-30	35	31	40	31	32	36	25	42	28	42	37	40	21	34	23

Ejercicios de repaso acumulado

En los ejercicios 1 a 5, consulte la siguiente lista de tiempos de demora en la salida (minutos) de los vuelos de American Airlines del aeropuerto JFK en Nueva York al aeropuerto LAX en Los Ángeles. Suponga que los datos son muestras seleccionadas al azar de poblaciones más grandes.

Vuelo 3	22	−11	7	0	−5	3	−8	8
Vuelo 19	19	−4	−5	−1	−4	73	0	1
Vuelo 21	18	60	142	−1	−11	−1	47	13

1. Exploración de los datos Incluya las unidades apropiadas en todas las respuestas.

a. Encuentre la media para cada uno de los tres vuelos.

b. Encuentre la desviación estándar para cada uno de los tres vuelos.

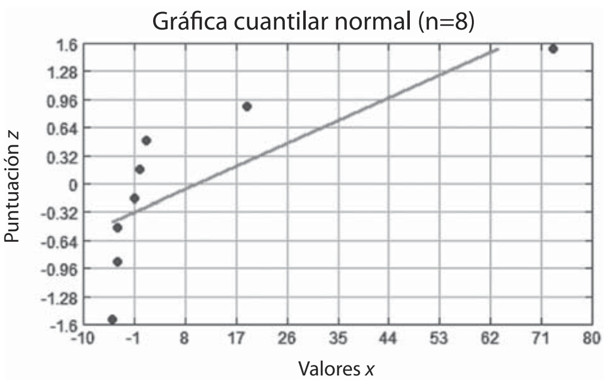
c. Encuentre la varianza para cada uno de los tres vuelos.

d. ¿Hay algún valor atípico obvio?

e. ¿Cuál es el nivel de medición de los datos (nominal, ordinal, de intervalo, de razón)?

2. Comparación de dos medias Tratando los datos como muestras de poblaciones más grandes, pruebe la afirmación de que existe una diferencia entre el tiempo medio de demora en la salida para el vuelo 3 y el vuelo 21.

3. Gráfica cuantilar normal La gráfica cuantilar normal adjunta se obtuvo de los tiempos de demora en la salida del vuelo 19. ¿Qué nos dice esta gráfica?



4. Intervalo de confianza Use los tiempos de demora en la salida del vuelo 3 y construya una estimación del intervalo de confianza del 95% para la media poblacional. Escriba una breve declaración que interprete el intervalo de confianza.

5. ANOVA La pantalla de XLSTAT que se presenta a continuación resulta de usar el análisis de varianza de un factor con las tres muestras.

a. ¿Cuál es la hipótesis nula?

b. Asumiendo un nivel de significancia de 0.05, ¿qué conclusión indican los resultados mostrados?

XLSTAT

Source	DF	Sum of squares	Mean squares	F	Pr > F
Q1	2	4263.0833	2131.5417	1.9104	0.1729

6. Monedas de ¢25 Suponga que los pesos de las monedas de ¢25 acuñadas después de 1964 se distribuyen normalmente con una media de 5.670 g y una desviación estándar de 0.062 g (con base en las especificaciones de Mint de EE.UU.).

a. Encuentre la probabilidad de que una moneda de ¢25 seleccionada al azar pese entre 5.600 g y 5.700 g.

b. Si se seleccionan al azar 25 monedas de ¢25, encuentre la probabilidad de que su peso promedio sea mayor a 5.675 g.

c. Encuentre la probabilidad de que cuando se seleccionan ocho monedas de ¢25 aleatoriamente, todas pesen menos de 5.670 g.

d. Si una máquina expendedora está diseñada para aceptar monedas de ¢25 con pesos por encima del percentil 10, P_{10}, encuentre el peso que separa las monedas aceptables de aquellas que no son aceptables.

7. Encuesta de prioridad del trabajo *USA Today* informó sobre una encuesta de Adecco Staffing aplicada a 1000 adultos seleccionados al azar. Entre los encuestados, 20% eligieron los beneficios de salud como lo más importante de su trabajo.

a. ¿Cuál es la cantidad de encuestados que eligieron los beneficios de salud como lo más importante de su trabajo?

b. Elabore una estimación del intervalo de confianza del 95% para la proporción de todos los adultos que eligen los beneficios de salud como lo más importante de su trabajo.

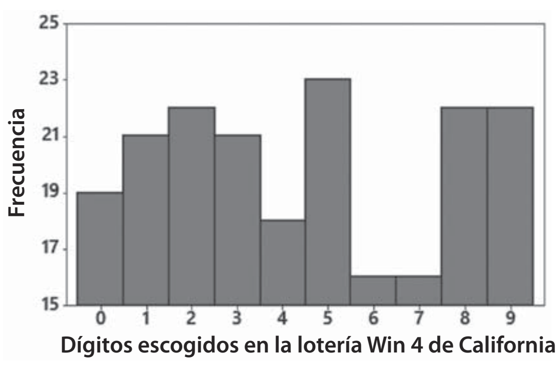
c. Con base en el resultado del inicio (b), ¿podemos concluir con seguridad que la verdadera proporción es diferente de 1/4? ¿Por qué?

8. Lotería Win 4 A continuación se muestra un histograma de dígitos seleccionados en la lotería Win 4 de California. Cada sorteo incluye la selección al azar (con reemplazo) de cuatro dígitos entre 0 y 9 inclusive.

a. ¿Qué es fundamentalmente erróneo en la gráfica?

b. ¿La pantalla bosqueja una distribución normal? ¿por qué sí o por qué no? ¿Cuál debería ser la forma del histograma?

c. Identifique las frecuencias y luego pruebe la afirmación de que los dígitos se seleccionan de una población en la que todos los dígitos son igualmente probables. ¿Hay algún problema con la lotería?



Proyecto de tecnología

¿El peso cambia con la edad? Consulte el conjunto de datos 1 "Datos corporales" en el apéndice B y utilice los pesos de los hombres divididos en tres rangos de edad diferentes de 18 a 25, 26 a 40 y 41 a 80. Pruebe la afirmación de que los hombres en esos tres grupos de edad tienen el mismo peso medio.

Clasificación Un desafío en este proyecto es identificar los pesos de los hombres en los tres grupos de edad. Primero, use la función de *clasificación* de su tecnología para ordenar todas las columnas usando el *género* como base para la clasificación. A continuación, puede eliminar todas las filas que representan mujeres. Luego, clasifique todas las columnas usando la *edad* como la base para ordenar. Entonces será mucho más fácil identificar los pesos en los diferentes grupos de edad.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: La brecha de edad/género

El conjunto de datos 18 "Citas rápidas" en el apéndice B incluye datos de un estudio de citas rápidas, y el ejercicio de repaso 4 "Citas rápidas" incluye algunos de esos datos en una tabla bidireccional con diferentes grupos de edad para hombres y mujeres. Use los métodos de este capítulo para analizar las calificaciones de atractivo en el conjunto de datos, con el fin de determinar si existe una brecha de género relacionada con los grupos de edad. ¿Las mujeres mayores parecen sentirse más atraídas por hombres más jóvenes? ¿Los hombres mayores parecen sentirse más atraídos por las mujeres más jóvenes? ¿O no hay diferencia?

Actividades en equipo

1. Actividad fuera de clase Los puntajes de Flesch Reading Ease y los puntajes de Flesch-Kincaid Grade Level miden la legibilidad de un texto. Algunos programas, como Microsoft Word, incluyen características que le permiten obtener puntajes de legibilidad de manera automática. Divídanse en grupos de tres o cuatro estudiantes. Utilizando al menos tres muestras de escritura diferentes, como el *New York Times*, el *USA Today* y el *Onion*, obtenga lo puntajes de legibilidad para diez muestras de texto de cada fuente. Use los métodos de este capítulo para determinar si hay alguna diferencia.

2. Actividad en clase Divida la clase en tres grupos. Un grupo debe registrar la frecuencia del pulso de cada miembro mientras él o ella permanece sentado. El segundo grupo debe registrar la frecuencia del pulso de cada miembro mientras él o ella está de pie. El tercer grupo debe registrar la frecuencia del pulso de cada miembro inmediatamente después de que él o ella se pare y se siente 10 veces. Analice los resultados. ¿Qué indican éstos?

3. Actividad en clase Pida a cada alumno de la clase que calcule la longitud del aula. Especifique que la longitud es la distancia entre el pizarrón y la pared opuesta. En la misma hoja de papel, cada alumno también debe escribir su género (masculino/femenino) y especialidad. Luego divídalos en grupos de tres o cuatro, y use los datos de toda la clase para responder las siguientes preguntas:

- ¿Hay una diferencia significativa entre la estimación media de los hombres y la estimación media de las mujeres?
- ¿Hay suficiente evidencia para rechazar la igualdad de las estimaciones medias para las diferentes especialidades? Describa cómo se categorizaron las especialidades.
- ¿Una interacción entre el género y la especialidad tiene un efecto en la longitud estimada?
- ¿El género parece tener un efecto en la longitud estimada?
- ¿La especialidad parece tener un efecto en la longitud estimada?

4. Actividad fuera de clase Biographyonline.net incluye información sobre las vidas de artistas, políticos, científicos, actores y otros personajes notables. Diseñe y realice un estudio observacional que comience con la elección de muestras de grupos selectos, seguida de una comparación de la vida de las personas de los diferentes grupos. ¿Algún grupo en particular parece tener una esperanza de vida diferente a la de los demás grupos? ¿Puede usted explicar tales diferencias?

5. Actividad fuera de clase Divídanse en grupos de tres o cuatro estudiantes. Cada grupo debe encuestar a otros estudiantes en la misma universidad pidiéndoles que identifiquen su especialidad y género. Puede incluir otros factores, como empleo (ninguno, tiempo parcial, tiempo completo) y edad (menor de 21, 21-30, más de 30). Para cada sujeto encuestado, determine la cantidad de seguidores de Twitter o amigos de Facebook.

- ¿El género parece tener un efecto en la cantidad de seguidores/amigos?
- ¿La especialidad tiene un efecto en la cantidad de seguidores/amigos?
- ¿Una interacción entre el género y la especialidad tiene un efecto en la cantidad de seguidores/amigos?

13

PRUEBAS NO PARAMÉTRICAS

PROBLEMA DEL CAPÍTULO

¿Los mejores televisores cuestan más?

La tabla 13-1, en la página siguiente, incluye rangos y costos (en cientos de dólares) de televisores LCD (con pantalla de cristal líquido) de al menos 60 pulgadas (según datos de *Consumer Reports*). Los rangos se basan en "puntuaciones generales" determinados por *Consumer Reports*, los menores números de rango corresponden a los "mejores" televisores con puntuaciones generales más altas. Entre los televisores incluidos en la tabla 13-1, el mejor televisor tiene un rango de 1 y cuesta $2300. ¿Hay una correlación entre los rangos y el costo? Si es así, ¿parece que los mejores televisores cuestan más? ¿Se obtiene lo que se paga?

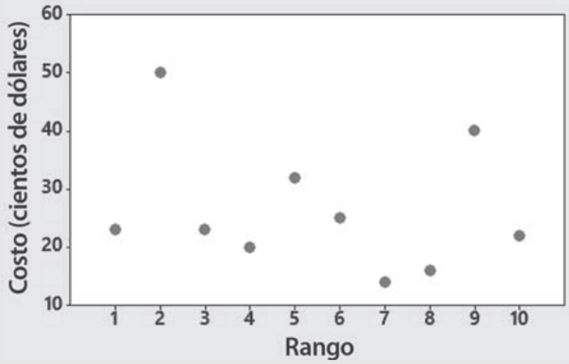
Sería inteligente comenzar el análisis con una exploración básica de los datos. Como queremos abordar el problema de la correlación, trazamos el diagrama de dispersión que se muestra en la página siguiente. Resulta claro que no existe un patrón en línea recta distintivo, por lo que no parece haber una correlación lineal. Podríamos ir más allá de este juicio subjetivo y proceder a calcular un coeficiente de correlación lineal r, pero consideremos la

continúa

naturaleza de los datos. Específicamente, los rangos simplemente identifican un orden y en realidad no miden ni cuentan nada. En lugar de usar el método de correlación lineal de la sección 10-1, podemos emplear el método de correlación de rangos descrito en la sección 13-6. Así podremos proporcionar resultados objetivos que son mejores que un juicio subjetivo.

TABLA 13-1 Rangos y costos de televisores LCD

Rango de calidad	1	2	3	4	5	6	7	8	9	10
Costo (cientos de dólares)	23	50	23	20	32	25	14	16	40	22



OBJETIVOS DEL CAPÍTULO

Los objetivos del capítulo son:

13-1 Conceptos básicos de las pruebas no paramétricas

- Desarrollar la capacidad de describir la diferencia entre las pruebas paramétricas y las pruebas no paramétricas.
- Identificar las ventajas y desventajas de las pruebas no paramétricas.
- Saber que las pruebas no paramétricas suelen ser menos *eficientes* que las pruebas paramétricas correspondientes.
- Desarrollar la capacidad de convertir datos en *rangos*.

13-2 Prueba del signo

- Desarrollar la capacidad de realizar una prueba del signo para afirmaciones que involucren pares de datos muestrales, que impliquen datos nominales o que traten de la mediana de una población.

13-3 Prueba de rangos con signo de Wilcoxon para datos pareados

- Desarrollar la capacidad de aplicar la prueba de rangos con signo de Wilcoxon para datos muestrales que constan de pares relacionados.

13-4 Prueba de la suma de rangos de Wilcoxon para dos muestras independientes

- Desarrollar la capacidad de aplicar la prueba de la suma de rangos de Wilcoxon para datos muestrales de dos poblaciones independientes.

13-5 Prueba de Kruskal-Wallis para tres o más muestras

- Desarrollar la capacidad de aplicar la prueba de Kruskal-Wallis para datos muestrales de tres o más poblaciones independientes.

13-6 Correlación de rangos

- Desarrollar la capacidad de calcular el valor del coeficiente de correlación del rango r_s, y utilizarlo para determinar si existe una correlación entre dos variables.

13-7 Prueba de rachas para aleatoriedad

- Desarrollar la capacidad de usar la prueba de rachas para aleatoriedad, para determinar si los datos muestrales ocurren en una secuencia aleatoria.

13-1 Conceptos básicos de las pruebas no paramétricas

Este capítulo presenta métodos de pruebas *no paramétricas*, que no tienen los requisitos más estrictos de las correspondientes pruebas paramétricas, las cuales se basan en muestras de poblaciones con parámetros específicos como μ o σ.

DEFINICIONES

Las **pruebas paramétricas** tienen requisitos sobre la distribución de las poblaciones involucradas; las **pruebas no paramétricas** (o **sin distribución**) no requieren que las muestras provengan de poblaciones con distribuciones normales o cualquier otra distribución en particular.

Terminología engañosa El término *prueba sin distribución* indica correctamente que una prueba no requiere una distribución particular. El término *pruebas no paramétricas* es engañoso en el sentido de que sugiere que las pruebas no se basan en un parámetro, pero hay algunas pruebas no paramétricas que se basan en un parámetro como la mediana. Debido al uso generalizado del término *prueba no paramétrica*, usamos esa terminología, pero la definimos como una prueba que no requiere una distribución particular. (El autor prefiere el término *prueba sin distribución*, pero no estaba de primero en la fila cuando se repartieron las definiciones).

Ventajas y desventajas

Ventajas de las pruebas no paramétricas

1. Debido a que las pruebas no paramétricas tienen requisitos menos rígidos que las pruebas paramétricas, se pueden aplicar a una variedad más amplia de situaciones.
2. Las pruebas no paramétricas se pueden aplicar a más tipos de datos que las pruebas paramétricas. Por ejemplo, las pruebas no paramétricas se pueden usar con datos que constan de rangos, y pueden emplearse con datos categóricos, como los géneros de los encuestados.

Desventajas de las pruebas no paramétricas

1. Las pruebas no paramétricas tienden a desperdiciar información porque los datos numéricos exactos a menudo se reducen a una forma cualitativa. Por ejemplo, con la prueba no paramétrica del signo (sección 13-2), las pérdidas de peso de las personas que hacen dieta se registran simplemente como signos negativos, y las magnitudes reales de las pérdidas de peso se omiten.
2. Las pruebas no paramétricas no son tan *eficientes* como las pruebas paramétricas, por lo que una prueba no paramétrica generalmente necesita pruebas más sólidas (como una muestra más grande o mayores diferencias) para rechazar una hipótesis nula.

Eficiencia de las pruebas no paramétricas Cuando se satisfacen los requisitos de las distribuciones de población, las pruebas no paramétricas son generalmente menos eficientes que sus correspondientes pruebas paramétricas. Por ejemplo, la sección 13-6 presenta el concepto de *correlación de rangos*, que tiene una calificación de eficiencia de 0.91 cuando se compara con la correlación lineal en la sección 10-1. Esto significa que, con todo lo demás igual, el método no paramétrico de correlación de rangos en la sección 13-6 requiere 100 observaciones muestrales para lograr los mismos resultados que 91 observaciones muestrales analizadas a través de la correlación lineal paramétrica en la sección 10-1, asumiendo que se satisfacen los requisitos más estrictos para utilizar la prueba paramétrica. En la tabla 13-2 se listan las pruebas no paramétricas junto con la prueba paramétrica correspondiente y la calificación de **eficiencia**. Dicha tabla muestra que varias pruebas no paramétricas tienen índices de eficiencia superiores a 0.90, por lo que la menor eficiencia podría no ser un factor importante al elegir entre pruebas paramétricas y no paramétricas. Sin embargo, dado que las pruebas paramétricas tienen calificaciones de eficiencia más altas que sus homólogas no paramétricas, generalmente es mejor usar las pruebas paramétricas cuando se satisfacen los supuestos requeridos.

TABLA 13-2 Eficiencia: Comparación de pruebas paramétricas y no paramétricas

Aplicación	Prueba paramétrica	Prueba no paramétrica	Calificación de eficiencia de prueba no paramétrica con poblaciones normales
Pares relacionados de datos muestrales	prueba *t*	Prueba del signo o prueba de rangos con signo de Wilcoxon	0.63 0.95
Dos muestras independientes	prueba *t*	Prueba de la suma de rangos de Wilcoxon	0.95
Tres o más muestras independientes	Análisis de varianza (prueba *F*)	Prueba de Kruskal-Wallis	0.95
Correlación	Correlación lineal	Prueba de correlación de rangos	0.91
Aleatoriedad	Sin prueba paramétrica	Prueba de rachas	Sin base para comparación

Rangos

Las secciones 13-2 a 13-5 usan métodos basados en rangos, definidos de la siguiente manera.

DEFINICIÓN

Los datos se *ordenan* cuando se organizan según algún criterio, como del más pequeño al más grande o del mejor al peor. Un **rango** es un número asignado a un elemento muestral individual de acuerdo con su lugar en la lista ordenada. Al primer elemento se le asigna un rango de 1, al segundo elemento se le asigna un rango de 2 y así sucesivamente.

Manejo de rangos empatados Si se produce un empate en los rangos, un procedimiento muy común es encontrar la media de los rangos involucrados en el empate y luego asignar este rango medio a cada uno de los elementos empatados, como en el siguiente ejemplo.

EJEMPLO 1 Manejo de rangos empatados

Los números 4, 5, 5, 5, 10, 11, 12 y 12 tienen rangos de 1, 3, 3, 3, 5, 6, 7.5 y 7.5, respectivamente. La siguiente tabla ilustra el procedimiento para manejar los empates.

Datos ordenados	Rango preliminar	Rango
4	1	**1**
5 5 5	2 3 La media es 3 4	**3** **3** **3**
10	5	**5**
11	6	**6**
12 12	7 8 La media es 7.5	**7.5** **7.5**

13-2 Prueba del signo

Concepto clave En esta sección se presenta la *prueba del signo*, que implica la conversión de los valores de datos a signos positivos y negativos, para después realizar pruebas con la finalidad de determinar si cualquiera de los signos se produce significativamente más a menudo que el otro signo.

DEFINICIÓN

La **prueba del signo** es una prueba no paramétrica (sin distribución) que utiliza signos positivos y negativos para evaluar diferentes afirmaciones, entre las que se incluyen las siguientes:

1. Afirmaciones que involucran pares de datos muestrales
2. Afirmaciones que implican datos nominales con dos categorías
3. Afirmaciones sobre la mediana de una sola población

Concepto básico de la prueba del signo La idea esencial que subyace a la prueba del signo es analizar las frecuencias de los signos positivos y negativos para determinar si son significativamente diferentes. Por ejemplo, considere los resultados de los ensayos clínicos del método XSORT para la selección de género. Entre las 726 parejas que usaron el método XSORT para tratar de tener una niña, 668 parejas tuvieron una niña. ¿Son *significativas* 668 niñas en 726 nacimientos? El sentido común debería sugerir que 668 niñas en 726 nacimientos son significativas, pero ¿qué pasa con 365 niñas en 726 nacimientos? ¿o 400 niñas en 726 nacimientos? La prueba del signo nos permite determinar cuándo tales resultados son significativos. La figura 13-1 resume el procedimiento para la prueba del signo.

Por consistencia y simplicidad, usaremos un dato estadístico de prueba basado en la cantidad de veces que ocurre el signo *menos frecuente*.

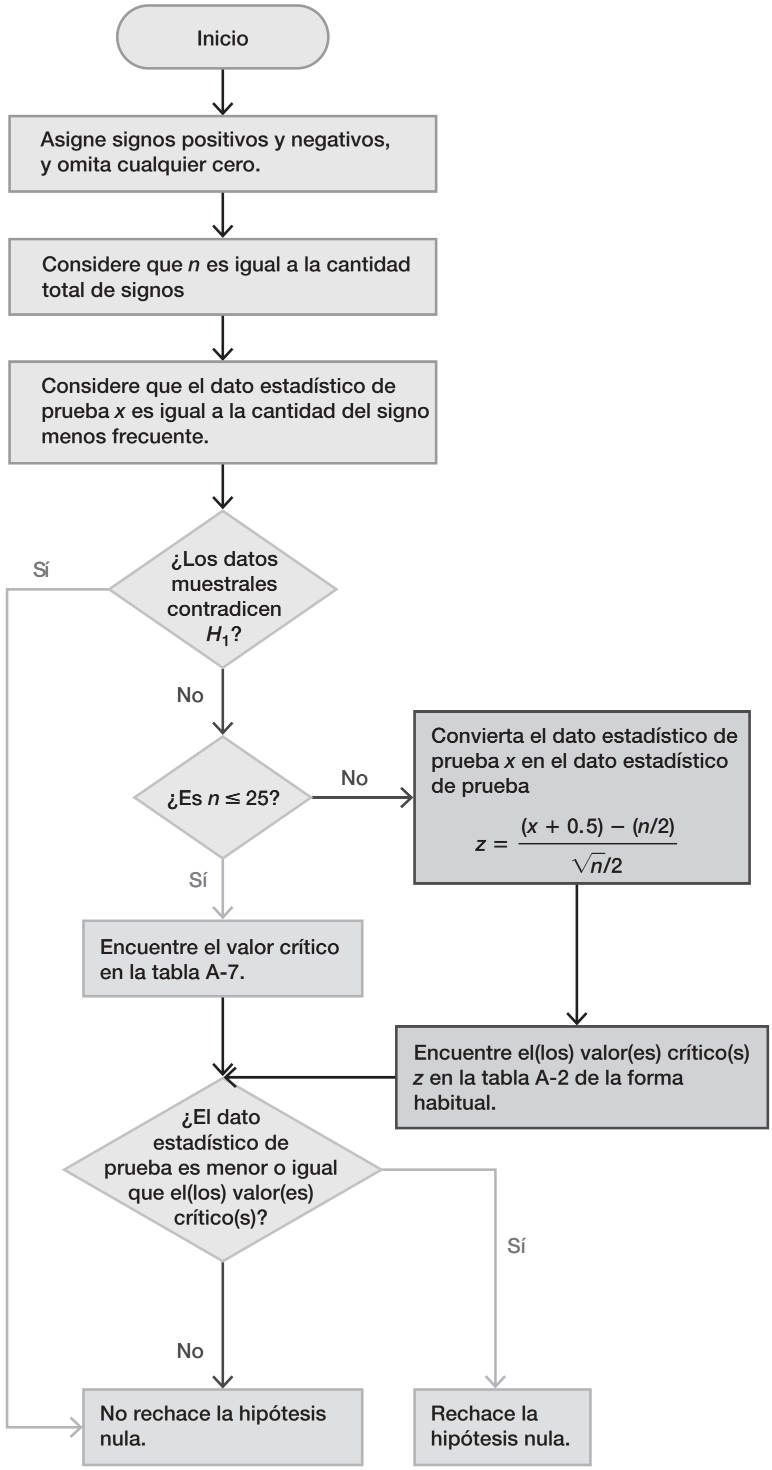


FIGURA 13-1 Procedimiento de la prueba del signo

ELEMENTOS CLAVE

Prueba del signo

Objetivo

Usar signos positivos y negativos para evaluar una afirmación que se encuentre en una de las siguientes tres categorías:

1. Pares relacionados

- Reste el segundo valor de cada par del primero, registre el signo de la diferencia y omita cualquier 0.

2. Datos nominales con dos categorías

- Represente cada miembro de una categoría con un signo positivo y cada miembro de la otra categoría con uno negativo.

3. Mediana de una población individual

- Reste la mediana de cada valor muestral, registre el signo de la diferencia y omita cualquier 0.

Notación

x = la cantidad de veces que ocurre el signo *menos frecuente*
n = el número total de signos positivos y negativos combinados

Requisitos

Los datos muestrales son una muestra aleatoria simple.

Nota: No es necesario que los datos muestrales provengan de una población con una distribución particular, por ejemplo una distribución normal.

Dato estadístico de prueba

Si $n \leq 25$: El dato estadístico de prueba es x = la cantidad de veces que ocurre el signo menos frecuente.
Si $n > 25$: El dato estadístico de prueba es

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

Valores *P*

A menudo, los valores P se obtienen mediante el uso de la tecnología, o se pueden encontrar usando el dato estadístico de prueba z.

Valores críticos

1. Si $n \leq 25$, los valores críticos x se encuentran en la tabla A-7.
2. Si $n > 25$, los valores críticos z se encuentran en la tabla A-2.

Sugerencia: Como x o z se basan en el signo *menos* frecuente, todas las pruebas unilaterales se tratan como si fueran pruebas de cola izquierda.

PRECAUCIÓN Al utilizar la prueba del signo en una prueba de una cola, evite sacar la conclusión incorrecta cuando un signo se produzca significativamente más a menudo o con menos frecuencia que el otro signo, y los datos muestrales *contradigan* la hipótesis alternativa. Una muestra de 7% del nacimiento de niños nunca se puede usar para apoyar la afirmación de que el nacimiento de un varón *ocurre más del* 50% de las veces, como en el ejemplo 1.

EJEMPLO 1 **Datos que contradicen la hipótesis alternativa**

Entre las 945 parejas que utilizaron el método de selección de género XSORT, 66 tuvieron niños, por lo que la proporción muestral de niños es 66/945 o 0.0698 (según datos del Genetics & IVF Institute). Considere la afirmación de que el método XSORT para la selección de género *aumenta* la probabilidad de bebés varones, de modo que la probabilidad de tener un niño sea $p > 0.5$. Esta afirmación de $p > 0.5$ se convierte en la hipótesis alternativa.

Usando el sentido común, vemos que con una proporción muestral de niños de 0.0698, nunca podemos apoyar una afirmación de que $p > 0.5$. (Necesitaríamos una proporción muestral de niños *mayor* de 0.5 por una cantidad significativa). Aquí, la proporción muestral de 66/945, o 0.0698, *contradice* la hipótesis alternativa porque no es mayor que 0.5.

INTERPRETACIÓN

Una hipótesis alternativa nunca puede ser respaldada con datos que la contradicen. La prueba del signo mostrará que 66 niños en 945 nacimientos son significativos, pero es significativo en la dirección incorrecta. Nunca podemos apoyar una afirmación de que $p > 0.5$ con una proporción muestral de 66/945, o 0.0698, que es *menor que* 0.5.

SU TURNO Ejercicio 3 “Contradicción de H_1”.

Afirmaciones sobre pares relacionados

Al utilizar la prueba del signo con datos que son pares relacionados, convertimos los datos brutos en signos positivos y negativos de la siguiente manera:

1. Reste cada valor de la segunda variable del valor correspondiente de la primera variable.
2. Registre sólo el *signo* de la diferencia encontrada en el paso 1. Excluya los *empates* borrando cualquier par relacionado en el que ambos valores sean iguales.

El concepto principal que subyace a este uso de la prueba del signo es el siguiente:

Si los dos conjuntos de datos tienen medianas iguales, la cantidad de signos positivos debería ser aproximadamente igual a la cantidad de signos negativos.

EJEMPLO 2 **¿Hay una diferencia de género en las edades de las mejores actrices y los mejores actores?**

La tabla 13-3 (del conjunto de datos 14 “Edades de ganadores del Oscar” en el apéndice B) lista las edades de los ganadores del premio Oscar a la mejor actriz y al mejor actor. Las edades se relacionan por el año en que se otorgaron los premios. Use la prueba del signo con los datos muestrales de la tabla 13-3 para probar la afirmación de que no hay diferencia entre las edades de las mejores actrices y los mejores actores.

TABLA 13-3 Edades de las mejores actrices y mejores actores

Mejor actriz	28	63	29	41	30	41	28	26	29	29
Mejor actor	62	52	41	39	49	41	44	51	54	50
Signo de diferencia	−	+	−	+	−	0	−	−	−	−

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS El único requisito de la prueba del signo es que los datos muestrales son una muestra aleatoria simple, y ese requisito se cumple.

Si no hay diferencia entre las edades de las mejores actrices y los mejores actores, la cantidad de signos positivos y negativos debe ser aproximadamente igual. En la tabla 13-3 tenemos 2 signos positivos, 7 signos negativos y 1 diferencia de 0. Desechamos la diferencia de 0 y procedemos con los 2 signos positivos y 7 negativos. La prueba del signo nos dice si los números de signos positivos y negativos son aproximadamente iguales.

La hipótesis nula es la afirmación de que no hay diferencia entre las edades de las mejores actrices y los mejores actores, y la hipótesis alternativa es la afirmación de que existe una diferencia.

H_0: No hay diferencia. (La mediana de las diferencias es igual a 0).

H_1: Hay una diferencia. (La mediana de las diferencias no es igual a 0).

Siguiendo el procedimiento de la prueba del signo resumido en la figura 13-1, $n = 9$ (el número total de signos positivos y negativos) y $x = 2$ (el número de signos del tipo menos frecuente, o el menor de 2 y 7).

Los datos muestrales no contradicen H_1, porque hay una diferencia entre los 2 signos positivos y los 7 signos negativos. Los datos muestrales muestran una diferencia, y debemos continuar con la prueba para determinar si esa diferencia es significativa.

La figura 13-1 muestra que con $n = 9$, debemos proceder a encontrar el valor crítico de la tabla A-7. Nos referimos a la tabla A-7, donde el valor crítico de 1 se encuentra para $n = 9$ y $\alpha = 0.05$ en dos colas.

Como $n \leq 25$, el dato estadístico de prueba es $x = 2$ (y no convertimos x en un puntaje z). Con un dato estadístico de prueba de $x = 2$ y un valor crítico x de 1, no podemos rechazar la hipótesis nula de que no hay diferencia. (Consulte la nota 2 incluida con la tabla A-7: “Rechace la hipótesis nula si el número de signos del tipo menos frecuente (x) es menor o igual que el valor de la tabla”. Como $x = 2$ *no* es menor o igual que el valor crítico de 1, no podemos rechazar la hipótesis nula). No hay evidencia suficiente para justificar el rechazo de la afirmación de que la mediana de las diferencias es igual a 0.

INTERPRETACIÓN

Concluimos que no hay evidencia suficiente para rechazar la afirmación de que no hay diferencias entre las edades de las mejores actrices y los mejores actores.

SU TURNO Resuelva el ejercicio 5 “Citas rápidas: atributos”.

Afirmaciones que implican datos nominales con dos categorías

En el capítulo 1, definimos los datos nominales como aquellos que constan sólo de nombres, etiquetas o categorías. La naturaleza de los datos nominales limita los cálculos que pueden realizarse, pero es posible identificar la *proporción* de los datos muestrales que pertenecen a una categoría particular, y podemos probar las afirmaciones sobre la proporción de población correspondiente p. El siguiente ejemplo usa datos nominales que consisten en géneros (niñas/niños). La prueba del signo se usa representando a las niñas con signos positivos (+) y a los niños con signos negativos (−). (Honestamente, esos signos se eligieron de manera arbitraria).

EJEMPLO 3 Selección de género

El Genetics & IVF Institute realizó un ensayo clínico de sus métodos para la selección del género de los bebés. Antes de que los ensayos clínicos concluyeran, 879 de 945 bebés nacidos de padres que usaban el método XSORT para la selección del género fueron niñas. Use la prueba del signo y un nivel de significancia de 0.05 para probar la afirmación de que este método de selección del género es eficaz para aumentar la probabilidad de tener una niña.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS El único requisito es que la muestra sea una muestra aleatoria simple. Según el diseño de este experimento, podemos suponer que los datos muestrales son una muestra aleatoria simple. ✓

Sea p la proporción de niñas en la población. La afirmación de que las niñas son más probables con el método XSORT se puede expresar como $p > 0.5$, por lo que las hipótesis nula y alternativa son las siguientes:

continúa

H_0: $p = 0.5$ (la proporción de niñas es igual a 0.5)

H_1: $p > 0.5$ (las niñas son más probables)

Si se representan a las niñas con signos positivos (+) y los niños por signos negativos (−), tenemos 879 signos positivos y 66 signos negativos. Utilizando el procedimiento de la prueba del signo resumido en la figura 13-1, el dato estadístico de prueba x es el menor valor entre 879 y 66, por lo que $x = 66$ niños. *En lugar de tratar de determinar si 879 niñas es un valor suficientemente grande como para ser significativamente alto, procedemos con el objetivo equivalente de tratar de determinar si 66 niños son suficientemente pocos como para ser significativamente bajos, por lo que tratamos la prueba como una prueba de la* cola izquierda.

Los datos muestrales no contradicen la hipótesis alternativa porque la proporción muestral de niñas es 879/945 o 0.930. que es mayor que 0.5, como en la hipótesis alternativa anterior. Continuando con el procedimiento de la figura 13-1, observamos que el valor de $n = 945$ es mayor que 25, por lo que el dato estadístico de prueba $x = 66$ se convierte (usando una corrección para la continuidad) en el dato estadístico de prueba z de la siguiente manera:

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

$$= \frac{(66 + 0.5) - \left(\frac{945}{2}\right)}{\frac{\sqrt{945}}{2}} = -26.41$$

Valor *P* Podríamos usar el dato estadístico de prueba $z = -26.41$ para encontrar el valor P de cola izquierda de 0.0000 (Tabla: 0.0001). Ese bajo valor de P nos hace rechazar la hipótesis nula.

Valor crítico Con $\alpha = 0.05$ en una prueba de cola izquierda, el valor crítico es $z = -1.645$. La figura 13-2 muestra que el dato estadístico de prueba $z = -26.41$ está en la región crítica limitada por $z = -1.645$, por lo que rechazamos la hipótesis nula de que la proporción de niñas es igual a 0.5.

Hay suficiente evidencia muestral para respaldar la afirmación de que las niñas son más probables con el método XSORT.

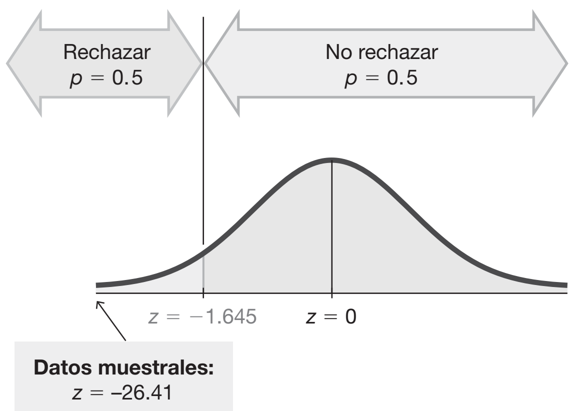


FIGURA 13-2 Prueba de efectividad del método de selección de género XSORT

INTERPRETACIÓN

El método XSORT de selección de género parece estar asociado con un aumento en la probabilidad de una niña, por lo que este método parece ser efectivo (pero esta prueba de hipótesis no prueba que el método XSORT sea la *causa* de dicho aumento).

SU TURNO Resuelva el ejercicio 9 “Encuesta de células madre”.

Afirmaciones sobre la mediana de una sola población

El siguiente ejemplo ilustra el procedimiento para usar la prueba del signo al probar una afirmación sobre la mediana de una sola población. Vea cómo los signos negativos y positivos se basan en el valor declarado de la mediana.

EJEMPLO 4 Temperaturas corporales

El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye las temperaturas corporales medidas en adultos. Use las 106 temperaturas indicadas para las 12 AM del día 2 con la prueba del signo para probar la afirmación de que la mediana es menor a 98.6 °F. De los 106 sujetos, 68 tuvieron temperaturas por debajo de 98.6 °F, 23 por encima de 98.6 °F, y 15 iguales a 98.6 °F.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS El único requisito es que la muestra sea una muestra aleatoria simple. Según el diseño de este experimento, suponemos que los datos muestrales pertenecen a una muestra aleatoria simple. ✓

La afirmación de que la mediana es menor a 98.6 °F es la hipótesis alternativa, mientras que la hipótesis nula es la afirmación 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)

De acuerdo con el procedimiento descrito en la figura 13-1, utilizamos un signo negativo para representar cada temperatura por debajo de 98.6 °F, y un signo positivo para cada temperatura superior a 98.6 °F. Descartamos los 15 valores de datos de 98.6, ya que dan como resultado diferencias iguales a cero. Tenemos 68 signos negativos y 23 signos positivos, por lo que $n = 91$ y $x = 23$ (el número de signos menos frecuentes). Los datos muestrales no contradicen la hipótesis alternativa, porque la mayoría de las 91 temperaturas están por debajo de 98.6 °F. El valor de n es mayor que 25, por lo que convertimos el dato estadístico de prueba x en el dato estadístico de prueba z:

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

$$= \frac{(23 + 0.5) - \left(\frac{91}{2}\right)}{\frac{\sqrt{91}}{2}} = -4.61$$

Valor *P* En esta prueba de cola izquierda, el dato estadístico de prueba $z = -4.61$ arroja un valor P de 0.0000 (Tabla: 0.0001). Debido a que ese valor P es tan pequeño, rechazamos la hipótesis nula.

Valor crítico En esta prueba de cola izquierda con $\alpha = 0.05$, use la tabla A-2 para obtener el valor crítico z de -1.645. En la figura 13-3 de la página siguiente vemos que el dato estadístico de prueba $z = -4.61$ está dentro de la región crítica, por lo tanto, se rechaza la hipótesis nula.

continúa

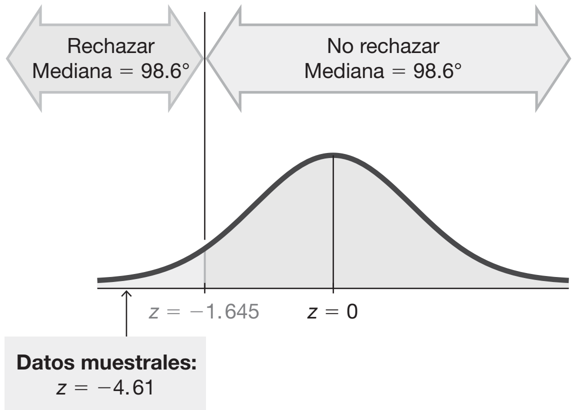


FIGURA 13-3 Prueba de la afirmación de que la mediana es inferior a 98.6 °F

INTERPRETACIÓN

Existe suficiente evidencia muestral para respaldar la afirmación de que la temperatura corporal media de los adultos sanos es menor a 98.6 °F. No es igual a 98.6 °F, como comúnmente se cree.

SU TURNO Resuelva el ejercicio 13 "Magnitudes de terremotos".

En el ejemplo 4, la prueba del signo de la afirmación de que la mediana está por debajo de 98.6 °F da como resultado un dato estadístico de prueba de $z = -4.61$ y un valor *P* de 0.00000202. Sin embargo, una prueba paramétrica de la afirmación de que $\mu < 98.6$ °F da como resultado un dato estadístico de prueba $t = -6.611$ con un valor *P* de 0.000000000813. Debido a que el valor *P* de la prueba del signo no es tan bajo como el valor *P* de la prueba paramétrica, vemos que la prueba del signo no es tan sensible como la prueba paramétrica. Ambas pruebas conducen al rechazo de la hipótesis nula, pero la prueba del signo no considera que los datos muestrales sean tan extremos, en parte porque la prueba del signo usa sólo información sobre la *dirección* de los datos, sin tomar en cuenta las *magnitudes* de los valores de los datos. La siguiente sección presenta la prueba Wilcoxon de los rangos con signo, que en gran medida supera esta dificultad.

Justificación del dato estadístico de prueba utilizado cuando $n > 25$ Cuando se encuentran valores críticos para la prueba del signo, utilizamos la tabla A-7 sólo para n hasta 25. Cuando $n > 25$, el dato estadístico de prueba z se basa en una aproximación normal a la distribución de probabilidad binomial con $p = q = 1/2$. En la sección 6-6 vimos que la aproximación normal a la distribución binomial es aceptable cuando tanto $np \geq 5$ como $nq \geq 5$. En la sección 5-2 vimos que $\mu = np$ y $\sigma = \sqrt{npq}$ para distribuciones de probabilidad binomiales. Debido a que esta prueba del signo asume que $p = q = 1/2$, cumplimos los requisitos previos $np \geq 5$ y $nq \geq 5$ siempre que $n \geq 10$. Además, suponiendo que $p = q = 1/2$, obtenemos $\mu = np = n/2$ y $\sigma = \sqrt{npq} = \sqrt{n/4} = \sqrt{n}/2$, por lo que la puntuación z estándar

$$z = \frac{x - \mu}{\sigma}$$

se convierte en

$$z = \frac{x - \left(\dfrac{n}{2}\right)}{\dfrac{\sqrt{n}}{2}}$$

Reemplazamos x por $x + 0.5$ como una corrección para la continuidad. Es decir, los valores de x son discretos, pero como utilizamos una distribución de probabilidad continua, un valor discreto como 10 está representado por el intervalo de 9.5 a 10.5. Como x representa el signo menos frecuente, actuamos de manera conservadora al referirnos sólo a $x + 0.5$; obtenemos el dato estadístico de prueba z que se muestra a continuación y en el recuadro de elementos clave.

$$z = \frac{(x + 0.5) - \left(\frac{n}{2}\right)}{\frac{\sqrt{n}}{2}}$$

CENTRO DE TECNOLOGÍA

Prueba del signo

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Sign Test** en el menú desplegable.
3. *Número conocido de signos:* Seleccione **el número dado de signos**, elija el formato para la afirmación, ingrese el nivel de significancia y los números de signos positivos y negativos.

   Pares de valores: Seleccione **Given Pairs of Values**, elija el formato para la afirmación, ingrese el nivel de significancia y seleccione las dos columnas de datos a incluir.
4. Haga clic en **Evaluate**.

Minitab

Minitab requiere una sola columna de valores.

Pares relacionados: *Ingrese una columna que contenga las diferencias.*

Datos nominales en dos categorías: *Ingrese un 1 para cada valor de una categoría y −1 para cada valor de la otra categoría. Ingrese 0 para el valor declarado de la mediana.*

Valores individuales que se probarán con la mediana declarada: *Ingrese los valores muestrales en una sola columna.*

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Nonparametrics** en el menú desplegable y elija **1-Sample Sign** en el submenú.
3. En *Variables*, seleccione la columna que contiene los datos a analizar.
4. Seleccione **Test Median** e ingrese el valor mediano declarado.
5. Elija el formato de la hipótesis alternativa.
6. Haga clic en **OK**.

StatCrunch

StatCrunch puede probar la afirmación de que una sola lista de valores proviene de una población con una mediana igual a algún valor especificado.

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Nonparametrics** en el menú desplegable y **Sign Test** en el submenú.
3. Seleccione la columna que contiene los datos a analizar.
4. Seleccione **Hypothesis test for median** y para H_0 ingrese el valor de la mediana declarada. Para H_A seleccione el formato deseado.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

La calculadora TI-83/84 Plus no tiene una función dedicada a la prueba del signo, pero la función binomcdf *de la calculadora se puede usar para encontrar el valor P para una prueba del signo.*

1. Presione **2ND** y luego las teclas **VARS** para acceder al menú *DISTR* (distribuciones).
2. Seleccione **binomcdf** y haga clic en **ENTER**.
3. Ingrese los valores para los ensayos *n*, *p* y *x* para completar el comando **binomcdf(*n*,*p*,*x*)**. Para los ensayos, ingrese la cantidad total de signos positivos y negativos. Para *p* ingrese **0.5.** Para *x* introduzca el número de signos del tipo menos frecuente.
4. Presione **ENTER**. El resultado es la probabilidad de obtener *x* o menos ensayos. *Duplique este valor para las pruebas de dos colas.*

SUGERENCIA: El resultado final es el valor *P*, por lo tanto, rechace la hipótesis nula si el valor *P* es menor o igual que el nivel de significancia. De lo contrario, no se puede rechazar la hipótesis nula.

CENTRO DE TECNOLOGÍA *continuación*

Prueba del signo

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la lista de opciones y luego haga clic en **Nonparametric tests**.
2. Seleccione **Comparison of two samples** del menú desplegable.
3. Ingrese el rango de datos para cada muestra en los *recuadros de las muestras 1 y 2*. Marque la casilla **Columns labels** si el rango de datos incluye etiquetas.
4. Seleccione **Paired samples** en el *formato de datos*.
5. Marque solamente la opción **sign test**.
6. Haga clic en la pestaña **Options**.
7. En Hipótesis Alternativa seleccione **Sample 1 – Sample 2 ≠ D**. Confirme que la *diferencia hipotética* (*D*) es **0**.
8. Ingrese un nivel de significancia y marque la casilla de **Exact p-value**.
9. Haga clic en **OK**.

Excel

Excel no tiene una función dedicada a la prueba del signo, pero puede usarse para encontrar el valor P para una prueba del signo.

1. Haga clic en el botón **Insert Function f_x**, seleccione la categoría **Statistical**, elija la función **BINOM.DIST** y haga clic en **OK**.
2. Para *Number_s* ingrese la cantidad de veces que aparece el signo *menos frecuente*. Para *Trials* ingrese la cantidad total de signos positivos y negativos. Para *probability-s*, ingrese **0.5**. Para *Cumulative*, ingrese **1** para “Verdadero”.
3. Haga clic en **OK**. Se mostrará el valor *P* de una sola cola. *Duplique este valor para pruebas de dos colas.*

SUGERENCIA: El resultado final es el valor *P*, por lo tanto, rechace la hipótesis nula si el valor *P* es menor o igual que el nivel de significancia. De lo contrario, no se puede rechazar la hipótesis nula.

13-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Prueba del signo para Novatos 15 La siguiente tabla lista algunos de los pesos (en kg) del conjunto de datos 6 “Novatos 15” del apéndice B. Esos pesos se midieron de estudiantes universitarios en septiembre y más tarde en abril de su primer año. Supongamos que planeamos usar la prueba del signo para probar la afirmación de que no hay diferencia entre los pesos de septiembre y los pesos de abril. ¿Qué requisitos deben cumplirse para esta prueba? ¿Existe algún requisito de que las poblaciones deban tener una distribución normal o cualquier otra distribución específica? ¿En qué sentido es esta prueba de signo una “prueba libre sin distribución”?

Peso de septiembre (kg)	67	53	64	74	67	70	55	74	62	57
Peso de abril (kg)	66	52	68	77	67	71	60	82	65	58

2. Identificación de signos Para la prueba del signo descrita en el ejercicio 1, identifique el número de signos positivos, el número de signos negativos, el número de empates, el tamaño de muestra n que se usa para la prueba del signo y el valor del dato estadístico de prueba.

3. Contradicción de H_1 Un paso importante en la realización de la prueba de signos es determinar si los datos muestrales contradicen la hipótesis alternativa H_1. Para la prueba del signo descrita en el ejercicio 1, identifique la hipótesis nula y la hipótesis alternativa, y explique por qué los datos muestrales contradicen o no contradicen la hipótesis alternativa.

4. Eficiencia de la prueba del signo Consulte la tabla 13-2 en la página 600 e identifique la eficiencia de la prueba del signo. ¿Qué nos dice ese valor sobre la prueba del signo?

Pares relacionados. ***En los ejercicios 5 a 8, use la prueba del signo para los datos que consisten en pares relacionados.***

5. Citas rápidas: Atributos A continuación se listan calificaciones de “atributos” dadas por parejas que participaron en una sesión de citas rápidas. Cada calificación de atributo es la suma de las calificaciones de cinco atributos (sinceridad, inteligencia, diversión, ambición, intereses compartidos). Las calificaciones listadas son del conjunto de datos 18 “Citas rápidas” en el apéndice B. Use un nivel de significancia de 0.05 para probar la afirmación de que existe una diferencia entre las calificaciones de atributos femeninos y las calificaciones de atributos masculinos.

Calificación de hombre dada por mujer	29	38	36	37	30	34	35	23	43
Calificación de mujer dada por hombre	36	34	34	33	31	17	31	30	42

6. Citas rápidas: Atractivo A continuación se listan calificaciones de “atractivo” (1 = nada atractivo, 10 = extremadamente atractivo) dadas por parejas que participaron en una sesión de citas rápidas. Los valores listados pertenecen al conjunto de datos 18 “Citas rápidas”. Use un nivel de significancia de 0.05 para probar la afirmación de que existe una diferencia entre las calificaciones de atractivo femenino y las calificaciones de atractivo masculino.

Calificación de hombre dada por mujer	4	8	7	7	6	8	6	4	2	5	9.5	7
Calificación de mujer dada por hombre	6	8	7	9	5	7	5	4	6	8	6	5

7. Citas rápidas: Atributos Repita el ejercicio 5 usando todas las calificaciones de atributos en el conjunto de datos 18 “Citas rápidas” del apéndice B.

8. Citas rápidas: Atractivo Repita el ejercicio 6 utilizando todas las calificaciones de atractivo en el conjunto de datos 18 “Citas rápidas” del apéndice B.

Datos nominales. ***En los ejercicios 9 a 12, use la prueba del signo para la afirmación que involucra datos nominales.***

9. Encuesta de células madre *Newsweek* realizó una encuesta en la que se les preguntó a los encuestados si “están a favor o en contra de utilizar dinero de los impuestos federales para financiar la investigación médica utilizando células madre obtenidas de embriones humanos”. De los encuestados, 481 estuvieron a favor, 401 se opusieron y 120 no estaban seguros. Use un nivel de significancia de 0.01 para probar la afirmación de que no hay diferencia entre las proporciones de quienes están a favor y en contra.

10. Negligencia médica En un estudio de 1228 demandas por negligencia médica seleccionadas al azar, se descubrió que 856 de ellas fueron canceladas o descartadas (según los datos de *Physicians Insurers Association of America*). Use un nivel de significancia de 0.01 para probar la afirmación de que existe una diferencia entre la tasa de demandas por negligencia médica que van a juicio y la tasa de dichas demandas que se cancelan o descartan.

11. Nacimientos Una muestra aleatoria de 860 nacimientos en el estado de Nueva York incluyó 426 niños y 434 niñas. Use un nivel de significancia de 0.05 para evaluar la afirmación de que al nacer los bebés, los niños y las niñas son igualmente probables.

12. Regla del tiempo extra en el fútbol americano Antes de que la regla del tiempo extra en la Liga Nacional de Fútbol americano se cambiara en 2011, entre 460 juegos con tiempo extra, 252 fueron ganados por el equipo que ganó el lanzamiento de la moneda al comienzo del tiempo extra. Usando un nivel de significancia de 0.05, pruebe la afirmación de que el lanzamiento de la moneda es justo en el sentido de que ninguno de los equipos tiene una ventaja al ganarlo. ¿El lanzamiento de la moneda parece ser justo?

Conjuntos de datos del apéndice B. ***En los ejercicios 13 a 16, consulte el conjunto de datos indicado en el apéndice B y use la prueba del signo para la afirmación sobre la mediana de una población.***

13. Magnitudes de terremotos Consulte en el conjunto de datos 21 “Terremotos” del apéndice B las magnitudes de terremotos. Use un nivel de significancia de 0.01 para probar la afirmación de que la mediana es igual a 2.00.

14. Profundidades de terremotos Consulte en el conjunto de datos 21 “Terremotos” del apéndice B las profundidades de terremotos (km). Use un nivel de significancia de 0.01 para probar la afirmación de que la mediana es igual a 5.0 km.

15. Prueba del peso mediano de monedas de ¢25 Consulte en el conjunto de datos 29 "Pesos de monedas" del apéndice B los pesos (g) de monedas de ¢25 seleccionadas al azar, que se acuñaron después de 1964. Se supone que las monedas de esta denominación tienen un peso mediano de 5.670 g. Use un nivel de significancia de 0.01 para probar la afirmación de que la mediana es igual a 5.670 g. ¿Las monedas de ¢25 parecen estar acuñadas de acuerdo con las especificaciones?

16. Old Faithful Consulte en el conjunto de datos 23 "Old Faithful" del apéndice B los intervalos de tiempo antes de las erupciones del géiser Old Faithful. Use un nivel de significancia de 0.05 para probar la afirmación de que esos tiempos pertenecen a una población con una mediana de 90 minutos.

13-2 Más allá de lo básico

17. 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 −). Un segundo método es tratar la mitad de los 0 como signos positivos y la mitad como signos negativos. (Si el número de ceros es impar, excluya uno para que se puedan dividir por igual). Con un tercer método, en las pruebas de dos colas, la mitad de los ceros es positiva y la mitad negativa; en las pruebas de una cola, todos los ceros son positivos o negativos, lo que sea que apoye la hipótesis nula. Repita el ejemplo 4 "Temperaturas corporales" usando el segundo y el tercer métodos para manejar los empates. ¿Los diferentes métodos conducen a estadísticos de prueba, valores P y conclusiones muy distintos?

18. Determinación de valores críticos En la tabla A-7 se listan los valores críticos para opciones limitadas de α. Use la tabla A-1 para agregar una nueva columna en la tabla A-7 (de $n = 1$ a $n = 8$) que represente un nivel de significancia de 0.03 en una cola o 0.06 en dos colas. Para cualquier n particular, use $p = 0.5$, porque la prueba del signo requiere la suposición de que P(signo positivo) = P(signo negativo) = 0.5. La probabilidad de x o menos signos similares es la suma de las probabilidades para valores que incluyen x.

13-3 Prueba de rangos con signo de Wilcoxon para datos pareados

Concepto clave En esta sección se presenta la *prueba de rangos con signo de Wilcoxon*, que inicia con la conversión de los datos muestrales en rangos. Esta prueba se puede usar para las dos aplicaciones que se describen en la siguiente definición.

DEFINICIÓN

La **prueba de rangos con signo de Wilcoxon** es una prueba no paramétrica que utiliza rangos para las siguientes aplicaciones:

1. Probar una afirmación de que una población de datos pareados tiene la propiedad de que los pares relacionados tienen diferencias con una mediana igual a cero
2. Probar una afirmación de que una sola población de valores individuales tiene una mediana igual a algún valor declarado

Cuando se prueba un valor declarado de una mediana para una población de valores individuales, creamos datos pareados relacionando cada valor muestral con la mediana declarada, por lo que se usa el mismo procedimiento para las dos aplicaciones anteriores.

Afirmaciones que involucran datos pareados

La prueba del signo (sección 13-2) se puede usar con datos pareados, pero esta prueba sólo usa los signos de las diferencias. Al usar rangos en vez de signos, la prueba de rangos con signo de Wilcoxon tiene en cuenta las magnitudes de las diferencias, por lo que incluye y utiliza más información que la prueba del signo y, por lo tanto, tiende a arrojar conclusiones que reflejan de mejor manera la verdadera naturaleza de los datos.

ELEMENTOS CLAVE

Prueba de rangos con signo de Wilcoxon

Objetivo: Utilizar la prueba de rangos con signo de Wilcoxon para la siguiente prueba:

- **Datos pareados:** Probar la afirmación de que una población de datos pareados posee la propiedad de que los pares relacionados tienen diferencias con una mediana igual a cero.
- **Una población de valores individuales:** Probar la afirmación de que una población tiene una mediana igual a algún valor declarado. (Al emparejar cada valor muestral con la mediana declarada, trabajamos de nuevo con datos pareados).

Notación

T = la menor de las siguientes dos sumas:

1. La suma de los rangos positivos de las diferencias d distintas de cero.
2. El valor absoluto de la suma de los rangos negativos de las diferencias d distintas de cero.

(Los detalles para evaluar T se dan en el procedimiento que se describe después de este recuadro de elementos clave).

Requisitos

1. Los datos son una muestra aleatoria simple.
2. La población de diferencias tiene una distribución que es aproximadamente *simétrica*, lo que significa que la mitad izquierda de su histograma es más o menos una imagen especular de su mitad derecha. (Para una muestra de datos pareados, obtenga diferencias restando el segundo valor del primer valor en cada par, para una muestra de valores individuales, obtenga diferencias restando el valor de la mediana declarada de cada valor muestral).

Nota: No es obligatorio que los datos tengan una distribución normal

Dato estadístico de prueba

Si $n \leq 30$, el dato estadístico de prueba es T.

Si $n > 30$, el dato estadístico de prueba es $$z = \frac{T - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}}$$

Valores *P*

Con frecuencia los valores P son proporcionados por la tecnología, o se pueden encontrar usando el dato estadístico de prueba y la tabla A-2.

Valores críticos

1. Si $n \leq 30$, el valor T crítico se encuentra en la tabla A-8.
2. Si $n > 30$, los valores z críticos se encuentran en la tabla A-2.

El siguiente procedimiento requiere que usted ordene los datos y luego asigne rangos. Cuando se trabaja con conjuntos de datos más grandes, la ordenación y la clasificación se vuelven tediosas, pero se puede utilizar la tecnología para automatizar ese proceso. Los diagramas de tallo y hojas también pueden ser muy útiles en la clasificación de datos.

Procedimiento de rangos con signo de Wilcoxon Para ver cómo se aplican los siguientes pasos, consulte la muestra de datos pareados que se listan en las dos primeras filas de la tabla 13-4 en la página siguiente. Suponga que queremos probar la hipótesis nula de que los datos pareados pertenecen a una población de pares relacionados con diferencias que tienen una mediana igual a cero.

TABLA 13-4 Edades de las mejores actrices y los mejores actores

Mejor actriz	28	63	29	41	30	41	28	26	29	29
Mejor actor	62	52	41	39	49	41	44	51	54	50
d (diferencia)	−34	+11	−12	+2	−19	0	−16	−25	−25	−21
Rango de [d]	9	2	3	1	5	×	4	7.5	7.5	6
Rango con signo	−9	+2	−3	+1	−5	×	−4	−7.5	−7.5	−6

Procedimiento de los rangos con signo de Wilcoxon

Paso 1: Para cada par de datos, encuentre la diferencia d restando el segundo valor del primer valor. Deseche cualquier pareja que tenga una diferencia de 0.

EJEMPLO: La tercera fila de la tabla 13-4 lista las diferencias encontradas restando las edades de los mejores actores de las edades de las mejores actrices, y la diferencia de 0 se ignorará en los siguientes pasos.

Paso 2: *No tome en cuenta los signos de las diferencias*, luego ordene las diferencias de menor a mayor y reemplace las diferencias por el valor de rango correspondiente (como se describe en la sección 13-1). Cuando las diferencias tengan el mismo valor numérico, asígneles la media de los rangos involucrados en el empate.

EJEMPLO: La cuarta fila de la tabla 13-4 muestra los rangos de los valores de $|d|$. Si se descarta la diferencia de 0, el valor más pequeño de $|d|$ es 2, por lo que se le asigna el rango de 1. El siguiente valor más pequeño de $|d|$ es 11, por lo que se le asigna un rango de 2, y así sucesivamente. (Empate: Hay dos valores $|d|$ de 25, por lo que encontramos la media de los rangos de 7 y 8, que es 7.5. A cada uno de esos valores $|d|$ se le asigna un rango de 7.5).

Paso 3: Adjunte a cada rango el signo de la diferencia de la que proviene. Es decir, inserte los signos que se descartaron en el paso 2.

EJEMPLO: La fila inferior de la tabla 13-4 lista los mismos rangos encontrados en la cuarta fila, pero se insertan los signos de las diferencias que se muestran en la tercera fila.

Paso 4: Encuentre la suma de los rangos que son positivos. También encuentre el valor absoluto de la suma de los rangos negativos.

EJEMPLO: La fila inferior de la tabla 13-4 lista los rangos con signo. La suma de los rangos positivos es $2 + 1 = 3$. La suma de los rangos negativos es $(-9) + (-3) + (-5) + (-4) + (-7.5) + (-7.5) + (-6) = -42$ y el valor absoluto de esta suma es 42. Las dos sumas de rangos son 3 y 42.

Paso 5: Sea T la *menor* de las dos sumas encontradas en el paso 4. Se puede usar cualquiera de las dos, pero para un procedimiento simplificado seleccionamos arbitrariamente la menor de las dos sumas.

EJEMPLO: Los datos de la tabla 13-4 dan como resultado las sumas de rangos de 3 y 42, por lo que la menor de esas dos sumas es 3.

Paso 6: Sea n la cantidad de pares de datos para los cuales la diferencia d no es 0.

EJEMPLO: Los datos en la tabla 13-4 tienen 9 diferencias que no son 0, entonces $n = 9$.

Paso 7: Determine el dato estadístico de prueba y los valores críticos según el tamaño de muestra, como se muestra en el recuadro de elementos clave anterior.

EJEMPLO: Para los datos de la tabla 13-4, el dato estadístico de prueba es $T = 3$. El tamaño de la muestra es $n = 9$, por lo que el valor crítico se encuentra en la tabla A-8. Usando un nivel de significancia de 0.05 con una prueba de dos colas, el valor crítico obtenido en la tabla A-8 es 6.

Paso 8: Como conclusión, rechace la hipótesis nula si los datos muestrales conducen a un dato estadístico de prueba que se encuentra en la región crítica; es decir, el dato estadístico de prueba es menor o igual que el valor o los valores críticos. De lo contrario, no se puede rechazar la hipótesis nula.

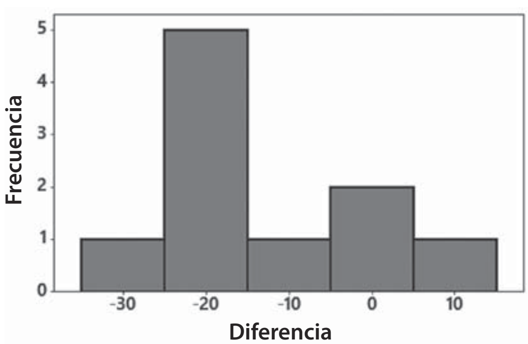
EJEMPLO: Si el dato estadístico de prueba es T (en lugar de z), rechace la hipótesis nula si T es menor o igual que el valor crítico. No puede rechazar la hipótesis nula si T es mayor que el valor crítico. Para la muestra de datos pareados en las dos primeras filas de la tabla 13-4, $T = 3$ y el valor crítico es 6, rechazamos la hipótesis nula de que los datos pareados son de una población de pares relacionados con diferencias que tienen una mediana igual a cero

EJEMPLO 1 Edades de las mejores actrices y los mejores actores

Las primeras dos filas de la tabla 13-4 incluyen las edades de las mejores actrices y de los mejores actores (del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B). Los datos se relacionan de acuerdo con el año en que se ganaron los premios. Use los datos muestrales en las primeras dos filas de la tabla 13-4 para probar la afirmación de que no hay diferencia entre las edades de las mejores actrices y los mejores actores. Utilice la prueba de rangos con signo de Wilcoxon con un nivel de significancia de 0.05.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los datos son una muestra aleatoria simple. (2) En la pantalla adjunta se muestra el histograma de las diferencias en la tercera fila de la tabla 13-4. El lado izquierdo de la gráfica debe ser aproximadamente una imagen especular del lado derecho, que no parece ser el caso. Pero con sólo 10 diferencias, la discrepancia entre los lados izquierdo y derecho no es demasiado extrema, por lo que consideraremos que este requisito se cumple.



La hipótesis nula es la afirmación de que no hay diferencia entre las edades de las mejores actrices y los mejores actores, y la hipótesis alternativa es la afirmación de que existe una diferencia.

H_0: No hay diferencia. (La mediana de las diferencias es igual a 0).

H_1: Hay una diferencia. (La mediana de las diferencias no es igual a 0).

Dato estadístico de prueba Debido a que estamos utilizando la prueba de rangos con signo de Wilcoxon, el dato estadístico de prueba se calcula utilizando el procedimiento de ocho pasos presentado anteriormente en esta sección. Tales pasos incluyen ejemplos que ilustran el cálculo del dato estadístico de prueba con los datos muestrales de la tabla 13-4, y el resultado es el dato estadístico de prueba de $T = 3$.

Valor crítico El tamaño de muestra es $n = 9$, por lo que el valor crítico se encuentra en la tabla A-8. Utilizando un nivel de significancia de 0.05 con una prueba de dos colas, se encuentra que el valor crítico en la tabla A-8 es 6.

continúa

Conclusión La tabla A-8 incluye una nota que establece que debemos rechazar la hipótesis nula si el dato estadístico de prueba T es menor o igual que el valor crítico. Como el dato estadístico de prueba de $T = 3$ es menor o igual que el valor crítico de 6, rechazamos la hipótesis nula.

INTERPRETACIÓN

Concluimos que parece haber una diferencia entre las edades de las mejores actrices y los mejores actores.

SU TURNO Resuelva el ejercicio 5 "Citas rápidas: atributos".

Afirmaciones sobre la mediana de una sola población

La prueba de rangos con signo de Wilcoxon también se puede usar para evaluar una afirmación de que una sola población tiene algún valor declarado de la mediana. Los procedimientos anteriores se pueden usar con un simple ajuste:

Al probar una afirmación sobre la mediana de una sola población, cree datos pareados relacionando cada valor muestral con el valor declarado de la mediana. Se puede utilizar el procedimiento descrito anteriormente.

EJEMPLO 2 Temperaturas corporales

El conjunto de datos 3 "Temperaturas corporales" en el apéndice B incluye temperaturas corporales medidas en adultos. Use las 106 temperaturas indicadas para las 12 AM del día 2 con la prueba de rangos con signo de Wilcoxon, para probar la afirmación de que la mediana es menor a 98.6 °F. Use un nivel de significancia de 0.05.

Statdisk

```
Num Unequal pairs: 91

Using Approximation
Test Statistic, T: 661.0000
Mean, μ: 2093
Standard Deviation: 252.6589
Test Statistic, z: -5.6677
Critical z: ±1.959962
```

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) El diseño del experimento que condujo a los datos en el conjunto de datos 3 justifica el tratamiento de la muestra como una muestra aleatoria simple. (2) Se cumple el requisito de una distribución de las diferencias aproximadamente simétrica, porque un histograma de esas diferencias es aproximadamente simétrico.

Al emparejar cada valor muestral individual con la mediana de 98.6 °F, estamos trabajando con datos pareados. En el margen se muestra la pantalla de Statdisk que muestra el dato estadístico de prueba $T = 661$, que se convierte en el dato estadístico de prueba $z = -5.67$. (La pantalla es de una prueba de dos colas; para esta prueba de cola izquierda, el valor crítico es -1.645). El dato estadístico de prueba de $z = -5.67$ arroja un valor P de 0.000, por lo que rechazamos la hipótesis nula de que la población de diferencias entre las temperaturas corporales y la mediana declarada de 98.6 °F es cero. Hay pruebas suficientes para respaldar la afirmación de que la temperatura corporal media es inferior a 98.6 °F. Esta es la misma conclusión obtenida a partir de la prueba del signo en el ejemplo 4 de la sección 13-2.

SU TURNO Resuelva el ejercicio 9 "Magnitudes de terremoto".

Justificación: En el ejemplo 1, los rangos sin signo de 1 a 9 tienen un total de 45, por lo que si no hay diferencias significativas, cada uno de los dos totales de los rangos con signo debe ser de alrededor de 45 ÷ 2 o 22.5. Es decir, los rangos negativos y los rangos positivos deberían dividirse como 22.5-22.5 o algo cercano, por ejemplo 24-21. La tabla A-8, de valores críticos, muestra que para el nivel de significancia de 0.05 con 9 pares de datos, el valor crítico es 6; por lo que una división de 6-39 representa una desviación significativa de la hipótesis nula, y cualquier división que esté más alejada también representará una desviación significativa de la hipótesis nula. Por el contrario, las divisiones como 7-38 no representan desviaciones significativas de una división 22.5-22.5, y no justificarían el rechazo de la hipótesis nula. La prueba de rangos con signo de Wilcoxon se basa en el total de menor rango, por lo que en lugar de analizar los dos números que constituyen la división, consideramos sólo el número más bajo.

La suma de todos los rangos $1 + 2 + 3 + \cdots + n$ es igual a $n(n + 1)/2$. Si esta suma de rangos debe dividirse equitativamente entre dos categorías (positiva y negativa), cada uno de los dos totales debe estar cerca de $n(n + 1)/4$, que es la mitad de $n(n + 1)/2$. El reconocimiento de este principio nos ayuda a comprender el dato estadístico de prueba utilizado cuando $n > 30$.

CENTRO DE TECNOLOGÍA

Prueba de rangos con signo de Wilcoxon

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Wilcoxon Tests** en el menú desplegable y **Wilcoxon (Matched Pairs)** en el submenú.
3. Ingrese un nivel de significancia y seleccione las dos columnas de datos a incluir.
4. Haga clic en **Evaluate**.

Minitab

1. Cree una columna que integre las diferencias entre los pares combinados. Para hacer esto, ingrese los datos pareados en las columnas *C1* y *C2*, seleccione **Edit-Command Line Editor** e ingrese el comando **Let C3 = C1 − C2.**
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Nonparametrics** en el menú desplegable y seleccione **1-Sample Wilcoxon** en el submenú.
4. En *Variables*, seleccione la columna que contiene las diferencias entre los pares relacionados (**C3**).
5. Seleccione **Test Median** e ingrese el valor de la mediana **0**.
6. Elija el formato de la hipótesis alternativa.
7. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Nonparametrics** en el menú desplegable y **Wilcoxon Signed Ranks** del submenú.
3. Elija **Paired** y seleccione las columnas que contienen los datos pareados que se analizarán.
4. Seleccione **Hypothesis test for median** y para H_0 ingrese **0** en el valor de la mediana declarada. Para H_A seleccione el formato deseado.
5. Haz clic en **Compute!**

Calculadora TI-83/84 Plus

Requiere los programas *SRTEST* y *ZZRANK* (disponibles en www.pearsonenespañol.com/triola)

1. Cree una lista de diferencias entre los valores de los pares relacionados. Para hacer esto, ingrese los datos pareados en las listas *L1* y *L2* y almacene las diferencias en *L3* ingresando *L1* − *L2* STO› *L3*.
2. Presione PRGM, seleccione **SRTEST** y presione ENTER dos veces.
3. Para **DATA** = ingrese el nombre de la lista que contiene las diferencias (*L3*) y presione ENTER.
4. Se mostrará el tamaño de muestra (*N*), la suma de los rangos positivos ($T+$) y la suma de los rangos negativos ($T-$). Presione ENTER para ver la media y la desviación estándar. Presione ENTER nuevamente para ver la puntuación *z*.
5. Si $n \le 30$, obtenga el valor *T* crítico de la tabla A-8. Si $n > 30$, obtenga los valores *z* críticos de la tabla A-2.

SUGERENCIA: El nombre de lista *L1* (y *L2* ... *L6*) se puede ingresar rápidamente presionando 2ND 1.

Excel

Complemento XLSTAT (requerido)

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Nonparametric tests**.
2. Seleccione **Comparison of two samples** del menú desplegable.
3. Ingrese el rango de datos para cada muestra en los recuadros de *las muestras 1 y 2*. Marque la casilla de **Column labels** si el rango de datos incluye etiquetas.
4. Seleccione **Paired samples** bajo el *formato de datos*.
5. Marque solamente la opción de **Wilcoxon signed-rank test**.
6. Haga clic en la pestaña **Options**.
7. En *Hipótesis alternativa*, seleccione el formato deseado. Confirme que la *diferencia hipotética* (*D*) es **0**.
8. Ingrese un nivel de significancia y marque la casilla de **Exact p-value**.
9. Haga clic en **OK**.

13-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Prueba de rangos con signo de Wilcoxon para temperaturas corporales La siguiente tabla muestra la temperatura corporal de siete sujetos a las 8 AM y a las 12 AM (del conjunto de datos 3 “Temperaturas corporales” en el apéndice B). Los datos son pareados porque cada par de temperaturas se mide de la misma persona. Use la prueba de rangos con signo de Wilcoxon para probar la afirmación de que no hay diferencia entre las temperaturas corporales de las 8 AM y las 12 AM.

a. ¿Qué requisitos deben cumplirse para esta prueba?

b. ¿Existe algún requisito de que las muestras deben ser de poblaciones que tienen una distribución normal o cualquier otra distribución específica?

c. ¿En qué sentido es esta prueba del signo una “prueba sin distribución”?

Temperatura (°F) a las 8 AM	98.0	97.6	97.2	98.0	97.0	98.0	98.2
Temperatura (°F) a las 12 AM	97.0	98.8	97.6	98.0	97.7	98.8	97.6

2. Temperaturas corporales Para los datos pareados que se listan en el ejercicio 1, identifique los siguientes componentes utilizados en la prueba de rangos con signo de Wilcoxon:

a. Diferencias d

b. Los rangos correspondientes a los valores distintos de cero de $|d|$

c. Los rangos con signos

d. La suma de los rangos positivos y la suma de los valores absolutos de los rangos negativos

e. El valor del dato estadístico de prueba T

f. El valor crítico de T (suponiendo un nivel de significancia de 0.05 en una prueba de ninguna diferencia entre las temperaturas corporales a las 8 AM y las 12 AM)

3. Prueba del signo contra Prueba de los rangos con signo de Wilcoxon Con base en los datos del ejercicio 1, podemos realizar pruebas para determinar la diferencia entre las temperaturas corporales a las 8 AM y a las 12 AM mediante la prueba del signo o la prueba de los rangos con signo de Wilcoxon. ¿En qué sentido la prueba de rangos con signo de Wilcoxon incorpora y usa más información que la prueba del signo?

4. Eficiencia de la prueba de rangos con signo de Wilcoxon Consulte la tabla 13-2 en la página 600 e identifique la eficiencia de la prueba de rangos con signo de Wilcoxon. ¿Qué nos dice ese valor sobre la prueba?

Uso de la prueba de rangos con signo de Wilcoxon. ***En los ejercicios 5 a 8, consulte los datos muestrales para los ejercicios dados en la sección 13-2 de la página 611. Use la prueba de rangos con signo de Wilcoxon para probar la afirmación de que los pares relacionados tienen diferencias que provienen de una población con una mediana igual a cero. Use un nivel de significancia de 0.05.***

5. Ejercicio 5 “Citas rápidas: atributos”

6. Ejercicio 6 “Citas rápidas: atractivo”

7. Ejercicio 7 “Citas rápidas: atributos”

8. Ejercicio 8 “Citas rápidas: atractivo”

En los ejercicios 9 a 12, consulte los datos muestrales de los ejercicios dados en la sección 13-2 de la página 611. Use la prueba de rangos con signo de Wilcoxon para probar la afirmación sobre la mediana de una población.

9. Ejercicio 13 “Magnitudes de terremotos”

10. Ejercicio 14 “Profundidades de terremotos”

11. Ejercicio 15 “Prueba del peso de monedas de ¢25”

12. Ejercicio 16 “Old Faithful”

13-3 Más allá de lo básico

13. Suma de rangos El ejercicio 12 usa el conjunto de datos 23 “Old Faithful” del apéndice B y el tamaño de muestra es 250.

a. Si tenemos muestras de datos pareados con 250 diferencias distintas de cero, ¿cuáles son los valores más pequeños y más grandes posibles de T?

b. Si tenemos muestras de datos pareados con 250 diferencias distintas de cero, ¿cuál es el valor esperado de T si la población consiste en datos pareados con diferencias que tienen una mediana de 0?

c. Si tenemos datos muestrales pareados con 250 diferencias distintas de cero y la suma de los rangos positivos es 1234, encuentre el valor absoluto de la suma de los rangos negativos.

d. Si tenemos datos muestrales pareados con n diferencias distintas de cero y una de las dos sumas de rangos es k, busque una expresión para la otra suma de rangos.

13-4 Prueba de la suma de rangos de Wilcoxon para dos muestras independientes

Concepto clave En esta sección se describe la prueba de la *suma de rangos de Wilcoxon*, que usa rangos de valores de dos muestras *independientes* para probar la hipótesis nula de que las muestras provienen de poblaciones que tienen medianas iguales. La prueba de suma de rangos de Wilcoxon es equivalente a la **prueba *U* de Mann-Whitney** (vea el ejercicio 13), que se incluye en algunos libros de texto y tecnologías (como Minitab, StatCrunch y XLSTAT). Esta es la idea básica que subyace a la prueba de la suma de rangos de Wilcoxon: si se extraen dos muestras de poblaciones idénticas y los valores individuales se *clasifican* como una sola colección combinada de valores, los rangos alto y bajo deberían coincidir 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, tenemos una indicación de que las dos poblaciones tienen medianas diferentes.

A diferencia de las pruebas t paramétricas para dos muestras independientes de la sección 9-2, la prueba de la suma de rangos de Wilcoxon no requiere poblaciones distribuidas normalmente y puede usarse con datos en el nivel ordinal de medición, como datos que constan de rangos. En la tabla 13-2 observamos que la prueba de la suma de rangos de Wilcoxon tiene una calificación de eficiencia de 0.95 cuando se compara con la prueba paramétrica. Como esta prueba tiene una calificación de eficiencia tan alta e implica cálculos más fáciles, a menudo se prefiere a la prueba t paramétrica, incluso cuando se cumple el requisito de normalidad.

PRECAUCIÓN No confunda la prueba de la suma de rangos de Wilcoxon para dos muestras *independientes* con la prueba de rangos con signo de Wilcoxon para datos pareados. Use el mnemónico “Impuesto Sobre la Renta” para que ISR le recuerde “**I**ndependiente: **S**uma de **R**angos”.

DEFINICIÓN

La prueba de la **suma de rangos de Wilcoxon** es una prueba no paramétrica que utiliza rangos de los datos muestrales de dos poblaciones independientes para probar esta hipótesis nula:

H_0: Dos muestras independientes provienen de poblaciones con medianas iguales.

(La hipótesis alternativa H_1 puede ser cualquiera de las siguientes tres posibilidades: las dos poblaciones tienen *diferentes* medianas, o la primera población tiene una mediana *mayor que* la mediana de la segunda población, o la primera población tiene una mediana *menor que* la mediana de la segunda población).

ELEMENTOS CLAVE

Prueba de la suma de rangos de Wilcoxon

Objetivo

Usar la prueba de la suma de rangos de Wilcoxon con muestras de dos poblaciones independientes para las siguientes hipótesis nula y alternativa:

H_0: Las dos muestras provienen de poblaciones con medianas iguales.

H_1: La mediana de la primera población es diferente (o mayor que, o menor que) la mediana de la segunda población.

Notación

n_1 = tamaño de la muestra 1
n_2 = tamaño de la muestra 2
R_1 = suma de los rangos para la muestra 1
R_2 = suma de los rangos para la muestra 2
R = igual que R_1 (suma de los rangos para la muestra 1)

μ_R = media de los valores R muestrales que se espera cuando las dos poblaciones tienen medianas iguales
σ_R = desviación estándar de los valores R muestrales que se espera cuando dos poblaciones tienen medianas iguales

Requisitos

1. Hay dos muestras aleatorias simples independientes.
2. Cada una de las dos muestras tiene más de 10 valores. (Para muestras con 10 o menos valores, existen tablas especiales disponibles en libros de referencia, como *CRC Standard Probability and Statistics Tables and Formulae*, publicado por CRC Press).

Nota: No existe un requisito de que las dos poblaciones tengan una distribución normal o cualquier otra distribución en particular.

Dato estadístico de prueba

$$z = \frac{R - \mu_R}{\sigma_R}$$

donde $\mu_R = \dfrac{n_1(n_1 + n_2 + 1)}{2}$ y $\sigma_R = \sqrt{\dfrac{n_1 n_2(n_1 + n_2 + 1)}{12}}$

n_1 = tamaño de la muestra a partir de la cual se encuentra la suma de rangos R
n_2 = tamaño de la otra muestra
R = suma de los rangos de la muestra con tamaño n_1

Valores *P*

Los valores P se pueden encontrar utilizando la tecnología o mediante el dato estadístico de prueba z y la tabla A-2.

Valores críticos

Los valores críticos se pueden encontrar en la tabla A-2 (porque el dato estadístico de prueba se basa en la distribución normal).

Procedimiento para encontrar el valor del dato estadístico de prueba

Para ver cómo se aplican los siguientes pasos, consulte los datos muestrales listados en la tabla 13-5. Los datos provienen del conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B.

Paso 1: Combine temporalmente las dos muestras en una muestra grande, luego reemplace cada valor muestral con su rango. (El valor más bajo obtiene un rango de 1, el siguiente valor más bajo obtiene un rango de 2, y así sucesivamente. Si los valores empatan, asígneles la media de los rangos implicados en el empate. Consulte la sección 13-1 para obtener una descripción de los rangos y el procedimiento para manejar los empates).

EJEMPLO: En la tabla 13-5, se muestran los rangos de las 27 evaluaciones de cursos por parte de los estudiantes entre paréntesis. El rango de 1 se asigna al valor muestral más bajo de 2.9, el rango de 2 se asigna al siguiente valor más bajo de 3.1, y el rango de 3 se asigna al siguiente valor más bajo de 3.3. El siguiente valor más bajo es 3.4, que ocurre cuatro veces, por lo que encontramos la media de los rangos de 4, 5, 6, 7, que es 5.5, y asignamos el rango de 5.5 a cada uno de esos cuatro valores empatados.

Paso 2: Encuentre la suma de los rangos para cualquiera de las dos muestras.

EJEMPLO: En la tabla 13-5, la suma de los rangos de la primera muestra es 159.5. (Es decir, $R_1 = 20.5 + 20.5 + 23.5 + ... + 3 = 159.5$).

Paso 3: Calcule el valor del dato estadístico de prueba z como se muestra en el recuadro de elementos clave anterior, donde cualquiera de las muestras se puede usar como "muestra 1". (Si ambos tamaños de muestra son mayores que 10, entonces la distribución muestral de R es aproximadamente normal con media μ_R y desviación estándar σ_R, y el dato estadístico de prueba es como se muestra en el recuadro de elementos clave anterior).

EJEMPLO: Los cálculos de μ_R, σ_R y z se muestran en el ejemplo 1 que sigue a continuación.

TABLA 13-5
Evaluaciones de cursos por los estudiantes

Profesora	Profesor
4.3 **(20.5)**	4.5 **(25.5)**
4.3 **(20.5)**	3.7 **(8)**
4.4 **(23.5)**	4.2 **(17.5)**
4.0 **(13.5)**	3.9 **(11)**
3.4 **(5.5)**	3.1 **(2)**
4.7 **(27)**	4.0 **(13.5)**
2.9 **(1)**	3.8 **(9.5)**
4.0 **(13.5)**	3.4 **(5.5)**
4.3 **(20.5)**	4.5 **(25.5)**
3.4 **(5.5)**	3.8 **(9.5)**
3.4 **(5.5)**	4.3 **(20.5)**
3.3 **(3)**	4.4 **(23.5)**
	4.1 **(16)**
	4.2 **(17.5)**
	4.0 **(13.5)**
$n_1 = 12$	$n_2 = 15$
$R_1 = 159.5$	$R_2 = 218.5$

EJEMPLO 1 Calificaciones de evaluación de cursos para profesores y profesoras

La tabla 13-5 lista las evaluaciones de los cursos impartidos por profesoras y profesores (del conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B). Use un nivel de significancia de 0.05 para evaluar la afirmación de que los profesores de ambos sexos tienen la misma calificación mediana en la evaluación de sus cursos.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los datos muestrales son dos muestras aleatorias simples e independientes. (2) Los tamaños de muestra son 12 y 15; por lo tanto, ambos tamaños de muestra son mayores que 10. Se cumplen los requisitos.

Las hipótesis nula y alternativa son las siguientes:

H_0: La mediana de las calificaciones en la evaluación de cursos impartidos por profesoras es igual a la mediana de las calificaciones en la evaluación de cursos impartidos por profesores.

H_1: La calificación mediana en la evaluación de cursos impartidos por profesoras es diferente de la calificación mediana en la evaluación de cursos impartidos por profesores.

Clasifique la lista combinada de las 27 clasificaciones de cursos, comenzando con un rango de 1 (asignado al valor más bajo de 2.9). Los rangos correspondientes a los valores muestrales individuales se muestran entre paréntesis en la tabla 13-5. R expresa la suma de los rangos para la muestra elegida como muestra 1. Si elegimos las evaluaciones de los cursos de profesoras como muestra 1, obtenemos

$$R = 20.5 + 20.5 + 23.5 + ... + 3 = 159.5$$

Debido a que hay calificaciones en la evaluación de cursos para 12 profesoras, tenemos $n_1 = 12$. Además, $n_2 = 15$ porque hay calificaciones en la evaluación de cursos para 15 profesores. Ahora es posible encontrar los valores de μ_R, σ_R y el dato estadístico de prueba z de la siguiente manera.

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{12(12 + 15 + 1)}{2} = 168$$

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}} = \sqrt{\frac{(12)(15)(12 + 15 + 1)}{12}} = 20.4939$$

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{159.5 - 168}{20.4939} = -0.41$$

La prueba tiene dos colas porque un gran valor positivo de z indicaría que se encuentran desproporcionadamente más rangos superiores en la Muestra 1, y un gran valor negativo de z indicaría que se encuentran desproporcionadamente más rangos inferiores en la muestra 1. En cualquier caso, tendríamos una fuerte evidencia en contra de la afirmación de que las dos muestras provienen de poblaciones con medianas iguales.

La importancia del dato estadístico de prueba z se puede tratar como en capítulos anteriores. Estamos probando (con $\alpha = 0.05$) la hipótesis de que las dos poblaciones tienen medianas iguales, por lo que tenemos una prueba de dos colas.

Valor *P* Si se usa la puntuación z no redondeada, el valor P es 0.678, por lo que no podemos rechazar la hipótesis nula de que las poblaciones de profesoras y profesoras tienen la misma calificación mediana en la evaluación de sus cursos.

Valores críticos Si usamos los valores críticos de $z = \pm 1.96$, vemos que el dato estadístico de prueba $z = -0.41$ *no* cae dentro de la región crítica, por lo que no podemos rechazar la hipótesis nula de que las poblaciones de profesoras y profesoras tienen la misma calificación mediana en la evaluación de sus cursos.

INTERPRETACIÓN

No hay evidencia suficiente para justificar el rechazo de la afirmación de que las profesoras y los profesores tienen la misma calificación mediana en la evaluación de sus cursos. De acuerdo con los datos muestrales disponibles, parece que las profesoras y los profesores imparten cursos que son calificados de manera similar.

SU TURNO Resuelva el ejercicio 5 "Profesores evaluados por sus alumnos".

En el ejemplo 1, si intercambiamos los dos conjuntos de valores muestrales y consideramos que la evaluación del curso de los profesores es la primera muestra, entonces $R = 218.5$, $\mu_R = 210$, $\sigma_R = 20.4939$ y $z = 0.41$, por lo que la conclusión es exactamente la misma.

Statdisk

Total Number of Values: 93
Rank Sum 1: 1619.5000
Rank Sum 2: 2751.5000

Mean, μ: 1880
Standard Deviation: 128.8669
Test Statistic, z: -2.0215
Critical z: ±1.959962

EJEMPLO 2 Calificaciones en la evaluación de cursos para profesores y profesoras

En el ejemplo 1 se usan 27 evaluaciones de cursos por parte de estudiantes, provenientes del conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B. Repetir el ejemplo 1 usando las calificaciones en la evaluación de cursos para los 93 profesores del conjunto de datos 17 no sería muy divertido. Los tamaños de muestra más grandes fomentan el uso de la tecnología. Si utilizamos Statdisk para repetir el ejemplo 1 usando las calificaciones en las evaluaciones de cursos para los 93 profesores del conjunto de datos 17, obtenemos la pantalla adjunta. Podemos ver que el dato estadístico de prueba es $z = -2.02$ (redondeado). El dato estadístico de prueba no redondeado de $z = -2.0215$ puede usarse para encontrar que el valor P en esta prueba de dos colas es 0.043. Además, el dato estadístico de prueba cae en la región crítica limitada por los valores críticos de -1.96 y 1.96. Rechazamos la hipótesis nula de las medianas iguales. Con base en la muestra más amplia de 93 docentes, parece que las profesoras y los profesores tienen calificaciones medianas diferentes en la evaluación de sus cursos.

SU TURNO Resuelva el ejercicio 9 "Profesores evaluados por sus alumnos".

CENTRO DE TECNOLOGÍA

Prueba de la suma de rangos de Wilcoxon

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior
2. Seleccione de **Wilcoxon Tests** en el menú desplegable y **Wilcoxon (Independent Samples)** en el submenú.
3. Ingrese un nivel de significancia y seleccione las dos columnas de datos a incluir.
4. Haga clic en **Evaluate**.

Minitab

1. Ingrese los dos conjuntos de datos muestrales en las columnas *C1* y *C2*.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Nonparametrics** en el menú desplegable y elija **Mann-Whitney** en el submenú.
4. Para la *primera muestra*, seleccione **C1** y para la *segunda muestra*, seleccione **C2**.
5. Ingrese el nivel de confianza (95.0 corresponde a un nivel de significancia de $\alpha = 0.05$).
6. En la *alternativa*, elija el formato de la hipótesis alternativa (*no igual* corresponde a una prueba de hipótesis de dos colas).
7. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Nonparametrics** en el menú desplegable y **Mann-Whitney** en el submenú.
3. Seleccione las columnas que se utilizarán para las dos muestras.
4. Seleccione **Hypothesis test for m1-m2** y para H_0 ingrese el valor de la diferencia declarada. Para H_A seleccione el formato deseado.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

Requiere los programas *RSTEST* y *ZZRANK* (disponibles en www.pearsonenespañol.com/triola)

1. Ingrese los dos conjuntos de datos muestrales en la lista *L1* y *L2*.
2. Presione PRGM, seleccione **RSTEST** y presione ENTER dos veces.
3. Para **GROUP A** = ingrese **L1** y presione ENTER. Para **GROUP B** = ingrese **L2** y presione ENTER.
4. La suma de rangos *R*, la media, la desviación estándar y el dato estadístico de prueba *z* se calcularán en función de la muestra con el menor número de valores. Presione ENTER nuevamente para obtener el dato estadístico de prueba *z*. Encuentre el valor crítico consultando la tabla A-2 o usando la función *normalcdf* como se describe en el Centro de tecnología de la sección 6-1.

SUGERENCIA: El nombre de lista *L1* (y *L2* ... *L6*) se puede ingresar rápidamente presionando 2ND 1

Excel

Complemento XLSTAT (requerido)

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Nonparametric tests**.
2. Seleccione **Comparison of two samples** del menú desplegable.
3. Ingrese el rango de datos para cada muestra en los cuadros de *Muestra 1 y 2*. Marque la casilla de **Column labels** si el rango de datos incluye etiquetas.
4. Seleccione **One column per sample** en el *formato de datos*.
5. Marque solamente la opción de **Mann-Whitney test**.
6. Haga clic en la pestaña **Options**.
7. En la hipótesis alternativa seleccione **Sample 1 − Sample 2 ≠ D**. Confirme que la *diferencia hipotética* (*D*) es **0**.
8. Ingrese un nivel de significancia y marque la casilla del **Exact p-value**.
9. Haga clic en **OK**.

SUGERENCIA: Dado que XLSTAT utiliza un procedimiento diferente al descrito en esta sección, los resultados pueden ser algo diferentes, especialmente para muestras pequeñas.

13-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Profesores evaluados por sus alumnos El ejemplo 1 en esta sección utilizó ejemplos de evaluaciones de cursos, y la siguiente tabla lista evaluaciones de estudiantes a sus profesoras y profesores (del conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B). ¿Se cumplen los requisitos para usar la prueba de la suma de rangos de Wilcoxon? ¿Por qué sí o por qué no?

Profesora	3.9	3.4	3.7	4.1	3.7	3.5	4.4	3.4	4.8	4.1	2.3	4.2	3.6	4.4
Profesor	3.8	3.4	4.9	4.1	3.2	4.2	3.9	4.9	4.7	4.4	4.3	4.1		

2. Suma de rangos Después de clasificar la lista combinada de evaluaciones a profesores que se dieron en el ejercicio 1, encuentre la suma de los rangos para las profesoras.

3. ¿Qué estamos probando? Consulte los datos muestrales del ejercicio 1. Suponiendo que usamos la prueba de la suma de rangos de Wilcoxon con esos datos, identifique la hipótesis nula y todas las posibles hipótesis alternativas.

4. Eficiencia Consulte la tabla 13-2 en la página 600 e identifique la eficiencia de la prueba de la suma de rangos de Wilcoxon. ¿Qué nos dice ese valor sobre la prueba?

Prueba de la suma de rangos de Wilcoxon. ***En los ejercicios 5 a 8, use la prueba de la suma de rangos de Wilcoxon.***

5. Profesores evaluados por sus alumnos Use los datos muestrales dados en el ejercicio 1 y pruebe la afirmación de que las calificaciones en la evaluación de las profesoras tienen la misma mediana que las calificaciones en la evaluación de los profesores. Use un nivel de significancia de 0.05.

6. Radiación en dientes de leche A continuación se listan las cantidades de estroncio-90 (en milibequelios, o mBq, por gramo de calcio) en una muestra aleatoria simple de dientes de leche obtenida de residentes de Pensilvania y residentes de Nueva York nacidos después de 1979 (según datos de "An Unexpected Rise in Strontium-90 in U.S. Deciduous Teeth in the 1990s", de Mangano *et al*., *Science of the Total Environment*). Use un nivel de significancia de 0.05 para evaluar la afirmación de que la cantidad mediana de estroncio 90 en los residentes de Pensilvania es la misma que la mediana de los residentes de Nueva York.

Pensilvania	155	142	149	130	151	163	151	142	156	133	138	161
Nueva York	133	140	142	131	134	129	128	140	140	140	137	143

7. Ensayos clínicos con Lipitor Los siguientes datos muestrales son cambios en los niveles de colesterol LDL en ensayos clínicos con Lipitor (atorvastatina). Se afirmó que Lipitor tenía un efecto sobre el colesterol LDL. (Los datos se basan en los resultados proporcionados en un memorando de Parke-Davis por David G. Orloff, MD, el líder del equipo médico que realizó los ensayos clínicos con Lipitor. Pfizer se negó a proporcionar los valores de los datos originales al autor). Los valores negativos representan disminuciones en colesterol LDL. Use un nivel de significancia de 0.05 para probar la afirmación de que, para aquellas personas tratadas con 20 mg de Lipitor y las tratadas con 80 mg de Lipitor, los cambios en el colesterol LDL tienen la misma mediana. ¿Qué sugieren los resultados?

Grupo tratado con 20 mg de Lipitor:													
−28	−32	−29	−39	−31	−35	−25	−36	−35	−26	−29	−34	−30	
Grupo tratado con 80 mg de Lipitor:													
−42	−41	−38	−42	−41	−41	−40	−44	−32	−37	−41	−37	−34	−31

8. Bloqueo en exámenes En un estudio realizado con estudiantes que se bloquearon durante exámenes, se estudió la disposición de las preguntas del examen por su efecto sobre la ansiedad. Los siguientes puntajes son medidas de una "prueba de ansiedad debilitante" (según datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4). ¿Hay suficiente evidencia para respaldar la afirmación de que las dos muestras provienen de poblaciones con diferentes medianas? ¿Hay suficiente evidencia para respaldar la afirmación de que la disposición de las preguntas de examen tiene un efecto en la puntuación? Use un nivel de significancia de 0.01.

Preguntas ordenadas de 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

Preguntas ordenadas de 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

Conjuntos de datos del apéndice B. ***En los ejercicios 9 a 12, consulte el conjunto de datos indicado en el apéndice B y use la prueba de la suma de rangos de Wilcoxon.***

9. Profesores evaluados por sus alumnos Repita el ejercicio 5 "Profesores evaluados por sus alumnos" utilizando todas las evaluaciones de los alumnos a sus profesores listadas en el conjunto de datos 17 "Evaluaciones de cursos" en el apéndice B.

10. ¿Los hombres hablan tanto como las mujeres? Consulte el conjunto de datos 24 "Las palabras cuentan" en el apéndice B y use los conteos de palabras medidos en hombres de la tercera columna y los conteos de palabras medidas en mujeres de la cuarta columna. Use un nivel de significancia de 0.01 para probar la afirmación de que, contrario a un mito popular, la mediana de las cantidades de palabras que pronuncian los hombres en un día es la misma que la mediana de las cantidades de palabras que pronuncian las mujeres diariamente.

11. IQ y exposición al plomo El conjunto de datos 7 "IQ y plomo" en el apéndice B lista las puntuaciones de IQ *completas* para una muestra aleatoria de sujetos con niveles de plomo "medio" en su sangre y otra muestra aleatoria de sujetos con niveles de plomo "altos" en su sangre. Use un nivel de significancia de 0.05 para probar la afirmación de que los sujetos con niveles de plomo medios tienen una mediana más alta de sus puntuaciones de IQ completas que los sujetos con niveles altos de plomo. ¿Parece que el nivel de plomo afecta las puntuaciones de IQ completas?

12. IQ y exposición al plomo El conjunto de datos 7 "IQ y plomo" en el apéndice B lista las *puntuaciones* de IQ del desempeño para una muestra aleatoria de sujetos con bajos niveles de plomo en la sangre y otra muestra aleatoria de sujetos con altos niveles de plomo en la sangre. Use un nivel de significancia de 0.05 para probar la afirmación de que los sujetos con bajos niveles de plomo tienen una mediana más alta de sus puntuaciones de IQ del desempeño que aquellos con niveles altos de plomo. ¿La exposición al plomo parece tener un efecto adverso?

13-4 Más allá de lo básico

13. Uso de la prueba *U* de Mann-Whitney La prueba U de Mann-Whitney es equivalente a la prueba de la suma de rangos de Wilcoxon para muestras independientes, en el sentido de que ambas se aplican a las mismas situaciones y siempre conducen 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

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R$$

y R es la suma de los rangos para la muestra 1. Use las calificaciones en la evaluación de cursos realizada por estudiantes de la tabla 13-5, página 621, para encontrar el dato estadístico de prueba para la prueba U de Mann-Whitney. Compare este valor con el dato estadístico de prueba encontrado mediante la prueba de la suma de rangos de Wilcoxon.

14. Determinación de valores críticos Supongamos que tenemos dos tratamientos (A y B) que producen resultados cuantitativos, y sólo tenemos dos observaciones para el tratamiento A y dos observaciones para el tratamiento B. No podemos usar la prueba de los rangos con signo de Wilcoxon expuesta en la presente sección, porque ninguno de los tamaños de muestra es mayor que 10.

Rango				Suma de rangos para el tratamiento A
1	2	3	4	
A	A	B	B	3

a. Complete la tabla adjunta listando las cinco filas correspondientes a los otros cinco resultados posibles e ingrese las sumas de rangos correspondientes para el tratamiento A.

b. Liste los valores posibles de R y sus probabilidades correspondientes. (Suponga que las filas 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.

13-5 Prueba de Kruskal-Wallis para tres o más muestras

Concepto clave En esta sección se describe la *prueba de Kruskal-Wallis*, que usa *rangos* de datos de tres o más muestras aleatorias simples e independientes para probar la hipótesis nula de que las muestras provienen de poblaciones con la misma mediana.

En la sección 12-1 se describió el análisis de varianza (ANOVA) de un factor como un método para probar la hipótesis nula de que tres o más poblaciones tienen la misma *media*, pero el ANOVA requiere que todas las poblaciones involucradas tengan distribuciones normales. La prueba de Kruskal-Wallis para *medianas* iguales no tiene tal requisito, por lo que es una prueba sin distribución o no paramétrica.

DEFINICIÓN

La **prueba de Kruskal-Wallis** (también llamada **prueba *H***) es una prueba no paramétrica que utiliza rangos de muestras aleatorias simples combinadas de tres o más poblaciones independientes para probar la hipótesis nula de que las poblaciones tienen la misma mediana. (La hipótesis alternativa es la afirmación de que las poblaciones tienen medianas que no son todas iguales).

Al aplicar la prueba de Kruskal-Wallis calculamos el dato estadístico de prueba H, el cual tiene una distribución que se puede aproximar por la distribución ji cuadrada siempre que cada muestra tenga al menos cinco observaciones. (Para realizar un repaso rápido de las características clave de la distribución ji cuadrada, consulte la sección 7-3).

El dato estadístico de prueba H mide la varianza de las sumas de rangos $R_1, R_2, \ldots, R_k$ de las diferentes muestras. Si los rangos se distribuyen de manera uniforme entre los grupos muestrales, entonces H debería ser un número relativamente pequeño. Si las muestras son muy diferentes, 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 conducen al rechazo de la hipótesis nula de que las muestras provienen de poblaciones idénticas. *Por lo tanto, la prueba de Kruskal-Wallis es una prueba de cola derecha.*

ELEMENTOS CLAVE

Prueba de Kruskal-Wallis

Objetivo

Utilizar la prueba de Kruskal-Wallis con muestras aleatorias simples de tres o más poblaciones independientes para las siguientes hipótesis nula y alternativa:

H_0: Las muestras provienen de poblaciones con la misma mediana.
H_1: Las muestras provienen de poblaciones con medianas que no son todas iguales.

Notación

N = número total de observaciones en todas las muestras combinadas
k = número de muestras diferentes
R_1 = suma de rangos para la muestra 1
n_1 = número de observaciones en la muestra 1

Para la muestra 2, la suma de rangos es R_2 y el número de observaciones es n_2, y se usa una notación similar para las muestras restantes.

Requisitos

1. Tenemos al menos tres muestras aleatorias simples e independientes.
2. Cada muestra tiene al menos cinco observaciones. (Si las muestras tienen menos de cinco observaciones, consulte las tablas especiales de valores críticos, como *CRC Standard Probability and Statistics Tables and Formulae* publicadas por CRC Press).

Nota: No existe un requisito de que las poblaciones tengan una distribución normal o cualquier otra distribución particular.

Dato 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 *P*

Con frecuencia, los valores P se obtienen mediante el uso de la tecnología. Al usar el dato estadístico de prueba H y el número de grados de libertad $(k - 1)$, se puede emplear la tabla A-4 con el fin de encontrar un rango de valores para el valor P.

Valores críticos

1. La prueba es *de cola derecha* y se pueden encontrar valores críticos a partir de la tecnología o de la distribución ji cuadrada en la tabla A-4.
2. $\text{gl} = k - 1$ (donde gl es el número de grados de libertad y k es la cantidad de muestras diferentes).

Procedimiento para encontrar el valor del dato estadístico de prueba *H* Para ver cómo se aplican los siguientes pasos, consulte los datos muestrales de la tabla 13-6 en la página siguiente. La tabla 13-6 incluye sólo algunos de los datos del Conjunto de datos 7 "IQ y plomo" en el apéndice B. Este conjunto de datos abreviado resulta más adecuado para ilustrar el método de la prueba de Kruskal-Wallis.

Paso 1: Combine temporalmente todas las muestras en una muestra grande y asigne un rango a cada valor muestral. (Clasifique los valores de menor a mayor, y en caso de empates, asigne a cada observación la media de los rangos involucrados).

EJEMPLO: en la tabla 13-6, los números entre paréntesis son los rangos del conjunto combinado de datos. El rango de 1 se asigna al valor más bajo de 64, el rango de 2 se asigna al siguiente valor más bajo de 78, y así sucesivamente. En caso de empate, a cada uno de los valores empatados se le asigna la media de los rangos involucrados en el empate.

Paso 2: Para cada muestra, encuentre la suma de los rangos y determine el tamaño de muestra.

EJEMPLO: En la tabla 13-6, la suma de los rangos para la primera muestra es 86, la suma de los rangos para la segunda muestra es 50.5, y la suma de los rangos para la tercera muestra es 53.5.

Paso 3: Calcule H usando los resultados del paso 2 y el dato estadístico de prueba y notación dados en el recuadro de elementos clave anterior.

EJEMPLO: El dato estadístico de prueba se calculó en el ejemplo 1.

TABLA 13-6 Puntuaciones de IQ del desempeño (rangos entre paréntesis)

Nivel bajo de plomo en la sangre	Nivel medio de plomo en la sangre	Nivel alto de plomo en la sangre
85 (**6.5**)	78 (**2**)	93 (**10**)
90 (**8.5**)	97 (**12.5**)	100 (**15.5**)
107 (**18.5**)	107 (**18.5**)	97 (**12.5**)
85 (**6.5**)	80 (**4**)	79 (**3**)
100 (**15.5**)	90 (**8.5**)	97 (**12.5**)
97 (**12.5**)	83 (**5**)	
101 (**17**)		
64 (**1**)		
$n_1 = 8$	$n_2 = 6$	$n_3 = 5$
$R_1 = 86$	$R_2 = 50.5$	$R_3 = 53.5$

EJEMPLO 1 Efecto del plomo en la puntuación del IQ

En la tabla 13-6 se listan las puntuaciones de IQ del desempeño (no verbales) de muestras de sujetos con niveles de plomo en la sangre bajo, medio y alto (del conjunto de datos 7 "IQ y plomo" del apéndice B). Use un nivel de significancia de 0.05 para probar la afirmación de que las tres muestras de puntuaciones de IQ del desempeño provienen de poblaciones con medianas que son todas iguales.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Cada una de las tres muestras es una muestra independiente aleatoria simple. (2) Cada tamaño de muestra es de al menos 5. Se cumplen los requisitos. ✓

Las hipótesis nula y alternativa son:

H_0: La población de sujetos con baja exposición al plomo, la población con exposición media al plomo y la población con alta exposición al plomo tienen puntuaciones de IQ del desempeño con la misma mediana.

H_1: Las tres poblaciones de puntuaciones de IQ del desempeño tienen tres medianas que no son todas iguales.

Dato estadístico de prueba Primero combine todos los datos muestrales y clasifíquelos, luego encuentre la suma de los rangos para cada categoría. En la tabla 13-6, los rangos se muestran entre paréntesis junto a los valores muestrales originales. Enseguida, encuentre el tamaño de muestra (n) y la suma de rangos (R) para cada muestra. Esos valores se presentan en la parte inferior de la tabla 13-6. Debido a que el número total de observaciones es 19, tenemos $n = 19$. Ahora podemos evaluar el dato estadístico de prueba de la siguiente manera:

$$H = \frac{12}{N(N+1)}\left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right) - 3(N+1)$$

$$= \frac{12}{19(19+1)}\left(\frac{86^2}{8} + \frac{50.5^2}{6} + \frac{53.5^2}{5}\right) - 3(19+1)$$

$$= 0.694$$

Debido a que cada muestra tiene al menos cinco observaciones, la distribución de H es aproximadamente una distribución ji cuadrada con $k - 1$ grados de libertad. El número de muestras es $k = 3$, entonces tenemos $3 - 1 = 2$ grados de libertad.

Valor *P* Con $H = 0.694$ y gl $= 2$, la tabla A-4 muestra que el valor P es mayor que 0.10. Mediante el uso de tecnología, obtenemos el valor $P = 0.707$. Debido a que el valor P es mayor que el nivel de significancia de 0.05, no podemos rechazar la hipótesis nula de las medianas poblacionales iguales.

Valor crítico Consulte la tabla A-4 para encontrar el valor crítico de 5.991, que corresponde a 2 grados de libertad y un nivel de significancia de 0.05 (con un área de 0.05 en la cola derecha). La figura 13-4 muestra que el dato estadístico de prueba $H = 0.694$ no cae dentro de la región crítica limitada por 5.991, por lo que no podemos rechazar la hipótesis nula de las medianas poblacionales iguales.

La figura 13-4 muestra el dato estadístico de prueba $H = 0.694$ y el valor crítico $\chi^2 = 5.991$. (La distribución ji cuadrada tiene la forma general que se muestra en la figura 13-4 cuando el número de grados de libertad es 1 o 2). El dato estadístico de prueba no se encuentra en la región crítica, por lo que no podemos rechazar la hipótesis nula de las medianas iguales

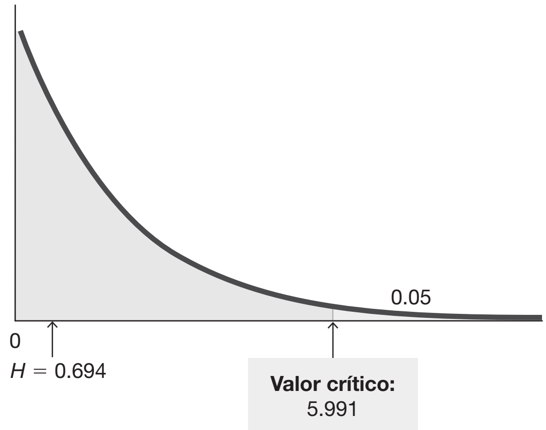


FIGURA 13-4 Distribución ji cuadrada para el ejemplo 1

INTERPRETACIÓN

No hay evidencia suficiente para rechazar la afirmación de que las tres muestras de puntuaciones de IQ del desempeño provienen de poblaciones con medianas que son todas iguales. Las medianas poblacionales no parecen ser significativamente diferentes.

SU TURNO Resuelva el ejercicio 5 "Citas rápidas".

Justificación: El dato estadístico de prueba H de Kruskal-Wallis es la versión del rango del dato estadístico de prueba F utilizado en el análisis de varianza discutido en el capítulo 12. Cuando tratamos con los rangos R en lugar de los valores originales x, muchos componentes están predeterminados. Por ejemplo, la suma de todos los rangos se puede expresar 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

$$\overline{R}_i = \frac{R_i}{n_i} \qquad \text{y} \qquad \overline{\overline{R}} = \frac{\Sigma R_i}{\Sigma n_i}$$

combina las varianzas ponderadas de los rangos para producir el dato estadístico de prueba H dado aquí, y esta expresión para H es algebraicamente equivalente a la expresión proporcionada anteriormente como el dato estadístico de prueba.

CENTRO DE TECNOLOGÍA

Prueba de Kruskal-Wallis

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Kruskal-Wallis Test** en el menú desplegable.
3. Ingrese un nivel de significancia y seleccione las columnas que se incluirán en el análisis.
4. Haga clic en **Evaluate**.

Minitab

1. Liste todos los datos muestrales en la columna *C1* e identifique la muestra (usando nombres o números) para el valor correspondiente en una segunda columna *C2*.
   - Para los datos de la tabla 13-6 de esta sección, ingrese los 19 valores muestrales en *C1* y en *C2* ingrese ocho *1s*, seguidos por seis *2*s y cinco *3*s.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Nonparametrics** en el menú desplegable y seleccione **Kruskal-Wallis** en el submenú.
4. Para la *respuesta*, seleccione la columna **C1** y para el *factor* seleccione la columna **C2**.
5. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Nonparametrics** en el menú desplegable y **Kruskal-Wallis** en el submenú.
3. Seleccione las columnas que se utilizarán en el análisis.
4. Marque **Adjust for ties** en la *Opciones*.
5. Haga clic en **Compute!**

SUGERENCIA: Si marca el ajuste para empates*, obtendrá resultados más precisos que pueden ser diferentes de los que se muestran en esta sección del libro.*

Calculadora TI-83/84 Plus

Requiere los programas *KWTEST* y *ZZRANK* (disponibles en www.pearsonenespañol.com/triola)

1. Los datos deben ingresarse como columnas en la *Matriz A*:

   Introducción manual de datos: presione 2ND luego x^{-1} para ir al menú *MATRIX*, seleccione **EDIT** en el menú superior, seleccione **[A]** y presione ENTER. Ingrese el número de filas y columnas necesarias, presione ENTER y proceda a ingresar los valores muestrales.

   Uso de listas existentes: Las listas se pueden combinar y almacenar en la *Matriz A*. Presione 2ND luego x^{-1} para ir al menú *MATRIX*, seleccione **MATH** en el menú superior y seleccione el elemento **List → matr**. Ingrese los nombres de la lista seguidos por el nombre de la matriz **[A]**, todos separados por comas. *Importante:* El nombre de la matriz debe ingresarse presionando 2ND luego x^{-1}, seleccionando **[A]**, y presionando ENTER. El siguiente es un resumen de los comandos utilizados para crear una matriz a partir de tres listas (*L1*, *L2*, *L3*): **List → matr(L1, L2, L3, [A])**.
2. Presione PRGM, seleccione **KWTEST** y presione ENTER dos veces. Se proporcionará el valor del dato estadístico de prueba *H* y la cantidad de grados de libertad. Consulte la tabla A-4 para encontrar el valor crítico.

SUGERENCIA: Si las muestras tienen diferentes tamaños, algunas de las entradas de la matriz serán ceros. Si alguno de los valores de datos originales es cero, agregue una constante conveniente a todos los valores muestrales para que no haya ceros entre los valores de datos originales.

Excel

Complemento XLSTAT (requerido)

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Nonparametric tests**.
2. Seleccione **Comparison of k samples** del menú desplegable.
3. En el recuadro de *muestras*, ingrese el rango de datos para los valores muestrales. Si el rango incluye etiquetas, marque la casilla de **Column labels**.
4. Seleccione **One column per sample** en el *formato de los datos*.
5. Marque solamente la opción de **Kruskal-Wallis**.
6. Haga clic en la pestaña **Options**.
7. Ingrese un nivel de significancia y marque la casilla de **Asymptotic p-value**.
8. Haga clic en **OK**.

13-5 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Citas rápidas En la parte superior de la página siguiente se muestran las calificaciones de los atributos de los hombres, otorgadas por las mujeres que participaron en eventos de citas rápidas (del conjunto de datos 18 “Citas rápidas” en el apéndice B). Al utilizar la prueba de Kruskal-Wallis, debemos clasificar todos los datos combinados, para después determinar la suma de los rangos para cada muestra. Encuentre la suma de los rangos para cada una de las tres muestras.

Edad 20-22	38	42	30.0	39	47	43	33	31	32	28
Edad 23-26	39	31	36.0	35	41	45	36	23	36	20
Edad 27-29	36	42	35.5	27	37	34	22	47	36	32

2. Requisitos Supongamos que queremos utilizar los datos del ejercicio 1 con la prueba de Kruskal-Wallis. ¿Se satisfacen los requisitos? Explique.

3. Notación Para los datos proporcionados en el ejercicio 1, identifique los valores de n_1, n_2, n_3 y N.

4. Eficiencia Consulte la tabla 13-2 en la página 600 e identifique la eficiencia de la prueba de Kruskal-Wallis. ¿Qué nos dice ese valor sobre la prueba?

Uso de la prueba de Kruskal-Wallis. ***En los ejercicios 5 a 8, use la prueba de Kruskal-Wallis.***

5. Citas rápidas Utilice los datos muestrales del ejercicio 1 para probar la afirmación de que las mujeres de los diferentes grupos de edad otorgan calificaciones de atributos con la misma mediana. Use un nivel de significancia de 0.05.

6. Arsénico en el arroz A continuación se listan las cantidades de arsénico en muestras de arroz integral de tres estados. Las cantidades se dan en microgramos de arsénico y todas las muestras tienen el mismo tamaño de porción. Los datos provienen de la Administración de Alimentos y Medicamentos de Estados Unidos. Use un nivel de significancia de 0.01 para probar la afirmación de que las tres muestras provienen de poblaciones con la misma mediana.

Arkansas	4.8	4.9	5.0	5.4	5.4	5.4	5.6	5.6	5.6	5.9	6.0	6.1
California	1.5	3.7	4.0	4.5	4.9	5.1	5.3	5.4	5.4	5.5	5.6	5.6
Texas	5.6	5.8	6.6	6.9	6.9	6.9	7.1	7.3	7.5	7.6	7.7	7.7

7. Mediciones de accidentes automovilísticos Utilice las siguientes medidas listadas de desaceleración del cofre (en *g*, donde *g* es la fuerza de la gravedad) a partir de muestras de automóviles pequeños, medianos y grandes. (Estos valores provienen del conjunto de datos 19 "Pruebas de colisión de automóviles" en el apéndice B). Use un nivel de significancia de 0.05 para probar la afirmación de que las diferentes categorías de tamaño tienen la misma desaceleración media del cofre en la prueba de colisión estándar. ¿Los datos sugieren que los autos más grandes son más seguros?

Pequeños	44	39	37	54	39	44	42
Medios	36	53	43	42	52	49	41
Grandes	32	45	41	38	37	38	33

8. Consumo de combustible en carretera A continuación se listan los consumos de combustible en carretera (mi/gal) para automóviles categorizados de acuerdo con su tamaño: pequeño, mediano y grande (del conjunto de datos 20 "Mediciones de automóviles" en el apéndice B). Utilizando un nivel de significancia de 0.05, pruebe la afirmación de que las tres categorías de tamaño tienen el mismo consumo mediano de combustible en carretera. ¿El tamaño de un automóvil parece afectar el consumo de combustible en la carretera?

Pequeños	28	26	23	24	26	24	25
Medios	28	31	26	30	28	29	31
Grandes	34	36	28	40	33	35	26

Conjuntos de datos del apéndice B. ***En los ejercicios 9 a 12, use la prueba de Kruskal-Wallis con los datos del apéndice B.***

9. Tiempos de servicio de comida rápida en la hora de la cena El conjunto de datos 25 "Comida rápida" en el apéndice B lista los tiempos de servicio (en segundos) para cenas en los establecimientos de comida rápida McDonald's, Burger King y Wendy's. Utilizando un nivel de significancia de 0.05, pruebe la afirmación de que las tres muestras provienen de poblaciones con la misma mediana.

10. Fumadores pasivos y activos El conjunto de datos 12 "Fumadores pasivos y activos" en el apéndice B lista los niveles de cotinina medidos en una muestra de sujetos que fuman, otra muestra de sujetos que no fuman pero están expuestos al humo de tabaco ambiental y una tercera muestra de sujetos que no fuman y no están expuestos al humo de tabaco ambiental. La cotinina se produce cuando

continúa

el cuerpo absorbe nicotina. Use un nivel de significancia de 0.01 para probar la afirmación de que las tres muestras provienen de poblaciones con la misma mediana. ¿Qué sugieren los resultados sobre un fumador que argumenta que absorbe tanta nicotina como las personas que no fuman?

11. Pesos al nacer El conjunto de datos 4 "Nacimientos" en el apéndice B lista el peso al nacer de bebés nacidos en el Albany Medical Center, el Bellevue Hospital en la ciudad de Nueva York, el Olean General Hospital y el Strong Memorial Hospital en Rochester, Nueva York. Use un nivel de significancia de 0.05 para probar la afirmación de que los cuatro diferentes hospitales tienen distintas medianas de los pesos al nacer.

12. Galletas con chispas de chocolate Consulte el Conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B y utilice los conteos de chispas de chocolate de los tres tipos diferentes de galletas Chips Ahoy. Use un nivel de significancia de 0.01 para probar la afirmación de que los tres tipos de galletas tienen la misma cantidad mediana de chispas de chocolate.

13-5 Más allá de lo básico

13. Corrección de empates para el dato estadístico de prueba *H* Al usar la prueba de Kruskal-Wallis, hay un factor de corrección que debe aplicarse siempre que haya muchos empates: Divida H por

$$1 - \frac{\Sigma T}{N^3 - N}$$

Primero combine todos los datos muestrales en una lista, y luego, en esa lista combinada, identifique los diferentes grupos de valores muestrales que están empatados. Para cada grupo individual de observaciones empatadas, identifique el número de valores muestrales que están empatados y designe ese número como t, luego calcule $T = t^3 - t$. Enseguida, sume los valores T para obtener ΣT. El valor de N es el número total de observaciones en todas las muestras combinadas. Use este procedimiento para encontrar el valor corregido de H en el ejemplo 1 de esta sección en la página 628. ¿El valor corregido de H difiere sustancialmente del valor encontrado en el ejemplo 1?

13-6 Correlación de rangos

Concepto clave Esta sección describe el método no paramétrico de la prueba de *correlación de rangos*, que usa *rangos* de datos pareados para probar una asociación entre dos variables. En la sección 10-1, se usaron datos de muestras pareadas para calcular valores del coeficiente de correlación lineal r, pero en esta sección usamos *rangos* como la base para calcular el coeficiente de correlación de rangos r_s. Como en el capítulo 10, debemos comenzar un análisis de datos pareados mediante la exploración con un diagrama de dispersión, a fin de identificar cualquier patrón en los datos, así como los valores atípicos.

DEFINICIÓN

La **prueba de correlación de rangos** (o **prueba de correlación de rangos de Spearman**) es una prueba no paramétrica que utiliza rangos de datos muestrales que constan de pares relacionados. Se usa para probar una asociación entre dos variables.

Usamos la notación r_s para el coeficiente de correlación de rangos a fin de no confundirlo con el coeficiente de correlación lineal r. El subíndice *s no* se refiere a una desviación estándar; se usa en honor a Charles Spearman (1863-1945), quien originó el método de la correlación de rangos. De hecho, r_s se denomina frecuentemente **coeficiente de correlación de rangos de Spearman**. Los componentes clave de la prueba de correlación de rangos se proporcionan en el siguiente recuadro de elementos clave, y el procedimiento se resume en la figura 13-5 de la página 634.

ELEMENTOS CLAVE

Correlación de rangos

Objetivo

Calcular el coeficiente de correlación de rangos r_s y usarlo para probar una asociación entre dos variables. Las hipótesis nula y alternativa son:

H_0: $\rho_s = 0$ (No existe correlación)
H_1: $\rho_s \neq 0$ (Existe correlación)

Notación

r_s = coeficiente de correlación de rangos para datos muestrales pareados (r_s es un estadístico muestral)

ρ_s = coeficiente de correlación de rangos para todos los datos de la población (ρ_s es un parámetro poblacional)

n = número de pares de datos muestrales

d = diferencia entre rangos para los dos valores dentro de un par individual

Requisitos

1. Los datos pareados son una muestra aleatoria simple.
2. Los datos son rangos o se pueden convertir en rangos.

Nota: A diferencia de los métodos paramétricos de la sección 10-1, *no* existe el requisito de que los pares de datos muestrales tengan una distribución normal bivariada (como se describió en la sección 10-1). *No* hay ningún requisito de una distribución normal para ninguna población.

Dato estadístico de prueba

Dentro de cada muestra, primero convierta los datos en *rangos*, luego encuentre el valor exacto del coeficiente de correlación de rangos r_s mediante el uso de la fórmula 10-1:

FÓRMULA 10-1

$$r_s = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

Dato estadístico de prueba más simple si no hay empates: Después de convertir los datos de cada muestra en rangos, si no hay empates entre los rangos para la primera variable y no hay empates entre los rangos para la segunda variable, el valor exacto del dato estadístico de prueba se puede calcular utilizando la fórmula 10-1 o la siguiente fórmula relativamente simple; pero probablemente sea más fácil emplear la fórmula 10-1 con tecnología:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

Valores *P*

En ocasiones, los valores *P* pueden obtenerse mediante el uso de la tecnología, pero utilícelos solamente si resultan de la correlación de rangos de Spearman. (No use valores *P* resultantes de las pruebas de correlación *lineal*; consulte la "precaución" en la parte superior de la página 635).

Valores críticos

1. Si $n \leq 30$, los valores críticos se encuentran en la tabla A-9.
2. Si $n > 30$, valores críticos de r_s se encuentran usando la fórmula 13-1.

FÓRMULA 13-1

$$r_s = \frac{\pm z}{\sqrt{n - 1}} \text{ (valores críticos para } n > 30)$$

donde el valor de z corresponde al nivel de significancia. (Por ejemplo, si $\alpha = 0.05$, $z = 1.96$).

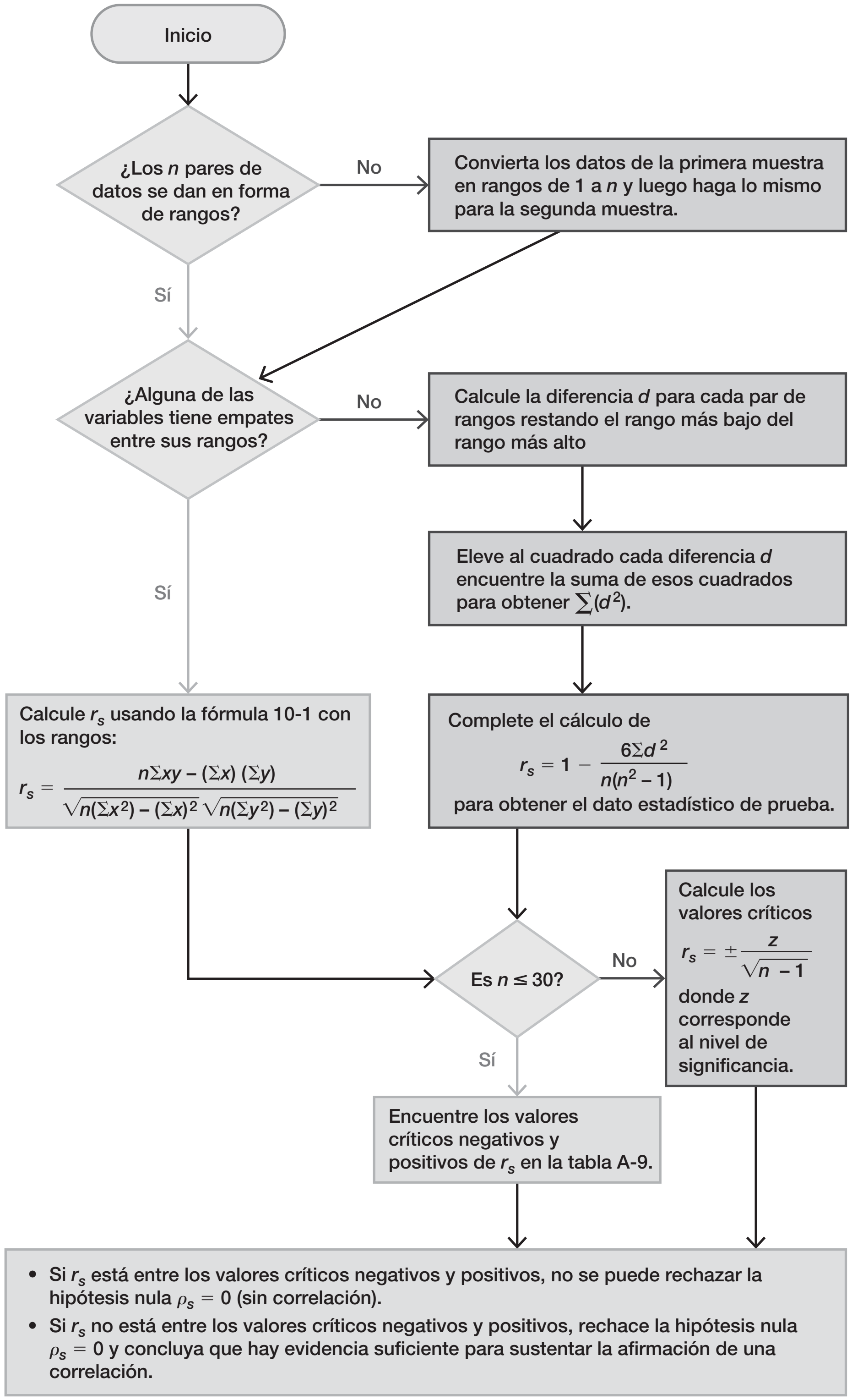


FIGURA 13-5 Procedimiento de correlación de rangos para la prueba H_0: $\rho_s = 0$

PRECAUCIÓN *No use valores P de la correlación lineal para los métodos de la correlación de rangos.* Al trabajar con datos que tienen empates entre rangos, el coeficiente de correlación de rangos r_s puede calcularse usando la fórmula 10-1. Es posible emplear la tecnología en lugar de los cálculos manuales, pero los valores *P* mostrados para la correlación *lineal* no se aplican a los métodos de la correlación de *rangos*.

Ventajas de la correlación de rangos: La correlación de rangos tiene las siguientes ventajas sobre los métodos paramétricos analizados en el capítulo 10:

1. La correlación de rangos se puede usar con datos pareados que son rangos o se pueden convertir en rangos. A diferencia de los métodos paramétricos del capítulo 10, el método de correlación de rangos no requiere una distribución normal para ninguna población.
2. La correlación de rangos se puede usar para detectar algunas (no todas) las relaciones que no son lineales.

Desventaja de la correlación de rangos: Eficiencia Una desventaja menor de la correlación de rangos es su índice de eficiencia de 0.91, como se describe en la sección 13-1. Esta calificación de eficiencia muestra que, con todas las demás circunstancias iguales, el método no paramétrico de la correlación de rangos requiere 100 pares de datos de muestra para lograr los mismos resultados que con 91 pares de observaciones muestrales analizadas mediante el método paramétrico, suponiendo que se cumplen los requisitos más estrictos del método paramétrico.

EJEMPLO 1 ¿Los mejores televisores cuestan más?

La tabla 13-1 del problema del capítulo lista los rangos y costos (en cientos de dólares) de televisores con pantallas LCD de al menos 60 pulgadas (según datos de *Consumer Reports*). Encuentre el valor del coeficiente de correlación de rangos y úselo para determinar si hay suficiente evidencia para respaldar la afirmación de una correlación entre calidad y precio. Use un nivel de significancia de 0.05. Según el resultado, ¿parece que puede obtenerse una mejor calidad si se gasta más?

TABLA 13-1 Rangos y costos de televisores LCD

Rango de calidad	1	2	3	4	5	6	7	8	9	10
Costo (cientos de dólares)	23	50	23	20	32	25	14	16	40	22

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Los datos muestrales son una muestra aleatoria simple de los televisores que se probaron. Los datos son rangos o se pueden convertir a rangos. ✓

Los rangos de calidad son enteros consecutivos y no provienen de una población que se distribuye normalmente, por lo que usamos el coeficiente de correlación de rangos para probar una relación entre calidad y precio. Las hipótesis nula y alternativa son:

H_0: $\rho_s = 0$ (*No* existe una correlación entre calidad y precio).

H_1: $\rho_s \neq 0$ (Existe una correlación entre calidad y precio).

Siguiendo el procedimiento de la figura 13-5, comenzamos convirtiendo los costos en la tabla 13-1 en sus rangos correspondientes que se muestran en la tabla 13-7 de la página siguiente. Al costo más bajo de $1400 en la tabla 13-1 se le asigna un rango de 1. Debido a que los costos quinto y sexto están empatados en $2300, asignamos el rango de 5.5 a cada uno de ellos (donde 5.5 es el promedio de los rangos 5 y 6). Los rangos correspondientes a los costos de la tabla 13-1 se muestran en la segunda fila de la tabla 13-7.

continúa

TABLA 13-7 Rangos de datos de la tabla 13-1.

Rango de calidad	1	2	3	4	5	6	7	8	9	10
Rango de costo	5.5	10	5.5	3	8	7	1	2	9	4

Debido a que hay empates entre los rangos, debemos usar la fórmula 10-1 para encontrar que el coeficiente de correlación de rangos r_s es igual a -0.274.

Fórmula 10-1

$$r_s = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

$$= \frac{10(280) - (55)(55)}{\sqrt{10(385) - (55)^2}\sqrt{10(384.5) - (55)^2}} = -0.274$$

Ahora consultamos la tabla A-9 para encontrar los valores críticos de $+0.648$ (con base en $\alpha = 0.05$ y $n = 10$). Debido a que el dato estadístico de prueba $r_s = -0.274$ está entre los valores críticos de -0.648 y 0.648, no podemos rechazar la hipótesis nula. No hay evidencia suficiente para respaldar la afirmación de una correlación entre calidad y costo. Según los datos muestrales proporcionados, parece que al pagar más no necesariamente se obtiene una mejor calidad.

SU TURNO Resuelva el ejercicio 7 “Chocolate y premios Nobel”.

EJEMPLO 2 Caso de muestra grande

Consulte las medidas de presión arterial sistólica y diastólica de 147 mujeres seleccionadas al azar en el conjunto de datos 1 “Datos corporales” del apéndice B y use un nivel de significancia de 0.05 para probar la afirmación de que en el caso de las mujeres existe una correlación entre las presiones arteriales sistólica y diastólica.

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS Los datos son una muestra aleatoria simple y se pueden convertir en rangos. ✓

Dato estadístico de prueba El valor del coeficiente de correlación de rangos es $r_s = 0.354$, que se puede encontrar utilizando la tecnología.

Valores críticos Debido a que hay 147 pares de datos, tenemos $n = 147$. Como n es mayor que 30, encontramos los valores críticos a partir de la fórmula 13-1 en lugar de la tabla A-9. Con $\alpha = 0.05$ en dos colas, consideramos que $z = 1.96$ para obtener los valores críticos de -0.162 y 0.162, como se muestra a continuación.

$$r_s = \frac{\pm z}{\sqrt{n-1}} = \frac{\pm 1.96}{\sqrt{147-1}} = \pm 0.162$$

El dato estadístico de prueba de $r_s = 0.354$ no está entre los valores críticos de -0.162 y 0.162, por lo que rechazamos la hipótesis nula de $r_s = 0$. Hay suficiente evidencia para respaldar la afirmación de que, en el caso de las mujeres, existe una correlación entre la presión arterial sistólica y la presión arterial diastólica.

SU TURNO Resuelva el ejercicio 13 “Chocolate y premios Nobel”.

Detección de patrones no lineales En ocasiones, los métodos de *correlación de rangos* nos permiten detectar relaciones que no podemos identificar con los métodos de correlación lineal del capítulo 10. Vea el siguiente diagrama de dispersión que muestra un patrón de puntos en forma de S que sugiere que hay una correlación entre x y y. Los métodos del capítulo 10 resultan en el coeficiente de correlación lineal de $r = 0.590$ y valores críticos de

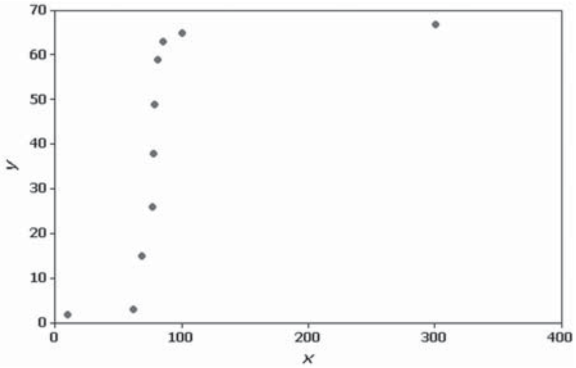
Vínculo directo entre el tabaquismo y el cáncer

Cuando encontramos una correlación estadística entre dos variables, debemos ser sumamente cuidadosos para evitar el error de concluir que existe un vínculo de causa y efecto. La industria tabacalera ha enfatizado una y otra vez que la correlación no implica causalidad. Sin embargo, el doctor David Sidransky, de la Universidad Johns Hopkins y otros investigadores encontraron un vínculo físico directo que implica mutaciones de un gen específico entre los fumadores. El análisis molecular de los cambios genéticos permite a los investigadores determinar si el consumo de cigarrillos es la causa de un cáncer. (Vea “Association Between Cigarette Smoking and Mutation of the p53 Gene in Squamous-Cell Carcinoma of the Head and Neck” de Brennan, Boyle, *et al*., *New England Journal of Medicine*, vol. 332, núm. 11). Aunque los métodos estadísticos no permiten determinar que fumar causa cáncer, es posible utilizar métodos estadísticos para identificar una asociación para que entonces los investigadores sean capaces de visualizar una demostración física.

±0.632, lo que sugiere que no hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre x y y. Pero si usamos los métodos de esta sección, obtendremos $r_s = 1$ y valores críticos de ±0.648, lo que sugiere que hay suficiente evidencia para respaldar la afirmación de una correlación entre x y y.

Con la correlación de rangos, a veces podemos detectar relaciones que no son lineales.

Patrón no lineal



CENTRO DE TECNOLOGÍA

Correlación de rangos

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Analysis** en el menú superior.
2. Seleccione **Rank Correlation** en el menú desplegable.
3. Ingrese un nivel de significancia y seleccione las dos columnas de datos que se incluirán.
4. Haga clic en **Evaluate**.

Minitab

1. Ingrese los datos pareados en las columnas *C1* y *C2*.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Basic Statistics** en el menú desplegable y elija **Correlation** en el submenú.
4. Seleccione las columnas que se incluirán en *Variables*.
5. Seleccione **Spearman rho** para el *método* y marque la casilla **Display p-values**.
6. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Nonparametrics** en el menú desplegable y **Spearman Correlation** en el submenú.
3. Seleccione las columnas que se utilizarán en el análisis.
4. En la *pantalla*, marque **Two-sided *P*-value**.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

La calculadora TI-83/84 Plus no está diseñada para calcular la correlación de rangos, pero podemos reemplazar cada valor con su rango correspondiente y calcular el valor del coeficiente de correlación lineal r.

1. Reemplace cada valor muestral con su rango correspondiente e ingrese los rangos pareados en las listas *L1* y *L2*.
2. Presione STAT, luego seleccione **TESTS** en el menú superior.
3. Seleccione **LinRegTTest** en el menú y presione ENTER.
4. Ingrese los nombres de lista para las variables x y y. Ingrese **1** para *Freq* y para β & ρ seleccione $\neq$ **0** para probar la hipótesis nula de no correlación.
5. Seleccione **Calculate** y presione ENTER. Como el cálculo de r se realiza utilizando rangos, el valor que se muestra como r es en realidad el coeficiente de correlación de rangos r_s. No tome en cuenta el valor P porque utiliza los métodos del capítulo 10, no los de esta sección.

CENTRO DE TECNOLOGÍA *continuación*

Correlación de rangos

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Excel

Complemento XLSTAT

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **Correlation/Association tests**.
2. Seleccione **Correlation tests** en el menú desplegable.
3. Para **Observations/variables table**, ingrese el rango de celdas para los valores de datos. Si el rango de datos incluye una etiqueta de datos, marque la casilla de **Variable labels**.
4. Para el *tipo de correlación*, seleccione **Spearman**.
5. Ingrese el nivel de significancia deseado.
6. Haga clic en **OK**. El coeficiente de correlación de rangos se muestra en la *matriz de correlación*. Si el valor mostrado está en negritas, podemos rechazar la afirmación de que no hay correlación.

Excel

Excel no tiene una función que calcule el coeficiente de correlación de rangos a partir de los valores muestrales originales, pero se puede usar el siguiente procedimiento.

1. Reemplace cada uno de los valores muestrales originales con su rango correspondiente.
2. Haga clic en **Insert function** f_x, seleccione la categoría **Statistical**, seleccione la función **CORREL** y haga clic en **OK**.
3. En *Array1* ingrese el rango de datos para la primera variable. En *Array2* ingrese el rango de datos para la segunda variable.
4. Haga clic en **OK** para obtener el coeficiente de correlación de rangos r_s.

13-6 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Regresión Si los métodos de esta sección se usan con datos muestrales pareados, y la conclusión es que hay suficiente evidencia para respaldar la afirmación de una correlación entre las dos variables, ¿podemos usar los métodos de la sección 10-2 para encontrar la ecuación de regresión que se puede utilizar en las predicciones? ¿Por qué sí o por qué no?

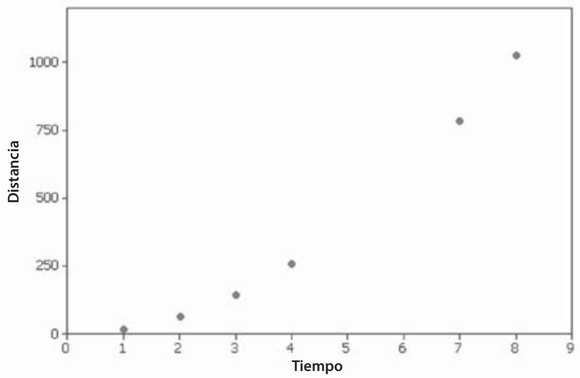
2. Nivel de medición ¿Cuál de los niveles de medición (nominal, ordinal, de intervalo, de razón) describe datos que no se pueden usar con los métodos de correlación de rangos? Explique.

3. Notación ¿Qué representan r, r_s, ρ y ρ_s? ¿Por qué se usa el subíndice s? ¿El subíndice s representa la misma desviación estándar introducida en la sección 3-2?

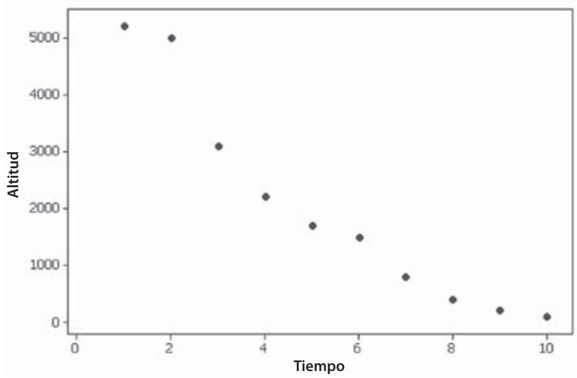
4. Eficiencia Consulte la tabla 13-2 en la página 600 e identifique la eficiencia de la prueba de correlación de rangos. ¿Qué nos dice ese valor sobre la prueba?

En los ejercicios 5 y 6, use el diagrama de dispersión para encontrar el valor del coeficiente de correlación de rangos r_s y los valores críticos correspondientes a un nivel de significancia de 0.05, utilizado para probar la hipótesis nula de $\rho_s = 0$. Determine si existe una correlación.

5. Datos de distancia/tiempo para un objeto que se deja caer



6. Datos de altitud/tiempo para una aeronave que desciende



Pruebas para el rango de correlación. ***En los ejercicios 7 a 12, use el coeficiente de correlación de rangos para probar una correlación entre las dos variables. Use un nivel de significancia de α = 0.05.***

7. Chocolate y premios Nobel La siguiente tabla lista el consumo de chocolate (kg *per cápita*) y el número de ganadores del Premio Nobel (por cada 10 millones de personas) para varios países (del conjunto de datos 16 en el apéndice B). ¿Existe una correlación entre el consumo de chocolate y la proporción de premios Nobel? ¿Cómo podría explicarse tal correlación?

Chocolate	11.6	2.5	8.8	3.7	1.8	4.5	9.4	3.6	2.0	3.6	6.4
Nobel	12.7	1.9	12.7	3.3	1.5	11.4	25.5	3.1	1.9	1.7	31.9

8. Edades de las mejores actrices y los mejores actores A continuación se listan edades de las mejores actrices y los mejores actores en el momento en que ganaron los premios Oscar (del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B). ¿Estos datos sugieren que existe una correlación entre las edades de las mejores actrices y los mejores actores?

Actriz	61	32	33	45	29	62	22	44	54
Actor	45	50	48	60	50	39	55	44	33

9. Pizza y el metro La "conexión de la pizza" es el principio de que el precio de una rebanada de pizza en la ciudad de Nueva York es siempre aproximadamente igual a la tarifa del metro. Use los datos que se listan a continuación para determinar si existe una correlación entre el costo de una rebanada de pizza y la tarifa del metro.

Año	1960	1973	1986	1995	2002	2003	2009	2013	2015
Costo de pizza	0.15	0.35	1.00	1.25	1.75	2.00	2.25	2.30	2.75
Tarifa del metro	0.15	0.35	1.00	1.35	1.50	2.00	2.25	2.50	2.75
IPC	30.2	48.3	112.3	162.2	191.9	197.8	214.5	233.0	237.2

10. IPC y el metro Utilice los datos de IPC/metro del ejercicio anterior para probar la correlación entre el IPC (índice de precios al consumidor) y la tarifa del metro.

11. Medición de focas a partir de fotografías A continuación se listan los anchos generales (en cm) de focas medidas a partir de fotografías y el peso de las focas (en kg). Los datos se basan en "Mass Estimation of Weddell Seals Using Techniques of Photogrammetry", de R. Garrott de la Universidad Estatal de Montana. El objetivo del estudio fue determinar si el peso de las focas podría determinarse a partir de fotografías aéreas. ¿Hay suficiente evidencia para concluir que existe una correlación entre los anchos generales y el peso de las focas?

Ancho general (cm)	7.2	7.4	9.8	9.4	8.8	8.4
Peso (kg)	116	154	245	202	200	191

12. Grillos y temperatura Se estudió la asociación entre la temperatura y el número de veces que un grillo chirría en 1 minuto. A continuación se listan los números de chirridos en 1 minuto y las temperaturas correspondientes en grados Fahrenheit (con base en datos de *The Songs of Insects* de George W. Pierce, Harvard University Press). ¿Existe evidencia suficiente para concluir que existe una relación entre el número de chirridos en 1 minuto y la temperatura?

Chirridos en 1 minuto	882	1188	1104	864	1200	1032	960	900
Temperatura (°F)	69.7	93.3	84.3	76.3	88.6	82.6	71.6	79.6

Conjuntos de datos del apéndice B. ***En los ejercicios 13 a 16, use los datos en el apéndice B para evaluar la correlación de rangos con un nivel de significancia de 0.05.***

13. Chocolate y premios Nobel Repita el ejercicio 7 usando todos los datos combinados de chocolate/Nobel en el conjunto de datos 16 "Premios Nobel y Chocolate" en el apéndice B.

14. Edades de las mejores actrices y los mejores actores Repita el ejercicio 8 usando todas las edades pareadas del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B.

15. Presión arterial Consulte las medidas de las presiones arteriales sistólica y diastólica de 153 hombres seleccionados al azar en el conjunto de datos 1 "Datos corporales" del apéndice B y use un nivel de significancia de 0.05 para probar la afirmación de que, en el caso de los hombres, existe una correlación entre la presión arterial sistólica y presión arterial diastólica.

16. IQ y volumen cerebral Consulte el conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B y pruebe la correlación entre el volumen cerebral (cm^3) y la puntuación del IQ.

13-6 Más allá de lo básico

17. Determinación de valores críticos Una alternativa al uso de la tabla A-9 para encontrar valores críticos de la correlación de rangos es calcularlos mediante la siguiente aproximación:

$$r_s = \pm\sqrt{\frac{t^2}{t^2 + n - 2}}$$

Aquí, t es el valor crítico t de la tabla A-3 correspondiente al nivel de significancia deseado y $n - 2$ grados de libertad. Use esta aproximación para encontrar los valores críticos r_s del ejercicio 15 "Presión sanguínea". ¿Cómo se comparan los valores críticos resultantes con los valores críticos que se encontrarían al usar la fórmula 13-1 de la página 633?

13-7 Prueba de rachas para aleatoriedad

Concepto clave En esta sección se describe la *prueba de rachas para aleatoriedad*, que se utiliza para determinar si una secuencia de datos muestrales tiene un orden aleatorio. Esta prueba requiere un criterio para categorizar cada valor de datos en una de dos categorías separadas, y analiza las *rachas* de esas dos categorías para determinar si parecen ser el resultado de un proceso aleatorio, o si las rachas sugieren que el orden de los datos no es aleatorio.

DEFINICIONES

Después de caracterizar cada valor de datos en una de dos categorías separadas, una **racha** es una secuencia de datos que tiene la misma característica; la secuencia está precedida y seguida de datos con una característica diferente o sin datos en absoluto.

La **prueba de rachas** utiliza el número de rachas en una secuencia de datos muestrales para probar la aleatoriedad en el orden de los datos.

ELEMENTOS CLAVE

Prueba de rachas para aleatoriedad

Objetivo

Aplicar la prueba de rachas para aleatoriedad a una secuencia de datos muestrales para probar la aleatoriedad en el *orden* de los datos. Use las siguientes hipótesis nula y alternativa:

H_0: Los datos están en un orden aleatorio.

H_1: Los datos están en un orden que no es aleatorio.

Notación

n_1 = número de elementos en la secuencia que tienen una característica particular. (La característica elegida para n_1 es arbitraria).
n_2 = número de elementos en la secuencia que tienen la otra característica
G = número de rachas

Requisitos

1. Los datos muestrales se organizan de acuerdo con algún esquema de ordenamiento, como el orden en que se obtuvieron los valores muestrales.
2. Cada valor de datos se puede categorizar en una de *dos* categorías (como masculino/femenino).

Dato estadístico de prueba y valores críticos

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 dato estadístico de prueba, los valores críticos y los criterios de decisión son los siguientes:

- **Dato estadístico de prueba:** número de rachas G
- **Valores críticos de G:** Use la tabla A-10.
- **Criterios de decisión:** Rechace la aleatoriedad si el número de rachas G es tal que
  - $G \leq$ menor valor crítico encontrado en la tabla A-10.
  - o $G \geq$ mayor valor crítico encontrado en la tabla A-10.

Para muestras grandes o $\alpha \neq 0.05$: Si $n_1 > 20$ o $n_2 > 20$ o $\alpha \neq 0.05$, el dato estadístico de prueba, los valores críticos y los criterios de decisión son los siguientes:

- **Dato estadístico de prueba:** $z = \dfrac{G - \mu_G}{\sigma_G}$

  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:** Use la tabla A-2.
- **Criterios de decisión:** Rechace la aleatoriedad si el dato estadístico de prueba z es tal que
  - $z \leq$ puntuación z crítica negativa (como -1.96)
  - o $z \geq$ puntuación z crítica positiva (como 1.96).

PRECAUCIÓN La prueba de rachas para aleatoriedad se basa en el *orden* en que ocurren los datos; *no* se basa en la *frecuencia* de los datos. Por ejemplo, una secuencia de 3 hombres y 20 mujeres podría parecer aleatoria, pero la cuestión de si 3 hombres y 20 mujeres constituyen una muestra *sesgada* (con un número desproporcionadamente mayor de mujeres) no se aborda mediante la prueba de rachas.

Rachas de suerte en los deportes

Existe la creencia de que los atletas suelen tener "rachas de suerte"—es decir, periodos breves de éxito extraordinario. El psicólogo Amos Tversky, de la Universidad de Stanford, y otros investigadores utilizaron la estadística para analizar los miles de tiros de los 76 de Filadelfia en una temporada completa y la mitad de otra. Encontraron que el número de "rachas de suerte" no difería de lo que se esperaría en pruebas aleatorias, donde el resultado de cada prueba es independiente de cualquier resultado previo. Es decir, la probabilidad de hacer una anotación no depende de las anotaciones o fallas anteriores.

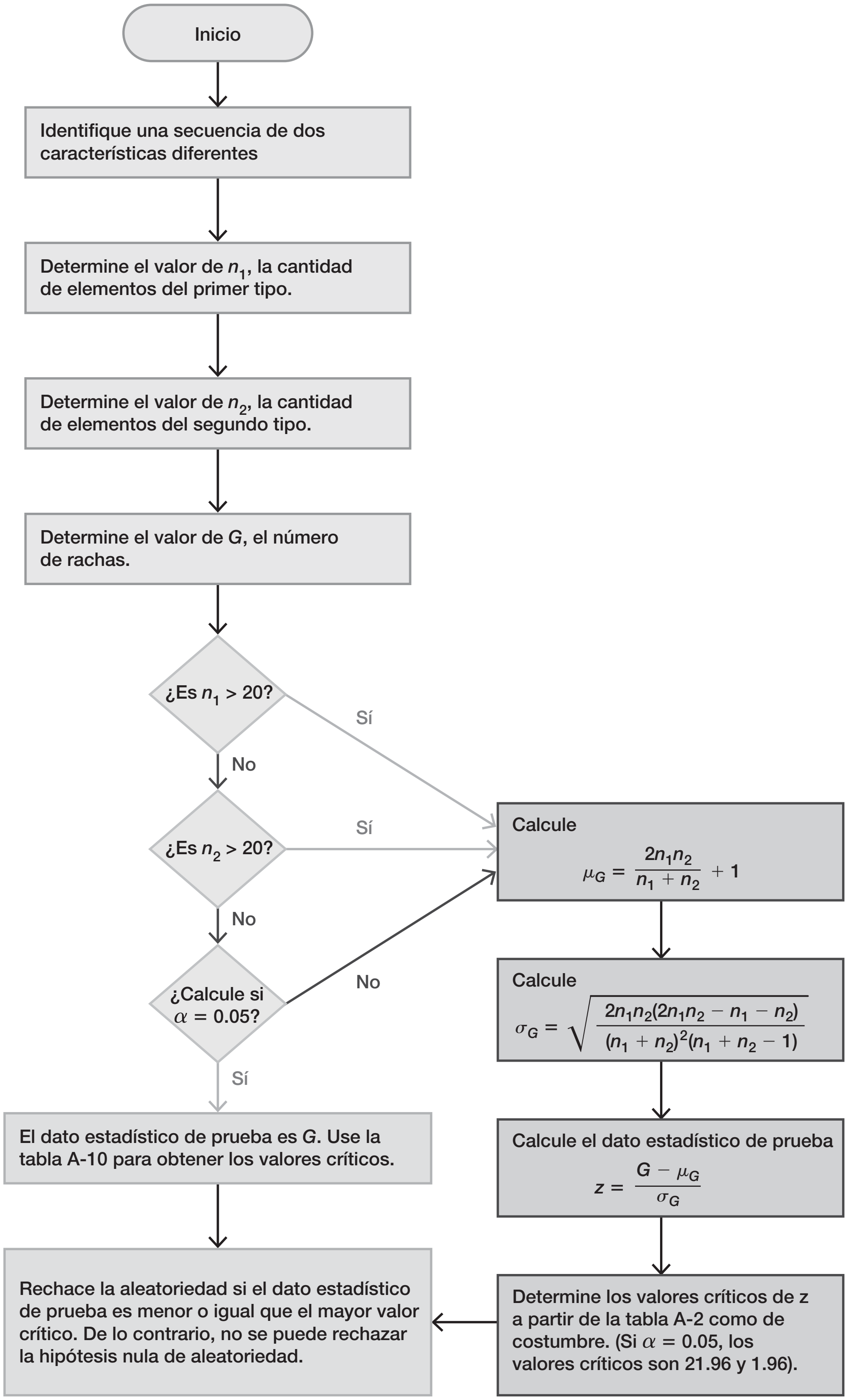


FIGURA 13-6 Procedimiento de la prueba de rachas para aleatoriedad

Principio fundamental de la prueba de rachas

La idea clave que subyace a la prueba de rachas es:

Rechace la aleatoriedad si el número de rachas es muy *bajo* o muy *alto.*

- Ejemplo: La secuencia de géneros FFFFFMMMMM no es aleatoria porque tiene sólo 2 rachas, por lo que el número de rachas es muy *bajo*.
- Ejemplo: La secuencia de géneros FMFMFMFMFM no es aleatoria porque tiene 10 rachas, que es una cantidad muy *alta.*

Los criterios exactos para determinar si un número de rachas es muy alto o muy bajo se encuentran en el recuadro de elementos clave. El procedimiento para la prueba de aleatoriedad se resume también en la figura 13-6.

EJEMPLO 1 Muestra pequeña: Partidos políticos de los presidentes

A continuación se listan los partidos políticos recientes de los presidentes de Estados Unidos. La letra R representa a un presidente republicano y la letra D representa a un presidente demócrata. Use un nivel de significancia de 0.05 para probar la aleatoriedad en la secuencia.

R R D D R D D R R D R R D R D

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los datos están ordenados. (2) Cada valor de datos se clasifica en una de dos categorías distintas (Republicano/Demócrata). Los requisitos se cumplen. ✓

Seguiremos el procedimiento resumido en la figura 13-6. Se ha identificado la secuencia de dos características (republicano/demócrata). Ahora debemos encontrar los valores de n_1, n_2 y el número de rachas G. La secuencia se muestra a continuación con el espaciado ajustado para identificar de mejor manera las diferentes rachas.

$$\underbrace{\text{RR}}_{\text{1a. racha}}\quad\underbrace{\text{DD}}_{\text{2a. racha}}\quad\underbrace{\text{R}}_{\text{3a. racha}}\quad\underbrace{\text{DD}}_{\text{4a. racha}}\quad\underbrace{\text{RR}}_{\text{5a. racha}}\quad\underbrace{\text{D}}_{\text{6a. racha}}\quad\underbrace{\text{RR}}_{\text{7a. racha}}\quad\underbrace{\text{D}}_{\text{8a. racha}}\quad\underbrace{\text{R}}_{\text{9a. racha}}\quad\underbrace{\text{D}}_{\text{10a. racha}}$$

La presentación anterior muestra que hay 8 presidentes republicanos y 7 presidentes demócratas, y el número de rachas es 10. Representamos esos resultados con la siguiente notación.

n_1 = número de presidentes republicanos = 8

n_2 = número de presidentes demócratas = 7

G = número de rachas = 10

Como $n_1 < 20$ y $n_2 < 20$ y el nivel de significancia es $\alpha = 0.05$, el dato estadístico de prueba es $G = 10$ (el número de rachas), y consultamos la tabla A-10 para encontrar los valores críticos de 4 y 13. Dado que $G = 10$ no es menor o igual que el valor crítico de 4, ni es mayor o igual que el valor crítico de 13, *no rechazamos la aleatoriedad*. No hay evidencia suficiente para rechazar la aleatoriedad en la secuencia de los partidos políticos de los presidentes más recientes.

¿La reproducción del iPod es realmente aleatoria?

En *The Guardian*, Steven Levy escribió acerca de una entrevista realizada a Steve Jobs, quien era director de Apple en ese momento, en la que le planteó el siguiente dilema: "Tengo un problema con mi iPod. La función de reproducción aleatoria no parece realmente aleatoria. Algunos artistas aparecen demasiado y algunos no aparecen nunca". Según Jeff Robbin, directivo del equipo de desarrollo de iTunes, "se trata de una aleatoriedad inequívoca". El matemático John Allen Paulos comentó que "a menudo interpretamos e imponemos patrones sobre eventos que son aleatorios". Levy también afirma que, cuando pensamos que la función de reproducción aleatoria del iPod no es aleatoria, el problema reside en nuestra percepción. Nuestra mente percibe patrones y tendencias que en realidad no existen. Con frecuencia escuchamos rachas de canciones consecutivas del mismo artista, y creemos que esto no se debe al azar; pero, con una verdadera aleatoriedad, este tipo de rachas consecutivas son mucho más probables de lo esperado.

La percepción incorrecta de ausencia de aleatoriedad provocó que Apple introdujera una característica de "reproducción inteligente" en una nueva versión de iTunes: ésta permite que los usuarios tengan control sobre las sucesiones de varias canciones del mismo artista, para evitar rachas consecutivas del mismo cantante. Según declaró Steve Jobs en aquella ocasión: "Lo estamos haciendo menos aleatorio para que se perciba más aleatorio".

EJEMPLO 2 Muestra grande: Aleatoriedad de los nacimientos

El conjunto de datos 4 "Nacimientos" en el apéndice B lista los sexos de 100 nacimientos en el Bellevue Hospital Center. Consideremos la secuencia de los géneros listados a continuación, donde 1: niño y 0: niña. (La lista completa de 100 géneros se puede ver en el conjunto de datos 4). Use un nivel de significancia de 0.05 para probar la afirmación de que la secuencia es aleatoria.

1 0 1 0 0 0 0 1 . . . 1 0 1

SOLUCIÓN

VERIFICACIÓN DE REQUISITOS (1) Los datos están ordenados. (2) Cada valor de datos se clasifica en una de dos categorías distintas (niño/niña). Los requisitos se cumplen. ✓

Las hipótesis nula y alternativa son:

H_0: La secuencia es aleatoria.

H_1: La secuencia no es aleatoria.

Un examen de la secuencia de 100 géneros da como resultado los siguientes valores:

n_1 = número de niños = 51

n_2 = número de niñas = 49

G = número de rachas = 59

Dado que $n_1 > 20$, necesitamos calcular el dato estadístico de prueba z, por lo que primero debemos evaluar μ_G y σ_G de la siguiente manera:

$$\mu_G = \frac{2n_1n_2}{n_1 + n_2} + 1 = \frac{2(51)(49)}{51 + 49} + 1 = 50.98$$

$$\sigma_G = \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

$$= \sqrt{\frac{(2)(51)(49)[2(51)(49) - 51 - 49]}{(51 + 49)^2(51 + 49 - 1)}} = 4.972673$$

Ahora encontramos el dato estadístico de prueba:

$$z = \frac{G - \mu_G}{\sigma_G} = \frac{59 - 50.98}{4.972673} = 1.61$$

Podemos usar el dato estadístico de prueba $z = 1.61$ para determinar el valor P de 0.1074, que es mayor que el nivel de significancia de $\alpha = 0.05$, por lo que no podemos rechazar la hipótesis nula de aleatoriedad.

Además, como 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 dato estadístico de prueba $z = 1.61$ no cae dentro de la región crítica, por lo que de nuevo no podemos rechazar la hipótesis nula de aleatoriedad.

INTERPRETACIÓN

No tenemos suficiente evidencia para rechazar la aleatoriedad de la secuencia de 100 nacimientos en el Bellevue Hospital Center (del conjunto de datos 4 "Nacimientos" en el apéndice B).

SU TURNO Realice el ejercicio 9 "Prueba de aleatoriedad de las victorias en el Súper Bowl".

Prueba de aleatoriedad por encima y por debajo de la media o la mediana

Podemos evaluar la aleatoriedad en la forma en que los datos numéricos fluctúan por encima o por debajo de una media o una mediana. Para probar la aleatoriedad por encima y por debajo de la mediana, por ejemplo, use los datos muestrales para hallar el valor de la mediana, luego reemplace cada valor individual con la letra *E* si está por *encima* de la mediana y reemplácelo por *D* si está *debajo* de la mediana. Elimine cualquier valor que sea igual a la mediana. (Es útil escribir las Es y las Ds directamente encima o debajo de los números que representan, porque esto facilita la verificación y también reduce la posibilidad de tener un número equivocado de letras). Después de encontrar la secuencia de letras *E* y *D*, podemos proceder a aplicar la prueba de rachas como se describió anteriormente.

Los economistas usan la prueba de rachas para aleatoriedad por encima y por debajo de la mediana para identificar tendencias o ciclos. Una tendencia económica ascendente contendría un predominio de *D*s al comienzo y *E*s al final, por lo que el número de corridas sería muy pequeño. Una tendencia descendente tendría *E*s dominando al principio y *D*s al final, con un pequeño número de rachas. Un patrón cíclico produciría una secuencia que cambia sistemáticamente, por lo que el número de rachas tendería a ser grande.

CENTRO DE TECNOLOGÍA

Prueba de rachas para aleatoriedad

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Determine la cantidad de elementos en la primera categoría, la cantidad de elementos en la segunda categoría y cuente el número de rachas.
2. Haga clic en **Analysis** en el menú superior.
3. Seleccione **Runs Tests** en el menú desplegable.
4. Ingrese un nivel de significancia y los valores para el número de rachas, el número de elementos en la primera categoría (*Elemento 1*) y el número de elementos en la segunda categoría (*Elemento 2*).
5. Haga clic en **Evaluate**.

Minitab

Minitab hará una prueba de rachas solamente con una secuencia de datos numéricos. El manual y el libro de trabajo del laboratorio estudiantil de Minitab proporciona información adicional sobre cómo eludir esta restricción.

1. Ingrese los datos numéricos en la columna *C1*.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Nonparametrics** en el menú desplegable y elija **Runs Test** en el submenú.
4. Seleccione la columna **C1** en V*ariables*.
5. Seleccione probar por encima y por debajo de la media o ingrese un valor que desee utilizar.
6. Haga clic en **OK**.

StatCrunch

No disponible

Calculadora TI-83/84 Plus

No disponible

Excel

No disponible

13-7 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

En los ejercicios 1 a 4, use la siguiente secuencia de afiliaciones a partidos políticos de los recientes presidentes de Estados Unidos, donde R representa a los republicanos y D representa a los demócratas.

R R R R D R D R R R D R R R D D R D D R R D R R D R D

1. Prueba de sesgo ¿Se puede usar la prueba de rachas para mostrar que la proporción de republicanos es significativamente mayor que la proporción de demócratas?

2. Notación Identifique los valores de n_1, n_2 y G que se usarían en la prueba de rachas para aleatoriedad.

3. Prueba de rachas Si usamos un nivel de significancia de 0.05 para probar la aleatoriedad, ¿cuáles son los valores críticos de la tabla A-10? Con base en esos valores y el número de rachas del ejercicio 2, ¿qué debería concluir sobre la aleatoriedad?

4. ¿Buena muestra? Dada la secuencia de datos, si no rechazamos la aleatoriedad, ¿se deduce que el método de muestreo es adecuado para los métodos estadísticos? Explique.

Uso de la prueba de rachas para aleatoriedad. ***En los ejercicios 5 a 10, use la prueba de rachas con un nivel de significancia de α = 0.05. (Todos los datos están ordenados por fila).***

5. Muertes en cumplimiento de la ley A continuación se listan las cifras de víctimas mortales del cumplimiento de la ley durante 20 años recientes y consecutivos. Primero encuentre la media, identifique cada valor por encima de la media (*E*) o por debajo de la media (*D*), luego pruebe la aleatoriedad por encima y por debajo de la media. ¿Hay una tendencia?

183	140	172	171	144	162	241	159	150	165
163	156	192	148	125	161	171	126	107	117

6. Dígitos pares e impares en Pi Un artículo del *New York Times* sobre el cálculo de las posiciones 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 muestran las primeras 25 posiciones decimales de π. Pruebe la aleatoriedad de la forma en que los dígitos impares (I) y pares (P) ocurren en la secuencia. Con base en el resultado, ¿la afirmación del *New York Times* parece ser precisa?

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

7. Lotería de reclutamiento En 1970, se utilizó una lotería para determinar quién sería reclutado en el ejército de Estados Unidos. Las 366 fechas en el año se colocaron en cápsulas individuales, se mezclaron, y luego se seleccionaron las cápsulas para identificar las fechas de nacimiento de los hombres que se reclutarían primero. A continuación se listan los primeros 30 resultados. Pruebe la aleatoriedad antes y después de la mitad del año, que es el 1 de julio.

Sep. 14	Abr. 24	Dic. 30	Feb. 14	Oct. 18	Sep. 6	Oct. 26	Sep. 7	Nov. 22
Dic. 6	Ago. 31	Dic. 7	Jul. 8	Abr. 11	Jul. 12	Dic. 29	Ene. 15	Sep. 26
Nov. 1	Jun. 4	Ago. 10	Jun. 26	Jul. 24	Oct. 5	Feb. 19	Dic. 14	Jul. 21
Jun. 5	Mar. 2	Mar. 31						

8. Periódicos Los expertos en medios afirman que los periódicos impresos diarios están disminuyendo debido a la llegada de Internet. A continuación se listan los números de periódicos impresos en Estados Unidos durante una secuencia reciente de años. Primero encuentre la mediana, luego pruebe la aleatoriedad de los números por encima y por debajo de la mediana. ¿Qué sugieren los resultados?

1611	1586	1570	1556	1548	1533	1520	1509	1489	1483	1480
1468	1457	1456	1457	1452	1437	1422	1408	1387	1382	

Prueba de rachas con muestras grandes. ***En los ejercicios 9 a 12, use la prueba de rachas con un nivel de significancia de α = 0.05. (Todos los datos están ordenados por fila).***

9. Prueba de aleatoriedad de las victorias en el Súper Bowl A continuación se listan las designaciones de la conferencia a la que pertenecen los equipos que ganaron el Súper Bowl, donde N expresa un equipo de la NFC y A uno de la AFC. ¿Los resultados sugieren que alguna de las conferencias es superior?

N	N	A	A	A	N	A	A	A	A	A	N	A	A	A	N	N	A	N	N	N	N	N	N
N	N	N	N	N	N	A	A	N	A	A	N	A	A	A	A	N	A	N	N	N	A	N	A

10. Victorias de la Serie Mundial de Béisbol Pruebe la afirmación de que la secuencia de victorias en la Serie Mundial de los equipos de la Liga Americana y la Liga Nacional es aleatoria. En la página siguiente se muestran los resultados recientes, con A = Liga Americana y N = Liga Nacional.

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	A	N	A	A	N	A	N	A	N	N	N	A	N										

11. Mercado de valores: pruebas de aleatoriedad por encima y por debajo de la mediana A continuación se listan los valores anuales más altos del promedio industrial Dow Jones para una secuencia de años recientes. Encuentre la mediana, luego pruebe la aleatoriedad por debajo y por encima de la mediana. ¿Qué sugiere el resultado sobre el mercado de valores como una consideración de inversión?

969	995	943	985	969	842	951	1036	1052	892	882	1015
1000	908	898	1000	1024	1071	1287	1287	1553	1956	2722	2184
2791	3000	3169	3413	3794	3978	5216	6561	8259	9374	11568	11401
11350	10635	10454	10855	10941	12464	14198	13279	10580	11625	12929	13589
16577	18054										

12. Aleatoriedad de los nacimientos Repita el ejemplo 2 "Aleatoriedad de los nacimientos" utilizando los nacimientos del conjunto de datos 4 "Nacimientos" en el apéndice B, para los sexos de los nacimientos en el Strong Memorial Hospital en lugar del Bellevue Center Hospital. Nuevamente use un nivel de significancia de 0.05 para probar la afirmación de que la secuencia de géneros es aleatoria.

13-7 Más allá de lo básico

13. Determinación de valores críticos

a. Usando todos los elementos A, A, A, B, B, B, B, B, B, lista las 84 diferentes secuencias posibles.

b. Encuentra el número de rachas para cada una de las 84 secuencias.

c. Use los resultados de los incisos (a) y (b) para encontrar sus propios valores críticos para *G*.

d. Compare sus resultados con los que se presentan en la tabla A-10.

Examen rápido del capítulo

1. Citas rápidas Algunos de los métodos no paramétricos de este capítulo utilizan rangos de datos. Encuentre los rangos correspondientes a las siguientes clasificaciones de atractivo (1 = Nada atractivo, 10 = Extremadamente atractivo) de hombres dadas por mujeres que participaron en un evento de citas rápidas (del conjunto de datos 18 "Citas rápidas"): 5, 7, 7, 8, 7.

2. Eficiencia ¿Qué significa cuando decimos que la prueba de correlación de rangos tiene una calificación de eficiencia de 0.91 cuando se compara con la prueba paramétrica de correlación lineal?

3. Pruebas no paramétricas

a. ¿Cuál de los siguientes términos se usa a veces en lugar de "prueba no paramétrica": *prueba de normalidad*; *prueba de anormalidad*; *prueba sin distribución*; *último testamento*; *prueba de paciencia*?

b. ¿Por qué el término que es la respuesta al inciso (a) es mejor que "prueba no paramétrica"?

4. Longitud del pie/estatura A continuación se listan las longitudes de los pies (cm) y las estaturas (cm) de los hombres del conjunto de datos 2 "Pies y estaturas" en el apéndice B. ¿Qué método de estadística no paramétrica se debe utilizar? ¿Qué característica de los datos se investiga con esta prueba?

Longitud del pie	27.8	25.7	26.7	25.9	26.4	29.2	26.8	28.1	25.4	27.9
Estatura	180.3	175.3	184.8	177.8	182.3	185.4	180.3	175.3	177.8	185.4

5. Longitud del pie/Estatura Al analizar los datos pareados del ejercicio 4, ¿los valores P y las conclusiones de la prueba no paramétrica y la prueba paramétrica son siempre los mismos?

6. Longitud del pie/Estatura Para los datos muestrales dados en el ejercicio 4, identifique al menos una ventaja de usar la prueba no paramétrica adecuada sobre la prueba paramétrica.

7. Prueba de rachas Suponga que utilizamos la prueba de aleatoriedad por encima y por debajo de la media para el valor monetario de Google al final de cada año durante 20 años y encontramos que $G = 2$. ¿Qué nos dice ese valor sobre el valor monetario de Google?

8. Prueba de rachas Considere datos muestrales que consisten en los géneros de delincuentes acusados de piratear los sistemas informáticos de las empresas. Determine si las siguientes afirmaciones son verdaderas o falsas.

a. Si la prueba de rachas sugiere que los datos muestrales ocurren en un orden aleatorio, se deduce que los datos se han seleccionado al azar.

b. Si la prueba de rachas sugiere que los datos muestrales ocurren en un orden aleatorio, entonces no hay una diferencia significativa entre las proporciones de hombres y mujeres.

9. Prueba del signo y prueba de los rangos con signo de Wilcoxon ¿Cuál es una gran ventaja de la prueba de los rangos con signo de Wilcoxon sobre la prueba del signo al analizar datos que constan de pares relacionados?

10. ¿Qué prueba? Tres jueces diferentes califican a los mismos cantantes en una escala de 0 a 10. ¿Qué método de este capítulo se puede utilizar para determinar si uno de los jueces es más severo o indulgente que los demás, según lo indica una calificación mediana que es significativamente diferente a la de los demás?

Ejercicios de repaso

Uso de pruebas no paramétricas. ***En los ejercicios 1 a 10, use un nivel de significancia de 0.05 con la prueba indicada. Si no se especifica ninguna prueba particular, use la prueba no paramétrica adecuada de este capítulo***

1. Estrés laboral e ingresos A continuación se listan las puntuaciones de estrés laboral y los salarios anuales medianos (miles de dólares) para varios empleos, incluyendo a los bomberos, pilotos de líneas aéreas, oficiales de policía y profesores universitarios (con base en datos de las "Puntuaciones de estrés laboral" de CareerCast.com). ¿Estos datos sugieren que existe una correlación entre el estrés laboral y el ingreso anual? ¿Parece que los trabajos con más estrés tienen salarios más altos?

Estrés	71.59	60.46	50.82	6.94	8.10	50.33	49.2	48.8	11.4
Salario mediano	45.6	98.4	57.0	69.0	35.4	46.1	42.5	37.1	31.2

2. Presidentes, papas y monarcas A continuación se enumeran los años que vivieron los presidentes, papas y monarcas británicos después de su toma de posesión, elección o coronación, respectivamente. Supongamos que los datos son muestras seleccionadas al azar de poblaciones más grandes. Pruebe la afirmación de que las tres muestras provienen de poblaciones con la misma mediana.

Presidentes	10	29	26	28	15	23	17	25	0	20	4	1	24	16	12
	4	10	17	16	0	7	24	12	4	18	21	11	2	9	36
	12	28	3	16	9	25	23	32							
Papas	2	9	21	3	6	10	18	11	6	25	23	6	2	15	32
	25	11	8	17	19	5	15	0	26						
Monarcas	17	6	13	12	13	33	59	10	7	63	9	25	36	15	

3. Serie mundial Las últimas 110 Series Mundiales de Béisbol terminaron con 63 victorias para equipos de la Liga Americana y 47 victorias para equipos de la Liga Nacional. Use la prueba del signo para probar la afirmación de que en cada Serie Mundial, el equipo de la Liga Americana tiene una probabilidad de 0.5 de ganar.

4. Lotería de California A continuación se listan los primeros dígitos consecutivos de la Lotería de California Daily 4. Pruebe la aleatoriedad de los enteros pares e impares. ¿La lotería parece estar funcionando como debería?

5	2	2	8	4	8	8	7	1	0	6	4	1	5	1	5	5	3	1	4	1	5	0	0	3	9	6	6	3	7

5. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se les pidió a las mujeres que calificaran el atractivo de sus citas masculinas, y una muestra de los resultados se detalla a continuación (1 = Nada atractivo, 10 = Extremadamente atractivo). Use la prueba del signo para probar la afirmación de que la muestra proviene de una población con una mediana igual a 5.

5	8	3	8	6	10	3	7	9	8	5	5	6	8	8	7	3	5	5	6	8	7	8	8	8	7

6. Citas rápidas Repita el ejercicio anterior usando la prueba de los rangos con signo de Wilcoxon.

7. Old Faithful A continuación se listan los intervalos de tiempo (min) entre las erupciones del géiser Old Faithful. Los tiempos "recientes" están dentro de los últimos años, y los "pasados" son de 1995. Pruebe la afirmación de que las dos muestras provienen de poblaciones con la misma mediana. ¿La conclusión cambia con un nivel de significancia de 0.01?

Reciente	78	91	89	79	57	100	62	87	70	88	82	83	56	81	74	102	61
Pasado (1995)	89	88	97	98	64	85	85	96	87	95	90	95					

8. Tarifas de aerolíneas A continuación se listan los costos (en dólares) de ocho vuelos diferentes desde Nueva York (JFK) a San Francisco para Virgin America, US Airways, United Airlines, JetBlue, Delta, American Airlines, Alaska Airlines y Sun Country Airlines. (Cada par de costos es para el mismo vuelo). Use la prueba del signo para probar la afirmación de que no hay diferencia en el costo entre los vuelos programados con 1 día de anticipación y los programados con 30 días de anticipación. ¿Cuál parece ser una sabia estrategia de programación?

Vuelos programados con un día de antelación	584	490	584	584	584	606	628	717
Vuelos programados con 30 días de antelación	254	308	244	229	284	509	394	258

9. Tarifas de aerolíneas Consulte los mismos datos del ejercicio anterior. Use la prueba de rangos con signo de Wilcoxon para probar la afirmación de que las diferencias entre las tarifas para vuelos programados con 1 día de anticipación y los programados con 30 días de anticipación tienen una mediana igual a 0. ¿Qué sugieren los resultados?

10. Clasificación de universidades por estudiantes y *US News & World Report* Cada año, la revista *US News & World Report* publica clasificaciones de las universidades con base en estadísticas como las tasas de admisión, las tasas de graduación, los tamaños de clase, la proporción de profesores y estudiantes, los salarios de los docentes y las calificaciones dadas por administradores. Los economistas Christopher Avery, Mark Glickman, Caroline Minter Hoxby y Andrew Metrick pusieron en práctica un método alternativo de analizar las opciones universitarias de 3240 estudiantes de alto rendimiento en su último año de la escuela preparatoria. Examinaron las universidades que ofrecieron la admisión junto con las universidades a las que los estudiantes decidieron asistir. La siguiente tabla lista las clasificaciones de una muestra pequeña de universidades. Encuentre el valor del coeficiente de correlación de rangos y úselo para determinar si existe una correlación entre las clasificaciones de los estudiantes y las clasificaciones de la revista.

Rangos de estudiantes	1	2	3	4	5	6	7	8
Rangos de *US News & World Report*	1	2	5	4	7	6	3	8

Ejercicios de repaso acumulado

En los ejercicios 1 a 3, use los datos que se dan a continuación. Los valores son los tiempos de demora en la salida (minutos) para los vuelos de American Airlines de Nueva York a Los Ángeles. Los valores negativos corresponden a vuelos que salieron temprano.

Vuelo 1 (min)	−2	−1	−2	2	−2	0	−2	−3
Vuelo 19 (min)	19	−4	−5	−1	−4	73	0	1
Vuelo 21 (min)	18	60	142	−1	−11	−1	47	13

1. Demoras en la salida de vuelos Compare las tres muestras usando medias, medianas y desviaciones estándar.

2. Prueba de normalidad Use los tiempos de demora en la salida para el vuelo 19 y prueba la normalidad usando una gráfica cuantilar normal.

3. Tiempos de demora en la salida Use una prueba no paramétrica para probar la afirmación de que las tres muestras provienen de poblaciones con el mismo tiempo mediano de demora en la salida.

4. Pruebas de drogas Hay una tasa de 3.9% de resultados positivos en las pruebas de drogas que se realizan a trabajadores en Estados Unidos (según datos de *Quest Diagnostics*). Suponiendo que este estadístico se basa en una muestra con un tamaño de 2000, elabore una estimación del intervalo de confianza del 95% para el *porcentaje* de resultados positivos en la prueba de drogas. Escriba una breve declaración que interprete el intervalo de confianza.

5. Pruebas de drogas Utilice los datos del ejercicio anterior y pruebe la afirmación de que la proporción de resultados positivos en las pruebas de drogas entre los trabajadores de Estados Unidos es mayor al 3%. Use un nivel de significancia de 0.05.

6. Aleatoriedad Consulte las siguientes edades de los presidentes electos de Estados Unidos el día que tomaron posesión (del conjunto de datos 15 "Presidentes" en el apéndice B). Pruebe la aleatoriedad por encima y por debajo de la media. ¿Los resultados sugieren una tendencia a la alza o a la baja?

57	61	57	57	58	57	61	54	68	49	64	48	65	52	46	54	49	47
55	54	42	51	56	55	51	54	51	60	62	43	55	56	52	69	64	46
54	47																

7. Tamaño de muestra Los avances en la tecnología están afectando dramáticamente diferentes aspectos de nuestras vidas. Por ejemplo, la cantidad de periódicos diarios impresos está disminuyendo debido al fácil acceso a las noticias en Internet y por la televisión. Para ayudar a abordar estos problemas, queremos estimar el porcentaje de adultos en Estados Unidos que usan una computadora al menos una vez al día. Encuentre el tamaño de muestra necesario para estimar ese porcentaje. Suponga que queremos un 95% de confianza en que el porcentaje muestral se encuentre dentro de dos puntos porcentuales del porcentaje poblacional real.

8. Media y mediana En un año reciente, los jugadores del equipo de béisbol de los Yankees de Nueva York tenían sueldos con una media de $7,052,129 y una mediana de $2,500,000. Explique por qué la media y la mediana pueden ser tan diferentes.

9. Miedo a las alturas Entre los lectores de un sitio web de *USA Today*, 285 eligieron responder a la siguiente pregunta: "¿Le teme a las alturas en los rascacielos?" Entre los que decidieron responder, 46% respondió "sí" y el 54% respondió "no". Use un nivel de significancia de 0.05 para probar la afirmación de que la mayoría de la población no le teme a las alturas en los rascacielos. ¿Qué hay de erróneo en esta prueba de hipótesis?

10. Teléfonos celulares y accidentes: análisis de un informe de periódico En un artículo de *Associated Press*, se informó que los investigadores "seleccionaron al azar a 100 automovilistas de Nueva York que habían sufrido un accidente y 100 que no habían tenido un accidente. De los que sufrieron un accidente, 13.7 por ciento poseía un teléfono celular, mientras que sólo 10.6 por ciento de los conductores sin accidentes tenían un teléfono en el automóvil". ¿Qué es erróneo en estos resultados?

Proyecto de tecnología

Los intentos anteriores para identificar o contactar vida inteligente extraterrestre han implicado esfuerzos para enviar mensajes de radio con información sobre nosotros, los terrícolas. El doctor Frank Drake, de la Universidad de Cornell, desarrolló un mensaje de radio de este tipo que podría transmitirse como una serie de pulsos y silencios. Los pulsos y los silencios se pueden considerar como 1s y 0s. A continuación se presenta un mensaje que consta de 77 entradas de 0s y 1s. Si factorizamos 77 en los números primos 7 y 11 y luego hacemos una cuadrícula de 11 $\times$ 7 y ponemos un punto en las posiciones correspondientes a un pulso de 1, podemos obtener una imagen simple de algo. Supongamos que la secuencia de 77 entradas de 1s y 0s se envía como un mensaje de radio que es interceptado por vida extraterrestre con suficiente inteligencia como para haber estudiado este libro. Si el mensaje de radio se prueba usando los métodos de este capítulo, ¿la secuencia parecerá ser "ruido aleatorio" o se identificará como un patrón que no es aleatorio? Además, trace la imagen representada por los dígitos e identifíquela.

0	0	1	1	1	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0
1	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	1	0	0
0	0	1	1	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
1	0	0	0	0	1	0	1	0	0	0	0	1	0							

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿El reclutamiento fue aleatorio?

El 1 de diciembre de 1969, durante la Guerra de Vietnam, se utilizó una lotería para determinar quién sería reclutado por el ejército de Estados Unidos, pero la lotería generó una gran controversia. Las diferentes fechas en un año se colocaron en 366 cápsulas individuales. Primero, se colocaron las 31 cápsulas de enero en una caja, luego se agregaron las 29 cápsulas de febrero y los dos meses se mezclaron. Enseguida se agregaron las 31 cápsulas de marzo y se mezclaron los tres meses. Este proceso continuó hasta que se incluyeron todos los meses. La primera cápsula seleccionada fue el 14 de septiembre, por lo que los hombres nacidos en esa fecha fueron reclutados primero. La lista adjunta muestra las 366 fechas en el orden de selección de prioridad. Esta lista puede descargarse en www.pearsonenespañol.com/triola.

Análisis de resultados

a. Use la prueba de rachas para probar la aleatoriedad de la secuencia por encima y por debajo de la mediana de 183.5.

b. Use la prueba de Kruskal-Wallis para probar la afirmación de que los 12 meses tienen números de prioridad extraídos de la misma población.

c. Calcule las 12 medias mensuales. Luego grafique esas 12 medias. (Donde la escala horizontal muestre los 12 meses y la escala vertical varíe de 100 a 260). Observe cualquier patrón que sugiera que los números de prioridad 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 declaración que explique por qué cree que fue o no fue justo. Si decidió que esta lotería fue injusta, describa un proceso para seleccionar los números de la lotería que hubiera sido justo.

Día	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
1	305	86	108	32	330	249	93	111	225	359	19	129
2	159	144	29	271	298	228	350	45	161	125	34	328
3	251	297	267	83	40	301	115	261	49	244	348	157
4	215	210	275	81	276	20	279	145	232	202	266	165
5	101	214	293	269	364	28	188	54	82	24	310	56
6	224	347	139	253	155	110	327	114	6	87	76	10
7	306	91	122	147	35	85	50	168	8	234	51	12
8	199	181	213	312	321	366	13	48	184	283	97	105
9	194	338	317	219	197	335	277	106	263	342	80	43
10	325	216	323	218	65	206	284	21	71	220	282	41
11	329	150	136	14	37	134	248	324	158	237	46	39
12	221	68	300	346	133	272	15	142	242	72	66	314

continúa

Día	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
13	318	152	259	124	295	69	42	307	175	138	126	163
14	238	4	354	231	178	356	331	198	1	294	127	26
15	17	89	169	273	130	180	322	102	113	171	131	320
16	121	212	166	148	55	274	120	44	207	254	107	96
17	235	189	33	260	112	73	98	154	255	288	143	304
18	140	292	332	90	278	341	190	141	246	5	146	128
19	58	25	200	336	75	104	227	311	177	241	203	240
20	280	302	239	345	183	360	187	344	63	192	185	135
21	186	363	334	62	250	60	27	291	204	243	156	70
22	337	290	265	316	326	247	153	339	160	117	9	53
23	118	57	256	252	319	109	172	116	119	201	182	162
24	59	236	258	2	31	358	23	36	195	196	230	95
25	52	179	343	351	361	137	67	286	149	176	132	84
26	92	365	170	340	357	22	303	245	18	7	309	173
27	355	205	268	74	296	64	289	352	233	264	47	78
28	77	299	223	262	308	222	88	167	257	94	281	123
29	349	285	362	191	226	353	270	61	151	229	99	16
30	164		217	208	103	209	287	333	315	38	174	3
31	211		30		313		193	11		79		100

Actividades en equipo

1. Actividad fuera de clase La mitad de los estudiantes debe inventar los resultados de 200 lanzamientos de monedas y la otra mitad debe recolectar los resultados de 200 lanzamientos reales de una moneda. Luego use la prueba de rachas para determinar si los resultados parecen ser aleatorios.

2. Actividad en clase Use la distribución de asientos existente en su clase y aplique la prueba de rachas para determinar si los estudiantes están organizados aleatoriamente de acuerdo con el género. Después de registrar la disposición de los asientos, el análisis se puede hacer en subgrupos de tres o cuatro estudiantes.

3. Actividad en clase Divídanse en grupos de 8 a 12 personas. *Mida* la estatura y el alcance de los brazos de cada miembro del grupo. Para medir el alcance de los brazos, el sujeto debe pararse con los brazos extendidos, como las alas de un avión. Divida las siguientes tareas entre subgrupos de tres o cuatro personas.

a. Utilice la correlación de rangos con los datos muestrales pareados para determinar si existe una correlación entre la estatura y el alcance.

b. Use la prueba del signo para probar la diferencia entre las dos variables.

c. Use la prueba de los rangos con signo de Wilcoxon para probar la diferencia entre las dos variables.

4. Actividad en clase Realice la actividad 3 usando el pulso en lugar del alcance de los brazos. Mida los pulsos contando el número de latidos en 1 minuto.

5. Actividad fuera de clase Divídanse en grupos de tres o cuatro estudiantes. Investigue la relación entre dos variables recolectando sus propios datos muestrales pareados y utilizando los métodos de la sección 13-6 para determinar si existe una correlación. Temas sugeridos:

- ¿Existe una correlación entre el sabor y el costo de las diferentes marcas de galletas con chispas de chocolate (o bebidas de cola)? (El sabor se puede medir en una escala numérica, como de 1 a 10).
- ¿Existe una correlación entre los salarios de los jugadores profesionales de béisbol (o basquetbol o fútbol) y sus logros en la temporada (como el promedio de bateo o los puntos anotados)?
- ¿Existe una correlación entre las tasas de consumo de combustible de los automóviles y sus pesos?
- ¿Existe una correlación entre la longitud de los pies de los hombres (o de las mujeres) y sus estaturas?
- ¿Existe una correlación entre los promedios de calificaciones de los estudiantes y el tiempo que pasan mirando televisión?
- ¿Existe una correlación entre las estaturas de los padres (o las madres) y las estaturas de sus primogénitos (o primogénitas)?

6. Actividad fuera de clase Consulte el proyecto "De los datos a la decisión" de este capítulo, que incluye el análisis de la lotería de 1970 utilizada para el reclutamiento de hombres en el ejército de Estados Unidos. Debido a que los resultados de 1970 generaron inquietudes sobre la aleatoriedad de la selección de los números de prioridad, diseñe un nuevo procedimiento para generar los 366 números de prioridad. Use su propio procedimiento para generar los 366 números y pruebe sus resultados usando las técnicas sugeridas en los incisos (a), (b) y (c) del proyecto "De los datos a la decisión". ¿Cómo se comparan sus resultados con los obtenidos en 1970? ¿Su proceso de selección aleatoria parece ser mejor que el utilizado en 1970? Escriba un reporte que describa claramente el proceso que diseñó. También incluya sus análisis y conclusiones.

7. Actividad fuera de clase Divídanse en grupos de tres o cuatro personas. Encueste a otros estudiantes pidiéndoles que identifiquen su especialidad y género. Para cada sujeto encuestado, determine su cantidad de seguidores en Twitter o amigos en Facebook. Use los datos muestrales para responder las siguientes preguntas:

- ¿Los números de seguidores en Twitter o amigos en Facebook parecen ser iguales para ambos sexos?
- ¿Los números de seguidores en Twitter o amigos en Facebook parecen ser iguales para las diferentes especialidades?

8. Actividad en clase Divídanse en grupos de 8 a 12 personas. Mida la estatura de cada miembro del grupo, también mida la altura de su ombligo, que es la altura desde el piso hasta el ombligo de la persona. Use el coeficiente de correlación de rangos para determinar si existe una correlación entre la estatura y la altura del ombligo.

9. Actividad en clase Divídanse en grupos de tres o cuatro personas. El apéndice B incluye muchos conjuntos de datos aún no utilizados con los métodos de este capítulo. Busque en el apéndice B algunas variables de interés, luego analícelas usando los métodos adecuados de estadística no paramétrica. Indique sus conclusiones y trate de identificar aplicaciones prácticas.

14

CONTROL ESTADÍSTICO DE PROCESOS

PROBLEMA DEL CAPÍTULO

Seguridad aérea: ¿la producción de altímetros para aviones está fuera de control estadístico?

La compañía Orange Avionics fabrica instrumentos y dispositivos de navegación utilizados en aviones, incluyendo los altímetros que proporcionan a los pilotos lecturas de sus alturas sobre el nivel del mar. La precisión de los altímetros es importante porque los pilotos dependen de ellos para mantener alturas seguras por encima de montañas, torres y edificios, así como una separación vertical adecuada con respecto a otros aviones. La precisión de los altímetros es especialmente importante cuando los pilotos se acercan al aterrizaje sin poder ver el suelo. Ha habido fallecimientos de pilotos y pasajeros en accidentes causados por lecturas incorrectas del altímetro que llevaron a los pilotos a creer que estaban a salvo sobre el suelo cuando en realidad volaban en forma peligrosamente baja.

Debido a que estos dispositivos son tan importantes para la seguridad operativa de la aviación, su exactitud está cuidadosamente regulada por las regulaciones gubernamentales. La Regulación 91.411 de la Administración Federal de Aviación (FAA, por sus siglas en inglés) exige pruebas periódicas a los altímetros de las aeronaves, los cuales deben cumplir con especificaciones incluidas en la Parte 43 del Apéndice E de las reglamentaciones de la FAA. Una de esas especificaciones es que un altímetro debe dar una lectura con un error de no más de 30 pies cuando se prueba para una altitud de 2000 pies.

En la compañía Orange Avionics, se seleccionan cinco altímetros al azar de su producción en cada uno de 20 días de producción consecutivos; la tabla 14-1 lista los errores (en pies) cuando se prueban en una cámara de presión que simula una altitud de 2000 pies. El día 1, por ejemplo, las lecturas de altitud reales para los cinco altímetros seleccionados son 1999, 2007, 2007, 1999 y 2002 pies, por lo que los errores correspondientes (en pies) son −1, 7, 7, −1 y 2.

En este capítulo, evaluamos este proceso de fabricación del altímetro al analizar el comportamiento de los errores a lo largo del tiempo. Veremos cómo se pueden usar los métodos estadísticos para monitorear un proceso de fabricación con el objetivo de identificar y corregir problemas graves. Además de ayudar a las empresas a mantenerse en el negocio, los métodos estadísticos pueden afectar nuestra seguridad de manera muy significativa.

TABLA 14-1 Errores de altímetros para avión (en pies)

Disponible para descarga en www.pearsonenespañol.com/triola

Día	Error (pies)					Media $\bar{x}$	Mediana	Rango R	Desviación estándar s
1	−1	7	7	−1	2	2.8	2	8	4.02
2	9	2	3	−1	−5	1.6	2	14	5.18
3	5	0	4	1	3	2.6	3	5	2.07
4	0	10	1	4	3	3.6	3	10	3.91
5	10	8	5	16	3	8.4	8	13	5.03
6	6	9	3	−6	−16	−0.8	3	25	10.18
7	14	4	2	14	8	8.4	8	12	5.55
8	20	9	4	−4	28	11.4	9	32	12.72
9	−12	−4	21	−5	−5	−1.0	−5	33	12.71
10	−21	−12	7	2	19	−1.0	2	40	15.76
11	2	20	12	−3	−13	3.6	2	33	12.86
12	−1	9	5	12	−4	4.2	5	16	6.69
13	−16	−21	−14	6	7	−7.6	−14	28	13.13
14	−21	−27	−3	−1	−9	−12.2	−9	26	11.37
15	−27	0	−37	−15	4	−15.0	−15	41	17.42
16	0	29	18	−27	25	9.0	18	56	22.99
17	29	−24	9	20	48	16.4	20	72	26.73
18	1	30	8	−15	13	7.4	8	45	16.47
19	2	−5	−23	−25	−4	−11.0	−5	27	12.19
20	3	−46	31	−20	−3	−7.0	−3	77	28.50

OBJETIVOS DEL CAPÍTULO

En este capítulo se presentan métodos para elaborar e interpretar las gráficas de control que se usan comúnmente para monitorear las características cambiantes de datos a lo largo del tiempo. Una gráfica de control incluye una línea central, un límite de control superior y un límite de control inferior. Se puede usar una gráfica de control para determinar si un proceso es estadísticamente estable (o si está bajo control estadístico) con sólo una variación natural y sin patrones, ciclos o puntos inusuales.

Los objetivos del capítulo son:

14-1 Gráficas de control para la variación y la media

- Desarrollar la capacidad de elaborar una gráfica de rachas.
- Desarrollar la capacidad de elaborar una gráfica de control para R (el rango).
- Desarrollar la capacidad de elaborar una gráfica de control para $\bar{x}$.
- Identificar los criterios fuera de control y aplicarlos para determinar si los datos del proceso están bajo control estadístico.

14-2 Gráficas de control para atributos

- Desarrollar la capacidad de elaborar una gráfica de control para p, la proporción correspondiente a algún atributo, como ser defectuoso.
- Identificar los criterios fuera de control y aplicarlos para determinar si los datos de los atributos están bajo control estadístico con sólo la variación natural y sin patrones, ciclos o puntos inusuales.

14-1 Gráficas de control para la variación y la media

Concepto clave En esta sección se presentan las gráficas de rachas, las gráficas R y las gráficas $\bar{x}$ como herramientas que permiten monitorear las características de los datos a lo largo del tiempo. Podemos usar dichas gráficas para determinar si un proceso es estadísticamente estable (o si está bajo control estadístico).

Datos de proceso

La siguiente definición describe formalmente el tipo de datos que se considerarán en este capítulo.

DEFINICIÓN

Los **datos de proceso** son datos dispuestos de acuerdo con una secuencia temporal. Son medidas de una característica de bienes o servicios que resultan de alguna combinación de equipo, personas, materiales, métodos y condiciones.

EJEMPLO 1 Errores de los altímetros de avión como datos de proceso

La tabla 14-1 incluye datos de proceso que consisten en los errores medidos (en pies) de altímetros para avión. Debido a que los valores de la tabla 14-1 están organizados de acuerdo con la hora en que fueron seleccionados, son datos de proceso.

Es importante saber que algunas empresas se han declarado en quiebra porque permitieron que sus procesos de fabricación se deterioraran sin un monitoreo constante. Esta sección presenta tres herramientas comúnmente usadas para monitorear los datos de proceso: gráficas de rachas, gráficas R y gráficas $\bar{x}$. Comenzamos con las gráficas de rachas.

Gráfica de rachas

Una gráfica de rachas es una de varias herramientas que se usan con frecuencia para monitorear un proceso con el fin de asegurar que las características deseadas no cambien. Una gráfica de rachas es básicamente igual a una gráfica de series de tiempo, que se estudió en la sección 2-3.

DEFINICIÓN

Una **gráfica de rachas** representa una secuencia de valores de datos *individuales* a través del tiempo. Un eje (generalmente el eje vertical) se utiliza para los valores de datos, y el otro eje (usualmente el eje horizontal) se usa para la secuencia de tiempo.

EJEMPLO 2 Gráfica de rachas para los errores en altímetros de avión

Trate los 100 errores en altímetros de avión de la tabla 14-1 como una cadena de 100 mediciones consecutivas y elabore una gráfica de rachas usando un eje vertical para los errores y un eje horizontal para identificar el orden cronológico de dichos errores.

SOLUCIÓN

En la figura 14-1 se muestra la gráfica de rachas generada por Minitab para los datos de la tabla 14-1. La escala horizontal identifica el número de muestra, por lo que el número 1 corresponde al primer error del altímetro, el número 2 corresponde al segundo error, y así sucesivamente. La escala vertical representa los errores medidos.

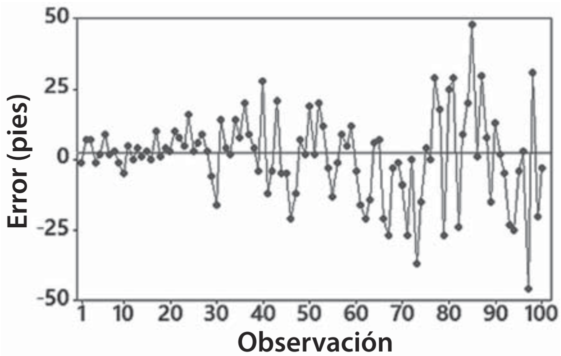


FIGURA 14-1 Gráfica de rachas para los altímetros de avión

INTERPRETACIÓN

Examine la figura 14-1 e intente identificar cualquier *patrón*. Vemos que a medida que el tiempo avanza de izquierda a derecha, los puntos parecen mostrar una *mayor variación*. Si este patrón continúa, algunos errores se volverán inaceptablemente grandes. El patrón de incremento de la variación es un problema clásico en el control de calidad, y si no se reconoce, las empresas pueden llegar a desaparecer.

SU TURNO Resuelva el ejercicio 6 “Latas de Pepsi: Gráfica de rachas”.

DEFINICIÓN

Un proceso es **estadísticamente estable** (o está **bajo control estadístico**) si sólo tiene variación natural, sin patrones, ciclos o puntos inusuales.

Interpretación de las gráficas de rachas

Una gráfica de rachas sin un patrón obvio sugiere que los datos provienen de un proceso que es *estadísticamente estable*, y los datos pueden tratarse como si provinieran de una población con una media, una desviación estándar, una distribución y otras características constantes. En la figura 14-1 se muestra un patrón de variación creciente, y ése es uno de varios criterios para determinar que un proceso no es *estadísticamente estable* (o está fuera de control estadístico). La violación de *uno o más* de los siguientes criterios indica que un proceso no es estadísticamente estable o está fuera de control estadístico.

El efecto Flynn: Tendencia al alza en las puntuaciones de IQ

Un IQ (cociente de inteligencia; IQ, por sus siglas en inglés) se mide a partir de pruebas estándar de inteligencia. Una gráfica de rachas o una gráfica de control de las puntuaciones de IQ revelarían que éstas exhiben una tendencia a incrementarse, ya que las puntuaciones de IQ han aumentado de forma estable desde que comenzaron a utilizarse aproximadamente en 1930. Esta tendencia es mundial y es igual en los distintos tipos de pruebas de inteligencia, incluso en aquellas que se basan casi por completo en el razonamiento abstracto y no verbal, con mínima influencia de la cultura. Esta tendencia al alza se conoce como *efecto Flynn*, ya que el científico político James R. Flynn descubrió esta tendencia en sus estudios con reclutas del ejército estadounidense. La cantidad del incremento es muy sustancial: con base en la puntuación media del IQ de 100, se estima que esta era de aproximadamente 77 en 1920. Por lo tanto, el estudiante común actual es brillante, si se le compara con sus bisabuelos. Hasta ahora no está claro si la tendencia ascendente en las puntuaciones de IQ indica una población cada vez más inteligente o si hay problemas con los métodos utilizados para la prueba de cociente intelectual.

Mejora de la calidad de los autos al reducir la variación

Ford y Mazda estaban produciendo transmisiones similares que se suponía debían hacerse con las mismas especificaciones, pero pronto se hizo evidente que las transmisiones de Ford requerían muchas más reparaciones de garantía que las transmisiones de Mazda fabricadas en Japón. Los investigadores de Ford buscaron la causa de esto y descubrieron que sus transmisiones cumplían con las especificaciones requeridas, pero la *variación* en las transmisiones de Ford era mucho mayor que las de Mazda. Mazda estaba utilizando una mejor esmeriladora y más cara, pero el mayor costo se compensaba con menos reparaciones bajo garantía. Armado con estos importantes resultados, Ford realizó cambios y procedió no sólo a cumplir con las especificaciones requeridas sino también a mejorar la calidad reduciendo la variación. (Vea *Taguchi Techniques for Quality Engineering* de Phillip J. Ross).

- **Aumento de la variación:** A medida que avanza la gráfica de rachas de izquierda a derecha, la variación vertical de los puntos aumenta, por lo que los valores de datos correspondientes experimentan un aumento en la variación. (Vea la figura 14-1). 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 se consideran defectuosos.
- **Tendencia al alza:** Los puntos suben al avanzar de izquierda a derecha, por lo que los valores correspondientes aumentan con el tiempo.
- **Distancia hacia abajo:** Los puntos caen al avanzar de izquierda a derecha, por lo que los valores correspondientes disminuyen con el tiempo.
- **Cambio hacia arriba:** Los puntos cercanos al inicio son notablemente más bajos que los que están cerca del final, por lo que los valores correspondientes han cambiado hacia arriba.
- **Cambio hacia abajo:** Los puntos cercanos al inicio son notablemente más altos que los que están cerca del final, por lo que los valores correspondientes han cambiado hacia abajo.
- **Valor excepcional:** Hay un solo punto que es excepcionalmente alto o bajo.
- **Patrón cíclico:** Hay un ciclo que se repite.

Causas de la variación

Muchos métodos diferentes de control de calidad intentan reducir la variación en los productos o servicios. La variación en un proceso puede ser el resultado de dos tipos de causas como se define a continuación.

DEFINICIÓN

La **variación aleatoria** se debe al azar; es el tipo de variación inherente en cualquier proceso que no es capaz de producir cada bien o servicio exactamente de la misma manera en cada ocasión.

La **variación asignable** es el resultado de causas que pueden identificarse (como maquinaria defectuosa o empleados no capacitados).

Más adelante en el 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 útiles para monitorear la estabilidad de un proceso.

Gráficas de control

Debido a que Walter Shewhart introdujo las tablas de control en 1924, a veces se les denomina gráficas de Shewhart. Comenzamos con una definición básica.

DEFINICIÓN

Una **gráfica de control** (o **gráfica de Shewhart** o **gráfica de comportamiento del proceso**) de una característica del proceso (como la media o la variación) consiste en valores representados 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 de en medio representa un valor central de las medidas características, mientras que los límites de control son las fronteras utilizadas para separar e identificar cualquier punto que se considere *significativamente alto o significativamente bajo*.

Supondremos que la desviación estándar poblacional σ no se conoce puesto que ahora consideramos dos de varios tipos diferentes de *gráficas de control*:

1. Gráficas R (o gráficas de rangos) utilizadas para monitorear la variación.
2. Gráficas $\overline{x}$ utilizadas para monitorear las medias.

Cuando se usan gráficas de control para monitorear un proceso, es común considerar juntas las gráficas R y $\overline{x}$, porque un proceso estadísticamente inestable puede ser el resultado de una *variación* creciente, un cambio en las *medias* o ambos.

Gráfica de control para el monitoreo de la variación: La gráfica *R*

Una **gráfica *R*** (o **gráfica de rango**) es un diagrama donde se representan rangos muestrales en vez de valores muestrales individuales, y se usa para monitorear la *variación* en un proceso. Puede tener más sentido usar desviaciones estándar, pero las gráficas de rango son bastante efectivas para los casos donde el tamaño de las muestras (o subgrupos) es de 10 o menos. Si todas las muestras tienen un tamaño superior a 10, se recomienda el uso de una gráfica s en lugar de una R. (Vea el ejercicio 13). Además de graficar los valores de los rangos, incluimos una línea central ubicada en $\overline{R}$, que representa la media de todos los rangos muestrales, así como otra línea para indicar el límite de control inferior y una tercera línea para marcar el límite de control superior. A continuación se presenta un resumen de la notación y los componentes de la gráfica R.

ELEMENTOS CLAVE

Monitoreo de la variación del proceso: Gráfica de control para *R*

Objetivo

Elaborar una gráfica de control para R (o una "gráfica R") que se pueda usar para determinar si la *variación* de los datos del proceso está bajo control estadístico.

Requisitos

1. Los datos son datos del proceso que consisten en una secuencia de muestras del mismo tamaño n.
2. La distribución de los datos del proceso es esencialmente normal.
3. Los valores de datos muestrales individuales son independientes.

Notación

n = tamaño de cada muestra o *subgrupo*

$\overline{R}$ = media de los rangos muestrales (la suma de los rangos muestrales dividida por el número de muestras)

Gráfica

Puntos graficados: rangos muestrales (cada punto representa el rango para cada subgrupo)

Línea central: $\overline{R}$ (la media de los rangos muestrales)

Límite de control superior (LCS): $D_4\overline{R}$ (donde D_4 es una constante que se encuentra en la tabla 14-2)

Límite de control inferior (LCI): $D_3\overline{R}$ (donde D_3 es una constante que se encuentra en la tabla 14-2)

TABLA 14-2 Constantes de la gráfica de control

n: Número de observaciones en el subgrupo	Gráfica *R*		Gráfica $\bar{x}$		Gráfica *s*	
	D_3	D_4	A_2	A_3	B_3	B_4
2	0.000	3.267	1.880	2.659	0.000	3.267
3	0.000	2.574	1.023	1.954	0.000	2.568
4	0.000	2.282	0.729	1.628	0.000	2.266
5	**0.000**	**2.114**	**0.577**	**1.427**	**0.000**	**2.089**
6	0.000	2.004	0.483	1.287	0.030	1.970
7	0.076	1.924	0.419	1.182	0.118	1.882
8	0.136	1.864	0.373	1.099	0.185	1.815
9	0.184	1.816	0.337	1.032	0.239	1.761
10	0.223	1.777	0.308	0.975	0.284	1.716

Fuente: Adaptada del *Manual ASTM sobre la Presentación de Datos y el Análisis de las Gráficas de Control*,© 1976 ASTM, pp. 134-136. Reproducida con autorización de la Sociedad Estadounidense de Pruebas y Materiales.

Variación asignable costosa

El Mars Climate Orbiter fue lanzado por la NASA hacia Marte, pero se destruyó al volar demasiado cerca de su planeta de 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 estaban en las unidades inglesas de libras fuerza, pero el Laboratorio de Propulsión supuso que las unidades utilizadas eran "newtons" métricos en vez de libras. Quienes dirigían la nave espacial dieron consecutivamente cantidades erróneas de la fuerza para ajustar la posición de la nave. Los errores causados por la discrepancia fueron muy pequeños al principio, pero el error acumulado a lo largo de los meses de travesía de la nave espacial fue la causa de su fracaso.

En 1962 la nave espacial que transportaba al satélite *Mariner I* fue destruida por controladores en la Tierra, cuando se salió de curso debido a la falta de un signo menos en un programa de cómputo.

Los valores de D_4 y D_3 son constantes calculadas por expertos en control de calidad, y están destinadas a simplificar los cálculos. Los límites de control superior e inferior de $D_4\bar{R}$ y $D_3\bar{R}$ son valores aproximadamente equivalentes a los límites del intervalo de confianza del 99.7%. Por lo tanto, es muy poco probable que los valores de un proceso estadísticamente estable caigan fuera de esos límites. Si un valor cae más allá de los límites de control, es muy probable que el proceso no sea estadísticamente estable.

EJEMPLO 3 Gráfica *R* de errores en altímetros

Elabore una gráfica de control para *R* usando los errores en los altímetros que se presentan en la tabla 14-1. Use las muestras de tamaño $n = 5$ para cada uno de los 20 días de producción.

SOLUCIÓN

Consulte la tabla 14-1 en el problema del capítulo de la página 655 para ver la columna de rangos muestrales *R*. El valor de $\bar{R}$ es la media de esos 20 rangos muestrales, por lo que su valor se encuentra de la siguiente manera:

$$\bar{R} = \frac{8 + 14 + \cdots + 77}{20} = 30.65 \text{ pies}$$

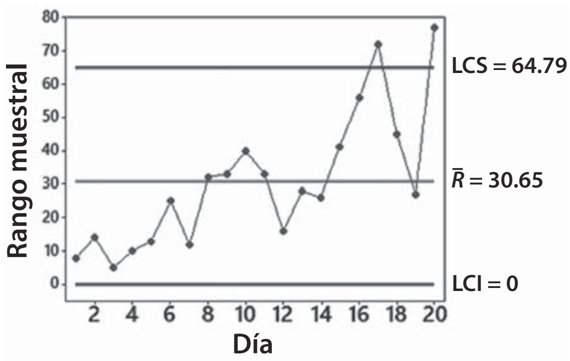
Por lo tanto, la línea central de nuestra gráfica *R* se ubica en $\bar{R} = 30.65$ pies. Para encontrar los límites de control superior e inferior, primero debemos encontrar los valores de D_3 y D_4. Con referencia a la tabla 14-2 para $n = 5$, obtenemos $D_4 = 2.114$ y $D_3 = 0.000$, por lo que los límites de control son como siguen:

$$\text{Límite de control superior (LCS): } D_4\bar{R} = (2.114)(30.65) = 64.79$$

$$\text{Límite de control inferior (LCI): } D_3\bar{R} = (0.000)(30.65) = 0.0000$$

Utilizando un valor de línea central de $\bar{R} = 30.65$ y límites de control de 64.79 y 0.0000, ahora procedemos a graficar los 20 rangos muestrales como 20 puntos individuales. El resultado se muestra en la página siguiente en la pantalla de Minitab de la gráfica *R*.

Gráfica *R*



SU TURNO Resuelva el ejercicio 7 “Latas de Pepsi: Gráfica *R*”.

Interpretación de gráficas de control

Al interpretar las gráficas de control, es importante tener la siguiente precaución:

PRECAUCIÓN 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 no corresponden a ninguna *especificación* de proceso que haya sido decretada por el fabricante.

Cuando se investiga la calidad de algún proceso, por lo general se deben considerar dos preguntas clave:

1. Con base en el comportamiento actual del proceso, ¿podemos concluir que éste se encuentra bajo control estadístico?
2. ¿Los productos o servicios del proceso cumplen con las especificaciones de diseño?

En el presente capítulo abordamos la primera pregunta, pero no la segunda; nos estamos enfocando en el comportamiento del proceso con el objetivo de determinar si éste se encuentra bajo control estadístico. Además, debemos comprender claramente los siguientes criterios específicos para determinar si un proceso está bajo control estadístico (o si es estadísticamente estable).

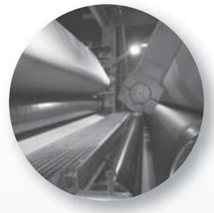
Criterios para determinar si un proceso está fuera de control

DEFINICIÓN

Un proceso **no es estadísticamente estable** o está **fuera de control estadístico** si se cumplen uno o más de los siguientes criterios para determinar si un proceso está fuera de control.

1. Hay un patrón, tendencia o ciclo que obviamente no es aleatorio.
2. Hay al menos un punto por encima del límite de control superior o al menos un punto por debajo del límite de control inferior.
3. *Regla de la racha de 8:* Hay al menos ocho puntos consecutivos por encima o por debajo de la línea central. (Con un proceso estadísticamente estable, hay una probabilidad de 0.5 de que un punto esté por encima o por debajo de la línea central, por lo que es muy poco probable que ocho puntos consecutivos estén todos sobre la línea central o debajo de ella).

¡No exageres!



La empresa Nashua Corp. tuvo problemas con su máquina para recubrimiento de papel y consideró gastar millones de dólares con la finalidad de reemplazarla. La máquina estaba funcionando bien y con un proceso estable, pero las muestras se empezaron a tomar con mucha frecuencia; con base en esos resultados, se le hicieron ajustes. Tales ajustes excesivos, denominados *alteración indebida*, causaron desviaciones de la distribución que hasta entonces había sido buena. El efecto fue un incremento en los defectos. Cuando el experto en estadística y control W. Edwards Deming estudió el proceso, recomendó que no se le hicieran ajustes, a menos que hubiera una señal de que el proceso había cambiado o se había vuelto inestable. La compañía funcionó mejor sin ajustes que con la alteración realizada.

Detección de sobornos con gráficas de control

Las gráficas de control se utilizaron para enviar a prisión a una persona que sobornaba a jugadores de jai alai en Florida para que perdieran. (Vea "Using Control Charts to Corroborate Bribery in Jai Alai", de Charnes y Gitlow, *The American Statistician*, vol. 49, núm. 4). El auditor de una cancha de jai alai notó que cantidades anormalmente grandes de dinero se jugaban en ciertos tipos de apuestas, y que algunos participantes no ganaban tanto como se esperaba cuando se realizaban tales apuestas. Se utilizaron gráficas R y $\bar{x}$ en la Corte como evidencia de patrones sumamente inusuales de apuestas. El estudio de las gráficas de control señala 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 especialista en estadística fue capaz de identificar una fecha en la cual la variación asignable parecía detenerse, y los fiscales supieron que se trataba de la fecha del arresto del sospechoso.

En este libro utilizaremos sólo los tres criterios mencionados anteriormente, pero algunas compañías usan criterios adicionales como los siguientes:

- Hay al menos seis puntos consecutivos, todos creciendo o decreciendo.
- Hay al menos 14 puntos consecutivos que alternan entre arriba y abajo (es decir, arriba, abajo, arriba, abajo, etcétera).
- Dos de tres puntos consecutivos se salen de los límites de control que están a 2 desviaciones estándar de la línea central.
- Cuatro de cinco puntos consecutivos se salen de los límites de control que están a 1 desviación estándar de la línea central.

EJEMPLO 4 Interpretación de la tabla *R* para errores en los altímetros

Examine la gráfica R que se muestra en la pantalla para el ejemplo 3 y determine si la variación del proceso está bajo control estadístico.

SOLUCIÓN

Podemos interpretar las gráficas de control para R aplicando los tres criterios anteriores para determinar si un proceso está fuera de control. Al aplicar los tres criterios a la pantalla de la gráfica R, concluimos que la variación en este proceso *no* está bajo control estadístico. Considerando los tres criterios utilizados, vemos que los primeros dos criterios para estar fuera de control se satisfacen: (1) Existe un patrón obvio de puntos crecientes. (2) Hay dos puntos por encima del límite de control superior. El tercer criterio no se cumple porque no hay ocho puntos consecutivos que estén por encima o por debajo de la línea central.

INTERPRETACIÓN

Concluimos que la *variación* (no necesariamente la media) del proceso está fuera de control estadístico.

SU TURNO Resuelva el ejercicio 7 "Latas de Pepsi: Gráfica *R*".

Gráfica de control para monitorear medias: La gráfica $\bar{x}$

Una **gráfica $\bar{x}$** es una representación de las medias muestrales, y se usa para monitorear el *centro* en un proceso. Además de representar las medias muestrales, incluimos una línea central ubicada en $\bar{\bar{x}}$, que expresa la media de todas las medias muestrales (igual a la media de todos los valores muestrales combinados), así como otra línea para el límite de control inferior y una tercera línea para el límite de control superior. Usando el método común en los negocios y la industria, los límites de control se basan en rangos en lugar de en las desviaciones estándar. (Consulte el ejercicio 14 para obtener una gráfica $\bar{x}$ basada en las desviaciones estándar).

ELEMENTOS CLAVE

Monitoreo de la media del proceso: Gráfica de control para $\bar{x}$

Objetivo

Elaborar una gráfica de control para $\bar{x}$ (o una gráfica $\bar{x}$) que pueda usarse para determinar si el *centro* de los datos del proceso está bajo control estadístico.

Requisitos

1. Los datos son datos de proceso que consisten en una secuencia de muestras del mismo tamaño n.
2. La distribución de los datos del proceso es esencialmente normal.
3. Los valores de datos muestrales individuales son independientes.

Notación

n = tamaño de cada muestra o *subgrupo*

$\bar{\bar{x}}$ = media de todas las medias muestrales (igual a la media de todos los valores muestrales combinados)

Gráfica

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 es una constante que se encuentra en la tabla 14-2)

Límite de control inferior (LCI): $\bar{\bar{x}} - A_2\bar{R}$ (donde A_2 es una constante que se encuentra en la tabla 14-2)

EJEMPLO 5 Gráfica $\bar{x}$ de errores en los altímetros

A partir de los errores en los altímetros de la tabla 14-1 en la página 655, con muestras de tamaño $n = 5$ para cada uno de 20 días, elabore una gráfica de control para $\bar{x}$. Con base en la gráfica de control para $\bar{x}$, determine si el proceso está bajo control estadístico.

SOLUCIÓN

Antes de graficar los 20 puntos correspondientes a los 20 valores de $\bar{x}$, primero debemos encontrar los valores de la línea central y los límites de control. Obtenemos

$$\bar{\bar{x}} = \frac{2.8 + 1.6 + \cdots - 7.0}{20} = 1.19 \text{ pies}$$

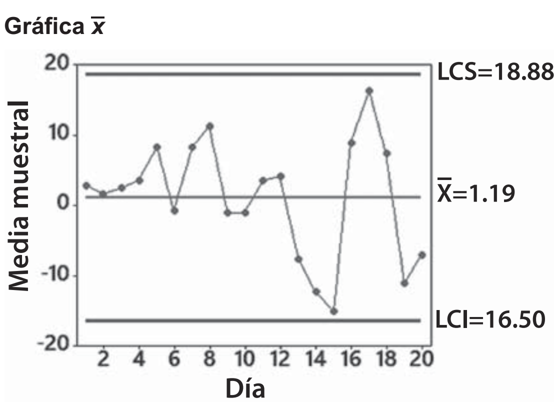
$$\bar{R} = \frac{8 + 14 + \cdots + 77}{20} = 30.65 \text{ pies}$$

Con referencia a la tabla 14-2 de la página 660, encontramos que para $n = 5$, $A_2 = 0.577$. Dado que ya se conocen los valores de $\bar{\bar{x}}$, A_2 y $\bar{R}$, ahora podemos evaluar los límites de control.

$$\text{Límite de control superior (LCS): } \bar{\bar{x}} + A_2\bar{R} = 1.19 + (0.577)(30.65) = 18.88 \text{ pies}$$

$$\text{Límite de control inferior (LCI) } \bar{\bar{x}} - A_2\bar{R} = 1.19 - (0.577)(30.65) = -16.50 \text{ pies}$$

La gráfica de control resultante para $\bar{x}$ será como se muestra en la pantalla de Minitab adjunta.



continúa

El alto costo de la baja calidad

Microsoft anunció que gastaría 1,150 millones de dólares para reparar las consolas del juego Xbox 360. Analistas de la industria estiman que alrededor de una tercera parte de las unidades están defectuosas y necesitan reparación. Muchos usuarios reportaron que la máquina se apagaba después de que tres luces rojas intermitentes aparecían en la consola.

Por otro lado, la Federal Drug Administration (FDA) recientemente llegó a un acuerdo en el que una compañía farmacéutica, Schering-Plough Corporation, pagaría la cantidad récord de 500 millones de dólares por no lograr corregir problemas en la producción de fármacos. Según un artículo del *New York Times*, de Melody Petersen, “Algunos de los problemas se relacionan con la falta de controles que identifican medicamentos defectuosos, mientras otros provienen de equipos demasiado viejos. Tales problemas están relacionados con unos 200 medicamentos, incluyendo el Claritin, el fármaco contra alergias que es el producto más vendido de Schering”.

INTERPRETACIÓN

Un examen de la gráfica $\bar{x}$ muestra que la media del proceso está fuera del control estadístico porque se cumple al menos uno de los tres criterios para estar fuera de control. Específicamente, se cumple el primer criterio porque hay un patrón de variación creciente, que también se identificó mediante la gráfica R.

SU TURNO Resuelva el ejercicio 8 “Latas de Pepsi: Gráfica $\bar{x}$”.

Al analizar los errores en los altímetros de la tabla 14-1 mediante una gráfica de rachas, una gráfica R y una gráfica $\bar{x}$, podemos ver que el proceso está fuera del control estadístico. Si el proceso de fabricación continúa sin correcciones, muchos de los altímetros no cumplirán con las reglamentaciones de la FAA y, en consecuencia, la compañía Orange Avionics sufrirá enormes pérdidas.

CENTRO DE TECNOLOGÍA

Gráficas de control

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Visite www.pearsonenespañol.com/triola para obtener instrucciones detalladas.

Minitab

Gráfica de rachas

1. Ingrese todos los datos muestrales en la columna *C1*.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Quality Tools** en el menú desplegable y seleccione **Run Chart** en el submenú.
4. En el cuadro de *columna única*, ingrese **C1**. En el tamaño de *subgrupo*, ingrese **1**.
5. Haga clic en **OK**.

Gráfica *R* y gráfica $\bar{x}$

1. Ingrese los valores muestrales secuencialmente en la columna *C1* o introdúzcalos en columnas o filas como en la tabla 14-1.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Control Charts** en el menú desplegable y seleccione **Variable Charts for Subgroups**
4. Seleccione **Xbar-R** del submenú.
5. Seleccione el formato de los datos ingresados en el paso 1 y luego seleccione las columnas adecuadas en el cuadro de entrada.
6. Ingrese el *tamaño del subgrupo* si todas las observaciones están en una columna.
7. Haga clic en el botón de **Xbar-R Options** y luego haga clic en la pestaña **Estimate** y seleccione **Rbar**.
8. Haga clic en **OK** dos veces.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Control Charts** en el menú desplegable y **X-bar**, **R** en el submenú.
3. Seleccione las columnas que se utilizarán.
4. Haga clic en **Compute!**
5. Haga clic en los botones de flecha para alternar entre las gráficas y los resultados numéricos.

Calculadora TI-83/84 Plus

La creación de gráficas de rachas, gráficas *R* y gráficas con una calculadora TI-83/84 Plus es posible pero no se recomienda debido a los complejos procedimientos requeridos y la información limitada que se muestra. Visite www.pearsonenespañol.com/triola para obtener instrucciones detalladas.

Excel

Complemento XLSTAT (*no disponible con todas las licencias de XLSTAT*)

1. Haga clic en la pestaña **XLSTAT** de la cinta de opciones y luego haga clic en **SPC**.
2. Seleccione **Subgroup charts** en el menú desplegable.
3. Para la *familia de gráficas*, haga clic en **Subgroup charts** y seleccione la **X-bar-R chart** en el *tipo de gráfico*.
4. Haga clic en la pestaña **General**.
5. En *formato de datos*, seleccione **Column** si los datos se listan en columnas separadas o **One column** si los datos están apilados en una sola columna.
6. Si se selecciona **One column**, ingrese el tamaño de muestra común a todas las entradas de datos en el *tamaño común de subgrupo*.
7. En el cuadro de *datos*, ingrese el rango de celdas de los datos. Si el rango de datos incluye una etiqueta, marque la casilla de **Column labels**.
8. Haga clic en la pestaña **Estimation** y seleccione **R bar**.
9. Haga clic en **OK** dos veces.

14-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

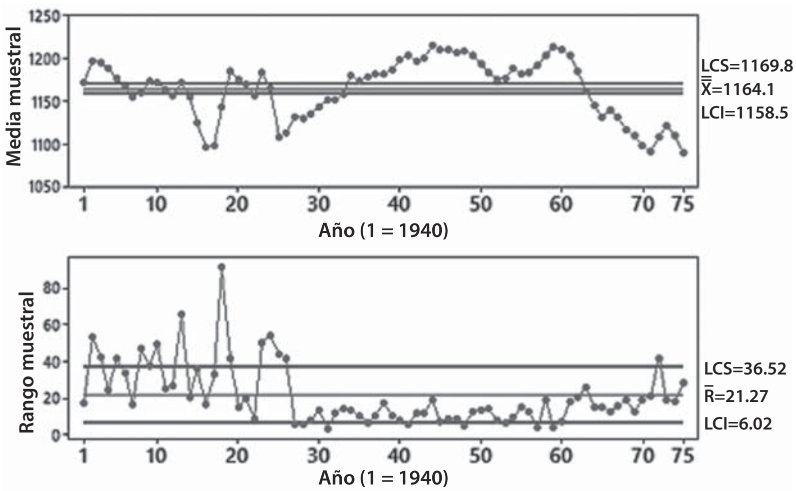
1. Requisito de la FAA La tabla 14-1 de la página 655 lista datos de proceso que consisten en los errores (en pies) de los altímetros para avión cuando se prueban para una altitud de 2000 pies, y la Administración Federal de Aviación (FAA por sus siglas en inglés) requiere que los errores sean como máximo de ± 30 pies. Si las gráficas de control $\overline{x}$ y R muestran que el proceso de fabricación de altímetros está bajo control estadístico, ¿eso indica que los altímetros tienen errores que son como máximo de ± 30 pies? ¿Por qué sí o por qué no?

2. Notación y terminología Considere los datos de proceso que consisten en las cantidades (onzas) de bebida de cola en latas de Coca-Cola regular seleccionadas al azar. El proceso debe ser monitoreado mediante gráficas de control $\overline{x}$ y R basadas en muestras de 50 latas seleccionadas al azar cada día durante 20 días consecutivos de producción. ¿Qué son las gráficas de control $\overline{x}$ y R, y qué denotan $\overline{\overline{x}}$, $\overline{R}$, LCS y LCI?

3. Elevaciones del lago Mead Muchas personas en Nevada, Arizona y California obtienen agua y electricidad del lago Mead y de la presa Hoover. En el ejercicio 4 se muestra una gráfica $\overline{x}$ (arriba) y una gráfica R (abajo) obtenidas mediante el uso de las elevaciones mensuales (pies) del lago Mead en la Presa Hoover (según los datos del Departamento del Interior de Estados Unidos). Las gráficas de control se basan en las 12 elevaciones mensuales para cada uno de los 75 años consecutivos y recientes. ¿Qué nos dice la gráfica $\overline{x}$ sobre el lago Mead?

4. Elevaciones del lago Mead ¿Qué nos dice la gráfica R sobre el lago Mead?

Minitab



Latas de Pepsi. ***En los ejercicios 5 a 8, consulte las cargas axiales (libras) en latas de Pepsi hechas de aluminio con un espesor de 0.0109 pulg., como se detalla en el conjunto de datos 30 "Latas de aluminio" del apéndice B. Una carga axial de una lata es el peso máximo admisible en su costado, y es importante tener una carga axial suficientemente alta para que la lata no se aplaste al presionar la tapa superior. Hay siete mediciones en cada uno de 25 días de producción. Si las 175 cargas axiales se dan en una columna, las primeras 7 son del primer día, las 7 siguientes son del segundo día, y así sucesivamente; de modo que el "tamaño del subgrupo" es 7.***

5. Latas de Pepsi: Notación Después de encontrar la media muestral y el rango muestral para cada uno de los 25 días, encuentre los valores de $\overline{\overline{x}}$ y $\overline{R}$. También determine los valores del LCI y el LCS para una gráfica R, después encuentre los valores del LCI y el LCS para una gráfica $\overline{x}$.

6. Latas de Pepsi: Grafica de rachas Trate las 175 cargas axiales como una cadena de medidas consecutivas y elabore un diagrama de rachas. ¿Qué sugiere el resultado?

7. Latas de Pepsi: Gráfica *R* Trate las siete mediciones de cada día como una muestra y elabore una gráfica R. ¿Qué sugiere el resultado?

8. Latas de Pepsi: Gráfica $\overline{x}$ Trate las siete mediciones de cada día como una muestra y elabore una gráfica. ¿Qué sugiere el resultado?

Monedas de ¢25. ***En los ejercicios 9 a 12, consulte la tabla adjunta de pesos (en gramos) de monedas de ¢25 acuñadas por el gobierno de Estados Unidos. Esta tabla está disponible para su descarga en www.pearsonenespañol.com/triola***

Día	Hora 1	Hora 2	Hora 3	Hora 4	Hora 5	$\bar{x}$	s	Rango
1	5.543	5.698	5.605	5.653	5.668	5.6334	0.0607	0.155
2	5.585	5.692	5.771	5.718	5.720	5.6972	0.0689	0.186
3	5.752	5.636	5.660	5.680	5.565	5.6586	0.0679	0.187
4	5.697	5.613	5.575	5.615	5.646	5.6292	0.0455	0.122
5	5.630	5.770	5.713	5.649	5.650	5.6824	0.0581	0.140
6	5.807	5.647	5.756	5.677	5.761	5.7296	0.0657	0.160
7	5.686	5.691	5.715	5.748	5.688	5.7056	0.0264	0.062
8	5.681	5.699	5.767	5.736	5.752	5.7270	0.0361	0.086
9	5.552	5.659	5.770	5.594	5.607	5.6364	0.0839	0.218
10	5.818	5.655	5.660	5.662	5.700	5.6990	0.0689	0.163
11	5.693	5.692	5.625	5.750	5.757	5.7034	0.0535	0.132
12	5.637	5.628	5.646	5.667	5.603	5.6362	0.0235	0.064
13	5.634	5.778	5.638	5.689	5.702	5.6882	0.0586	0.144
14	5.664	5.655	5.727	5.637	5.667	5.6700	0.0339	0.090
15	5.664	5.695	5.677	5.689	5.757	5.6964	0.0359	0.093
16	5.707	5.890	5.598	5.724	5.635	5.7108	0.1127	0.292
17	5.697	5.593	5.780	5.745	5.470	5.6570	0.1260	0.310
18	6.002	5.898	5.669	5.957	5.583	5.8218	0.1850	0.419
19	6.017	5.613	5.596	5.534	5.795	5.7110	0.1968	0.483
20	5.671	6.223	5.621	5.783	5.787	5.8170	0.2380	0.602

9. Monedas de ¢25: Notación Encuentre los valores de $\bar{\bar{x}}$ y $\bar{R}$. También determine los valores del LCI y el LCS para una gráfica R, luego encuentre los valores del LCI y LCS para una gráfica .

10. Monedas de ¢25: Gráfica *R* Trate las cinco mediciones de cada día como una muestra y elabore una tabla R. ¿Qué sugiere el resultado?

11. Monedas de ¢25: Gráfica $\bar{x}$ Trate las cinco mediciones de cada día como una muestra y elabore una gráfica $\bar{x}$. ¿Qué sugiere el resultado?

12. Monedas de ¢25: Gráfica de rachas Trate las 100 mediciones consecutivas de los 20 días como valores individuales y elabore una gráfica de rachas. ¿Qué sugiere el resultado?

14-1 Más allá de lo básico

13. Gráfica *s* En esta sección describimos las gráficas de control para R y $\bar{x}$ basadas en *rangos*. Las gráficas de control para monitorear la variación y el centro (la media) también se pueden basar en las *desviaciones estándar*. Una *gráfica s* para monitorear la variación se traza al graficar desviaciones estándar muestrales con una línea central en $\bar{s}$ (la media de las desviaciones estándar muestrales) y límites de control en $B_4\bar{s}$ y $B_3\bar{s}$, donde B_4 y B_3 se encuentran en la tabla 14-2 en la página 660 de esta sección. Elabore una gráfica para los datos de la tabla 14-1 en la página 655. Compare el resultado con la gráfica R del ejemplo 3 "Gráfica R de errores en altímetros".

14. Gráfica $\bar{x}$ basada en desviaciones estándar Una gráfica $\bar{x}$ basada en *desviaciones estándar* (y no en rangos) se construye graficando las medias muestrales con una línea central en $\bar{\bar{x}}$ y límites de control en $\bar{\bar{x}} + A_3\bar{s}$ y $\bar{\bar{x}} - A_3\bar{s}$, donde A_3 se encuentra en la tabla 14-2 de la página 660 y $\bar{s}$ es la media de las desviaciones estándar muestrales. Use los datos en la tabla 14-1 de la página 655 para elaborar una gráfica $\bar{x}$ basada en desviaciones estándar. Compare el resultado con la gráfica $\bar{x}$ en función de los rangos muestrales del ejemplo 5 "Gráfica $\bar{x}$ de errores en altímetros".

14-2 Gráficas de control para atributos

Concepto clave Esta sección presenta un método para elaborar una gráfica de control con el fin de monitorear la proporción p para algún *atributo*; por ejemplo, si un servicio o producto manufacturado es defectuoso o discordante. (Un bien o servicio es discordante si no cumple con las especificaciones o los requisitos dados. En ocasiones, los productos discordantes se descartan, se reparan o se "recuperan" y se venden a precios bajos). La gráfica de control se interpreta utilizando los mismos tres criterios dados en la sección 14-1 para determinar si el proceso es estadísticamente estable:

Criterios para determinar si un proceso está fuera de control

DEFINICIÓN

Un proceso **no es estadísticamente estable** o está **fuera de control estadístico** si se cumplen uno o más de los siguientes criterios:

1. Hay un patrón, una tendencia o un ciclo que claramente no es aleatorio.
2. Hay al menos un punto por encima del límite de control superior o al menos un punto por debajo del límite de control inferior.
3. *Regla de la racha de 8*: Hay al menos ocho puntos consecutivos por encima o por debajo de la línea central. (Con un proceso estadísticamente estable, existe una probabilidad de 0.5 de que un punto esté por encima o por debajo de la línea central, por lo que es muy poco probable que ocho puntos consecutivos estén todos sobre la línea central o debajo de ella).

Como en la sección 14-1, seleccionamos muestras de tamaño n en intervalos de tiempo regulares y graficamos los puntos en una gráfica secuencial con una línea central y límites de control. (Hay formas de tratar con muestras de diferentes tamaños, pero no las consideraremos aquí).

DEFINICIÓN

Una **gráfica de control para *p*** (o **gráfica *p***) es un diagrama de las proporciones de algún atributo (por ejemplo, si los productos son defectuosos) graficadas secuencialmente a través del tiempo e incluye una línea central, un límite de control inferior (LCI) y un límite de control superior (LCS).

La notación y los valores de la gráfica de control se resumen en el siguiente recuadro de elementos clave. En este recuadro, el atributo "defectuoso" puede reemplazarse por cualquier otro atributo relevante (de modo que cada elemento de la muestra pertenezca a una de dos categorías distintas).

ELEMENTOS CLAVE

Monitoreo de un atributo del proceso: Gráfica de control para *p*

Objetivo

Elabore una gráfica de control para p (o una "gráfica p") que se pueda usar para determinar si la proporción de algún atributo (por ejemplo, si los productos son defectuosos) de los datos del proceso se encuentran bajo control estadístico.

Requisitos

1. Los datos son datos de proceso que consisten en una secuencia de muestras del mismo tamaño n.
2. Cada elemento de la muestra pertenece a una de dos categorías (como defectuoso o no defectuoso).
3. Los valores de datos muestrales individuales son independientes.

continúa

Notación

$$\bar{p} = \text{estimación de la proporción de artículos defectuosos en el proceso}$$

$$= \frac{\text{número total de defectos encontrados entre todos los artículos muestreados}}{\text{número total de artículos muestreados}}$$

$$\bar{q} = \text{estimación de la proporción de elementos del proceso que no son defectuosos}$$

$$= 1 - \bar{p}$$

$$n = \text{tamaño de cada muestra o subgrupo individual}$$

Gráfica

Puntos representados: proporciones de las muestras individuales de tamaño n

Línea central: $\bar{p}$

Límite de control superior: $\bar{p} + 3\sqrt{\frac{\bar{p}\,\bar{q}}{n}}$ (Use 1 si este resultado es mayor que 1).

Límite de control inferior: $\bar{p} - 3\sqrt{\frac{\bar{p}\,\bar{q}}{n}}$ (Use 0 si este resultado es menor que 0).

PRECAUCIÓN Tenga en cuenta la siguiente distinción en los cálculos: Al evaluar $\bar{p}$, divida el número total de defectos entre el número *total* de elementos muestreados. Sin embargo, los cálculos para el LCS y el LCI requieren división por n, el tamaño de cada muestra individual.

Usamos $\bar{p}$ como línea central porque es la mejor estimación de la proporción de defectos del proceso. Las expresiones para los límites de control corresponden a los límites del intervalo de confianza del 99.7% para los intervalos de confianza descritos en la sección 7-1. (La sección 7-1 no incluyó ningún intervalo de confianza del 99.7%, pero la puntuación z utilizada para éste es $z = 2.97$, redondeada a 3 en las expresiones utilizadas para el LCI y el LCS en esta sección).

EJEMPLO 1 Altímetros para avión defectuosos

El problema del capítulo describe el proceso de fabricación de altímetros para aviones. La sección 14-1 incluye ejemplos de gráficas de control para monitorear los errores en las lecturas del altímetro. Se considera que un altímetro está defectuoso si no se puede calibrar o corregir para obtener lecturas precisas (a menos de 30 pies de la altitud real de 2000 pies). La compañía Orange Avionics fabrica altímetros en lotes de 100, y cada altímetro se prueba y se determina si es aceptable o defectuoso. A continuación se listan los altímetros defectuosos en lotes sucesivos de 100. Elabore una gráfica de control para la proporción p de altímetros defectuosos y determine si el proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios de un proceso fuera de control se aplica.

Defectos: 2 0 1 3 1 2 2 4 3 5 12 7

SOLUCIÓN

La línea central para la gráfica de control se encuentra con el valor de $\overline{p}$:

$$\overline{p} = \frac{\text{número total de defectos en todas las muestras combinadas}}{\text{número total de altímetros muestreados}},$$

$$= \frac{2 + 0 + 1 + \cdots + 7}{12 \cdot 100} = \frac{42}{1200} = 0.035$$

Debido a que $\overline{p} = 0.035$, se sigue que $\overline{q} = 1 - \overline{p} = 0.965$. Mediante el uso de $\overline{p} = 0.035$, $\overline{q} = 0.965$, y $n = 100$, encontramos los límites de control de la siguiente manera:

Límite de control superior:

$$\overline{p} + 3\sqrt{\frac{\overline{p}\,\overline{q}}{n}} = 0.035 + 3\sqrt{\frac{(0.035)(0.965)}{100}} = 0.090$$

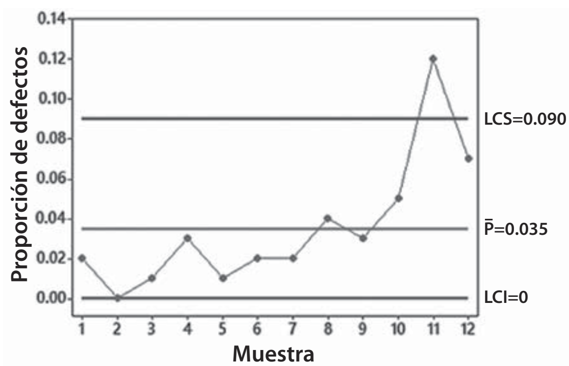
Límite de control inferior:

$$\overline{p} - 3\sqrt{\frac{\overline{p}\,\overline{q}}{n}} = 0.035 - 3\sqrt{\frac{(0.035)(0.965)}{100}} = -0.020$$

Debido a que el límite de control inferior es negativo, utilizamos en su lugar el valor de 0.

Una vez encontrados los valores para la línea central y los límites de control, podemos proceder a trazar la gráfica de control para las proporciones de altímetros defectuosos. La gráfica de control de Minitab para p se muestra en la pantalla adjunta.

Minitab



INTERPRETACIÓN

Podemos interpretar la gráfica de control para p considerando los tres criterios para un proceso fuera de control listados anteriormente en esta sección. Concluimos que el proceso está fuera del control estadístico por las siguientes razones: parece haber una tendencia al alza y hay un punto que se extiende más allá del límite de control superior. La compañía debe tomar medidas correctivas inmediatas.

SU TURNO Resuelva el ejercicio 5 "Monedas de un euro".

PRECAUCIÓN Los límites de control superior e inferior de una gráfica de control para una proporción p se basan en el comportamiento *real* del proceso, no en el comportamiento *deseado*. Los límites de control superior e inferior no están relacionados con ninguna de las *especificaciones* del proceso que pudiera haber establecido el fabricante.

Control de calidad en Perstorp

Perstorp Components, Inc. utiliza una computadora que genera automáticamente gráficas de control para monitorear el grosor del aislamiento del piso que fabrica para las camionetas Ford Ranger y Jeep Grand Cherokee. El costo de la computadora de $20,000 se compensó con los ahorros de $40,000 del primer año de operaciones, dinero que se empleaba anteriormente para generar gráficas de control de forma manual que aseguraban que el grosor del aislamiento cumpliera con las especificaciones establecidas entre 2.912 y 2.988 mm. Por medio de gráficas de control y de otros métodos de control de calidad, Perstorp redujo sus mermas en más de dos tercios.

CENTRO DE TECNOLOGÍA

Gráfica de control para *p*

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

Visite www.pearsonenespañol.com/triola para obtener instrucciones detalladas.

Minitab

1. Ingrese el número de defectos (o elementos con cualquier atributo particular) en la columna *C1*.
2. Haga clic en **Stat** en el menú superior.
3. Seleccione **Control Charts** en el menú desplegable y seleccione **Attribute Charts** y **P** del submenú.
4. En el cuadro de *variables*, ingrese la columna **C1**.
5. En los *tamaños de subgrupo* ingrese el tamaño de las muestras individuales.
6. Haga clic en **OK**.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Control Charts** en el menú desplegable y **p** del submenú.
3. Seleccione la columna que se utilizará, seleccione **Constant** e ingrese el tamaño de las muestras individuales.
4. Haga clic en **Compute!**
5. Haga clic en los botones de flecha para alternar entre la gráfica y los resultados numéricos.

Calculadora TI-83/84 Plus

No disponible.

Excel

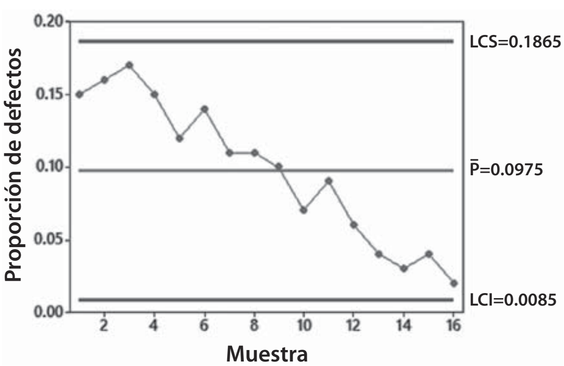
Complemento XLSTAT (*no disponible con todas las licencias XLSTAT*)

1. Haga clic en la pestaña **XLSTAT** en la cinta de opciones y luego haga clic en **SPC**.
2. Seleccione **Attribute charts** en el menú desplegable.
3. Para la *familia de gráficas*, haga clic en **Attribute charts** y seleccione **P chart** en el *tipo de gráfico*.
4. Haga clic en la pestaña **General**.
5. En el cuadro de *datos*, ingrese el rango de celdas de los datos. Si el rango de datos incluye una etiqueta, marque la casilla de **Column labels**.
6. En el cuadro del *tamaño de subgrupo común*, ingrese el tamaño de muestra que es común a todas las entradas de datos.
7. Haga clic en **OK**. Los resultados incluyen una gráfica *p*.

14-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Monedas de 25¢ Las especificaciones para una moneda de 25¢ requieren que sea 8.33% de níquel y 91.67% de cobre; debe pesar 5.670 g y tener un diámetro de 24.26 mm y un espesor de 1.75 mm; asimismo, debe tener 119 ranuras en el borde. Una moneda de 25¢ se considera defectuosa si se desvía sustancialmente de esas especificaciones. Se monitorea un proceso de producción, se registran los defectos y se obtiene la gráfica de control adjunta. ¿Este proceso parece estar bajo control estadístico? De lo contrario, identifique cualquier criterio de un proceso fuera de control que se cumpla. ¿Se está deteriorando el proceso de fabricación?



2. Notación La gráfica de control para el ejercicio 1 muestra un valor de $\overline{p} = 0.0975$. ¿Qué significa ese valor y cómo se obtiene? ¿Qué indican el LCS y el LCI?

3. Límites de control Al elaborar una gráfica de control para las proporciones de monedas de diez centavos defectuosas, se encuentra que el límite inferior de control es -0.00325. ¿Cómo se debe ajustar ese valor?

4. Monedas de un euro Después de elaborar una gráfica de control para las proporciones de monedas de un euro defectuosas, se concluye que el proceso está bajo control estadístico. ¿Esto implica que casi todas las monedas cumplen con las especificaciones deseadas? Explique.

Gráficas de control para *p*. ***En los ejercicios 5 a 12, use los datos de proceso dados para elaborar una gráfica de control para p. En cada caso, use los tres criterios presentados al principio de esta sección para determinar si el proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para un proceso fuera de control se aplica.***

5. Monedas de un euro Considere un proceso de acuñación de monedas con valor de un euro. A continuación se listan las cantidades de monedas defectuosas en lotes sucesivos de 10,000 monedas seleccionadas al azar en días consecutivos de producción.

32 21 25 19 35 34 27 30 26 33

6. Monedas de un euro Repita el ejercicio 5, suponiendo que el tamaño de cada lote es de 100 en lugar de 10,000. Compare la gráfica de control con la encontrada para el ejercicio 5. Comente sobre la calidad general del proceso de fabricación descrito en el ejercicio 5 en comparación con el proceso de fabricación descrito en este ejercicio.

7. Tabletas de aspirina Los frascos de tabletas de aspirina suelen tener etiquetas que indican que cada tableta contiene 325 mg de aspirina, pero las cantidades reales pueden variar entre 315 mg y 335 mg. Una tableta es defectuosa si tiene menos de 315 mg o más de 335 mg de aspirina. A continuación se listan los defectos encontrados en lotes de 1000 tabletas.

16 18 13 9 10 8 6 5 5 3

8. Desfibriladores Al menos una corporación fue demandada por la fabricación de desfibriladores cardíacos defectuosos. A continuación se listan las cantidades de desfibriladores defectuosos en lotes sucesivos de 1000. Elabore una gráfica de control para la proporción *p* de desfibriladores defectuosos y determine si el proceso está bajo control estadístico. Si no es así, identifique cuál de los tres criterios para un proceso fuera de control se aplica.

8 5 6 4 9 3 12 7 8 5 22 4 9 10 11 8 7 6 8 5

9. Tasa de votación En cada uno de los años recientes y consecutivos en los que hubo elecciones presidenciales, se seleccionaron aleatoriamente 1000 personas en edad de votar en Estados Unidos y se determinó el número de personas que realmente votaron, con los resultados que se listan a continuación. Comente sobre el comportamiento de la votación de la población.

631 619 608 552 536 526 531 501 551 491 513 553 568

10. Baterías de automóvil Las baterías de automóvil defectuosas son una molestia porque pueden dañar e incomodar a los conductores, incluso podrían ponerlos en peligro. Se considera que una batería de automóvil es defectuosa si falla antes de que expire la garantía. Los defectos se identifican cuando las baterías se devuelven según el programa de garantía. La corporación Powerco Battery fabrica baterías de automóvil en lotes de 250, y los números de defectos se listan a continuación para cada uno de 12 lotes consecutivos. ¿El proceso de fabricación requiere corrección?

3 4 2 5 3 6 8 9 12 14 17 20

11. Latas de bebidas de cola En cada uno de varios días consecutivos de producción de latas de refresco de cola, se prueban 50 latas y a continuación se lista el número de defectos por día. ¿Las proporciones de defectos parecen ser aceptables? ¿Qué acción se debería tomar?

8 7 9 8 10 6 5 7 9 12 9 6 8 7 9 8 11 10 9 7

12. Baterías de teléfonos inteligentes La compañía SmartBatt fabrica baterías para teléfonos inteligentes. A continuación listan las cantidades de defectos en lotes de 200 baterías seleccionadas al azar en cada uno de 12 días consecutivos de producción. ¿Qué acción se debería tomar?

5 7 4 6 3 10 10 13 4 15 4 19

14-2 Más allá de lo básico

13. Gráfica *np* Una variación de la gráfica de control para p es la **gráfica *np***, en la que se grafican los *números reales* de defectos en lugar de las *proporciones* de defectos. El gráfico np tiene una línea central con valor de $n\overline{p}$, y los límites de control tienen valores de $n\overline{p} + 3\sqrt{n\overline{p}\,\overline{q}}$ y $n\overline{p} - 3\sqrt{n\overline{p}\,\overline{q}}$. Las gráficas p y np difieren sólo en la escala de valores utilizada para el eje vertical. Elabore la gráfica np para el ejemplo 1 "Altímetros para avión defectuosos" en esta sección. Compare la gráfica np con la gráfica de control para p que se proporciona en la presente sección.

Examen rápido del capítulo

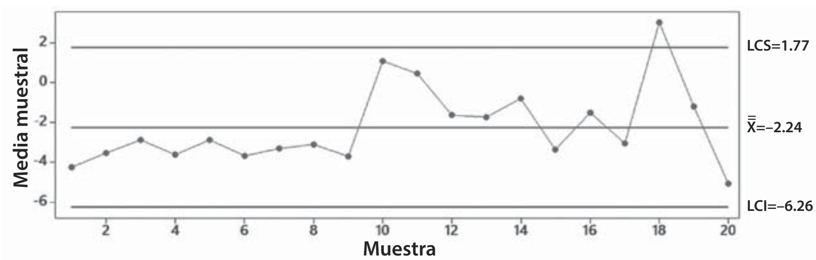
1. ¿Qué son los datos de *proceso*?

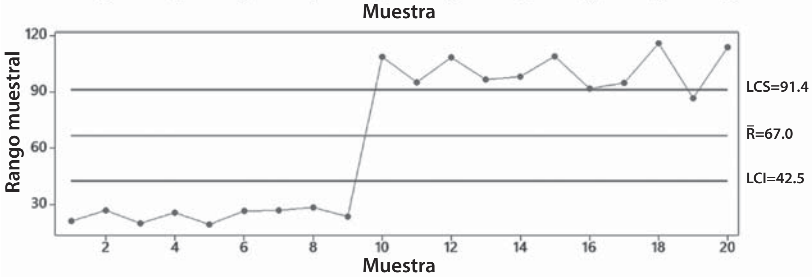
2. ¿Cuál es la diferencia entre la *variación aleatoria* y la *variación asignable*?

3. Identifique tres criterios específicos para determinar cuándo un proceso está fuera de control estadístico.

4. ¿Cuál es la diferencia entre una gráfica R y una gráfica $\overline{x}$?

En los ejercicios 5 a 8, use las siguientes dos gráficas de control que resultan de probar lotes de altímetros para avión recién fabricados, con 100 en cada lote. Los valores muestrales originales son errores (en pies) obtenidos al probar los altímetros en una cámara de presión que simula una altitud de 6000 pies. La Administración Federal de Aviación requiere un error de no más de 40 pies a esa altura.





5. ¿La variación del proceso está bajo control estadístico? ¿Por qué sí o por qué no?

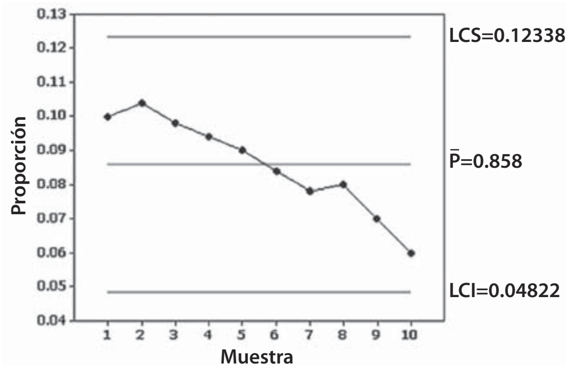
6. ¿Cuál es el valor de $\overline{R}$? En general, ¿cómo se obtiene un valor de $\overline{R}$?

7. ¿La media del proceso está bajo control estadístico? ¿Por qué sí o por qué no?

8. ¿Cuál es el valor de $\overline{\overline{x}}$? En general, ¿cómo se encuentra un valor de $\overline{\overline{x}}$?

9. Si la gráfica R y la gráfica $\overline{x}$ muestran que el proceso de fabricación de altímetros para avión está bajo control estadístico, ¿podemos concluir que los altímetros satisfacen el requisito de la Administración Federal de Aviación de tener errores no mayores a 40 pies cuando se prueban a una altitud de 6000 pies?

10. Examine la siguiente gráfica p para baterías de calculadora defectuosas y describa brevemente la acción que debe tomarse.



Ejercicios de repaso

Consumo de energía. ***Los ejercicios 1 a 5 se refieren a la cantidad de energía consumida en el hogar del autor. (La mayoría de los datos son reales, pero algunos se inventaron). Cada valor representa la energía consumida (kWh) en un periodo de dos meses. Considere que cada subgrupo consiste en las seis cantidades pertenecientes al mismo año. Los datos pueden descargarse de www.pearsonenespañol.com/triola.***

	Ene.-Feb.	Mar.-Abr.	Mayo-Jun.	Jul.-Ago.	Sept.-Oct.	Nov.-Dic.
Año 1	3637	2888	2359	3704	3432	2446
Año 2	4463	2482	2762	2288	2423	2483
Año 3	3375	2661	2073	2579	2858	2296
Año 4	2812	2433	2266	3128	3286	2749
Año 5	3427	578	3792	3348	2937	2774
Año 6	4016	3458	3395	4249	4003	3118
Año 7	5261	2946	3063	5081	2919	3360
Año 8	3853	3174	3370	4480	3710	3327

1. Consumo de energía: Notación Después de encontrar los valores de la media y el rango para cada año, encuentre los correspondientes a $\bar{\bar{x}}$ y $\bar{R}$. Después determine los valores del LCI y el LCS para una gráfica R y para una gráfica $\bar{x}$.

2. Consumo de energía: Gráfica *R* Considere que cada subgrupo consiste en los 6 valores pertenecientes a un año. Elabore 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 un proceso fuera de control conduce al rechazo de la variación estadísticamente estable.

3. Consumo de energía: Gráfica $\bar{x}$ Considere que cada subgrupo consiste en los 6 valores pertenecientes a un año. Elabore 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 un proceso fuera de control conduce al rechazo de una media estadísticamente estable.

4. Consumo de energía: Gráfica de rachas Elabore una gráfica de rachas para los 48 valores. ¿Parece haber un patrón que sugiera que el proceso no está bajo control estadístico?

5. Tiempos de servicio El restaurante Newport registra los tiempos (minutos) que transcurren antes de que se solicite a los clientes su pedido. Cada día, 50 clientes se seleccionan al azar y el pedido se considera defectuoso si demora más de tres minutos. A continuación se listan las cantidades de pedidos defectuosos en días consecutivos. Elabore una gráfica de control apropiada y determine si el proceso está bajo control estadístico. Si no es así, identifique qué criterios conducen al rechazo de la estabilidad estadística.

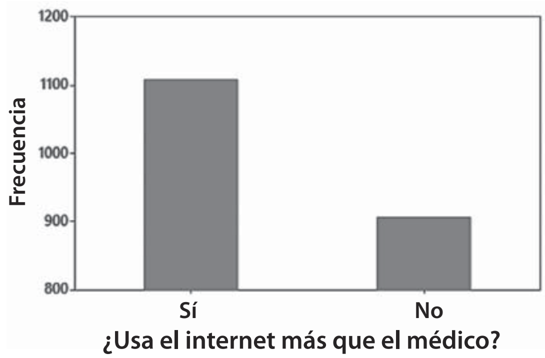
3 2 3 5 4 6 7 9 8 10 11 9 12 15 17

Ejercicios de repaso acumulado

1. Médicos en Internet: Intervalo de confianza En una encuesta aplicada a $n = 2015$ adultos, 1108 de ellos dijeron que conocían más a menudo los síntomas de enfermedades a través de Internet que por su médico (según una encuesta de MerckManuals.com). Use los datos para elaborar una estimación del intervalo de confianza del 95% para la proporción poblacional de todos los adultos que dicen que conocen más a menudo los síntomas de enfermedades a través de Internet que por su médico. ¿El resultado sugiere que la mayoría de los adultos conocen más a menudo los síntomas de enfermedades a través de Internet que por su médico?

2. Médicos en Internet: Prueba de hipótesis Utilice los resultados de la encuesta dados en el ejercicio 1 y use un nivel de significancia de 0.05 para evaluar la afirmación de que la mayoría de los adultos conocen más a menudo los síntomas de enfermedades a través de Internet que por su médico.

3. Médicos en Internet: Gráfica La gráfica adjunta fue creada para representar los resultados de la encuesta descrita en el ejercicio 1. ¿Es la gráfica algo engañosa? ¿Si es así, por qué?



4. Médicos en Internet: Probabilidad Según los datos de la encuesta proporcionados en el ejercicio 1, suponga que 55% de los adultos conocen más a menudo los síntomas de enfermedades a través de Internet que por su médico.

a. Encuentre la probabilidad de que tres adultos seleccionados al azar conozcan más a menudo los síntomas de enfermedades a través de Internet que por su médico.

b. Encuentre la probabilidad de que al seleccionar tres adultos aleatoriamente, al menos uno de ellos conozca más menudo los síntomas de enfermedades a través de Internet que por su médico.

5. Manchas solares y el DJIA A continuación se listan las cantidades anuales de manchas solares combinados con los valores anuales más altos del Promedio Industrial Dow Jones (DJIA, por sus siglas en inglés). Las cantidades de manchas solares se miden de acuerdo con los puntos oscuros en el Sol, y el índice DJIA mide el valor de ciertas acciones seleccionadas. Los datos son de años recientes y consecutivos. Use un nivel de significancia de 0.05 para probar una correlación lineal entre los valores del DJIA y las cantidades de manchas solares. ¿El resultado es sorprendente?

Manchas solares	45	31	46	31	50	48	56	38	65	51
DJIA	10,941	12,464	14,198	13,279	10,580	11,625	12,929	13,589	16,577	18,054

6. Manchas solares y el DJIA Utilice los datos del ejercicio 5 y encuentre la ecuación de la línea de regresión. Luego, determine el mejor valor predicho del DJIA en el año 2004, cuando el número de manchas solares fue de 61. ¿Cómo se compara el resultado con el valor real del DJIA de 10,855?

7. Estaturas Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, suponga que las estaturas de los hombres se distribuyen normalmente, con una media de 68.6 pulgadas y una desviación estándar de 2.8 pulgadas.

a. La Guardia Costera de Estados Unidos exige que los hombres tengan una altura entre 60 y 80 pulgadas. Encuentre el porcentaje de hombres que satisfacen ese requisito de estatura.

b. Encuentre la probabilidad de que 4 hombres elegidos al azar tengan estaturas con una media mayor a 70 pulg.

8. Sistemas de seguridad infantil defectuosos La compañía Tracolyte Manufacturing produce bastidores de plástico para asientos infantiles de automóvil. Durante cada semana de producción, el Departamento de Transporte selecciona y prueba 120 bastidores acerca de su cumplimiento de todas las regulaciones. Los bastidores se consideran defectuosos si no cumplen con todos los requisitos. A continuación se listan cantidades de bastidores defectuosos entre los 120 que se prueban cada semana. Use una gráfica de control para p con el fin de verificar que el proceso esté bajo control estadístico. Si no está bajo control, explique por qué no lo está.

3 2 4 6 5 9 7 10 12 15

9. Sistemas de seguridad infantil Use las cantidades de sistemas de seguridad infantil defectuosos que se proporcionan en el ejercicio 8. Encuentre la media, la mediana y la desviación estándar. ¿Qué característica importante de los datos muestrales se pasa por alto si exploramos los datos mediante el uso de esos estadísticos?

10. ¿Vale la pena declararse culpable? La tabla adjunta resume los datos muestrales seleccionados al azar de acusados en casos de robo en San Francisco (con base en datos de "Does It Pay to Plead Guilty? Differential Sentencing and the Functioning of the Criminal Courts", de Brereton y Casper, *Law and Society Review*, vol. 16, núm. 1). Todos los sujetos tenían condenas carcelarias previas. Use un nivel de significancia de 0.05 para probar la afirmación de que la sentencia (enviado a prisión o no enviado a prisión) es independiente de la declaración de culpabilidad. Si usted fuera un abogado que defiende a un acusado culpable, ¿estos resultados sugieren que debe alentar una declaración de culpabilidad?

	Se declara culpable	Se declara no culpable
Enviado a prisión	392	58
No enviado a prisión	564	14

Proyecto de tecnología

a. Simule el siguiente proceso durante 20 días: cada día, 200 calculadoras se fabrican con una tasa de defectos del 5% y la proporción de defectos se registra para cada uno de los 20 días. Las calculadoras de un día se simulan generando al azar 200 números, donde cada número se encuentra entre 1 y 100. Considere que un resultado de 1, 2, 3, 4 o 5 es un defecto, siendo aceptable de 6 a 100. Esto corresponde a una tasa de defectos del 5%. (*Sugerencia:* Consulte el proyecto de tecnología del capítulo 11 en la página 563 para obtener sugerencias sobre tecnología).

b. Elabore una gráfica p para la proporción de calculadoras defectuosas y determine si el proceso está bajo control estadístico. Como sabemos que el proceso es realmente estable con $p = 0.05$, la conclusión de que no es estable sería un error de tipo I; es decir, tendríamos una señal positiva falsa, lo que hace que creamos que el proceso debe ajustarse cuando, en realidad, debe dejarse como está.

c. El resultado del inciso (a) es una simulación de 20 días. Ahora simule otros 10 días de fabricación de calculadoras, pero modifique estos últimos 10 días para que la tasa de defectos sea de 10% en vez de 5%.

d. Combine los datos generados en los incisos (a) y (c) para representar un total de 30 días de resultados muestrales. Elabore una gráfica p para este conjunto combinado de datos. ¿El proceso está fuera de control estadístico? Si llegamos a la conclusión de que el proceso no está fuera de control estadístico, estaríamos cometiendo un error de tipo II; es decir, creemos que el proceso está bien cuando, de hecho, debe modificarse para corregir el cambio al 10% en la tasa de defectos.

DE LOS DATOS A LA DECISIÓN

Pensamiento crítico: ¿Están las cargas axiales bajo control estadístico? ¿El proceso de fabricación de latas funciona como debería?

Los ejercicios 5 a 8 en la sección 14-1 utilizaron datos de proceso de la fabricación de latas de aluminio con 0.0109 pulgadas de espesor. Consulte el conjunto de datos 30 "Latas de aluminio" en el apéndice B y realice un análisis de los datos de proceso para latas con un espesor de 0.0111 pulgadas. Los valores en el conjunto de datos son las cargas axiales medidas en las latas, y las tapas superiores se presionan en su sitio con presiones que varían entre 158 lb y 165 lb. Las 175 cargas axiales se dan en una columna, las primeras 7 son del primer día, las 7 siguientes son del segundo día, y así sucesivamente, de modo que el "tamaño del subgrupo" es 7.

Análisis de los resultados

En función de los datos de proceso dados, ¿debería la empresa tomar alguna medida correctiva? Escriba un informe que resuma sus conclusiones. Aborde no sólo la cuestión de la estabilidad estadística sino también la capacidad de las latas para resistir las presiones aplicadas cuando las tapas superiores se presionan en su sitio. También compare el comportamiento de las latas de 0.0111 pulgadas con el comportamiento de las latas de 0.0109 pulgadas, y recomiende qué espesor se debe utilizar.

Actividades en equipo

1. Actividad fuera de clase Colecte sus propios datos de proceso y analícelos utilizando los métodos de este capítulo. Sería ideal recopilar datos de un proceso de fabricación real, pero eso podría ser difícil. En su lugar, considere usar una simulación o hacer referencia a datos publicados. Obtenga una copia de los resultados de la computadora y escriba un breve informe que resuma sus conclusiones. A continuación damos algunas sugerencias:

• Dispare cinco tiros libres de básquetbol (o lance cinco hojas de papel arrugadas en una papelera) y anote el número de tiros acertados; luego repita este procedimiento 20 veces. Use una gráfico *p* para comprobar la estabilidad estadística en la proporción de tiros acertados.

• Mida su pulso contando la cantidad de veces que su corazón late en 1 minuto. Mida su pulso cuatro veces por hora durante varias horas, luego elabore las gráficas de control adecuadas. ¿Qué factores contribuyen a la variación aleatoria? ¿y a la variación asignable?

• Busque en Internet y registre el Promedio Industrial Dow Jones (DJIA) al cierre de cada día hábil de las últimas 12 semanas. Use gráficas de rachas y control para explorar la estabilidad estadística del DJIA. Identifique al menos una consecuencia práctica de tener este proceso estadísticamente estable e identifique al menos una consecuencia práctica de tener este proceso fuera de control estadístico.

• Encuentre la tasa de matrimonios por cada 10,000 habitantes durante varios años. Suponga que cada año se seleccionan y encuestan 10,000 personas al azar para determinar si están casados. Use una gráfica *p* para evaluar la estabilidad estadística de la tasa de matrimonios. (Otras tasas posibles: muerte, accidente fatal, crimen).

• Busque en Internet y encuentre las cantidades de carreras anotadas por su equipo de béisbol favorito en juegos recientes, luego use gráficas de rachas y control para explorar el control estadístico.

2. Actividad en clase Si el profesor puede proporcionar el número de faltas en cada clase, grupos de tres o cuatro estudiantes pueden analizarlo en cuanto a su estabilidad estadística y hacer recomendaciones basadas en las conclusiones.

3. Actividad fuera de clase Realice una investigación para encontrar una descripción del experimento del embudo de Deming, luego use un embudo y canicas para recopilar datos de las diferentes reglas para ajustar la ubicación del embudo. Elabore las gráficas de control apropiadas para las diferentes reglas de ajuste del embudo. ¿Qué ilustra el experimento del embudo? ¿A qué conclusión se llega?

15

LA ÉTICA EN ESTADÍSTICA

Concepto clave A pesar de que los métodos estadísticos nos otorgan un gran poder para comprender mejor el mundo en el que vivimos, éste también brinda oportunidades para usos que son fundamentalmente antiéticos. Es importante considerar algunos problemas éticos en la estadística; la presente sección considera tales aspectos relacionados con la recolección de datos, el análisis y la generación de informes.

I. Recolección de datos

En ocasiones, los usos poco éticos de la estadística comienzan con el proceso de recolección de datos. Ocasionalmente, los seres humanos son tratados de maneras que no son éticas y las muestras pueden sesgarse de formas que tampoco lo son.

Tratamiento humano durante la obtención de datos

Existen muchos estudios en los que la salud y la seguridad de los sujetos de una investigación se han visto comprometidas para que otros pudieran resultar beneficiados.

Estudios con prisioneros A principios de la década de 1970, se estimó que 90% de las investigaciones farmacéuticas en Estados Unidos se realizaban utilizando prisioneros como sujetos de prueba, a veces sin su conocimiento o consentimiento. En la actualidad, la Regla Común (que se describe más adelante en esta sección) protege a los sujetos humanos de investigación en Estados Unidos.

Experimento de Milgram Durante la década de 1960, Stanley Milgram llevó a cabo uno de los estudios psicológicos más infames. Se ordenó a los sujetos que administraran "descargas eléctricas" progresivamente más fuertes a actores que se encontraban en otra habitación. Estos actores no eran visibles para el sujeto y en realidad no recibían ninguna descarga. Los gritos grabados previamente se reproducían de modo que el sujeto pudiera escucharlos. El actor golpeaba la pared y suplicaba al sujeto que se detuviera.

Antes del experimento, los expertos estimaron que sólo 3% de los sujetos continuaría dando descargas después de escuchar los gritos y las súplicas para detenerse, pero 65% de los sujetos aumentaron las descargas hasta alcanzar el máximo de 375 voltios. Este experimento infligió una angustia extrema en los sujetos y ahora se considera poco ético y psicológicamente abusivo.

Sesgo del muestreo

Los investigadores pueden sesgar intencional o involuntariamente su estudio mediante el uso de una muestra parcial. El sesgo en el muestreo puede distorsionar seriamente los resultados y llevar a conclusiones erróneas o engañosas. En algunos casos, los investigadores usaron intencionalmente muestras sesgadas para garantizar que se obtuvieran los resultados deseados.

Selección arreglada El informe "¿Están los investigadores escogiendo los participantes en sus estudios de antidepresivos?" de la Escuela de Ciencias de la Salud en la University of Pittsburgh, describió los resultados de estudios clínicos realizados con el objetivo de obtener la aprobación de antidepresivos por parte de la Administración de Alimentos y Medicamentos. Sólo un pequeño porcentaje de los individuos deprimidos cumplió con los criterios requeridos para participar en el ensayo clínico y aquellos que calificaron tuvieron mejores resultados que otras personas deprimidas que no calificaron. En este caso, los criterios necesarios para participar en el ensayo clínico crearon un sesgo muestral que favorecía a los fabricantes de medicamentos que patrocinaron la investigación.

Sesgo de no respuesta El sesgo de no respuesta ocurre cuando las personas que no responden a una encuesta difieren de aquellos que sí la responden. Dos aspectos del tema de una encuesta que afectan las respuestas son la *prominencia* (si el tema es de interés o no para el encuestado) y la *deseabilidad social* (si el tema es amenazante o embarazoso, o si no lo es). Un estudio encontró que cuando un tema tiene un gran interés para los encuestados, las probabilidades de obtener una respuesta aumentan casi al doble. Otros estudios han encontrado que los comportamientos socialmente deseables, como el ejercicio y la buena nutrición, con frecuencia se declaran en exceso, mientras que los comportamientos indeseables, como fumar, no son aceptados. Para la investigación del alcohol, varios estudios han demostrado que los bebedores empedernidos tienen menos probabilidades de responder. (Vea "Non-response Bias in a Sample Survey on Alcohol Consumption", de V. M. Lahaut *et al*., *Alcohol and Alcoholism*).

Sesgo del entrevistador La manera en que se formula una pregunta puede afectar la respuesta. El autor recibió una encuesta por correo que formulaba las siguientes preguntas redactadas de una manera en que se pretendía afectar las respuestas:

- "¿Apoya usted la creación de una política nacional de seguridad social que sea administrada por burócratas en Washington, DC?"
- "¿Está usted a favor de crear un 'Cuerpo de Voluntarios Ciudadanos' financiado por el gobierno, que les pague a los jóvenes para que trabajen en iglesias y organizaciones benéficas, realizando el trabajo que ahora hacen los veteranos militares, y que reciban la misma paga y prestaciones que éstos?"

Sesgo voluntario Las personas que se ofrecen como voluntarias para participar constituyen una muestra de respuesta voluntaria y, a menudo, tienen características diferentes en comparación con los participantes que son seleccionados por quienes realizan la encuesta. Los efectos del sesgo voluntario han sido ampliamente estudiados para el tema de la sexualidad humana. "Volunteer Bias in Sexuality Research" de D. S. Strassberg *et al*., Departamento de Psicología de la University of Utah, encontró que, en comparación con los no voluntarios, los sujetos voluntarios presentaban una actitud más positiva hacia la sexualidad, menos culpa y más experiencia sexual.

II. Análisis

Con frecuencia, los resultados de los estudios de investigación se distorsionan por un análisis inadecuado. Aunque los datos muestrales pueden no ser concluyentes, ser insuficientes o apoyar un resultado indeseable, muchos investigadores han sucumbido a presiones para informar resultados significativos o "concluyentes" a través de análisis negligentes y poco éticos.

Datos falsificados Varias encuestas han intentado determinar la incidencia de la falsificación o fabricación de datos entre los científicos. "How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data", de D. Fanelli, *PLoS ONE*, vol. 4, núm. 5, encontró que casi 2% de los científicos admitieron haber fabricado, falsificado o modificado datos o resultados al menos una vez. Es probable que sea una estimación conservadora, dada la sensibilidad del tema y el potencial de sesgo del factor de deseabilidad social descrito anteriormente. Un ejemplo notable de la falsificación de datos se encuentra en el estudio original que sugiere un vínculo entre las vacunas infantiles y el autismo.

Andrew Wakefield publicó un estudio de 1998 que relacionaba la vacuna MMR con el autismo. En 2010, la *British Medical Journal* proporcionó pruebas de que el estudio de Wakefield se basó en datos falsificados. *The Lancet*, que publicó el estudio original en 1998, se retractó de éste y la licencia médica de Wakefield fue revocada. Mientras que algunos activistas del autismo aún defienden sus acciones, la comunidad científica ha condenado esta investigación por ser poco ética.

Métodos estadísticos inapropiados El uso inadecuado de métodos estadísticos puede conducir a resultados incorrectos y distorsionados, incluso si los datos son sólidos. "Statistical Errors in Manuscripts Submitted to *Biochemia Medica* Journal", de Simundic y Nikolac, *Biochemia Medica*, vol. 19, núm. 3, examinó 55 manuscritos que se enviaron para su publicación y encontró que 48 (87%) contenían al menos un error estadístico; 34 (62%) tuvieron una elección incorrecta de la prueba estadística; 22% tuvo una interpretación incorrecta de un valor *P*; y 75% incluyó el uso incorrecto de una prueba estadística para comparar tres o más grupos. La revisión por colegas es una defensa primaria contra el uso de métodos estadísticos inapropiados, pero no garantiza que tales errores estadísticos sean identificados y corregidos.

Elección de un nivel de significancia Una buena práctica consiste en seleccionar un nivel de significancia para una prueba de hipótesis y encontrar un valor *P antes* de recolectar los datos muestrales. Una mala práctica es ajustar posteriormente el nivel de significancia a fin de que los resultados parezcan significativos cuando no lo son en función del nivel de significancia original.

III. Ética e informes

Otras oportunidades para prácticas poco éticas ocurren con la interpretación de los resultados estadísticos y la decisión de cómo y dónde reportar los hallazgos. Los intereses personales y egoístas pueden influir en las conclusiones y recomendaciones finales.

Conflictos de interés

Autor fantasma Un autor médico fantasma es alguien que contribuye a la redacción de un estudio o artículo de investigación, pero no se le reconoce en el trabajo publicado. En los últimos años, varios litigios han revelado que las compañías farmacéuticas han contratado redactores profesionales para escribir artículos sobre ensayos clínicos y han pagado a los médicos para que acepten la autoría de tales artículos. Ni el papel de los autores fantasmas ni la compensación financiera para el autor reconocido se menciona en los trabajos.

Apoyo financiero La organización sin fines de lucro Fair Warning informó que durante varios años recientes, los hospitales invirtieron millones de dólares para comprar nuevos desfibriladores automáticos de corazón. Un comité de la Asociación Estadounidense del Corazón recomendó la actualización sin el beneficio de la investigación o los ensayos clínicos que demostraran que los nuevos dispositivos eran mejores que los dispositivos que se estaban utilizando. De acuerdo con *Fair Warning*, más de 25% de los miembros del comité "tenían vínculos comerciales con los fabricantes de los dispositivos" y "según una estimación, las deficiencias de los equipos automatizados ocasionan que cada año mueran cerca de 1000 pacientes adicionales por paro cardíaco en Estados Unidos". Debido a la aprobación en 2010 de una ley federal, ahora se requiere que las compañías divulguen los pagos por bienes y servicios brindados a médicos u hospitales de enseñanza. Si bien esta ley no *prohíbe* las relaciones financieras entre empresas y médicos u hospitales de enseñanza, sí requiere la divulgación de tales relaciones.

Reporte de resultados no significativos Un artículo reciente apareció con el titular "Después de 39 años, el apoyo a la pena de muerte disminuye en Estados Unidos". Una encuesta de Gallup encontró que el 61% de los estadounidenses estaba a favor de la pena de

Anonimato y confidencialidad

Una encuesta con *anonimato* se realiza si no se conocen las identidades de los encuestados. Una encuesta con *confidencialidad* se presenta si las identidades de los encuestados no se divulgan. Idealmente, las encuestas deben ser anónimas y confidenciales, pero eso no siempre es práctico o bueno. La confidencialidad puede dejarse de lado si se encuentra que algunos encuestados son un peligro para ellos mismos o para los demás. En tales casos, los encuestados deben ser informados con una declaración como la siguiente: "Toda la información que proporcione será confidencial, a menos que implique riesgos graves para usted o para otras personas".

El periódico *National Observer* ya no existe, pero en cierta ocasión contrató a una empresa para realizar una encuesta confidencial por correo. La encuesta se realizó con la promesa de que "cada respuesta individual se mantendría confidencial". Un suscriptor inteligente usó una luz ultravioleta para detectar un código de identificación único impreso en la encuesta en tinta invisible. Aquí, la confidencialidad se prometió y se cumplió, pero no ocurrió lo mismo con el anonimato. En lugar de usar tinta invisible, la empresa debería haber informado a los encuestados que su información no era anónima.

muerte, lo que implicaba una caída del 3% con respecto la tasa del año anterior. La descripción del "método de la encuesta" al final del artículo indicaba que "se puede decir con un 95% de confianza que el margen máximo de error en el muestreo es de 4%". Con un error de muestreo del 4%, una disminución del 3% en los encuestados que indicaron estaban a favor de la pena de muerte no es significativo. Realmente no sabemos si hubo una disminución, un aumento o ningún cambio en las opiniones sobre la pena de muerte.

IV. Cumplimiento de la ética

Para ayudar a evitar prácticas estadísticas poco éticas, muchas organizaciones han establecido directrices para sus miembros o para los autores cuyos trabajos publican. Por ejemplo, la Asociación Estadounidense de Estadística, ratificó sus "directrices éticas para la práctica estadística", que proporciona pautas éticas en ocho temas generales, que incluyen las Responsabilidades con los sujetos de investigación, las Responsabilidades en publicaciones y testimonios, y las Responsabilidades con otros estadísticos. El texto completo de las Directrices Éticas se puede encontrar en el sitio web www.amstat.org.

La **Regla Común** se ha establecido en Estados Unidos como el estándar de ética para la investigación biomédica y conductual que involucra seres humanos. Este es el estándar de referencia para cualquier investigación financiada por el gobierno y prácticamente todas las instituciones académicas. Para la regla común, los siguientes requisitos resultan fundamentales:

1. Que las personas que participan como sujetos en la investigación cubierta se seleccionen de forma equitativa y den su consentimiento por escrito completamente informado y totalmente voluntario; y
2. Que la investigación propuesta sea revisada por un grupo de supervisión independiente denominado Junta de Revisión Institucional (IRB, por sus siglas en inglés) y aprobada sólo si los riesgos para los sujetos se han minimizado y son razonables en relación con los beneficios anticipados, si los hubiera, para los sujetos, y la importancia del conocimiento que razonablemente se espera como resultado.

El texto detallado de la regla común se puede encontrar en el sitio web del Departamento de Salud y Servicios Humanos de Estados Unidos. En www.hhs.gov.

Como consumidores de datos, es importante tener un escepticismo saludable cuando se citen estadísticas para respaldar un hallazgo. Deberíamos hacernos preguntas como las siguientes:

- ¿Quién informa sobre el estudio?, ¿tiene algún interés financiero o personal en el resultado?
- ¿Cómo se obtuvieron los datos muestrales? ¿Cuál es el potencial de sesgo?
- ¿Estos resultados han sido revisados o replicados por colegas o por cualquier otra persona?
- ¿Qué detalles se proporcionan sobre los métodos utilizados?, ¿cuál es el margen de error, el intervalo de confianza, el valor *P*, el tamaño del efecto, etcétera?

Los investigadores siempre deben cumplir con los más altos estándares éticos en su trabajo. Todos debemos esforzarnos por comportarnos éticamente de la siguiente manera:

- Sea cabal y honesto acerca de los hallazgos, incluso si los resultados no son los esperados o deseados.
- Busque el consejo de profesionales cuando no esté seguro sobre qué análisis estadísticos son los apropiados.
- Siempre revele las relaciones financieras o cualquier otro interés en el resultado de la investigación.
- Dé crédito sólo a aquellas personas que hicieron una contribución significativa al estudio.
- Esté preparado y dispuesto a compartir los datos, métodos, análisis y resultados.
- Identifique claramente cualquier supuesto, limitación o pregunta pendiente en la investigación.

Puntos de discusión

1. ¿Es ético utilizar prisioneros como sujetos si comprenden los riesgos potenciales y han dado su consentimiento?, ¿es ético ofrecerles un incentivo para que den su consentimiento, como darles privilegios adicionales o reducirles la sentencia?
2. ¿El experimento de Milgram descrito en esta sección fue poco ético? ¿Por qué sí o por qué no?
3. Para proyectos de investigación que afectan a sujetos humanos, ¿el daño psicológico potencial es diferente del daño físico potencial?
4. ¿Es poco ético infectar intencionalmente a sujetos de prueba humanos?, ¿qué hay de los animales?
5. El sesgo voluntario está bien documentado en el tema de la sexualidad humana. ¿Cuáles son algunos otros temas que pueden tener un alto riesgo de sesgo voluntario?
6. Una reducción del sesgo muestral puede generar un resultado general menos positivo, que puede retrasar o detener la liberación de un medicamento del que un segmento de la población podría beneficiarse. ¿Es esto una compensación aceptable?
7. Un artículo sobre el tema de los errores estadísticos en la investigación médica listó 47 posibles fuentes de error. Intente identificar 5 de estas posibles fuentes de error.
8. ¿Qué es la autoría médica fantasma?, ¿esta práctica es poco ética?
9. ¿Deberían existir estándares comunes para reclamar la autoría de un estudio?, ¿qué debería incluirse en esos estándares?
10. ¿La ley que exige la revelación de pagos es suficiente para evitar que el apoyo financiero sesgue los resultados de los estudios?, ¿tiene usted otras recomendaciones?

APÉNDICE A Tablas

TABLA A-1 Probabilidades binomiales

		p													
n	x	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	x
2	0	.980	.903	.810	.640	.490	.360	.250	.160	.090	.040	.010	.003	0+	0
	1	.020	.095	.180	.320	.420	.480	.500	.480	.420	.320	.180	.095	.020	1
	2	0+	.003	.010	.040	.090	.160	.250	.360	.490	.640	.810	.903	.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	.063	.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	.063	.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	.313	.230	.132	.051	.008	.001	0+	2
	3	0+	.001	.008	.051	.132	.230	.313	.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 un valor de probabilidad positivo menor que 0.0005.

Puntuaciones *z* NEGATIVAS

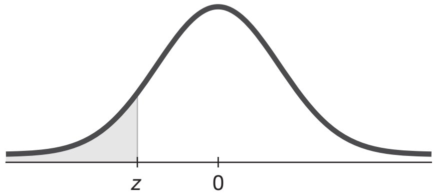


TABLA A-2 Distribución normal estándar (*z*): Área acumulada desde la IZQUIERDA

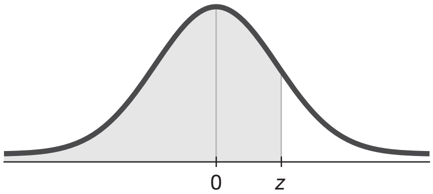
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.50 y menores	.0001									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051 *	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505 *	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

NOTA: Para valores de *z* por debajo de –3.49, use un área de 0.0001.

(*continúa*)

*Use los siguientes valores comunes que resultan de interpolación:

Puntuación *z*	Área
−1.645	0.0500
−2.575	0.0050



Puntuaciones z POSITIVAS

TABLA A-2 *(continuación)* Área acumulada desde la IZQUIERDA

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495 *	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949 *	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.50 y mayores	.9999									

NOTA: Para valores de *z* superiores a 3.49, use un área de 0.9999.

*Use los siguientes valores comunes que resultan de interpolación:

Puntuación *z*	Área
1.645	0.9500
2.575	0.9950

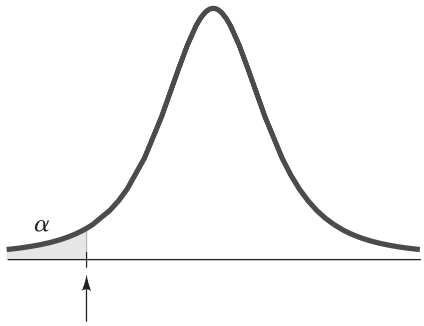
Valores críticos comunes

Nivel de confianza	Valor crítico
0.90	1.645
0.95	1.96
0.99	2.575

TABLA A-3 Distribución t: Valores t críticos

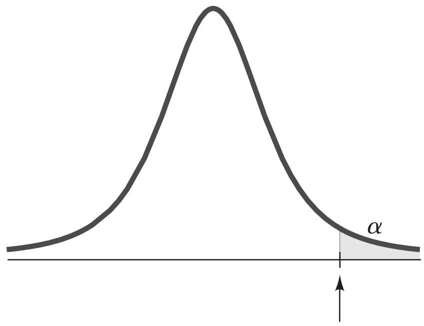
	Área en una cola				
	0.005	0.01	0.025	0.05	0.10
Grados de libertad	**Área en dos colas** 0.01	0.02	0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316
26	2.779	2.479	2.056	1.706	1.315
27	2.771	2.473	2.052	1.703	1.314
28	2.763	2.467	2.048	1.701	1.313
29	2.756	2.462	2.045	1.699	1.311
30	2.750	2.457	2.042	1.697	1.310
31	2.744	2.453	2.040	1.696	1.309
32	2.738	2.449	2.037	1.694	1.309
33	2.733	2.445	2.035	1.692	1.308
34	2.728	2.441	2.032	1.691	1.307
35	2.724	2.438	2.030	1.690	1.306
36	2.719	2.434	2.028	1.688	1.306
37	2.715	2.431	2.026	1.687	1.305
38	2.712	2.429	2.024	1.686	1.304
39	2.708	2.426	2.023	1.685	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
60	2.660	2.390	2.000	1.671	1.296
70	2.648	2.381	1.994	1.667	1.294
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
1000	2.581	2.330	1.962	1.646	1.282
2000	2.578	2.328	1.961	1.646	1.282
Grandes	2.576	2.326	1.960	1.645	1.282

Cola izquierda



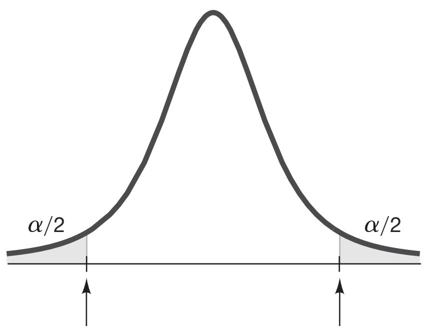
Valor t crítico
(negativo)

Cola derecha



Valor t crítico
(positivo)

Dos colas



Valor t crítico
(negativo)

Valor t crítico
(positivo)

TABLA A-4 Distribución Ji cuadrada (χ^2)

	Área a la *derecha* del valor crítico									
Grados de libertad	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

Fuente: Donald B. Owen, *Handbook of Statistical Tables.*

Grados de libertad

$n - 1$	**Intervalo de confianza o prueba de hipótesis** para una desviación estándar σ o varianza σ^2
$k - 1$	**Prueba de bondad de ajuste** con k categorías diferentes
$(r - 1)(c - 1)$	**Prueba de tabla de contingencia** con r filas y c columnas
$k - 1$	**Prueba de Kruskal-Wallis** con k muestras diferentes

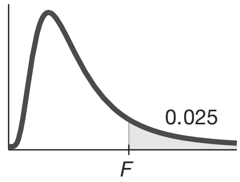


TABLA A-5 Distribución F ($\alpha = 0.025$ en la cola derecha)

Grados de libertad en el denominador (df_2)	Grados de libertad en el numerador (df_1)								
	1	2	3	4	5	6	7	8	9
1	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
2	38.506	39.000	39.165	39.248	39.298	39.331	39.335	39.373	39.387
3	17.443	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473
4	12.218	10.649	9.9792	9.6045	9.3645	9.1973	9.0741	8.9796	8.9047
5	10.007	8.4336	7.7636	7.3879	7.1464	6.9777	6.8531	6.7572	6.6811
6	8.8131	7.2599	6.5988	6.2272	5.9876	5.8198	5.6955	5.5996	5.5234
7	8.0727	6.5415	5.8898	5.5226	5.2852	5.1186	4.9949	4.8993	4.8232
8	7.5709	6.0595	5.4160	5.0526	4.8173	4.6517	4.5286	4.4333	4.3572
9	7.2093	5.7147	5.0781	4.7181	4.4844	4.3197	4.1970	4.1020	4.0260
10	6.9367	5.4564	4.8256	4.4683	4.2361	4.0721	3.9498	3.8549	3.7790
11	6.7241	5.2559	4.6300	4.2751	4.0440	3.8807	3.7586	3.6638	3.5879
12	6.5538	5.0959	4.4742	4.1212	3.8911	3.7283	3.6065	3.5118	3.4358
13	6.4143	4.9653	4.3472	3.9959	3.7667	3.6043	3.4827	3.3880	3.3120
14	6.2979	4.8567	4.2417	3.8919	3.6634	3.5014	3.3799	3.2853	3.2093
15	6.1995	4.7650	4.1528	3.8043	3.5764	3.4147	3.2934	3.1987	3.1227
16	6.1151	4.6867	4.0768	3.7294	3.5021	3.3406	3.2194	3.1248	3.0488
17	6.0420	4.6189	4.0112	3.6648	3.4379	3.2767	3.1556	3.0610	2.9849
18	5.9781	4.5597	3.9539	3.6083	3.3820	3.2209	3.0999	3.0053	2.9291
19	5.9216	4.5075	3.9034	3.5587	3.3327	3.1718	3.0509	2.9563	2.8801
20	5.8715	4.4613	3.8587	3.5147	3.2891	3.1283	3.0074	2.9128	2.8365
21	5.8266	4.4199	3.8188	3.4754	3.2501	3.0895	2.9686	2.8740	2.7977
22	5.7863	4.3828	3.7829	3.4401	3.2151	3.0546	2.9338	2.8392	2.7628
23	5.7498	4.3492	3.7505	3.4083	3.1835	3.0232	2.9023	2.8077	2.7313
24	5.7166	4.3187	3.7211	3.3794	3.1548	2.9946	2.8738	2.7791	2.7027
25	5.6864	4.2909	3.6943	3.3530	3.1287	2.9685	2.8478	2.7531	2.6766
26	5.6586	4.2655	3.6697	3.3289	3.1048	2.9447	2.8240	2.7293	2.6528
27	5.6331	4.2421	3.6472	3.3067	3.0828	2.9228	2.8021	2.7074	2.6309
28	5.6096	4.2205	3.6264	3.2863	3.0626	2.9027	2.7820	2.6872	2.6106
29	5.5878	4.2006	3.6072	3.2674	3.0438	2.8840	2.7633	2.6686	2.5919
30	5.5675	4.1821	3.5894	3.2499	3.0265	2.8667	2.7460	2.6513	2.5746
40	5.4239	4.0510	3.4633	3.1261	2.9037	2.7444	2.6238	2.5289	2.4519
60	5.2856	3.9253	3.3425	3.0077	2.7863	2.6274	2.5068	2.4117	2.3344
120	5.1523	3.8046	3.2269	2.8943	2.6740	2.5154	2.3948	2.2994	2.2217
∞	5.0239	3.6889	3.1161	2.7858	2.5665	2.4082	2.2875	2.1918	2.1136

(*continúa*)

TABLA A-5 *(continuación)* Distribución F ($\alpha = 0.025$ en la cola derecha)

Grados de libertad en el denominador (df_2)	Grados de libertad en el numerador (df_1)									
	10	12	15	20	24	30	40	60	120	∞
1	968.63	976.71	984.87	993.10	997.25	1001.4	1005.6	1009.8	1014.0	1018.3
2	39.398	39.415	39.431	39.448	39.456	39.465	39.473	39.481	39.490	39.498
3	14.419	14.337	14.253	14.167	14.124	14.081	14.037	13.992	13.947	13.902
4	8.8439	8.7512	8.6565	8.5599	8.5109	8.4613	8.4111	8.3604	8.3092	8.2573
5	6.6192	6.5245	6.4277	6.3286	6.2780	6.2269	6.1750	6.1225	6.0693	6.0153
6	5.4613	5.3662	5.2687	5.1684	5.1172	5.0652	5.0125	4.9589	4.9044	4.8491
7	4.7611	4.6658	4.5678	4.4667	4.4150	4.3624	4.3089	4.2544	4.1989	4.1423
8	4.2951	4.1997	4.1012	3.9995	3.9472	3.8940	3.8398	3.7844	3.7279	3.6702
9	3.9639	3.8682	3.7694	3.6669	3.6142	3.5604	3.5055	3.4493	3.3918	3.3329
10	3.7168	3.6209	3.5217	3.4185	3.3654	3.3110	3.2554	3.1984	3.1399	3.0798
11	3.5257	3.4296	3.3299	3.2261	3.1725	3.1176	3.0613	3.0035	2.9441	2.8828
12	3.3736	3.2773	3.1772	3.0728	3.0187	2.9633	2.9063	2.8478	2.7874	2.7249
13	3.2497	3.1532	3.0527	2.9477	2.8932	2.8372	2.7797	2.7204	2.6590	2.5955
14	3.1469	3.0502	2.9493	2.8437	2.7888	2.7324	2.6742	2.6142	2.5519	2.4872
15	3.0602	2.9633	2.8621	2.7559	2.7006	2.6437	2.5850	2.5242	2.4611	2.3953
16	2.9862	2.8890	2.7875	2.6808	2.6252	2.5678	2.5085	2.4471	2.3831	2.3163
17	2.9222	2.8249	2.7230	2.6158	2.5598	2.5020	2.4422	2.3801	2.3153	2.2474
18	2.8664	2.7689	2.6667	2.5590	2.5027	2.4445	2.3842	2.3214	2.2558	2.1869
19	2.8172	2.7196	2.6171	2.5089	2.4523	2.3937	2.3329	2.2696	2.2032	2.1333
20	2.7737	2.6758	2.5731	2.4645	2.4076	2.3486	2.2873	2.2234	2.1562	2.0853
21	2.7348	2.6368	2.5338	2.4247	2.3675	2.3082	2.2465	2.1819	2.1141	2.0422
22	2.6998	2.6017	2.4984	2.3890	2.3315	2.2718	2.2097	2.1446	2.0760	2.0032
23	2.6682	2.5699	2.4665	2.3567	2.2989	2.2389	2.1763	2.1107	2.0415	1.9677
24	2.6396	2.5411	2.4374	2.3273	2.2693	2.2090	2.1460	2.0799	2.0099	1.9353
25	2.6135	2.5149	2.4110	2.3005	2.2422	2.1816	2.1183	2.0516	1.9811	1.9055
26	2.5896	2.4908	2.3867	2.2759	2.2174	2.1565	2.0928	2.0257	1.9545	1.8781
27	2.5676	2.4688	2.3644	2.2533	2.1946	2.1334	2.0693	2.0018	1.9299	1.8527
28	2.5473	2.4484	2.3438	2.2324	2.1735	2.1121	2.0477	1.9797	1.9072	1.8291
29	2.5286	2.4295	2.3248	2.2131	2.1540	2.0923	2.0276	1.9591	1.8861	1.8072
30	2.5112	2.4120	2.3072	2.1952	2.1359	2.0739	2.0089	1.9400	1.8664	1.7867
40	2.3882	2.2882	2.1819	2.0677	2.0069	1.9429	1.8752	1.8028	1.7242	1.6371
60	2.2702	2.1692	2.0613	1.9445	1.8817	1.8152	1.7440	1.6668	1.5810	1.4821
120	2.1570	2.0548	1.9450	1.8249	1.7597	1.6899	1.6141	1.5299	1.4327	1.3104
∞	2.0483	1.9447	1.8326	1.7085	1.6402	1.5660	1.4835	1.3883	1.2684	1.0000

Con base en datos de Maxine Merrington y Catherine M. Thompson, “Tables of Percentage Points of the Inverted Beta (F) Distribution”, *Biometrika 33* (1943): 80-84.

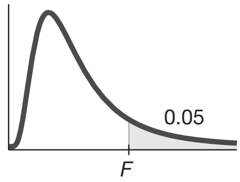


TABLA A-5 *(continuación)* Distribución F ($\alpha = 0.05$ en la cola derecha)

Grados de libertad en el denominador (df_2)	Grados de libertad en el numerador (df_1) 1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

(*continúa*)

TABLA A-5 *(continuación)* Distribución F ($\alpha = 0.05$ en la cola derecha)

Grados de libertad en el denominador (df_2)	Grados de libertad en el numerador (df_1)									
	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487	19.496
3	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

Con base en datos de Maxine Merrington y Catherine M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution", *Biometrika 33* (1943): 80-84.

TABLA A-6 Valores críticos del coeficiente de correlación de Pearson r

n	$\alpha = .05$	$\alpha = .01$
4	.950	.990
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: $\rho = 0$ (sin correlación) contra H_1: $\rho \neq 0$ (correlación), rechace H_0 si el valor absoluto de ρ es mayor o igual que el valor crítico en la tabla.

TABLA A-7 Valores críticos para la prueba del signo

	α			
	.005 (una cola)	.01 (una cola)	.025 (una cola)	.05 (una cola)
n	.01 (dos colas)	.02 (dos colas)	.05 (dos colas)	.10 (dos colas)
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, por lo que no se rechaza la hipótesis nula.
2. Rechace la hipótesis nula si el número de signos del tipo menos frecuente (x) es menor o igual que el valor de la tabla.
3. Para valores de n mayores que 25, se usa 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, por lo que no se rechaza la hipótesis nula.
2. Conclusiones:
   Rechace la hipótesis nula si el estadístico de prueba *T* es menor o igual que el valor crítico encontrado en esta tabla.
   No rechace la hipótesis nula si el estadístico de prueba *T* es mayor que el valor crítico encontrado en la tabla.

Con base en datos de *Some Rapid Approximate Statistical Procedures*, Copyright © 1949, 1964 Lederle Laboratories Division of American Cyanamid Company.

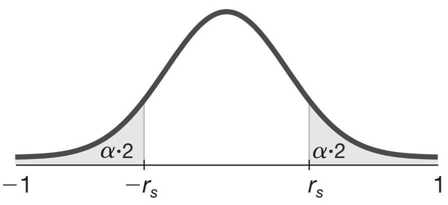


TABLA A-9 Valores críticos del coeficiente de correlación de rangos de Spearman r_s

n	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	.929
8	.643	.738	.833	.881
9	.600	.700	.783	.833
10	.564	.648	.745	.794
11	.536	.618	.709	.755
12	.503	.587	.678	.727
13	.484	.560	.648	.703
14	.464	.538	.626	.679
15	.446	.521	.604	.654
16	.429	.503	.582	.635
17	.414	.485	.566	.615
18	.401	.472	.550	.600
19	.391	.460	.535	.584
20	.380	.447	.520	.570
21	.370	.435	.508	.556
22	.361	.425	.496	.544
23	.353	.415	.486	.532
24	.344	.406	.476	.521
25	.337	.398	.466	.511
26	.331	.390	.457	.501
27	.324	.382	.448	.491
28	.317	.375	.440	.483
29	.312	.368	.433	.475
30	.306	.362	.425	.467

NOTAS:

1. Para $n > 30$ use $r_s \pm z/\sqrt{n-1}$, donde z corresponde al nivel de significancia. Por ejemplo, si $\alpha = 0.05$, entonces $z = 1.96$.
2. Si el valor absoluto del estadístico de prueba r_s es mayor o igual que el valor crítico positivo, entonces rechace H_0: $\rho_s = 0$ y concluya que hay suficiente evidencia para respaldar la afirmación de una correlación.

Con base en datos de *Biostatistical Analysis*, 4a. edición © 1999, de Jerrold Zar, Prentice Hall, Inc., Upper Saddle River, Nueva Jersey, y "Distribution of Sums of Squares of Rank Differences to Small Numbers with Individuals", *The Annals of Mathematical Statistics*, vol. 9, núm. 2.

TABLA A-10 Prueba de rachas para aleatoriedad: valores críticos para el número G de rachas

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

NOTAS:

1. Las entradas en esta tabla son los valores G críticos, suponiendo una prueba de dos colas con un nivel de significancia de $\alpha = 0.05$.
2. Rechace la hipótesis nula de aleatoriedad si se cumple alguna de las siguientes condiciones:
   - El número G de rachas es menor o igual que la entrada más pequeña en la tabla.
   - El número G de rachas es mayor o igual que la entrada más grande en la tabla.

Tomado de "Tables for Testing Randomness of Groupings in a Sequence of Alternatives", *The Annals of Mathematical Statistics*, vol. 14, núm. 1. Reproducido con autorización del Institute of Mathematical Statistics.

APÉNDICE B Conjuntos de datos

Los conjuntos completos de datos están disponibles en www.pearsonenespañol.com/triola

En este apéndice se listan sólo las primeras cinco filas de cada conjunto de datos (como referencia se anota el título en español aunque los datos y su descripción están en inglés). Los conjuntos de datos completos pueden descargarse en www.pearsonenespañol.com/triola para uso en varias tecnologías, entre las que se incluyen Excel, calculadoras TI-83/84 Plus y Minitab. Estos conjuntos de datos se incluyen con Statdisk, que está disponible de forma gratuita para los usuarios de este libro y se puede descargar en www.pearsonenespañol.com/triola.

Conjunto de datos 1: Datos corporales

Las mediciones y exámenes corporales son de 300 sujetos (las primeras cinco filas se muestran aquí). **EDAD** se da en años, para **GÉNERO** 1 = masculino y 0 = femenino, **PULSO** es la frecuencia cardíaca (latidos por minuto), **SISTÓLICA** es la presión arterial sistólica (mm Hg), **DIASTÓLICA** es la presión arterial diastólica (mm Hg), **HDL** es el colesterol HDL (mg/dL), **LDL** es el colesterol LDL (mg/dL), **BLANCOS** es el conteo de glóbulos blancos (1000 células/μl), **ROJOS** es el conteo de glóbulos rojos (millones de células/μL), **PLAQUETAS** es el conteo de plaquetas (1000 células/μL), **PESO** es el peso (kg), **ESTATURA** es la altura (cm), **CINTURA** es la circunferencia de la cintura (cm), **CIRC BRAZO** es la circunferencia del brazo (cm), e **IMC** es el índice de masa corporal (kg/m^2). Los datos provienen del Centro Nacional de Estadísticas de la Salud.

Lista de nombres para TI-83/84 (BODY): AGE, GENDR, PULSE, SYS, DIAS, HDL, LDL, WHITE, REDBC, PLATE, WT, HT, WAIST, ARMC, BMI

EDAD	GÉNERO (1 = M)	PULSO	SISTÓLICA	DIASTÓLICA	HDL	LDL	BLANCOS	PLAQUETAS	ROJOS	PESO	ESTATURA	CINTURA	CIRC BRAZO	IMC
43	0	80	100	70	73	68	8.7	4.80	319	98.6	172.0	120.4	40.7	33.3
57	1	84	112	70	35	116	4.9	4.73	187	96.9	186.0	107.8	37.0	28.0
38	0	94	134	94	36	223	6.9	4.47	297	108.2	154.4	120.3	44.3	45.4
80	1	74	126	64	37	83	7.5	4.32	170	73.1	160.5	97.2	30.3	28.4
34	1	50	114	68	50	104	6.1	4.95	140	83.1	179.0	95.1	34.0	25.9

Conjunto de datos 2: Pie y estatura

Las medidas del pie y la estatura son de 40 sujetos (las primeras cinco filas se muestran aquí). **SEXO** es el sexo del sujeto, **EDAD** es la edad en años, **LONGITUD PIE** es la longitud del pie (cm), **HUELLA ZAPATO** es la longitud del zapato (cm), **TAMAÑO ZAPATO** es el tamaño reportado del zapato, y **ESTATURA** es altura (cm) del sujeto.

Datos de Rohren, Brenda, "Estimation of Stature from Foot and Shoe Length: Applications in Forensic Science". Copyright © 2006. Reproducido con autorización de la autora. Brenda Rohren (MA, MFS, LIMHP, LADC, MAC) era una estudiante de posgrado en la Universidad Wesleyan de Nebraska cuando realizó la investigación y escribió el informe.

Lista de nombres para TI-83/84 (FOOTHT): FTSEX (1 = masculino), FTAGE, FTLN, SHOPT, SHOSZ, FHT

SEXO	EDAD	LONGITUD PIE	HUELLA ZAPATO	TAMAÑO ZAPATO	ESTATURA
M	67	27.8	31.3	11.0	180.3
M	47	25.7	29.7	9.0	175.3
M	41	26.7	31.3	11.0	184.8
M	42	25.9	31.8	10.0	177.8
M	48	26.4	31.4	10.0	182.3

Conjunto de datos 3: Temperaturas corporales

Las temperaturas corporales (en °F) son de 107 sujetos, tomadas en dos días consecutivos a las 8 AM y 12 AM (las primeras cinco filas se muestran aquí). **SEXO** es el género del sujeto, y **FUMA** indica si el sujeto fuma (Y) o no fuma (N). Datos proporcionados por el Dr. Steven Wasserman, el Dr. Philip Mackowiak, y el Dr. Myron Levine de la Universidad de Maryland.

Lista de nombres para TI-83/84 (BODYTEMP): D1T8, D1T12, D2T8, D2T12 (sin lista para SEX y SMOKE). **Los valores de datos faltantes se representan con 9999.**

SEXO	FUMA	DÍA 1-8 AM	DÍA 1-12 AM	DÍA 2-8 AM	DÍA 2-12 AM
M	Y	98.0	98.0	98.0	98.6
M	Y	97.0	97.6	97.4	—
M	Y	98.6	98.8	97.8	98.6
M	N	97.4	98.0	97.0	98.0
M	N	98.2	98.8	97.0	98.0

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 4: Nacimientos

Los datos son de 400 nacimientos (las primeras cinco filas se muestran aquí). Para **GÉNERO,** 1 = masculino y 0 = femenino. **DURACIÓN DE ESTANCIA** es en días, el **PESO AL NACER** está en gramos y los **CARGOS TOTALES** son en dólares.

Lista de nombres para TI-83/84 (BIRTHS): FLOS, MLOS, FBWT, MBWT, FCHRG, MCHRG [Se proporcionan listas separadas para niñas (F) y niños (M). Sin lista para FACILIT, INSURANCE ADMITTED y DISCHARGED]

INSTALACIÓN	SEGURO	GÉNERO (1 = M)	DURACIÓN DE ESTANCIA	ADMISIÓN	SALIDA	PESO AL NACER	CARGOS TOTALES
Albany Medical Center Hospital	Compañía de seguros	0	2	VIE	DOM	3500	13986
Albany Medical Center Hospital	Blue Cross	1	2	VIE	DOM	3900	3633
Albany Medical Center Hospital	Blue Cross	0	36	MIE	JUE	800	359091
Albany Medical Center Hospital	Compañía de seguros	1	5	LUN	SAB	2800	8537
Albany Medical Center Hospital	Compañía de seguros	1	2	VIE	DOM	3700	3633

Conjunto de datos 5: Estaturas familiares

Los datos de estatura son de 134 familias (las primeras cinco filas se muestran aquí). Las estaturas se dan en pulgadas. Sólo se incluyen familias con al menos un hijo de cada sexo, y se consideran las estaturas del primer hijo y la primera hija. Los datos provienen de una publicación de Francis Galton (1822-1911), quien desarrolló los conceptos de desviación estándar, línea de regresión y correlación entre dos variables.

Lista de nombres para TI-83/84 (FAMHT): DAD, MOM, SON1, DGHT1

PADRE	MADRE	PRIMER HIJO	PRIMERA HIJA
70.0	64.0	68.0	65.0
71.0	65.5	72.0	66.0
69.0	63.5	70.5	65.0
69.5	66.0	71.0	66.5
70.0	58.0	72.0	66.0

Conjunto de datos 6: Estudiantes de primer año 15

Se proporcionan pesos de 67 estudiantes universitarios (las primeras cinco filas se muestran aquí). **SEXO** es el sexo del sujeto, **PESO** es el peso en kilogramos, y el **IMC** es el índice de masa corporal medido. Las mediciones se realizaron en septiembre del primer año y luego en abril del primer año.

Los resultados se publican en Hoffman, D. J., Policastro, P., Quick, V. y Lee, S. K.: "Changes in Body Weight and Fat Mass of Men and Women in the First Year of College: A Study of the 'Freshman 15.'" *Journal of American College Health*, 1o. de julio, 2006, vol. 55, núm. 1, p. 41. Copyright © 2006. Reproducido con autorización.

Lista de nombres para TI-83/84 (FRESH15): WTSP, WTAPR, BMISP, BMIAP, (sin lista para SEX)

SEXO	PESO SEPT	PESO ABRIL	IMC SEPT	IMC ABRIL
M	72	59	22.02	18.14
M	97	86	19.70	17.44
M	74	69	24.09	22.43
M	93	88	26.97	25.57
F	68	64	21.51	20.10

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 7: IQ y plomo

Los datos son de 121 sujetos (las primeras cinco filas se muestran aquí). Los datos se midieron en niños en dos años consecutivos, y los niños vivían cerca de una fundición de plomo. **PLOMO** es el grupo de plomo en la sangre [1 = *nivel bajo de plomo* (niveles de plomo en la sangre < 40 microgramos/100 ml en ambos años), 2 = *nivel de plomo medio* (niveles de plomo en sangre ≥ 40 microgramos/100 ml en exactamente uno de dos años), 3 = *nivel alto de plomo* (nivel de plomo en la sangre ≥ 40 microgramos/100 ml en ambos años)]. **EDAD** es la edad en años, **SEXO** es el sexo del sujeto (1 = hombre, 2 = mujer). **AÑO1** es el nivel de plomo en sangre en el primer año y **AÑO2** es el nivel de plomo en sangre en el segundo año. **IQ VERB** se mide con la puntuación del IQ verbal. **IQ DESE** se mide con la puntuación del IQ de desempeño. **IQ COMP** se mide con la puntuación del IQ completo.

Los datos provienen de "Neuropsychological Dysfunction in Children with Chronic Low-Level Lead Absorption", de P. J. Landrigan, R. H. Whitworth, R. W. Baloh, N. W. Staehling, W. F Barthel, y B. F. Rosenblum, *Lancet*, vol. 1, núm. 7909.

Lista de nombres para TI-83/84 (IQLEAD): LEAD, IQAGE, IQSEX, YEAR1, YEAR2, IQV, IQP, IQF

PLOMO	EDAD	SEXO	AÑO1	AÑO2	IQ VERB	IQ DESE	IQ COMP
1	11	1	25	18	61	85	70
1	9	1	31	28	82	90	85
1	11	1	30	29	70	107	86
1	6	1	29	30	72	85	76
1	11	1	2	34	72	100	84

Conjunto de datos 8: IQ y tamaño del cerebro

Los datos son de 20 gemelos monocigóticos (idénticos) (las primeras cinco filas se muestran aquí). **PAR** identifica el conjunto de gemelos, **SEXO** es el género del sujeto (1 = masculino, 2 = femenino), **ORDEN** es el orden de nacimiento, **IQ** se mide con la puntuación del IQ completo, **VOL** es el volumen cerebral total (cm^3), **ÁREA** es el área total superficial del cerebro (cm^2), **CCSA** es el área superficial (cm^2) del cuerpo calloso (que conecta los hemisferios cerebrales izquierdo y derecho), **CIRC** es la circunferencia de la cabeza (cm) y **PESO** es el peso corporal (kg).

Datos proporcionados por M. J. Tramo, W. C. Loftus, T. A. Stukel, J. B. Weaver, M. S. Gazziniga. Consulte "Brain Size, Head Size, and IQ in Monozygotic Twins", *Neurology*, vol. 50.

Lista de nombres para TI-83/84 (IQBRAIN): PAIR, SEX, ORDER, IQ, VOL, AREA, CCSA, CIRC, BWT

PAR	SEXO(1 = M)	ORDEN	IQ	VOL	ÁREA	CCSA	CIRC	PESO
1	2	1	96	1005	1913.88	6.08	54.7	57.607
1	2	2	89	963	1684.89	5.73	54.2	58.968
2	2	1	87	1035	1902.36	6.22	53.0	64.184
2	2	2	87	1027	1860.24	5.80	52.9	58.514
3	2	1	101	1281	2264.25	7.99	57.8	63.958

Conjunto de datos 9: Mediciones de osos

Los datos provienen de 54 osos salvajes anestesiados (las primeras cinco filas se muestran aquí). **EDAD** se da en meses, **MES** es el mes de medición 1 = enero, **SEXO** está codificado con 0 = mujer y 1 = hombre, **LONGCAB** es la longitud de la cabeza (pulgadas), **ANCCAB** es el ancho de la cabeza (pulgadas), **CUELLO** es distancia alrededor del cuello (en pulgadas), **LONGITUD** es la longitud del cuerpo (pulgadas), **PECHO** es la distancia alrededor del pecho (pulgadas) y **PESO** se mide en libras. Los datos son de Gary Alt y Minitab, Inc.

Lista de nombres para TI-83/84 (BEARS): BAGE, BSEX, BHDLN, BHDWD, BNECK, BLEN, BCHST, BWGHT (sin lista para MONTH)

EDAD	MES	SEXO(1 = M)	LONGCAB	ANCCAB	CUELLO	LONGITUD	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	0	15.5	6.5	22.0	62.0	35.0	166

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 10: Muertes de manatíes

Se proporcionan datos anuales para Florida durante 24 años (las primeras cinco filas se muestran aquí). **MUERTES** es el número anual de muertes de manatíes causadas por embarcaciones, **BARCOS** es el número de embarcaciones de recreo registradas (decenas de miles), **POB** es la población de Florida (millones) y **TEMP AGUA** es la temperatura media anual del agua (en °F).

Lista de nombres para TI-83/84 (MANATEE): DEATH, BOATS, POP, WTEMP (sin lista para YEAR)

AÑO	MUERTES	BARCOS	POB	TEMP AGUA
1991	53	68	13.3	71.9
1992	38	68	13.5	70.4
1993	35	67	13.7	70.5
1994	49	70	14.0	71.7
1995	42	71	14.3	70.9

Conjunto de datos 11: Alcohol y tabaco en las películas

Los datos provienen de 50 películas animadas para niños (las primeras cinco filas se muestran aquí). **DURACIÓN** es la duración de la película en minutos, **TABACO** es el tiempo de uso de tabaco en segundos y **ALCOHOL** es el tiempo de consumo de alcohol en segundos. Los datos se basan en Goldstein, Adam O., Sobel, Rachel A., Newman, Glen R., "Tobacco and Alcohol Use in G-Rated Children's Animated Films". *Journal of the American Medical Association*, 24-31 de marzo, 1999, vol. 281, núm. 12, p. 1132. Copyright © 1999. Todos los derechos reservados.

Lista de nombres TI-83/84 (CHMOVIE): CHLEN, CHTOB, CHALC (sin lista para MOVIE y STUDIO)

PELÍCULA	ESTUDIO	DURACIÓN (MIN)	TABACO (SEG)	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

Conjunto de datos 12: Fumadores pasivos y activos

Los datos son de 120 sujetos (las primeras cinco filas se muestran aquí) en tres grupos: **FUMADOR** incluye a los sujetos que fuman, **ETS** incluye no fumadores expuestos al humo de tabaco ambiental, **NOETS** incluye no fumadores no expuestos al humo de tabaco ambiental. Todos los valores son niveles medidos de cotinina sérica (en ng/ml), un metabolito de la nicotina. (Cuando el cuerpo absorbe nicotina, se produce la cotinina). Los datos provienen del Departamento de Salud y Servicios Humanos de Estados Unidos, tomados del National Center for Health Statistics, Third National Health and Nutrition Examination Survey.

Lista de nombres para TI-83/84 (SMOKE): SMKR, ETS, NOETS

FUMADOR	ETS	NOETS
1	384	0
0	0	0
131	69	0
173	19	0
265	1	0

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 13: Contenido de cigarrillos

Los datos son de 75 cigarrillos (las primeras cinco filas se muestran aquí) en tres categorías: **KING** incluye cigarrillos tamaño *king* sin filtro, no mentolados y no *light*; **MENT** incluye cigarrillos mentolados de 100 mm de largo, filtrados y no *light*; y **100** incluye cigarrillos de 100 mm de longitud con filtro, no mentolados y no *light*. **ALQ** es la cantidad de alquitrán por cigarrillo (miligramos), **NICOTINA** es la cantidad de nicotina por cigarrillo (miligramos) y **CO** es la cantidad de monóxido de carbono por cigarrillo (miligramos). Los datos son de la Comisión Federal de Comercio.

Lista de nombres para TI-83/84 (CIGARET): KGTAR, KGNIC, KGCO, MNTAR, MNNIC, MNCO, FLTAR, FLNIC, FLCO

ALQ KING	NICOTINA KING	CO KING	ALQ MENT	NICOTINA MENT	CO MENT	ALQ 100	NICOTINA 100	CO 100
20	1.1	16	16	1.1	15	5	0.4	4
27	1.7	16	13	0.8	17	16	1.0	19
27	1.7	16	16	1.0	19	17	1.2	17
20	1.1	16	9	0.9	9	13	0.8	18
20	1.1	16	14	0.8	17	13	0.8	18

Conjunto de datos 14: Edades de ganadores del Óscar

Los datos son de 87 años (las primeras cinco filas se muestran aquí). Los valores de datos son edades (años) de actrices y actores en el momento en que ganaron el Óscar en las categorías de Mejor Actriz y Mejor Actor. Las edades se listan en orden cronológico por fila, por lo que cada fila tiene edades pareadas del mismo año. (*Nota:* En 1968 hubo un empate en la categoría de Mejor Actriz, y se usa la media de las dos edades; en 1932 hubo un empate en la categoría de Mejor Actor, y se usa la media de las dos edades).

Estos datos son sugeridos por el artículo "Ages of Oscar-Winning Best Actors and Actresses", de Richard Brown y Gretchen Davis, *Mathematics Teacher Magazine*. En ese artículo, el año de nacimiento del ganador del premio se restó del año de la ceremonia de premiación, pero las edades que figuran aquí se calculan a partir de la fecha de nacimiento del ganador y la fecha de la ceremonia de premiación.

Lista de nombres para TI-83/84 (OSCARS): OSCRF, OSCRM

ACTRICES	ACTORES
22	44
37	41
28	62
63	52
32	41

Conjunto de datos 15: Presidentes

Los datos provienen de 38 presidentes de Estados Unidos (las primeras cinco filas se muestran aquí). Los presidentes que asumieron el cargo como resultado de un asesinato o renuncia no están incluidos. **EDAD** es la edad en años en el momento de la investidura. **DÍAS** es la cantidad de días que el sujeto sirvió como presidente. **AÑOS** es el número de años vividos después de la primera investidura. **ESTATURA** es la altura (cm) del presidente. **ESTATURA OPO** es la altura (cm) del principal oponente para la presidencia.

Lista de nombres para TI-83/84 (POTUS): PRAGE, DAYS, YEARS, PRHT, HTOPP (no hay lista para PRESIDENT). **Los valores de los datos faltantes se representan por 9999.**

PRESIDENTE	EDAD	DÍAS	AÑOS	ESTATURA	ESTATURA OPO
Washington	57	2864	10	188	
J. Adams	61	1460	29	170	189
Jefferson	57	2921	26	189	170
Madison	57	2921	28	163	
Monroe	58	2921	15	183	

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 16: Premios Nobel y chocolate

Los datos provienen de 23 países (las primeras cinco filas se muestran aquí). **CHOCOLATE** incluye el consumo de chocolate (kg *per cápita*), **NOBEL** incluye el número de Premios Nobel (por cada 10 millones de personas), **POBLACIÓN** incluye la población (en millones) e **INTERNET** incluye la cantidad de usuarios de Internet por cada 100 personas.

Los datos son de "The Real Secret to Genius? Reading Between the Lines", de McClintock, Stangle y Cetinkaya-Rundel, *Chance*, vol. 27, núm. 1; y "Chocolate Consumption, Cognitive Function, and Nobel Laureates", de Franz Messerli, *New England Journal of Medicine*, vol. 367, núm. 16.

Lista de nombres para TI-83/84 (NOBEL): CHOC, NOBEL, POPUL, INTNT (sin lista para COUNTRY)

PAÍS	CHOCOLATE	NOBEL	POBLACIÓN	INTERNET
Australia	4.5	5.5	22	79.5
Austria	10.2	24.3	8	79.8
Bélgica	4.4	8.6	11	78.0
Brasil	2.9	0.1	197	45.0
Canadá	3.9	6.1	34	83.0

Conjunto de datos 17: Evaluaciones de cursos

Los datos provienen de 93 evaluaciones de cursos realizadas por estudiantes universitarios (las primeras cinco filas se muestran aquí). **EVAL CURSO** incluye la calificación media del curso, **EVAL PROF** incluye la calificación media del profesor, **EDAD PROF** incluye la edad del profesor en años, **TAMAÑO** incluye el número de evaluaciones registradas en cada curso, **BELLEZA PROF** incluye la calificación media de la belleza basada en la foto del profesor. Según los datos de Andrew Gelman y Jennifer Hill, 2007, "Replication Data for Data Analysis Using Regression Multilevel Hierarchical Models", http://hdl.handle.net/1902.1/10285.

Lista de nombres para TI-83/84 (EVALS): CEVAL, PEVAL, PAGE, SIZE, PBTY (sin lista para PROF, PROF GENDER, PROF PHOTO y CLASS LEVEL)

PROF	EVAL CURSO	EVAL PROF	GÉNERO PROF	EDAD PROF	TAMAÑO	BELLEZA PROF	FOTO PROF	NIVEL CLASE
1	4.3	4.7	mujer	36	24	5.000	color	superior
2	4.5	4.6	hombre	59	17	3.000	color	superior
3	3.7	4.1	hombre	51	55	3.333	color	superior
4	4.3	4.5	mujer	40	40	3.167	color	superior
5	4.4	4.8	mujer	31	42	7.333	color	superior

Conjunto de datos 18: Citas rápidas

Los datos son de 199 citas (las primeras cinco filas se muestran aquí). **DEC DE MUJ** es la decisión (1 = sí) de la mujer sobre tener una nueva cita, **EDAD MUJ** es la edad de la mujer, **ME GUSTA POR MUJ** es la calificación de cuánto "le gusta" el varón a la mujer (escala del 1 al 10), **ATRACT POR MUJ** es la calificación de "atractivo" dada por la mujer acerca del hombre (escala de 1 al 10), **ATRIB POR MUJ** es la suma de calificaciones de cinco atributos (sinceridad, inteligencia, diversión, ambición, intereses compartidos) dadas por la mujer al hombre. Los datos para hombres usan los descriptores correspondientes. Las clasificaciones de escala más alta corresponden a impresiones más positivas.

Datos replicados de *Data Analysis Using Regression and Multilevel/Hierarchical Models*, de Andrew Gelman y Jennifer Hill, Cambridge University Press.

Lista de nombres para TI-83/84 (DATE): DBYF, DBYM, AGEF, AGEM, LBYF, LBYM, ABYF, ABYM, ATBYF, ATBYM

DEC DE MUJ	DEC DE HOM	EDAD MUJ	EDAD HOM	ME GUSTA POR MUJ	ME GUSTA POR HOM	ATRACT POR MUJ	ATRACT POR HOM	ATRIB POR MUJ	ATRIB POR HOM
0	1	27	28	7	7	5	8	38	36
0	0	24	26	7	7	7	6	36	38
1	0	26	28	6	3	7	8	29	35
1	0	34	27	8	6	8	6	38	28
1	0	22	25	5	5	7	5	33	34

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 19: Pruebas de choques automovilísticos

Los datos provienen de 21 autos usados en pruebas de choque (las primeras cinco filas se muestran aquí). Los mismos automóviles se utilizan en el conjunto de datos 20. Los datos son mediciones de automóviles que se impactaron contra una barrera fija a 35 mi/h con un maniquí de pruebas de choque en el asiento del conductor. **HIC** es una medida de un criterio estándar de "lesión en la cabeza", **PECHO** es la desaceleración del pecho (en *g*, donde *g* es la fuerza de la gravedad), **FEMUR I** es la carga medida en el fémur izquierdo (en lb), **FEMUR D** es la carga medida en el fémur derecho (en lb), **TTI** es una medición del índice de trauma torácico lateral y **PELVIS** es la desaceleración de la pelvis (en *g*, donde *g* es la fuerza de la gravedad). Los datos provienen de la Oficina Nacional de Seguridad del Tráfico en Carreteras.

Lista de nombres para TI-83/84 (CRASH): HIC, CHEST, FEML, FEMR, TTI, PLVS (sin lista para CAR y SIZE)

AUTO	TAMAÑO	HIC	PECHO	FEMUR I	FEMUR D	TTI	PELVIS
Chev Aveo	Pequeño	371	44	1188	1261	62	71
Honda Civic	Pequeño	356	39	289	324	63	71
Mitsubishi Lancer	Pequeño	275	37	329	446	35	45
VW Jetta	Pequeño	544	54	707	1048	44	66
Hyundai Elantra	Pequeño	326	39	602	1474	58	71

Conjunto de datos 20: Mediciones de automóviles

Los datos provienen de 21 autos (las primeras cinco filas se muestran aquí). Se utilizan los mismos automóviles que en el conjunto de datos 19. Los datos son mediciones de automóviles que tienen transmisiones automáticas. **PESO** es el peso del automóvil (lb), **LONGITUD** es la longitud del automóvil (pulgadas), **FRENADO** es distancia de frenado (pies) desde 60 mi/h, **CILINDROS** es el número de cilindros, **DESPLAZAMIENTO** es el desplazamiento del motor (litros), **CIUDAD** es el consumo de combustible (mi/gal) para condiciones de conducción en la ciudad, **CARRETERA** es el consumo de combustible (mi/gal) para las condiciones de conducción en carretera y **GEI** es una medida de las emisiones de gases de efecto invernadero (en toneladas/año, expresadas como equivalentes de CO_2). Los datos provienen de la Oficina Nacional de Seguridad del Tráfico en Carreteras.

Lista de nombres para TI-83/84 (CARS): CWT, CLN, CBRK, CCYL, CDISP, CCITY, CHWY, CGHG (sin lista para CAR y SIZE)

AUTO	TAMAÑO	PESO	LONGITUD	FRENADO	CILINDROS	DESPLAZAMIENTO	CIUDAD	CARRETERA	GEI
Chev Aveo	Pequeño	2560	154	133	4	1.6	25	34	6.6
Honda Civic	Pequeño	2740	177	136	4	1.8	25	36	6.3
Mitsubishi Lancer	Pequeño	3610	177	126	4	2.0	22	28	7.7
VW Jetta	Pequeño	3225	179	137	4	2.0	29	40	6.4
Hyundai Elantra	Pequeño	2895	177	138	4	2.0	25	33	6.6

Conjunto de datos 21: Terremotos

Los datos son de 600 datos pareados (las primeras cinco filas se muestran aquí) de mediciones de magnitud/profundidad seleccionadas aleatoriamente de 10,594 terremotos registrados en un año en una ubicación en el sur de California. Sólo se usan terremotos con una magnitud de al menos 1.00. **MAGNITUD** es la magnitud medida en la escala de Richter y **PROFUNDIDAD** es la profundidad en km. La magnitud y la profundidad describen la fuente del terremoto. Los datos provienen del Centro de Datos de Terremotos del Sur de California.

Lista de nombres para TI-83/84 (QUAKE): MAG, DEPTH

MAGNITUD	PROFUNDIDAD
2.45	0.7
3.62	6.0
3.06	7.0
3.30	5.4
1.09	0.5

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 22: Tornados

Los datos son de 500 tornados (las primeras cinco filas se muestran aquí) ordenados cronológicamente. **MES** es el mes del tornado (1 = enero), **ESCALA F** es la escala Fujita de la intensidad del tornado, **MUERTES** es el número de muertes causadas por el tornado, **LONGITUD (MI)** es la distancia que el tornado recorrió en millas y **ANCHO (YD)** es el ancho del tornado en yardas. Los datos provienen del Servicio Meteorológico Nacional.

Lista de nombres para TI-83/84 (TORNADO): FSCAL, FATAL, TLEN, TWDTH (sin lista para YEAR y MONTH). **Los 10 valores faltantes de la escala F están representados por 9999.**

AÑO	MES	ESCALA F	MUERTES	LONGITUD (MI)	ANCHO (YD)
1950	5	2	0	3.60	100
1950	5	4	0	34.30	150
1950	6	1	0	0.20	10
1950	8	1	0	0.10	10
1951	6	–	0	0.10	10

Conjunto de datos 23: *Old Faithful*

Los datos son de 250 erupciones (las primeras cinco filas se muestran aquí) del géiser Old Faithful en el Parque Nacional Yellowstone. **INT ANTES** es el intervalo de tiempo (min) anterior a la erupción, **DURACIÓN** es el tiempo (seg) de la erupción, **INT DESPUÉS** es el intervalo de tiempo (min) después de la erupción, **ALTURA (ft)** es la altura de la erupción y **ERROR PRED** es el error (min) del tiempo predicho para la erupción. De acuerdo con datos de la Asociación de Observación y Estudio de Géiseres.

Lista de nombres para TI-83/84 (OLDFAITH): INTBF, DUR, INTAF, OFHT, PRED

INT ANTES (MIN)	DURACIÓN (SEG)	INT DESPUÉS (MIN)	ALTURA (FT)	ERROR PRED (MIN)
82	251	83	130	4
99	243	76	125	−13
88	250	86	120	−2
92	240	82	120	−6
86	243	87	130	−0

Conjunto de datos 24: Las palabras cuentan

Los datos provienen de conteos de las palabras habladas en un día por 396 sujetos masculinos (M) y femeninos (F) en seis grupos de muestras diferentes (las primeras cinco filas se muestran aquí). La columna **M1** expresa el conteo de palabras para los hombres en la muestra 1, **F1** es el conteo de las mujeres en la muestra 1, y así sucesivamente.

Muestra 1: Parejas reclutadas con edades comprendidas entre los 18 y los 29 años
Muestra 2: Estudiantes reclutados en clases introductorias de psicología, de 17 a 23 años
Muestra 4: Estudiantes reclutados en clases introductorias de psicología, de 17 a 22 años
Muestra 5: Estudiantes reclutados en clases introductorias de psicología, de 18 a 26 años
Muestra 6: Estudiantes reclutados en clases introductorias de psicología, de 17 a 23 años

Los resultados se publicaron en "Are Women Really More Talkative Than Men?" de Mehl, Vazire, Ramírez-Esparza, Slatcher, Pennebaker, *Science*, vol. 317, núm. 5834.

Lista de nombres para TI-83/84 (WORDS): M1, F1, M2, F2, M3, F3, M4, F4, M5, F5, M6, F6.

M1	F1	M2	F2	M3	F3	M4	F4	M5	F5	M6	F6
27531	20737	23871	16109	21143	6705	47016	11849	39207	15962	28408	15357
15684	24625	5180	10592	17791	21613	27308	25317	20868	16610	10084	13618
5638	5198	9951	24608	36571	11935	42709	40055	18857	22497	15931	9783
27997	18712	12460	13739	6724	15790	20565	18797	17271	5004	21688	26451
25433	12002	17155	22376	15430	17865	21034	20104		10171	37786	12151

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 25: Comida rápida

Los datos son de 400 observaciones (las primeras cinco filas se muestran aquí) de los tiempos de servicio en automóvil (seg) en diferentes restaurantes de comida rápida. Los tiempos comienzan cuando un vehículo se detiene en la ventanilla de pedido y finaliza cuando el vehículo sale de la ventanilla de entrega. Los tiempos de la comida se midieron entre las 11:00 AM y las 2:00 PM y los horarios de las cenas se midieron entre las 4:00 PM y las 7:00 PM. Datos recopilados por el autor.

Lista de nombre para TI-83/84 (FASTFOOD): MCDL, MCDD, BKL, BKD, WL, WD, DDL, DDD

COMIDA MCDONALDS	CENA MCDONALDS	COMIDA BURGER KING	CENA BURGER KING	COMIDA WENDYS	CENA WENDYS	COMIDA DUNKIN DONUTS	CENA DUNKIN DONUTS
107	84	116	101	466	56	86	181
139	121	131	126	387	82	201	50
197	119	147	153	368	120	179	177
209	146	120	116	219	116	131	107
281	266	126	175	177	121	126	68

Conjunto de datos 26: Pesos y volúmenes de bebidas de cola

Los datos son de 144 latas de bebidas cola (las primeras cinco filas se muestran aquí). **PESO** es el peso en libras y **VOL** es el volumen en onzas.

Lista de nombres para TI-83/84 (COLA): CRGWT, CRGVL, CDTWT, CDTVL, PRGWT, PRGVL, PDTWT, PDTVL

PESO COCA REG	VOL COCA REG	PESO DIET COKE	VOL DIET COKE	PESO PEPSI REG	VOL PEPSI REG	PESO DIET PEPSI	VOL PEPSI DIET
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

Conjunto de datos 27: Pesos de M&Ms

Los datos son de 100 pesos (gramos) de caramelos simples M&M (las primeras cinco filas se muestran aquí). Datos recopilados por el autor.

Lista de datos TI-83/84 (MM): RED, ORNG, YLLW, BROWN, BLUE, GREEN

ROJO	NARANJA	AMARILLO	CAFÉ	AZUL	VERDE
0.751	0.735	0.883	0.696	0.881	0.925
0.841	0.895	0.769	0.876	0.863	0.914
0.856	0.865	0.859	0.855	0.775	0.881
0.799	0.864	0.784	0.806	0.854	0.865
0.966	0.852	0.824	0.840	0.810	0.865

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 28: Galletas con chispas de chocolate

Los datos provienen de 170 galletas con chispas de chocolate (las primeras cinco filas se muestran aquí). Las marcas son Chips Ahoy regular, Chips Ahoy Chewy, Chips Ahoy reducidas en grasa, Keebler y Hannaford. Los valores son conteos de cantidades de chispas de chocolate en cada galleta. Datos recopilados por el autor.

Lista de nombres TI-83/84 (CHIPS): CAREG, CACHW, CARF, KEEB, HANNA.

CHIPS AHOY REG	CHIPS AHOY CHEWY	CHIPS AHOY RED GRASA	KEEBLER	HANNAFORD
22	21	13	29	13
22	20	24	31	15
26	16	18	25	16
24	17	16	32	21
23	16	21	27	15

Conjunto de datos 29: Pesos de monedas

Los datos son de 222 monedas (las primeras cinco filas se muestran aquí) que consisten en los pesos de las monedas (gramos). Los "centavos anteriores a 1983" se fabricaron después de los peniques indios y de trigo, y son 97% de cobre y 3% de zinc. Los "centavos posteriores a 1983" son 3% de cobre y 97% de zinc. Las "monedas de ¢25 de plata anteriores a 1964" son 90% de plata y 10% de cobre. Las "monedas de ¢25 posteriores a 1964" están hechas con una aleación de cobre y níquel.

Lista de nombres para TI-83/84(MONEDAS): CPIND, CPWHT, CPPRE, CPPST, CPCAN, CQPRE, CQPST, CDOL

PENIQUES DEL INDIO	PENIQUES DE TRIGO	PENIQUES ANTERIORES A 1983	PENIQUES POSTERIORES A 1983	PENIQUES CANADIENSES	CUARTOS ANTERIORES A 1964	CUARTOS POSTERIORES A 1964	MONEDAS DE DÓLAR
3.0630	3.1366	3.1582	2.5113	3.2214	6.2771	5.7027	8.1008
3.0487	3.0755	3.0406	2.4907	3.2326	6.2371	5.7495	8.1072
2.9149	3.1692	3.0762	2.5024	2.4662	6.1501	5.7050	8.0271
3.1358	3.0476	3.0398	2.5298	2.8357	6.0002	5.5941	8.0813
2.9753	3.1029	3.1043	2.4950	3.3189	6.1275	5.7247	8.0241

Conjunto de datos 30: Latas de aluminio

Los datos provienen de 350 latas (las primeras cinco filas se muestran aquí) que consisten en las cargas axiales máximas (libras) medidas. Las cargas axiales se aplican cuando las partes superiores de las latas son presionadas en su sitio. **LATAS 109** incluye latas que tienen un espesor de 0.0109 pulgadas, y **LATAS 111** incluye latas que tienen un espesor de 0.0111 pulgadas.

Lista de nombres TI-83/84 (CANS): CN109, CN111

LATAS 109	LATAS 111
270	287
273	216
258	260
204	291
254	210

(Los conjuntos de datos completos están disponibles en www.pearsonenespañol.com/triola)

Conjunto de datos 31: Peso de la basura

Los datos provienen de 62 hogares (las primeras cinco filas se muestran aquí) que consisten en los pesos (libras) de basura dispuesta en diferentes categorías. **TAMAÑO CASA** es el tamaño del hogar. Datos proporcionados por Masakuza Tani, del Proyecto de Basura, Universidad de Arizona.

Lista de nombres TI-83/84 (GARBAGE): HHSIZ, METAL, PAPER, PLAS, GLASS, FOOD, YARD, TEXT, OTHER, TOTAL

TAMAÑO CASA	METAL	PAPEL	PLÁSTICO	VIDRIO	COMIDA	PATIO	TEXTIL	OTRO	TOTAL
2	1.09	2.41	0.27	0.86	1.04	0.38	0.05	4.66	10.76
3	1.04	7.57	1.41	3.46	3.68	0.00	0.46	2.34	19.96
3	2.57	9.55	2.19	4.52	4.43	0.24	0.50	3.60	27.60
6	3.02	8.82	2.83	4.92	2.98	0.63	2.26	12.65	38.11
4	1.50	8.72	2.19	6.31	6.30	0.15	0.55	2.18	27.90

Conjunto de datos 32: Velocidades de transmisión de datos en aeropuertos

Los datos provienen de 50 aeropuertos (las primeras cinco filas se muestran aquí) que consisten en las velocidades de transmisión de datos (Mbps) de cuatro operadores de telefonía celular. Con base en datos de CNN.

Nombres de lista TI-83/84 (DATASPED): VRZN, SPRNT, ATT, TMOBL (sin lista para AIRPORT CODE)

CÓDIGO AEROPUERTO	VERIZON	SRINT	ATT	T-MOBILE
RSW	38.5	13.0	9.7	8.6
ORD	55.6	30.4	8.2	7.0
SNA	22.4	15.2	7.1	18.5
MEM	14.1	2.4	14.4	16.7
MKE	23.1	2.7	13.4	5.6

APÉNDICE C Sitios web y bibliografía de libros

Sitios web

Triola Stats: www.TriolaStats.com

Acceda a recursos digitales continuamente actualizados para la serie de estadística de Triola, incluidos los conjuntos de datos descargables, complementos de los libros de texto, videos instructivos en línea, el Blog de Triola y más.

Statdisk: www.Statdisk.org

Descargue el software estadístico gratuito Statdisk que está diseñado específicamente para este libro y contiene todos los conjuntos de datos del apéndice B. En la página de ayuda del sitio se proporciona información detallada sobre el uso de Statdisk.

StatCrunch: www.statcrunch.com

Libros

***Un asterisco denota un libro recomendado para su lectura. Los demás libros se recomiendan como textos de referencia.**

Bennett, D. 1998. *Randomness*. Cambridge, Mass.: Harvard University Press.

*Best, J. 2012. *Damned Lies and Statistics*. Berkeley, Calif.: University of California Press.

*Best, J. 2004. *More Damned Lies and Statistics*. Berkeley, Calif.: University of California Press.

*Campbell, S. 2004. *Flaws and Fallacies in Statistical Thinking*. Mineola, N.Y.: Dover Publications.

*Crossen, C. 1996. *Tainted Truth: The Manipulation of Fact in America*. Nueva York: Simon & Schuster.

*Freedman, D., R. Pisani, R. Purves y A. Adhikari. 2007. *Statistics*. 4a. ed. Nueva York: W. W. Norton & Company.

*Gonick, L. y W. Smith. 1993. *The Cartoon Guide to Statistics*. Nueva York: Harper Collins.

*Heyde, C. y E. Seneta, eds. 2001. *Statisticians of the Centuries*. Nueva York: Springer-Verlag.

*Hollander, M. y F. Proschan. 1984. *The Statistical Exorcist: Dispelling Statistics Anxiety*. Nueva York: Marcel Dekker.

*Holmes, C. 1990. *The Honest Truth About Lying with Statistics*. Springfield, Ill.: Charles C. Thomas.

*Hooke, R. 1983. *How to Tell the Liars from the Statisticians*. Nueva York: Marcel Dekker.

*Huff, D. 1993. *How to Lie with Statistics*. Nueva York: W. W. Norton & Company.

*Jaffe, A. y H. Spirer. 1998. *Misused Statistics*. New York: Marcel Dekker.

Kaplan, M. 2007. *Chances Are*. Nueva York: Penguin Group.

Kotz, S. y D. Stroup. 1983. *Educated Guessing—How to Cope in an Uncertain World*. Nueva York: Marcel Dekker.

Lapp, James. 2018. *Student Solutions Manual to Accompany Elementary Statistics*. 13a. ed. Boston: Pearson.

Mlodinow, L. 2009. *The Drunkard's Walk*. Nueva York: Vintage Books.

*Moore, D. y W. Notz. 2012. *Statistics: Concepts and Controversies*. 8a. ed. San Francisco: Freeman.

*Paulos, J. 2001. *Innumeracy: Mathematical Illiteracy and Its Consequences*. Nueva York: Hill and Wang.

*Reichmann, W. 1981. *Use and Abuse of Statistics*. Nueva York: Penguin.

*Rossman, A. y B. Chance. 2011. *Workshop Statistics: Discovery with Data*. 4a. ed. Emeryville, Calif.: Key Curriculum Press.

*Salsburg, D. 2001. *The Lady Tasting Tea: How Statistics Revolutionized the Twentieth Century*. Nueva York: W. H. Freeman.

Sheskin, D. 2011. *Handbook of Parametric and Nonparametric Statistical Procedures*. 5a. ed. Boca Raton, Fla.: CRC Press.

Simon, J. 1997. *Resampling: The New Statistics*. 2a. ed. Arlington, Va.: Resampling Stats.

*Stigler, S. 1986. *The History of Statistics*. Cambridge, Mass.: Harvard University Press.

Taleb, N. 2010. *The Black Swan*. 2a. ed. Nueva York: Random House.

Triola, M. 2018. *Minitab Student Laboratory Manual and Workbook*. 13a. ed. Boston: Pearson.

Triola, M. 2018. *Statdisk 13 Student Laboratory Manual and Workbook*. 13a. ed. Boston: Pearson.

Triola, M. y L. Franklin. 1994. *Business Statistics*. Boston: Addison-Wesley.

Triola, M., M. Triola y J. Roy. 2018. *Biostatistics for the Biological and Health Sciences*. 2a. ed. Boston: Pearson.

*Tufte, E. 2001. *The Visual Display of Quantitative Information*. 2a. ed. Cheshire, Conn.: Graphics Press.

Tukey, J. 1977. *Exploratory Data Analysis*. Boston: Pearson.

Vickers, A. 2009. *What Is a P-Value Anyway?* Boston: Pearson.

Whelan, C. 2013. *Naked Statistics*. Nueva York: W. W. Norton & Company.

Zwillinger, D. y S. Kokoska. 2000. *CRC Standard Probability and Statistics Tables and Formulae*. Boca Raton, Fla.: CRC Press.

APÉNDICE D

Respuestas a ejercicios de sección con número impar, respuestas a todos los exámenes rápidos, ejercicios de repaso y ejercicios de repaso acumulado de los capítulos.

Respuestas del capítulo 1

Sección 1-1

1. Los encuestados son una muestra de respuesta voluntaria o una muestra autoseleccionada. Debido a que las personas con fuertes intereses en el tema tienen más probabilidades de responder, es muy posible que sus respuestas no reflejen las opiniones o el comportamiento de la población en general.
3. La significancia estadística se indica al utilizar métodos estadísticos para llegar a la conclusión de que un tratamiento es efectivo, pero el sentido común podría sugerir que el tratamiento no implica una diferencia suficiente para justificar su uso o para ser práctico. Sí, es posible que un estudio tenga significancia estadística, pero no significancia práctica.
5. Sí, parece haber un potencial para crear un sesgo.
7. No, no parece haber un potencial para crear un sesgo.
9. La muestra es una muestra de respuesta voluntaria y tiene un gran potencial para ser defectuosa.
11. El método de muestreo parece ser correcto.
13. Con sólo 1% de posibilidades de obtener dichos resultados con un programa que no tiene ningún efecto, el programa parece tener significancia estadística. Además, debido a que la pérdida promedio de 22 libras parece sustancial, el programa también parece tener significancia práctica.
15. Debido a que hay 19% de posibilidades de obtener esa cantidad de niñas por casualidad, el método parece carecer de significancia estadística. El resultado de 1020 niñas en 2000 nacimientos (51% de niñas) es superior a la tasa de aproximadamente 50% que se espera por casualidad, pero no parece ser lo suficientemente alta como para tener una significancia práctica. No muchas parejas se molestarían en adoptar un procedimiento que aumenta la probabilidad de tener una niña de 50% a 51%.
17. Sí. Cada columna de temperaturas de 8 AM y 12 AM se registra para el mismo sujeto, por lo que cada par es correspondiente.
19. Los datos se pueden usar para abordar la cuestión de si existe una correlación entre las temperaturas corporales a las 8 AM y a las 12 AM. Además, los datos pueden usarse para determinar si existen diferencias entre las temperaturas corporales a las 8 AM y a las 12 AM.
21. No. Los recuentos de glóbulos blancos miden una cantidad diferente a los recuentos de glóbulos rojos, por lo que sus diferencias no tienen sentido.
23. No. El Centro Nacional de Estadísticas de la Salud no tiene ninguna razón para recopilar o presentar los datos de una manera parcial.
25. Es cuestionable que el patrocinador sea la Comisión de la Papa de Idaho y el vegetal favorito sea la papa.
27. La correlación o asociación entre dos variables no significa que una de las variables sea la causa de la otra. La correlación no implica causalidad. Claramente, el consumo de crema agria no está directamente relacionado de ninguna manera con las muertes en motocicletas.
29. a. 700 adultos b. 55%
31. a. 559.2 encuestados
    b. No. Debido a que el resultado es un conteo de encuestados entre las 1165 mujeres casadas o comprometidas que fueron encuestadas, el resultado debe ser un número entero.
    c. 559 encuestados d. 8%
33. Como una reducción del 100% eliminaría todo el tamaño, no es posible reducir el tamaño en un 100% o más.
35. Debido a que una reducción del 100% eliminaría toda la placa, no es posible reducirla en más del 100%.
37. La redacción de la pregunta es parcial y tiende a alentar respuestas negativas. El tamaño de muestra de 20 es demasiado pequeño. Los encuestados son autoseleccionados en lugar de ser seleccionados al azar por el periódico. Si 20 lectores responden, los porcentajes deben ser múltiplos de 5, por lo que el 87% y el 13% no son resultados posibles.

Sección 1-2

1. La población está compuesta por todos los adultos en Estados Unidos, y la muestra son los 2276 adultos encuestados. Debido a que el valor del 33% se refiere a la muestra, es un dato estadístico.
3. Sólo el inciso (a) describe datos discretos.
5. Estadístico 7. Parámetro 9. Estadístico
11. Parámetro 13. Continuo 15. Discreto
17. Discreto 19. Continuo 21. Ordinal
23. Nominal 25. De intervalo 27. Ordinal
29. Los números no son conteos o medidas de nada. Están en el nivel nominal de medición, y no tiene sentido calcular su promedio (media).
31. Las temperaturas están en el nivel de medición de intervalo. Debido a que no hay un punto de inicio natural con 0°F que represente “sin calor”, las razones como “dos veces” no tienen sentido, por lo que es incorrecto decir que el lugar de residencia del autor es dos veces más cálido que Auckland, Nueva Zelanda.
33. a. Continuo, porque el número de valores posibles es infinito y no contable
    b. Discreto, porque el número de valores posibles es finito
    c. Discreto, porque el número de valores posibles es finito
    d. Discreto, porque el número de valores posibles es infinito y contable

Sección 1-3

1. El estudio es un experimento porque los sujetos recibieron tratamientos.
3. Los tamaños de muestra grupales de 547, 550 y 546 son todos grandes para que los investigadores puedan ver los efectos del tratamiento con paracetamol.
5. La muestra parece ser una muestra de conveniencia. Al enviar por correo electrónico la encuesta a un grupo de usuarios de Internet fácilmente disponible, fue fácil obtener resultados. Aunque existe un potencial real para obtener un grupo de muestra que no es representativo de la población, las indicaciones de qué oído se utiliza para las llamadas de teléfono celular y qué mano es dominante no parecen ser factores que se distorsionarían mucho por un sesgo muestral.
7. Con 717 respuestas, la tasa de respuesta es del 14%, que parece ser bastante baja. En general, una tasa de respuesta muy baja crea un serio potencial para obtener una muestra sesgada que consiste en aquellos con un interés especial en el tema.

9. Sistemático 11. Aleatorio 13. Por conglomerados
15. Estratificado 17. Aleatorio 19. De conveniencia
21. Estudio observacional. La muestra es una muestra de conveniencia que consiste en sujetos que decidieron responder. Tales muestras de respuesta voluntaria tienen una alta probabilidad de no ser representativas de la población total, por lo que la muestra puede ser parcial. La pregunta fue publicada en una edición electrónica de un periódico, por lo que la muestra está sesgada desde el principio.
23. Experimento. Este experimento crearía una situación *extremadamente* peligrosa e ilegal que tiene un potencial real de causar lesiones o la muerte. Ya es bastante difícil conducir en la ciudad de Nueva York estando completamente sobrio.
25. Experimento. La muestra sesgada creada mediante el uso de conductores de la ciudad de Nueva York no puede arreglarse con una muestra más grande. La muestra más grande seguirá siendo una muestra sesgada que no es representativa de los conductores en Estados Unidos.
27. Estudio observacional. Los encuestados que han sido condenados por delitos mayores no responderán honestamente a la segunda pregunta. La encuesta sufriría un "sesgo de deseabilidad social" porque los sujetos tenderían a responder de manera que los encuestados los vieran de manera favorable.
29. Estudio prospectivo 31. Estudio transversal
33. Diseño de pares combinados 35. Diseño completamente aleatorizado
37. a. No es una muestra aleatoria simple, pero es una muestra aleatoria.
b. Muestra aleatoria simple y también muestra aleatoria.
c. No es una muestra aleatoria simple ni tampoco una muestra aleatoria.

Capítulo 1: Examen rápido

1. No. Los números no miden ni cuentan nada.
2. Nominal 3. Continuo 4. Datos cuantitativos
5. De razón 6. Estadístico 7. No
8. Estudio observacional
9. Los sujetos no sabían si estaban tomando aspirina o el placebo.
10. Muestra aleatoria simple

Capítulo 1: Ejercicios de repaso

1. El patrocinador de la encuesta tiene el potencial de beneficiarse de los resultados, lo que genera dudas sobre la objetividad de los resultados.
2. a. La muestra es una muestra de respuesta voluntaria, por lo que los resultados son cuestionables.
b. Estadístico c. Estudio observacional
3. Aleatorizado: los sujetos fueron asignados a los diferentes grupos a través de un proceso de selección aleatoria, por lo que tuvieron las mismas posibilidades de pertenecer a cada grupo. Doble ciego: los sujetos no sabían en cuál de los tres grupos se encontraban y las personas que evaluaron los resultados tampoco lo sabían.
4. No. La correlación no implica causalidad.
5. Sólo el inciso (c) es una muestra aleatoria simple.
6. Sí. Las dos preguntas dan la falsa impresión de que abordan cuestiones muy diferentes. La mayoría de las personas estaría a favor de defender el matrimonio, por lo que es probable que la primera pregunta reciba una cantidad sustancial de respuestas "sí". La segunda pregunta describe mejor el problema y es mucho más probable que los sujetos tengan respuestas variadas.
7. a. Discreto b. De razón
c. Las respuestas enviadas por correo serían una muestra de respuesta voluntaria, por lo que aquellos con opiniones fuertes o un mayor interés en los temas tienen más probabilidades de responder. Es muy posible que los resultados no reflejen las opiniones verdaderas de la población de todos los estudiantes universitarios a tiempo completo.
d. Estratificado e. Por conglomerados
8. a. Si no tienen grasa, tienen un 100% menos que cualquier otra cantidad con grasa, por lo que la cifra del 125% no puede ser correcta.
b. 686 c. 28%
9. a. Datos de intervalo; muestra sistemática
b. Datos nominales; muestra estratificada
c. Datos ordinales; muestra de conveniencia
10. Debido a que hay un 15% de posibilidades de obtener los resultados por casualidad, esos resultados podrían ocurrir fácilmente de esa manera, por lo que el método no parece tener significancia estadística. El resultado de 236 niñas en 450 nacimientos es una tasa del 52.4%, por lo que está por encima del 50% esperado por azar, pero no parece ser lo suficientemente alto como para tener una significancia práctica. El procedimiento no parece tener significancia estadística o significancia práctica.

Capítulo 1: Ejercicios de repaso acumulado

1. 3162.5 gramos. Todos los pesos terminan con 00, lo que sugiere que fueron redondeados de modo que los últimos dígitos tuvieran esa característica.
2. 0.015625 3. 16.00 es un valor significativamente alto.
4. −6.64 5. 1067
6. 575 gramos 7. 27343.75 gramos^2
8. 0.20 9. 0.00065536
10. 31,381,059,609 (o aproximadamente 31,381,060,000)
11. 78,364,164,096 (o alrededor de 78,364,164,000)
12. 0.000000531441

Respuestas del capítulo 2

Sección 2-1

1. La tabla resume 50 tiempos de servicio. No es posible identificar los valores exactos de todos los tiempos originales.
3.

Tiempo (seg)	Frecuencia relativa
60-119	14%
120-179	44%
180-239	28%
240-299	4%
300-359	10%

5. Ancho de clase: 10. Puntos medios de clase: 24.5, 34.5, 44.5, 54.5, 64.5, 74.5, 84.5. Límites de clase: 19.5, 29.5, 39.5, 49.5, 59.5, 69.5, 79.5, 89.5. Número: 87.
7. Ancho de clase: 100. Puntos medios de clase: 49.5, 149.5, 249.5, 349.5, 449.5, 549.5, 649.5. Límites de clase: −0.5, 99.5, 199.5, 299.5, 399.5, 499.5, 599.5, 699.5. Número: 153.
9. No. La frecuencia máxima está en la segunda clase en lugar de estar cerca del medio, por lo que las frecuencias por debajo del máximo no reflejan las que están por encima de éste.

11.

Duración (seg)	Frecuencia
125-149	1
150-174	0
175-199	0
200-224	3
225-249	34
250-274	12

13.

Tiempos de servicio (seg) para la comida en Burger King	Frecuencia
70-109	11
110-149	23
150-189	7
190-229	6
230-269	3

15. La distribución no parece ser una distribución normal.

Tiempos de servicio (seg) para la comida en Wendy's	Frecuencia
70-149	25
150-229	15
230-309	6
310-389	3
390-469	1

17. Debido a que hay desproporcionadamente más 0 y 5, parece que las estaturas se reportaron en lugar de medirse. Es probable que los resultados no sean muy precisos.

Último dígito	Frecuencia
0	9
1	2
2	1
3	3
4	1
5	15
6	2
7	0
8	3
9	1

19. Las actrices parecen ser generalmente más jóvenes que los actores.

Edad (años)	Actrices	Actores
20-29	33.3%	1.1%
30-39	39.1%	32.2%
40-49	16.1%	41.4%
50-59	3.4%	17.2%
60-69	5.7%	6.9%
70-79	1.1%	1.1%
80-89	1.1%	

21.

Edad (años) de la mejor actriz cuando ganó el Oscar	Frecuencia acumulada
Menos de 30	29
Menos de 40	63
Menos de 50	77
Menos de 60	80
Menos de 70	85
Menos de 80	86
Menos de 90	87

23. No. La frecuencia relativa más alta de 24.8% no es mucho más alta que las demás.

Reacción adversa	Frecuencia relativa
Dolor de cabeza	23.6%
Hipertensión	8.7%
Resp. Sup. infección del tracto	24.8%
Nasofaringitis	21.1%
Diarrea	21.9%

25. Sí, la distribución de frecuencia parece ser una distribución normal.

Presión arterial sistólica (mm Hg)	Frecuencia
80-99	11
100-119	116
120-139	131
140-159	34
160-179	7
180-199	1

27. Sí, la distribución de frecuencias parece ser una distribución normal.

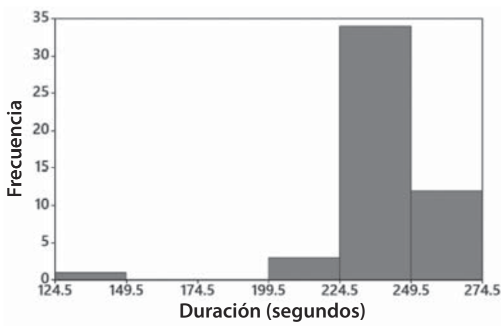
Magnitud	Frecuencia
1.00-1.49	19
1.50-1.99	97
2.00-2.49	187
2.50-2.99	147
3.00-3.49	100
3.50-3.99	38
4.00-4.49	8
4.50-4.99	4

29. Un valor atípico puede aumentar drásticamente el número de clases.

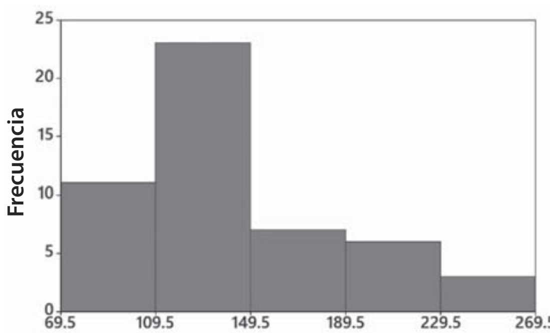
Peso (lb)	Con valor atípico	Sin valor atípico
200-219	6	6
220-239	5	5
240-259	12	12
260-279	36	36
280-299	87	87
300-319	28	28
320-339	0	
340-359	0	
360-379	0	
380-399	0	
400-419	0	
420-439	0	
440-459	0	
460-479	0	
480-499	0	
500-519	1	

Sección 2-2

1. El histograma debe tener forma de campana.
3. Con un conjunto de datos tan pequeño, la verdadera naturaleza de la distribución no se puede ver con un histograma.
5. 40
7. La forma de la gráfica no cambiaría. La escala vertical sería diferente, pero las alturas relativas de las barras serían las mismas.
9. Debido a que está lejos de tener una forma de campana, el histograma no parece representar datos de una población con distribución normal.

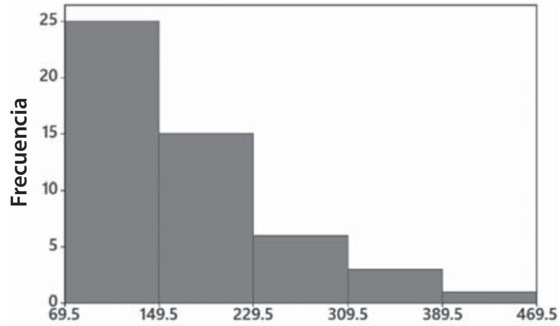


11. El histograma parece ser asimétrico a la derecha (o positivamente asimétrico).



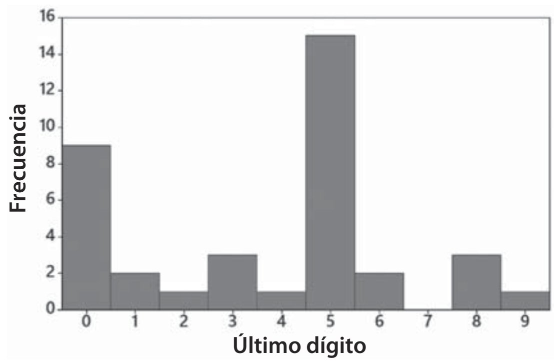
Tiempo de servicio en la comida en Burger King (segundos)

13. El histograma parece ser asimétrico a la derecha (o positivamente asimétrico).

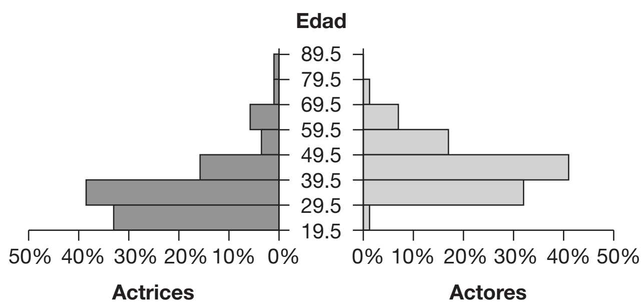


Tiempo de servicio en la comida en Wendy's (segundos)

15. Los dígitos 0 y 5 parecen surgir más a menudo que los otros dígitos, por lo que puede indicar que las estaturas se reportaron y no se midieron realmente. Esto sugiere que los datos pueden no ser muy útiles.

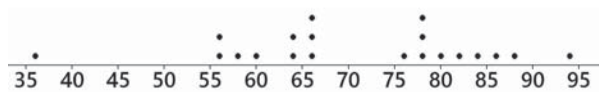


17. Las edades de las actrices son más bajas que las edades de los actores.



Sección 2-3

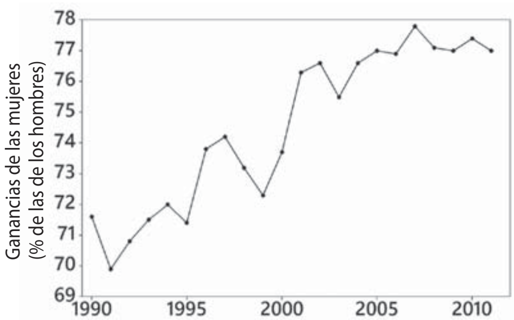
1. El conjunto de datos es demasiado pequeño para que una gráfica revele características importantes de los datos. Con un conjunto de datos tan pequeño, sería mejor simplemente listar los datos o colocarlos en una tabla.
3. No. Las gráficas deben trazarse de una manera que sea justa y objetiva. Se debe permitir que los lectores hagan sus propios juicios, en lugar de ser manipulados por gráficos engañosos.
5. El pulso de 36 latidos por minuto parece ser un valor atípico.



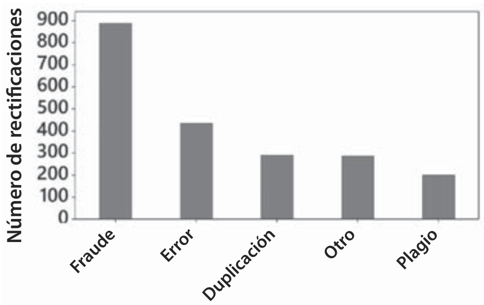
7. Los datos están ordenados de menor a mayor, como en 36, 56, 56, etcétera.

```
3 | 6
4 |
5 | 668
6 | 044666
7 | 6888
8 | 02468
9 | 4
```

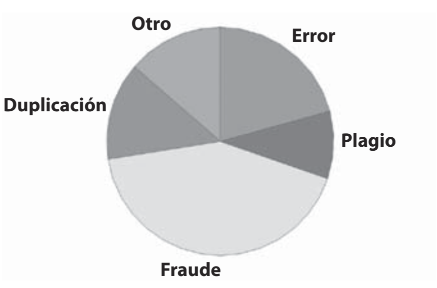
9. Hay una tendencia gradual al alza que parece estar estabilizándose en los últimos años. Una tendencia ascendente sería útil para que las ganancias de las mujeres fuesen iguales a las de los hombres.



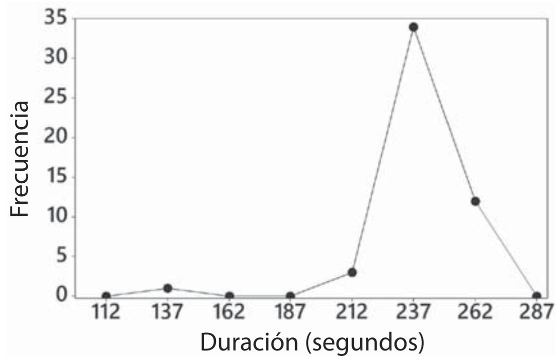
11. La mala conducta incluye fraude, duplicación y plagio, por lo que parece ser un factor importante.



13.



15. La distribución parece ser asimétrica a la izquierda (o negativamente asimétrica).

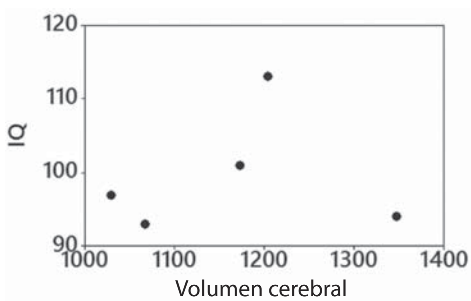


17. Debido a que la escala vertical comienza con una frecuencia de 200 en lugar de 0, la diferencia entre las respuestas “no” y “sí” es muy exagerada. La gráfica hace que parezca que alrededor de *cinco* veces más respondieron “no”, cuando la proporción es un poco menos de 2.5 a 1.
19. Los dos costos son de naturaleza unidimensional, pero los biberones son objetos tridimensionales. El costo de $4500 no es ni el doble del costo de $2600, pero los biberones hacen que parezca que el costo mayor es aproximadamente cinco veces el costo más bajo.
21.

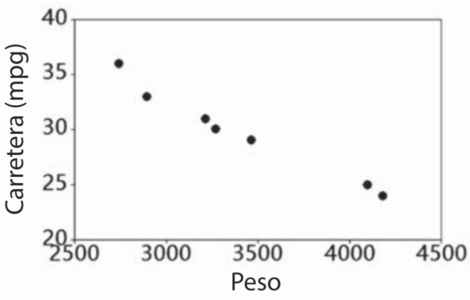
```
96. | 59
97. | 0001112333444
97. | 55666666788888999
98. | 00000000000002222233444444444444
98. | 5555666666666666666677777888888899
99. | 001244
99. | 56
```

Sección 2-4

1. El término *lineal* se refiere a una *línea* recta, y *r* mide qué tan bien se ajusta un diagrama de dispersión de la muestra de datos pareados a un patrón de línea recta.
3. Un diagrama de dispersión es una gráfica de datos cuantitativos pareados (x, y). Proporciona una imagen visual de los datos graficados como puntos, y dicha imagen es útil para permitirnos ver patrones en los datos y para reconocer si existe una correlación entre las dos variables.
5. No parece haber una correlación lineal entre el volumen cerebral y la puntuación de IQ.



7. Parece que hay una correlación lineal entre el peso y el consumo de combustible en carretera



9. Con $n = 5$ pares de datos, los valores críticos son ± 0.878. Debido a que $r = 0.127$ está entre -0.878 y 0.878, la evidencia no es suficiente para concluir que existe una correlación lineal.
11. Con $n = 7$ pares de datos, los valores críticos son ± 0.754. Como $r = -0.987$ está en la región de la cola izquierda debajo de -0.754, hay datos suficientes para concluir que hay una correlación lineal.
13. Debido a que el valor P no es pequeño (0.05 o menos), hay una alta probabilidad (83.9%) de obtener los resultados de la muestra cuando no hay correlación, por lo que la evidencia no es suficiente para concluir que hay correlación lineal.
15. Debido a que el valor P es pequeño (0.05 o menos), hay una pequeña posibilidad de obtener los resultados de la muestra cuando no hay correlación, por lo que hay suficiente evidencia para concluir que existe una correlación lineal.

Capítulo 2: Examen rápido

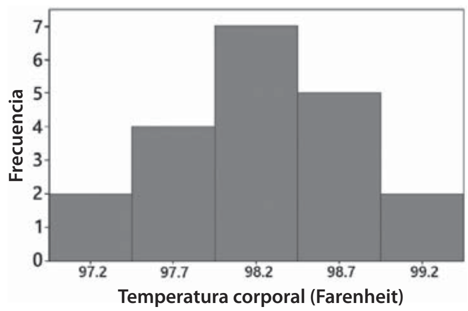
1. Ancho de clase: 3. No es posible identificar los valores de datos originales.
2. Límites de clase: 17.5 y 20.5. Límites de clase: 18 y 20.
3. 40 4. 19 y 19 5. Diagrama de Pareto
6. Histograma 7. Gráfica de dispersión
8. No, el término “distribución normal” tiene un significado diferente al del término “normal” que se usa en el habla común. Una distribución normal tiene forma de campana, pero los dígitos de la lotería seleccionados al azar tendrán una forma uniforme o plana.
9. Variación
10. Las barras del histograma comienzan relativamente bajas, aumentan hasta un máximo y luego disminuyen. Además, el histograma es simétrico, y la mitad izquierda es aproximadamente una imagen especular de la mitad derecha.

Capítulo 2: Ejercicios de repaso

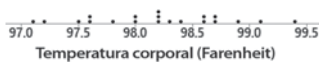
1.

Temperatura (°F)	Frecuencia
97.0-97.4	2
97.5-97.9	4
98.0-98.4	7
98.5-98.9	5
99.0-99.4	2

2. Sí, los datos parecen ser de una población con una distribución normal porque las barras comienzan bajas y alcanzan un máximo, luego disminuyen, y la mitad izquierda del histograma es aproximadamente una imagen especular de la mitad derecha.



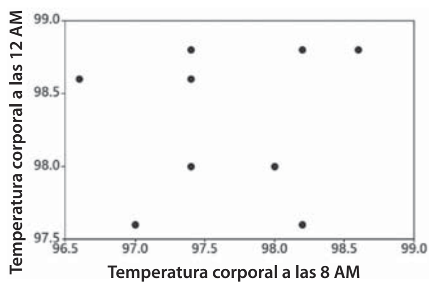
3. Al utilizar menos clases, el histograma ilustra de mejor manera la distribución.



4. No hay valores atípicos.

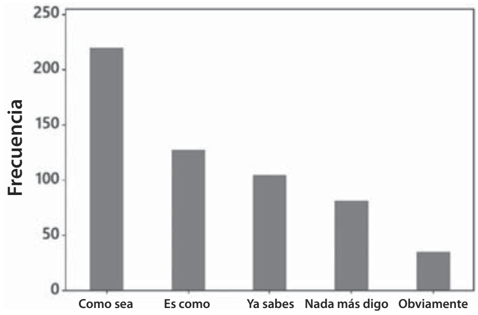
```
97.|125668
98.|002223466779
99.|14
```

5. No. No hay un patrón que sugiera que haya una relación.



6. a. Gráfica de series de tiempo b. Gráfica de dispersión
   c. Diagrama de Pareto
7. Una gráfica circular desperdicia tinta en componentes que no son datos; las gráficas circulares carecen de una escala apropiada; las gráficas circulares no muestran los tamaños relativos de los diferentes componentes, así como algunas otras gráficas, por ejemplo un diagrama de Pareto.

8. El diagrama de Pareto hace un mejor trabajo. Llama la atención sobre las palabras o frases más molestas y muestra los tamaños relativos de las diferentes categorías.

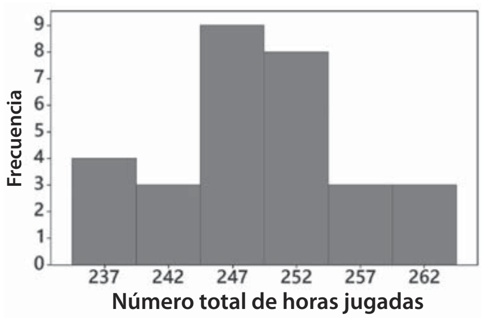


Capítulo 2: Ejercicios de repaso acumulado

1.

Horas totales	Frecuencia
235-239	4
240-244	3
245-249	9
250-254	8
255-259	3
260-264	3

2. a. 235 horas y 239 horas b. 234.5 horas y 239.5 horas
   c. 237 horas
3. La distribución está más cerca de ser una distribución normal que las otras.



4. Comience la escala vertical a una frecuencia de 2 en vez de a una frecuencia de 0.
5. Al observar las dimensiones del diagrama de tallo y hojas, podemos ver que la distribución se aproxima a una distribución normal.

```
23 | 6789
24 | 112555677889
25 | 00012233888
26 | 024
```

6. a. Continuo b. Cuantitativo c. De razón
   d. Muestra de conveniencia e. Muestra

Capítulo 3 Respuestas

Sección 3-1

1. El término *promedio* no se usa en estadística. Se debe emplear el término *media* para el resultado obtenido al sumar todos los valores de la muestra y dividir el total por el número de valores muestrales.
3. Usan diferentes métodos para proporcionar un valor (o valores) del centro o la mitad de la lista ordenada de datos.
5. $\bar{x} = 57.1$; mediana $= 60.0$; moda = ninguna; rango medio $= 53.0$. Los números de jersey son datos nominales que sólo reemplazan nombres, y no miden ni cuentan nada, por lo que los estadísticos resultantes no tienen sentido.
7. $\bar{x} = \$172.0$ millones; mediana $= \$160.0$ millones; moda $= \$150$ millones; rango medio $= \$200.0$ millones. Además del hecho de que todas las demás celebridades tienen patrimonios netos inferiores a los dados, no se puede conocer nada significativo sobre la población de patrimonios netos de todas las celebridades. Todos los números terminan en 0 o 5, y parecen ser estimaciones redondeadas.
9. $\bar{x} = 7.7$ huracanes; mediana $= 8.0$ huracanes; moda $= 8$ huracanes; rango medio $= 8.5$ huracanes. Los datos son datos de series de tiempo, pero las medidas de tendencia central no revelan nada sobre una tendencia que consista en un patrón de cambio a lo largo del tiempo.
11. $\bar{x} = \$1454.2$; mediana $= \$1500.0$; moda $= \$1500$; rango medio $= \$1375.0$. La muestra consiste en las "mejores compras" de televisores, por lo que no es una muestra aleatoria y no es probable que sea representativa de la población. El precio más bajo es un dato estadístico relevante para alguien que planea comprar uno de los televisores.
13. $\bar{x} = 32.6$ mg; mediana $= 39.5$ mg; moda $= 0$ mg; rango medio $= 27.5$ mg. Los estadounidenses consumen algunas marcas con mucha más frecuencia que otras, pero las 20 marcas se ponderan por igual en los cálculos, por lo que los estadísticos no son necesariamente representativos de la población de todas las latas de las mismas 20 marcas consumidas por los estadounidenses.
15. $\bar{x} = 9.45$ pulg.; mediana $= 9.30$ pulg.; moda $= 9.1$ pulg.; rango medio $= 9.50$ pulg. Debido a que las mediciones se realizaron en 1988, no son necesariamente representativas de la población actual de todas las mujeres del ejército.
17. $\bar{x} = \$365.3$; mediana $= \$200.0$; moda $= \$500$; rango medio $= \$1269.5$. Las cantidades de \$1500 y \$2500 parecen ser valores atípicos.
19. $\bar{x} = 2.8$ cigarrillos; mediana $= 0.0$ cigarrillos; moda $= 0$ cigarrillos; rango medio $= 25.0$ cigarrillos. Debido a que los sujetos seleccionados *informan* la cantidad de cigarrillos fumados, es muy posible que los datos no sean del todo precisos. ¿Y qué hay de esa persona que fuma 50 cigarrillos (o 2.5 paquetes) al día? ¿Qué está pensando?
21. Sistólica: $\bar{x} = 127.6$ mm Hg; mediana $= 124.0$ mm Hg. Diastólica: $\bar{x} = 73.6$ mm Hg; mediana $= 75.0$ mm Hg. Dado que las presiones arteriales sistólica y diastólica miden diferentes características, una comparación de las medidas de tendencia central no tiene sentido. Debido a que los datos coinciden, tendría más sentido investigar si existe una asociación o *correlación* entre las medidas de presión arterial sistólica y las medidas de presión arterial diastólica.

23. Hombres: $\bar{x} = 69.5$ latidos por minuto; mediana = 66.0 latidos por minuto. Mujeres: $\bar{x} = 82.1$ latidos por minuto; mediana = 84.0 latidos por minuto. Los pulsos de los hombres parecen ser más bajos que los de las mujeres.
25. $\bar{x} = 0.8$ y mediana = 1.0. Diez de los tornados carecen de mediciones en la escala F.
27. $\bar{x} = 98.20°F$; mediana = 98.40°F. Estos resultados sugieren que la media es inferior a 98.6°F.
29. $\bar{x} = 36.2$ años, que es igual a la media de 36.2 años obtenida al usar la lista original de valores.
31. $\bar{x} = 224.0$ (1000 células/μL). La media de la distribución de frecuencias es bastante cercana a la media de 224.3 (1000 células/μL) obtenida al usar la lista original de valores.
33. 3.14; sí
35. a. 90 latidos por minuto b. $n - 1$
37. 504 lb es un valor atípico. Mediana: 285.5 lb; media: 294.4 lb; 10% de la media recortada: 285.4 lb; 20% de la media recortada: 285.8 lb. Mediana, el 10% de la media recortada y el 20% de la media recortada son bastante similares, pero la media no recortada de 294.4 lb difiere de estas porque se ve fuertemente afectada por la inclusión del valor atípico.
39. 1.5290%
41. La mediana encontrada usando la expresión dada es 153.125 segundos, que se redondea a 153.1 segundos. Este valor difiere en 2.6 segundos de la mediana de 150.5 segundos encontrados al usar la lista original de tiempos de servicio.

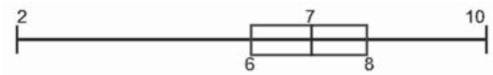
Sección 3-2

1. 119.0 cm^3 es bastante cercano al valor exacto de la desviación estándar de 124.9 cm^3.
3. 401.6577 kg^2
5. Rango = 92.0; $s^2 = 1149.5$; $s = 33.9$. Los números de jersey son datos nominales que sólo reemplazan nombres, y no miden ni cuentan nada, por lo que los estadísticos resultantes no tienen sentido.
7. Rango = \$100.0 millones; $s^2 = 1034.4$ (millones de dólares)^2; $s = \$32.2$ millones. Debido a que los datos son de las celebridades con el valor neto más alto, las medidas de variación no son típicas para todas las celebridades.
9. Rango = 13.0 huracanes; $s^2 = 10.2$ huracanes^2; $s = 3.2$ huracanes. Los datos pertenecen a una serie de tiempo, pero las medidas de variación no revelan nada sobre una tendencia que constituya un patrón de cambio a través del tiempo.
11. Rango = \$850.0; $s^2 = 61{,}117.4$ (dólares)^2; $s = \$247.2$. La muestra consiste en las "mejores compras" de televisores, por lo que no es una muestra aleatoria y no es probable que sea representativa de la población. Las medidas de variación probablemente no sean típicas de todos los televisores de 60 pulgadas o más.
13. Rango = 55.0 mg; $s^2 = 413.4$ mg^2; $s = 20.3$ mg. Los estadounidenses consumen algunas marcas con mucha más frecuencia que otras, pero las 20 marcas se ponderan por igual en los cálculos, por lo que los estadísticos no son necesariamente representativos de la población de todas las latas de las mismas 20 marcas consumidas por los estadounidenses.
15. Rango = 1.80 pulg.; $s^2 = 0.27$ pulg^2; $s = 0.52$ pulg. Debido a que las mediciones se realizaron en 1988, no son necesariamente representativas de la población actual de todas las mujeres del ejército.
17. Rango = \$2461.0; $s^2 = 290{,}400.4$ (dólares)^2; $s = \$538.9$. Las cantidades de \$1500 y \$2500 parecen ser valores atípicos, y es probable que tengan un gran efecto en las medidas de variación.
19. Rango = 50.0 cigarrillos; $s^2 = 89.7$ cigarrillos^2; $s = 9.5$ cigarrillos. Debido a que los sujetos seleccionados informan la cantidad de cigarrillos fumados, es muy posible que los datos no sean del todo precisos, por lo que los resultados podrían no reflejar los hábitos de tabaquismo entre los adultos de California.
21. Sistólica: 14.6%. Diastólica: 16.9%. La variación es aproximadamente la misma.
23. Hombres: 16.2%. Mujeres: 11.2%. Los pulsos de los hombres parecen variar más que los pulsos de las mujeres.
25. Rango = 4.0, $s^2 = 0.9$ y $s = 0.9$
27. Rango = 3.10°F; $s^2 = 0.39(°F)^2$; $s = 0.62°F$
29. 1.0, que está muy cerca de la $s = 0.9$ encontrada al usar todos los datos.
31. 0.78°F, que no es sustancialmente diferente de $s = 0.62°F$.
33. Los valores significativamente bajos son menores o iguales a 49.0 latidos por minuto, y los valores significativamente altos son mayores o iguales a 99.0 latidos por minuto. Un pulso de 44 latidos por minuto es significativamente bajo.
35. Los valores significativamente bajos son menores o iguales a 24.74 cm, y los valores significativamente altos son mayores o iguales a 29.90 cm. Una longitud de pie de 30 cm es significativamente alta.
37. $s = 12.7$ años difiere del valor exacto de 11.5 años por una cantidad algo grande.
39. $s = 68.4$ está algo lejos del valor exacto de 59.5.
41. a. 95% b. 68%
43. Al menos 89% de las mujeres tienen conteos de plaquetas dentro de 3 desviaciones estándar de la media. El mínimo es 58.9 y el máximo es 451.3.
45. a. 24.7 cigarrillos^2 b. 24.7 cigarrillos^2 c. 12.3 cigarrillos^2
    d. El inciso (b), porque las muestras repetidas producen variaciones que apuntan al mismo valor (24.7 cigarrillos^2) del mismo modo que la varianza poblacional. Use la división por $n - 1$.
    e. No. La media de las varianzas muestrales (24.7 cigarrillos^2) es igual a la varianza poblacional (24.7 cigarrillos^2), pero la media de las desviaciones estándar muestrales (3.5 cigarrillos) no es igual a la desviación estándar poblacional (5.0 cigarrillos).

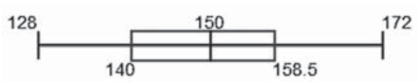
Sección 3-3

1. La estatura de James está 4.07 desviaciones estándar por encima de la media.
3. El diagrama de caja y bigotes inferior representa el peso de las mujeres, dado que muestra pesos que generalmente son más bajos.
5. a. 60.20 Mbps b. 3.76 desviaciones estándar
   c. $z = 3.76$
   d. La velocidad de datos de 77.8 Mbps es significativamente alta.
7. a. 38 latidos por minuto b. 3.04 desviaciones estándar
   c. $z = -3.04$
   d. El pulso de 36 latidos por minuto es significativamente bajo.
9. Los valores significativamente bajos son menores o iguales a 10.9; los valores significativamente altos son mayores o iguales a 31.3.
11. Los pesos significativamente bajos son menores o iguales a 5.51542 g, y los pesos significativamente altos son mayores o iguales a 5.76318 g.

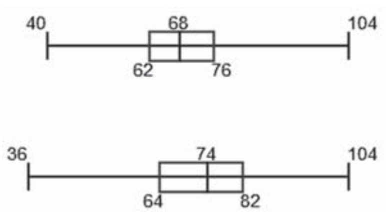
13. Con puntuaciones z de 10.83 y -16.83, la puntuación z de -16.83 está más alejada de la media, por lo que el hombre más bajo tiene una estatura más extrema.
15. Hombre: puntuación $z = -2.69$; mujer: puntuación $z = -2.18$. El hombre tiene el peso más extremo.
17. Percentil 58 19. percentil 34
21. 2.45 Mbps (Tecnología: Excel: 2.44 Mbps).
23. 3.8 Mbps (Tecnología: Minitab: 3.85 Mbps; Excel: 3.75 Mbps)
25. 1.6 Mbps 27. 0.5 Mbps
29. Resumen de 5 números: 2, 6.0, 7.0, 8.0, 10.



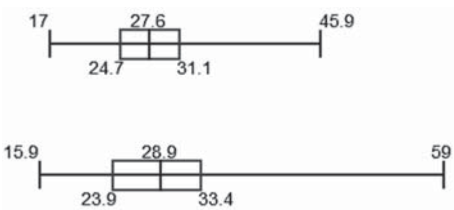
31. Resumen de 5 números: 128 mBq, 140.0 mBq, 150.0 mBq, 158.5 mBq, 172 mBq (Tecnología: Minitab produce $Q_1 = 139.0$ mBq y $Q_3 = 159.75$ mBq. Excel produce $Q_1 = 141.0$ mBq y $Q_3 = 157.25$ mBq).



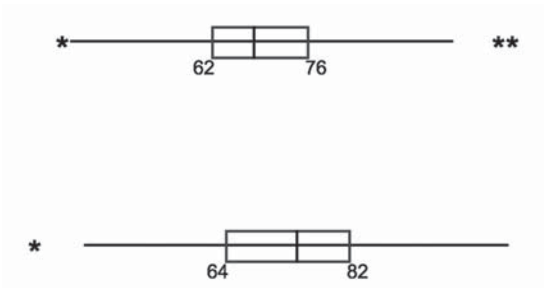
33. El diagrama de caja y bigotes superior representa a los hombres. Los hombres parecen tener pulsos ligeramente más bajos que las mujeres.



35. El diagrama de caja y bigotes superior representa los valores del IMC para los hombres. Los dos diagramas de caja no parecen ser muy diferentes, por lo que los valores del IMC de hombres y mujeres parecen ser aproximadamente los mismos, excepto por algunos valores de IMC muy altos para mujeres que hicieron que el diagrama de caja se extendiera más hacia la derecha.



37. El diagrama de caja y bigotes superior representa a los hombres. Los hombres parecen tener pulsos ligeramente más bajos que las mujeres. Valores atípicos para hombres: 40 latidos por minuto, 102 latidos por minuto, 104 latidos por minuto. Valores atípicos para las mujeres: 36 latidos por minuto.



Capítulo 3: Examen rápido

1. 6.8 horas 2. 7.0 horas
3. Dos modas: 7 horas, 8 horas 4. 1.7 horas^2
5. Sí, porque 0 horas es sustancialmente menor que todos los otros valores de datos.
6. -0.93 7. 75% o 60 horas de sueño
8. Mínimo, primer cuartil Q_1, segundo cuartil Q_2 (o mediana), tercer cuartil Q_3, máximo
9. 1.5 horas (a partir del rango/4) 10. $\bar{x}$, μ, s, σ, s^2, σ^2

Capítulo 3: Ejercicios de repaso

1. a. 1.0 min; b. 0.5 min; c. 1 min; d. 7.5 min; e. 29.0 min; f. 7.9 min; g. 61.8 min^2; h. -4.5 min; i. 2.5 min. (Tecnología: Minitab produce $Q_1 = -4.75$ min y $Q_3 = 3.25$ min. Excel produce $Q_1 = -4.25$ min y $Q_3 = 1.75$ min).
2. $z = -0.13$. El error de predicción de 0 min no es significativo porque su puntuación z está entre 2 y -2, por lo que está dentro de 2 desviaciones estándar de la media.
3. Resumen de 5 números: -7 min, -4.5 min, 0.5 min, 2.5 min, 22 min. (Tecnología: Minitab produce $Q_1 = -4.75$ min y $Q_3 = 3.25$ min. Excel produce $Q_1 = -4.25$ min y $Q_3 = 1.75$ min).

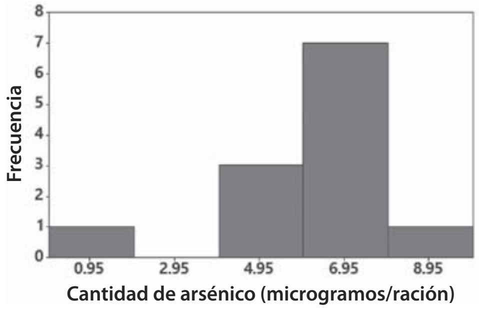
4. 23.0. Los números no miden ni cuentan nada. Se usan como reemplazos de los nombres de las categorías, por lo que los números se encuentran en el nivel nominal de medición. En este caso, la media es un estadístico sin sentido.
5. La mujer tiene el mayor peso relativo al nacer porque su puntuación z de 0.23 es mayor que la puntuación z de 0.19 para el hombre.
6. El valor atípico es 646. La media y la desviación estándar con el valor atípico incluido son $\bar{x} = 267.8$ y $s = 131.6$. Los estadísticos con el valor atípico excluido son $\bar{x} = 230.0$ y $s = 42.0$. Ambos estadísticos cambiaron en una cantidad sustancial, por lo que aquí el valor atípico tiene un efecto muy fuerte en la media y la desviación estándar.
7. El mínimo es 119 mm, el primer cuartil Q_1 es 128 mm, el segundo cuartil Q_2 (o mediana) es 131 mm, el tercer cuartil Q_3 es 135 mm y el máximo es 141 mm.
8. Con un mínimo de 117 segundos y un máximo de 256 segundos, $s \approx 34.8$ segundos, que es muy cercano a la desviación estándar de 33.7 segundos que se encuentra en la muestra más grande.

Capítulo 3: Ejercicios de repaso acumulado

1.

Arsénico (μg)	Frecuencia
0.0-1.9	1
2.0-3.9	0
4.0-5.9	3
6.0-7.9	7
8.0-9.9	1

2.



3. **1. | 5**
2. |
3. |
4. | 9
5. | 44
6. | 13679
7. | 38
8. | 2

4. a. 6.09 μg b. 6.45 μg c. 1.75 μg
d. 3.06 $(\mu)^2$ e. 6.70 μg
5. La escala vertical no comienza en 0, por lo que las diferencias entre los distintos resultados son exageradas.
6. No. Una distribución normal aparecería en un histograma como en forma de campana, pero el histograma no tiene forma de campana.

Respuestas del capítulo 4

Sección 4-1

1. $P(A) = 1/1000$, o $0.001.P(\overline{A}) = 999/1000$ o 0.999.
3. Inciso (*c*). 5. 0, 3/5, 1, 0.135
7. 1/9 o 0.111 9. Significativamente alto
11. Ni significativamente bajo ni significativamente alto
13. 1/2 o 0.5 15. 0.43
17. 1/10 o 0.1 19. 0
21. 5/555 o 0.00901. El empleador sufriría porque estaría en riesgo al contratar a alguien que use drogas.
23. 50/555 o 0.0901. Este resultado parece ser una estimación razonable de la tasa de prevalencia.
25. 879/945 o 0.930. Sí, la técnica parece ser efectiva.
27. 428/580 o 0.738; sí
29. 0.270. No, no es improbable que alguien no use los sitios de redes sociales.
31. a. café/café, café/azul, azul/café, azul/azul
b. 1/4 c. 3/4
33. 3/8 o 0.375 35. 4/16 o 1/4 o 0.25.
37. La alta probabilidad de 0.327 muestra que los resultados de la muestra pudieron haber ocurrido fácilmente por casualidad. Parece que no hay pruebas suficientes para concluir que las mujeres embarazadas puedan predecir correctamente el sexo de su bebé.
39. a. Al comparar el grupo de tratamiento de 200 mg con el grupo placebo, la baja probabilidad de menos de 0.049 indica que los resultados de la muestra no pudieron haber ocurrido fácilmente por casualidad. Parece que 200 mg de cafeína tiene un efecto en la memoria.
b. Al comparar el grupo de 300 mg con el grupo de 200 mg, la alta probabilidad de 0.75 indica que los resultados de la muestra podrían haber ocurrido fácilmente por casualidad. No hay pruebas suficientes para concluir que hay diferentes efectos del grupo de tratamiento de 300 mg y el grupo de tratamiento de 200 mg.
41. a. 9999: 1 b. 4999: 1
c. La descripción no es precisa. Las posibilidades en contra de ganar son 9999:1 y las posibilidades a favor de ganar son 1:9999, no 1:10,000.
43. a. \$5 b. 5:2
c. 772:228 o 193:57 o aproximadamente 3.39:1 (o aproximadamente 17:5)
d. Aproximadamente \$8.80 (en lugar del pago real de \$7.00)

Sección 4-2

1. $P(A)$ representa la probabilidad de seleccionar un adulto con ojos azules, y $P(\overline{A})$ representa la probabilidad de seleccionar un adulto que no tenga ojos azules.
3. Dado que las selecciones se realizan sin reemplazo, los eventos son dependientes. Debido a que el tamaño de la muestra de 1068 es menor al 5% del tamaño de la población de 15,524,971, las selecciones se pueden tratar como independientes (según la guía del 5% para cálculos engorrosos).
5. 0.74
7. $P(\overline{N}) = 0.670$, donde $P(\overline{N})$ es la probabilidad de seleccionar al azar a alguien con una respuesta distinta de "nunca".
9. 756/1118 o 0.676
11. 1020/1118 o 0.912; no disjuntos
13. a. 0.0200. Sí, los eventos son independientes.
b. 0.0199. Los eventos son dependientes, no independientes.
15. a. 0.779. Sí, los eventos son independientes.
b. 0.779. Los eventos son dependientes, no independientes.
17. 709/1118 o 0.634 19. 0.0156
21. a. 300 b. 154 c. 0.513
23. 0.990
25. a. 0.03 b. 0.0009 c. 0.000027
d. Al usar una unidad sin una copia de seguridad, la probabilidad de una falla total es 0.03, y con tres unidades de disco independientes, la probabilidad cae a 0.000027. Al cambiar de una unidad a tres, la probabilidad de falla total disminuye de 0.03 a 0.000027, y esa es una mejora sustancial en la confiabilidad. ¡Haga una copia de seguridad de sus datos!
27. 0.838. La probabilidad de 0.838 es alta, por lo que es factible que se acepte todo el lote, incluso si incluye muchos defectos.
29. a. 0.299
b. Usando la guía del 5% para cálculos engorrosos: 0.00239 [usando el resultado redondeado del inciso (a)] o 0.00238
31. a. 0.999775 b. 0.970225
c. La disposición en serie proporciona una mejor protección.
33. a. $P(A \text{ o } B) = P(A) + P(B) - 2P(A \text{ y } B)$
b. 691/1118 o 0.618

Sección 4-3

1. El evento de no tener al menos 1 defecto entre los 3 iPhones, lo que significa que los 3 iPhones son buenos.
3. La probabilidad de que la persona seleccionada sea un compañero de clase de la escuela secundaria, dado que la persona seleccionada es mujer.
5. 7/8 o 0.875 7. 0.982 9. 0.344
11. 0.965. La probabilidad es suficientemente alta como para que pueda estar razonablemente segura de obtener un defecto por su trabajo.
13. a. 27/43 o 0.628 b. 16/43 o 0.372
c. Parece que cuando los estudiantes tienen cuatro monedas de ¢25, es más probable que gasten el dinero que lo conserven.

15. a. 27/43 o 0.628 b. 12/46 o 0.261
 c. Parece que los estudiantes son más propensos a gastar el dinero cuando se les dan cuatro monedas de ¢25 que cuando se les da un billete de $1.
17. 2/1155 o 0.00173. Esta es la probabilidad de que la prueba haga parecer que el sujeto tiene hepatitis C cuando el sujeto no la tiene, por lo que es probable que el sujeto experimente estrés innecesario y pruebas adicionales.
19. 335/337 o 0.994. El resultado muy alto hace que la prueba parezca efectiva para identificar la hepatitis C.
21. a. 0.9991
 b. 0.999973. La regla de redondeo habitual para las probabilidades daría como resultado una probabilidad de 1.00, lo que indicaría incorrectamente que estamos seguros de tener al menos un disco duro en funcionamiento.
23. 0.490. La probabilidad no es baja, por lo que será necesario realizar más pruebas de las muestras individuales en 49% de las muestras combinadas.
25. 0.569

Sección 4-4

1. El símbolo ! es el símbolo factorial que representa el producto descendente de los números enteros, como en $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$. Seis personas pueden pararse en una fila de 720 maneras diferentes.
3. Debido a que se permite la repetición, los números se seleccionan *con reemplazo*, por lo que la regla de las combinaciones y las dos reglas de la permutación no se aplican. La regla del conteo por multiplicación puede usarse para mostrar que la cantidad de posibles resultados es $10 \cdot 10 \cdot 10 \cdot 10 = 10{,}000$.
5. 1/10,000 7. 1/171 9. 1/40,320
11. 1/254,251,200
13. 1/100,000,000. No, hay demasiadas posibilidades diferentes.
15. 1/1820 o 0.000549 17. 1/292,201,338
19. 1/100,000
21. Códigos de área: 800. Números de teléfono: 6,400,000,000. Sí. (Con una población total de aproximadamente 400,000,000, habría aproximadamente 16 números de teléfono por cada adulto y niño).
23. a. 5040 b. 210 c. 1/210
25. 40,320; 1/40,320 27. 653,837,184,000
29. 1/258,890,850. Hay *muchas más* posibilidades de ser alcanzado por un rayo.
31. Hay 62 diferentes caracteres posibles. El alfabeto requiere 26 caracteres y hay 10 dígitos, por lo que el sistema de código Morse es más que adecuado.
33. 128/2652 o 32/663 o 0.0483; 4.83% o aproximadamente 5%
35. 12 37. 2,095,681,645,538 (alrededor de 2 billones)

Capítulo 4: Examen rápido

1. 4/5 o 0.8 2. 0.8 3. 4/12 o 1/3 4. 0.548
5. La respuesta varía, pero la probabilidad debe ser baja, por ejemplo 0.001.
6. 0.0680 7. 0.727 8. 0.00874
9. 0.00459 10. 0.0131

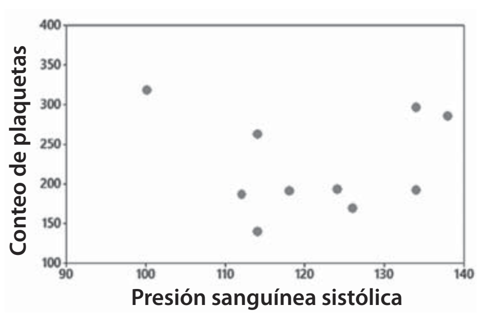
Capítulo 4: Ejercicios de repaso

1. 0.589 2. 0.657 3. 0.592 4. 0.832
5. 0.798 6. 0.347 7. 0.198
8. $\overline{A}$ es el evento de seleccionar un conductor y obtener a alguien que *no* estaba usando el cinturón de seguridad. $P(\overline{A}) = 0.411$
9. $\overline{A}$ es el evento de seleccionar un conductor y obtener a alguien que fue asesinado. $P(\overline{A}) = 0.445$.
10. 0.171
11. La respuesta varía, pero Forbes informa que aproximadamente 19% de los automóviles son negros, por lo que cualquier estimación entre 0.10 y 0.30 sería buena.
12. a. 0.25 b. 0.316
 c. No, no es improbable porque la probabilidad de 0.316 muestra que el evento ocurre con bastante frecuencia.
13. a. 1/365 b. 31/365
 c. La respuesta varía, pero probablemente es bastante pequeña, por ejemplo 0.01 o menos.
 d. Sí
14. 0.0335. No.
15. a. 1/237,336 b. 1/4 c. 1/4 d. 1/3,797,376
16. 1/962,598
17. a. 999/1000 o 0.999
 b. 999,999/1,000,000 o 0.999999
18. 1/342

Capítulo 4: Ejercicios de repaso acumulativos

1. a. 0.168 g/dL b. 0.160 g/dL c. 0.220 g/dL
 d. 0.260 g/dL e. 0.069 g/dL f. 0.005 $(g/dL)^2$
2. a. 0.09, 0.120, 0.160, 0.180, 0.35 (todo en unidades de g/dL). Valor atípico: 0.35 g/dL.
 b. 

 c.
```
.0 | 9
.1 | 113457788
.2 | 3
.3 | 5
```
3. a. 46% b. 0.460 c. Muestra estratificada
4. a. Muestra de conveniencia
 b. Si los estudiantes de la universidad son en su mayoría de una región circundante que incluye una gran proporción de un grupo étnico, los resultados pueden no reflejar la población general de Estados Unidos.
 c. 0.75 d. 0.64
5. La falta de un patrón de los puntos en el diagrama de dispersión sugiere que no parece haber una asociación entre la presión arterial sistólica y el recuento de plaquetas en la sangre.



6. a. 1/2,598,960 b. 1/28 c. 1/72,770,880

Capítulo 5 Respuestas

Sección 5-1

1. La variable aleatoria es x, que es el número de niñas en cuatro nacimientos. Los valores posibles de x son 0, 1, 2, 3 y 4. Los valores de la variable aleatoria x son numéricos.
3. $\Sigma P(x) = 0.063 + 0.250 + 0.375 + 0.250 + 0.063 = 1.001$. La suma no es exactamente 1 debido a un error de redondeo. La suma está suficientemente cerca de 1 para satisfacer el requisito. Además, la variable x es una variable aleatoria numérica y sus valores están asociados con probabilidades, y cada una de las probabilidades está entre 0 y 1 inclusive, según sea necesario. La tabla describe una distribución de probabilidad.
5. a. Variable aleatoria continua b. No es una variable aleatoria c. Variable aleatoria discreta d. Variable aleatoria continua e. Variable aleatoria discreta
7. Distribución de probabilidad con $\mu = 2.5$, $\sigma = 1.1$.
9. No es una distribución de probabilidad porque la suma de las probabilidades es 0.100, que no es 1 como se requiere. Las probabilidades muestran que Ted necesita trabajo.
11. Distribución de probabilidad con $\mu = 1.6$, $\sigma = 0.9$. (La suma de las probabilidades es 0.999, pero eso se debe a errores de redondeo).
13. No es una distribución de probabilidad porque las respuestas no son valores de una variable aleatoria numérica.
15. $\mu = 4.0$ niñas, $\sigma = 1.4$ niñas.
17. Un número significativamente alto de niñas son mayores o iguales a $\mu + 2\sigma$, y $\mu + 2\sigma = 4.0 + 2(1.4) = 6.8$ niñas. Debido a que 6 niñas no es mayor o igual a 6.8 niñas, no es un número significativamente alto de niñas.
19. a. 0.109 b. 0.144 c. Inciso (b)
 d. No, porque la probabilidad de 0.144 no es muy baja (menor o igual a 0.05).
21. $\mu = 1.5$ sonámbulos, $\sigma = 1.0$ sonámbulos
23. Un número significativamente alto de sonámbulos son mayores o iguales a $\mu + 2\sigma$, y $\mu + 2\sigma = 1.5 + 2(1.0) = 3.5$ sonámbulos. Debido a que 3 sonámbulos no son mayores o iguales a 3.5, 3 sonámbulos no son un número significativamente alto.
25. a. 0.363 b. 0.535
 c. La probabilidad del inciso (b).
 d. No, porque la probabilidad de 1 o menos sonámbulos es 0.535, que no es baja (no es menor o igual que 0.05).
27. a. 1000 b. 1/1000 c. \$ 499 d. −50¢
 e. Debido a que ambas apuestas tienen el mismo valor esperado de −50¢, ninguna apuesta es mejor que la otra.
29. a. −\$161 y \$99,839 b. −\$21
 c. Sí. El valor esperado para la compañía de seguros es de \$21, lo que indica que la compañía puede esperar hacer un promedio de \$21 por cada una de esas políticas.

Sección 5-2

1. El cálculo dado asume que los dos primeros consumidores se sienten cómodos con los drones y los últimos tres consumidores no se sienten cómodos con los drones, pero hay otros arreglos que consisten en dos consumidores que se sienten cómodos y tres que no. Las probabilidades correspondientes a esos otros arreglos también deberían incluirse en el resultado.
3. Debido a que las 30 selecciones se realizan sin reemplazo, son dependientes, no independientes. Con base en la directriz del 5% para cálculos engorrosos, las 30 selecciones se pueden tratar como independientes. (Las 30 selecciones constituyen el 3% de la población de 1009 respuestas, y 3% no es más del 5% de la población). La probabilidad se puede encontrar usando la fórmula de probabilidad binomial, pero requeriría la aplicación de esa fórmula 21 veces (o 10 veces si somos inteligentes), entonces sería mejor usar tecnología.
5. No es binomial. Cada uno de los pesos tiene más de dos resultados posibles.
7. Binomial.
9. No es binomial. Debido a que los senadores son seleccionados sin reemplazo, las selecciones no son independientes. (La directriz del 5% para cálculos engorrosos no debe utilizarse porque los 40 senadores seleccionados constituyen el 40% de la población de 100 senadores, y 40% supera a 5%).
11. Binomial. Aunque los eventos no son independientes, pueden tratarse como independientes mediante la aplicación de la directriz del 5%. El tamaño de muestra de 1019 no es más del 5% de la población de todos los adultos.
13. a. 0.128 b. IIC, ICI, CII; 0.128 para cada una
 c. 0.384
15. 0.0000819 (Tabla: 0+)
17. 0.797 (Tabla: 0.798)
19. 0.168
21. 0.147
23. 0.0889
25. 0.00000451. El resultado de 7 minorías es significativamente bajo. La probabilidad muestra que es muy poco probable que un proceso de selección aleatoria resulte en 7 o menos minorías. (El Tribunal Supremo rechazó la hipótesis de que el proceso fue aleatorio).
27. a. 0.00154 (Tabla: 0.002) b. 0.000064 (Tabla: 0+)
 c. 0.00160 (Tabla: 0.002)
 d. Sí, la pequeña probabilidad del inciso (c) sugiere que 5 es un número significativamente alto.
29. a. $\mu = 18.0$ niñas; $\sigma = 3.0$ niñas
 b. Los valores de 12.0 niñas o menos son significativamente bajos, los valores de 24.0 niñas o más son significativamente altos, y los valores entre 12.0 niñas y 24.0 niñas no son significativos.
 c. Sí, porque el resultado de 26 niñas es mayor o igual a 24.0 niñas. Un resultado de 26 niñas sugeriría que el método XSORT es efectivo.
31. a. $\mu = 7.5$ chícharos; $\sigma = 1.4$ chícharos
 b. Los valores de 4.7 o menos son significativamente bajos, los valores de 10.3 o más son significativamente altos, y los valores entre 4.7 y 10.3 no son significativos.
 c. No, porque el resultado de 9 chícharos con vainas verdes no es mayor o igual a 10.3.
33. 0.304; no.
35. 0.662. La probabilidad muestra que se aceptarán aproximadamente 2/3 de todos los envíos. Con aproximadamente 1/3 de los envíos rechazados, el proveedor haría bien en mejorar la calidad.
37. a. 8.7 y 23.3 (o 8.6 y 23.4 si usa la σ redondeada). Debido a que 19 se encuentra entre esos límites, no es significativamente bajo ni significativamente alto.
 b. 0.0736
 c. La probabilidad de 19 o más M&M verdes es 0.242.
 d. La probabilidad del inciso (c) es relevante. El resultado de 19 M&M verdes no es significativamente alto.
 e. Los resultados no proporcionan evidencia sólida contra la hipótesis de 16% para M&M verdes.

39. a. 238.3 y 287.2 (o 287.1 si usa la σ redondeada). Como 308 es mayor que 287.2, el valor de 308 es significativamente alto.
b. 0.0000369
c. La probabilidad de 308 o más es 0.000136.
d. La probabilidad del inciso (c) es relevante. El resultado de 308 votantes que votaron por el ganador es significativamente alto.
e. Los resultados sugieren que los votantes encuestados mintieron o tenían un recuerdo defectuoso de cómo votaron.

41. 0.0468 43. 0.132

Sección 5-3

1. $\mu = 535/576 = 0.929$, que es el número promedio de visitas por región. $x = 2$, porque queremos la probabilidad de que una región seleccionada aleatoriamente tenga exactamente 2 éxitos, y $e \approx 2.71828$, que es una constante utilizada en todas las aplicaciones de la fórmula 5-9.
3. Valores posibles de x: 0, 1, 2,... (sin límite superior). No es posible tener $x = 2.3$ llamadas en un día. x es una variable aleatoria discreta.
5. a. $P(5) = 0.158$
b. En 55 años, el número esperado de años con 5 huracanes es 8.7.
c. El valor esperado de 8.7 años es cercano al valor real de 8 años, por lo que la distribución de Poisson funciona bien aquí.
7. a. $P(7) = 0.140$
b. En 55 años, el número esperado de años con 7 huracanes es 7.7.
c. El valor esperado de 7.7 años es cercano al valor real de 7 años, por lo que la distribución de Poisson funciona bien aquí.
9. 11.6 nacimientos; 0.0643 (0.0649 usando la media redondeada). Hay menos de un 7% de probabilidad de obtener exactamente 15 nacimientos en un día determinado.
11. a. 62.2
b. 0.0155 (0.0156 usando la media redondeada)
13. a. 0.170
b. El número esperado está entre 97.9 y 98.2, dependiendo del redondeo.
c. El número esperado de regiones con 2 visitas está cerca de 93, que es el número real de regiones con 2 éxitos.
15. 0.9999876 o 0.9999877 (usando la media no redondeada). Muy alta posibilidad, o "casi seguro", de que ocurra al menos una muerte.
17. 0.0000178. No, es muy poco probable que se produzca al menos un premio mayor en 50 años.

Capítulo 5: Examen rápido

1. No, la suma de las probabilidades es 4/3 o 1.333, que es mayor que 1.
2. $\mu = 16.0$; $\sigma = 3.6$
3. Los valores son parámetros porque representan la media y la desviación estándar para la población de todos los que hacen suposiciones aleatorias para las 80 preguntas, no una muestra de resultados reales.
4. No. (Usando la regla práctica del rango, el límite que separa los valores significativamente altos es 23.2, pero 20 no es mayor o igual a 23.2. Si se usan probabilidades, la probabilidad de 20 o más respuestas correctas es 0.163, que no es baja).
5. Sí. (Usando la regla práctica del rango, el límite que separa los valores significativamente bajos es 8.8, y 8 es menor que 8.8. Si se usan probabilidades, la probabilidad de 8 o menos respuestas correctas es 0.0131, que es baja).
6. Sí. (La suma de las probabilidades es 0.999 y se puede considerar que es 1 debido a errores de redondeo).
7. 4.0 vuelos 8. 0.8 vuelo^2
9. 0+ indica que la probabilidad es un número positivo muy pequeño. No indica que es imposible que ninguno de los cinco vuelos llegue a tiempo.
10. 0.057

Capítulo 5: Ejercicios de repaso

1. 0.274
2. 0.999. No. Los cinco amigos no se seleccionan al azar de la población de adultos. Además, el hecho de que estén de vacaciones juntos sugiere que es más probable que su situación financiera incluya tarjetas de crédito.
3. $\mu = 3.7$ adultos y $\sigma = 1.0$ adulto.
4. Si se usa $\mu = 3.7$ adultos y $\sigma = 1.0$ adulto, los valores son significativamente altos si son iguales o mayores que $\mu + 2\sigma = 5.7$ adultos. El resultado de cinco adultos con tarjetas de crédito no es significativamente alto porque no es igual o mayor a 5.7 adultos. Además, la probabilidad de que los cinco adultos tengan tarjetas de crédito es 0.222, que no es baja (menor o igual a 0.05).
5. Si se usa $\mu = 3.7$ adultos y $\sigma = 1.0$ adulto, los valores son significativamente bajos si son iguales o menores que $\mu - 2\sigma = 1.7$ adultos. El resultado de 1 adulto con tarjeta de crédito es significativamente bajo porque es menor o igual a 1.7 adultos. Además, la probabilidad de que 1 o menos adultos tengan tarjetas de crédito es 0.0181, que es baja (menor o igual a 0.05).
6. No. Las respuestas no son *numéricas.*
7. No, porque $\Sigma P(x) = 0.9686$, pero la suma debe ser 1. (Hay un pequeño margen permitido para errores de redondeo, pero la suma de 0.9686 está demasiado lejos de 1).
8. Sí, distribución de probabilidad con $\mu = 4.6$ personas, $\sigma = 1.0$ personas. (La suma de las probabilidades es 0.999, pero eso se debe a errores de redondeo).
9. a. 236.0 cheques
b. $\mu = 236.0$ cheques y $\sigma = 12.8$ cheques
c. 210.4 cheques (o 210.3 cheques si se usa la σ no redondeada)
d. Sí, porque 0 es menor o igual a 210.4 cheques (o 210.3 cheques).
10 a. 7/365 o 0.0192 b. 0.981
c. 0.000182
d. No, porque el evento es muy raro. (Pero es posible que ocurra más de una muerte en un accidente automovilístico u otro evento similar, por lo que podría ser conveniente considerar un plan de contingencia).

Capítulo 5: Ejercicios de repaso acumulado

1. a. 9.6 lunas b. 5.0 lunas c. 0 lunas
d. 28.0 lunas e. 11.0 lunas f. 120.3 lunas^2
g. −12.4 lunas, 31.6 lunas
h. No, porque ninguno de los planetas tiene un número de lunas menor o igual a −12.4 lunas (lo cual es imposible, de todos modos) y ninguno de los planetas tiene un número de lunas igual o superior a 31.6 lunas.
i. De razón j. Discreto

2. a. 1/1000 o 0.001 b. 0.365
   c. 0.254 d. −50¢
3. a. 0.263 b. 0.342 c. 0.0688 d. 0.732 e. 0.658
4. a. 253
   b. Muestra: Los 347 profesionales de recursos humanos encuestados. Población: todos los profesionales de recursos humanos.
   c. El 73% es un dato estadístico porque es una medida basada en una muestra, no en toda la población.
5. No se muestra una escala vertical, pero una comparación de los números muestra que 7,066,000 es aproximadamente 1.2 veces el número de 6,000,000; sin embargo, la gráfica hace parecer que la meta de 7,066,000 personas es aproximadamente 3 veces el número de personas inscritas. La gráfica es engañosa en el sentido de que crea la falsa impresión de que las inscripciones reales están muy por debajo de la meta, que no es el caso. Fox News se disculpó por su gráfica y proporcionó una gráfica corregida.
6. a. 0.254 b. 0.255 (Tabla: 0.256)
   c. $\mu = 5.6$ adultos, $\sigma = 1.3$ adultos
   d. Sí. Usando la regla práctica del rango, 1 es menor que $\mu - 2\sigma = 3.0$; si se usan probabilidades, la probabilidad de 1 o menos es 0.00129 (Tabla: 0.001), que es baja, menor que 0.05.

Respuestas del capítulo 6

Sección 6-1

1. La palabra "normal" tiene un significado especial en las estadísticas. Se refiere a una distribución específica en forma de campana que se puede describir mediante la fórmula 6-1. Los dígitos de la lotería no tienen una distribución normal.
3. La media es $\mu = 0$ y la desviación estándar es $\sigma = 1$.
5. 0.4 7. 0.2 9. 0.6700 11. 0.6993 (Tabla: 0.6992)
13. 1.23 15. −1.45 17. 0.1093 19. 0.8997 21. 0.4013
23. 0.9772 25. 0.0214 (Tabla: 0.0215)
27. 0.0174 29. 0.9545 (Tabla: 0.9544)
31. 0.8413 (Tabla: 0.8412) 33. 0.999997 (Tabla: 0.9999)
35. 0.5000 37. 2.33 39. −2.05, 2.05
41. 1.28 43. 1.75 45. 68.27% (Tabla: 68.26%)
47. 99.73% (Tabla: 99.74%)
49. a. 2.28% b. 2.28% c. 95.45% (Tabla: 95.44%)

Sección 6-2

1. a. $\mu = 0$; $\sigma = 1$
   b. Las puntuaciones z son números sin unidades de medida.
3. La distribución normal estándar tiene una media de 0 y una desviación estándar de 1, pero una distribución normal no estándar tiene un valor diferente para uno o ambos parámetros.
5. 0.8849 7. 0.9053 9. 136 11. 69
13. 0.0115 (Tabla: 0.0116) 15. 0.6612 17. 24.9 pulg.
19. 20.9 pulg. y 26.1 pulg. No, 26 pulg. No son significativamente altos.
21. a. 72.11% (Tabla: 72.23%). Sí, alrededor de 28% de las mujeres no califican debido a su estatura.
    b. De 58.2 pulg. a 69.2 pulg.
23. a. 0.92% (Tabla: 0.90%). Debido a que pocos hombres pueden cumplir con el requisito de estatura, es probable que la mayoría de los personajes de Mickey Mouse sean mujeres.
    b. 64.0 pulg. a 68.6 pulg.
25. 0.2015 (Tabla: 0.2005). No, no es probable que la proporción de esquizofrénicos sea tan alta como 0.2005, o alrededor de 20%.
27. 55.67% (Tabla: 55.64%). Sí, alrededor del 44% de las mujeres quedan excluidas.
29. a. 0.1717 (Tabla: 0.1711) b. 2011.5 g (Tabla: 2011.4 g)
    c. Los pesos al nacer son significativamente bajos si son de 2011.4 g o menos, y poseen un "bajo peso al nacer" si tienen 2495 g o menos. Los pesos al nacer entre 2011.4 g y 2495 g tienen "bajo peso al nacer", pero no son significativamente bajos.
31. a. 0.0038; o se produjo un evento muy raro o el marido no es el padre.
    b. 240 días
33. a. La media es de 69.5817 (69.6 redondeados) latidos por minuto y la desviación estándar es de 11.3315 (11.3 redondeados) latidos por minuto. Un histograma confirma que la distribución es más o menos normal.
    b. 47.4 latidos por minuto; 91.8 latidos por minuto
35. a. 75; 12
    b. No, la conversión también debe explicar la variación.
    c. Calificación B: de 66.3 a 75.4 (Tabla: 66.2 a 75.4)
    d. Use un esquema como el que se proporciona en el inciso (c), porque la variación se incluye en el proceso de curvatura.

Sección 6-3

1. a. A largo plazo, las proporciones muestrales tendrán una media de 0.512.
   b. Las proporciones muestrales tenderán a tener una distribución que es aproximadamente normal.
3. Media muestral; varianza muestral; proporción muestral
5. No. La muestra no es una muestra aleatoria simple de la población de todos los nacimientos en todo el mundo. La proporción de niños nacidos en China es sustancialmente más alta que en otros países.
7. a. 4.7
   b.

Varianza muestral s^2	Probabilidad
0.0	3/9
0.5	2/9
8.0	2/9
12.5	2/9

   c. 4.7
   d. Sí. La media de la distribución muestral de las varianzas muestrales (4.7) es igual al valor de la varianza poblacional (4.7), por lo que las varianzas muestrales tienden al valor de la varianza poblacional.
9. a. 5
   b.

Mediana muestral	Probabilidad
4.0	1/9
4.5	2/9
5.0	1/9
6.5	2/9
7.0	2/9
9.0	1/9

   c. 6.0

d. No. La media de la distribución muestral de las medianas muestrales es 6.0, y no es igual al valor de la mediana poblacional (5.0), por lo que las medianas de las muestras no tienden al valor de la mediana poblacional.

11. a.

$\bar{x}$	Probabilidad
34	1/16
35	2/16
36	1/16
37.5	2/16
38.5	2/16
41	1/16
42.5	2/16
43.5	2/16
46	2/16
51	1/16

b. La media de la población es 40.5 y la media de la muestra también es 40.5.

c. Las medias muestrales tienden a la media poblacional. Las medias muestrales son buenas estimaciones de las medias poblacionales porque se centran en el valor de la media de la población en lugar de subestimarla o sobreestimarla sistemáticamente.

13. a.

Rango	Probabilidad
0	4/16
2	2/16
5	2/16
7	2/16
10	2/16
15	2/16
17	2/16

b. El rango de la población es 17, pero la media de los rangos muestrales es 7. Estos valores no son iguales.

c. Los rangos muestrales no se dirigen al rango poblacional de 17, por lo que los rangos muestrales no son buenos estimadores de los rangos poblacionales.

15.

Proporción de niñas	Probabilidad
0	0.25
0.5	0.50
1	0.25

Sí. La proporción de niñas en 2 nacimientos es 0.5, y la media de las proporciones muestrales es 0.5. El resultado sugiere que una proporción muestral es un estimador insesgado de una proporción poblacional.

17. a.

Proporción correcta	Probabilidad
0	16/25
0.5	8/25
1	1/25

b. 0.2

c. Sí. La distribución muestral de las proporciones de las muestras tiene una media de 0.2 y la proporción poblacional también es 0.2 (porque hay 1 respuesta correcta entre 5 elecciones). Sí, la media de la distribución muestral de las proporciones muestrales siempre es igual a la proporción poblacional.

19. La fórmula produce $P(0) = 0.25$, $P(0.5) = 0.5$ y $P(1) = 0.25$, que sí describe la distribución muestral de las proporciones muestrales. La fórmula es sólo una forma diferente de presentar la misma información en la tabla que describe la distribución muestral.

Sección 6-4

1. La muestra debe tener más de 30 valores, o debe haber evidencia de que la población de los promedios de las calificaciones de los estudiantes de estadística tiene una distribución normal.

3. $\mu_{\bar{x}}$ representa la media de todas las medias muestrales, y $\sigma_{\bar{x}}$ representa la desviación estándar de todas las medias muestrales. Para las muestras de 64 puntuaciones de IQ, $\mu_{\bar{x}} = 100$ y $\sigma_{\bar{x}} = 15/\sqrt{64} = 1.875$.

5. a. 0.6844
   b. 0.9726
   c. Debido a que la población original tiene una distribución normal, la distribución de las medias muestrales es una distribución normal para cualquier tamaño de muestra.

7. a. 0.1271 (Tabla: 0.1272)
   b. 0.2510
   c. Debido a que la población original tiene una distribución normal, la distribución de las medias muestrales es normal para cualquier tamaño de muestra.

9. 0.7030 (Tabla: 0.7019). El elevador no parece estar seguro porque hay un 70% de posibilidades de que se sobrecargue cuando lleva 27 hombres adultos.

11. a. 131
    b. 0.0000179 (Tabla: 0.0001)
    c. No. Es posible que los 4 sujetos tengan una media de 132, mientras que algunos de ellos tienen puntuaciones por debajo del requisito de Mensa de 131.

13. a. 140 libras
    b. 0.9999999998 (Tabla: 0.9999)
    c. 0.9458 (Tabla: 0.9463)
    d. La nueva capacidad de 20 pasajeros no parece ser suficientemente segura porque la probabilidad de sobrecarga es demasiado alta.

15. a. 0.0047
    b. 0.0000 (Tabla: 0.0001)
    c. El resultado del inciso (a) es relevante porque los asientos están ocupados por individuos.

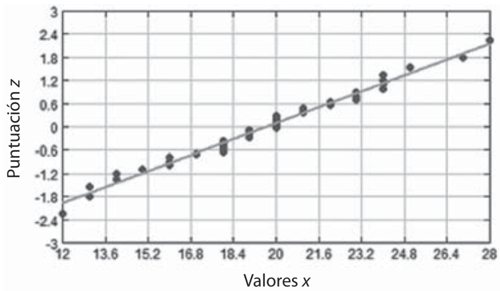
17. a. 0.5575 (Tabla: 0.5564)
    b. 0.9996 (Tabla: 0.9995)
    c. Inciso (a) porque los asientos de eyección serán ocupados por mujeres individuales, no por grupos de mujeres.

19. a. 0.8877 (Tabla 0.8869)
    b. 1.0000 cuando se redondea a cuatro lugares decimales (Tabla: 0.9999)
    c. La probabilidad del inciso (a) es más relevante porque muestra que 89% de los pasajeros varones no necesitarán doblarse. El resultado de la parte (b) nos da información sobre la media para un grupo de 100 hombres, pero no nos proporciona información útil sobre la comodidad y la seguridad de pasajeros varones individuales.
    d. Como los hombres generalmente son más altos que las mujeres, un diseño que se adapte a una proporción adecuada de hombres necesariamente incluirá una mayor proporción de mujeres.

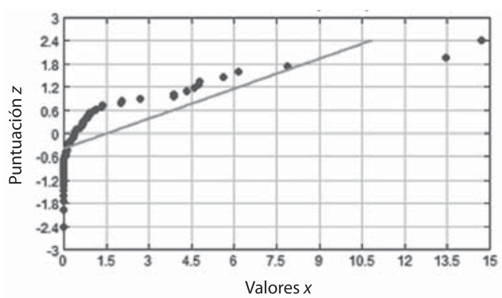
21. a. Sí. El muestreo es sin reemplazo y el tamaño de muestra de $n = 50$ es mayor que 5% del tamaño de población finita de 275. $\sigma_{\bar{x}} = 2.0504584$.
b. 0.5963 (Tabla: 0.5947)

Sección 6-5

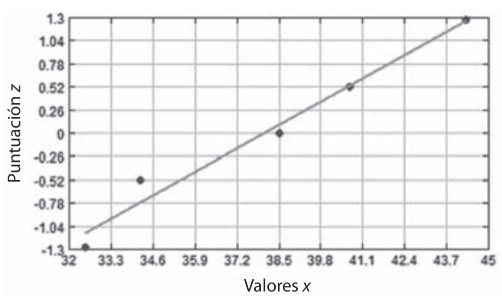
1. El histograma debe tener aproximadamente una forma de campana, y la gráfica cuantilar normal debe tener puntos que se aproximen a un patrón en línea recta.
3. Debemos verificar que la muestra proviene de una población que tiene una distribución normal. Podemos verificar la normalidad usando un histograma, identificando el número de valores atípicos y elaborando una gráfica cuantilar normal.
5. Normal. Los puntos están razonablemente cerca de un patrón en línea recta, y no hay otro patrón que no sea un patrón en línea recta.
7. No es normal. Los puntos no están razonablemente cerca de un patrón en línea recta, y parece que hay un patrón que no es un patrón en línea recta.
9. Normal 11. No normal
13. Normal



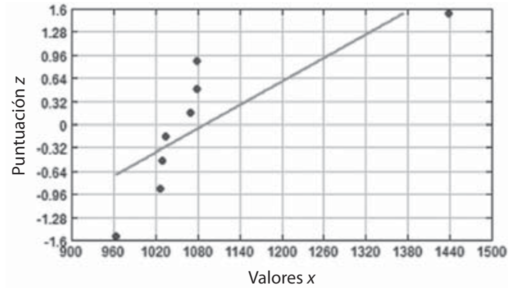
15. No normal



17. Normal. Los puntos tienen coordenadas (32.5, −1.28), (34.2, −0.52), (38.5, 0), (40.7, 0.52), (44.3, 1.28).



19. No normal. Los puntos tienen coordenadas (963, −1.53), (1027, −0.89), (1029, −0.49), (1034, −0.16), (1070, 0.16), (1079, 0.49), (1079, 0.89), (1439, 1.53)

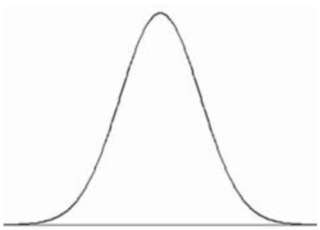


21. a. Sí b. Sí c. No

Sección 6-6

1. a. El área debajo (a la izquierda) de 502.5
b. El área entre 501.5 y 502.5
c. El área arriba (a la derecha) de 502.5
3. $p = 0.2$; $q = 0.8$; $\mu = 20$; $\sigma = 4$. El valor de $\mu = 20$ muestra que para las personas que hacen conjeturas al azar en las 100 preguntas, el número medio de respuestas correctas es 20. Para las personas que hacen 100 conjeturas al azar, la desviación estándar de $\sigma = 4$ es una medida de cuánto varían los números de respuestas correctas.
5. 0.1102 (Tabla: 0.1093)
7. La aproximación normal no debe ser utilizada.
9. Si se usa la aproximación normal: 0.2028 (Tabla: 0.2033); Tecnología que usa binomial: 0.2047. No, 20 no es un número significativamente bajo de autos blancos.
11. Si se usa la aproximación normal: 0.0549 (Tabla: 0.0542); Tecnología que usa binomial: 0.0513. La determinación de si 14 autos rojos es significativamente alto debe basarse en la probabilidad de 14 o más autos rojos, no la probabilidad de exactamente 14 autos rojos.
13. a. Si se usa la aproximación normal: 0.0212 (Tabla: 0.0217); Tecnología que usa binomial: 0.0209.
b. Si se usa la aproximación normal: 0.2012 (Tabla: 0.2005); Tecnología que usa binomial: 0.2006). El resultado de 231 decisiones revocadas no es significativamente alto.
15. a. Si se usa la aproximación normal: 0.0114 (Tabla: 0.0113); Tecnología que usa binomial: 0.0113.
b. El resultado de 109 es significativamente bajo.
17. a. Si se usa la aproximación normal: 0.2785 (Tabla: 0.2776); Tecnología que usa binomial: 0.2799.
b. El resultado de 705 chícharos con flores rojas no es significativamente alto.
c. El resultado de 705 chícharos con flores rojas no es una evidencia sólida contra el supuesto de Mendel de que 3/4 de los chícharos tendrán flores rojas.
19. a. Si se usa la aproximación normal: 0.0000 (Tabla: 0.0001); Tecnología que usa binomial: 0.0000.
b. Los resultados sugieren que las personas encuestadas no respondieron con precisión.
21. (1) 0.1723; (2) 0.1704; (3) 0.1726. No, las aproximaciones no están muy alejadas.

Capítulo 6: Examen rápido

1. 

2. $z = -1.34$ 3. 0.9983 4. 0.1546 (Tabla: 0.1547)
5. a. $\mu = 0$ y $\sigma = 1$
   b. $\mu_{\bar{x}}$ representa la media de todas las medias muestrales, y $\sigma_{\bar{x}}$ representa la desviación estándar de todas las medias muestrales.
6. 0.8092 (Tabla: 0.8106) 7. 0.6280 (Tabla: 0.6292)
8. 84.6 mm Hg (Tabla: 84.5 mm Hg)
9. 0.9568 (Tabla: 0.9564)
10. La gráfica cuantilar normal sugiere que los niveles de presión arterial diastólica de las mujeres se distribuyen normalmente.

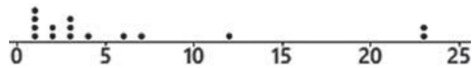
Capítulo 6: Ejercicios de repaso

1. a. 0.9382 b. 0.9382 c. 0.8983
   d. -0.67 e. 0.0668
2. a. 1.13% b. 63.8 pulg.
3. a. 1.42% (Tabla: 1.43%) b. 59.0 pulg.
4. a. Normal b. 100 c. $15/\sqrt{64} = 1.875$
5. a. Un estimador no sesgado es un dato estadístico que se dirige al valor del parámetro poblacional en el sentido de que la distribución muestral del estadístico tiene una media que es igual al parámetro correspondiente.
   b. Media; diferencia; proporción c. Verdadero
6. a. 88.77% (Tabla: 88.69%). Con alrededor del 11% de todos los hombres que necesitan agacharse, el diseño no parece ser adecuado, pero el monorriel Mark VI parece funcionar bastante bien en la práctica.
   b. 75.1 pulg.
7. a. Debido a que las mujeres son generalmente un poco más bajas que los hombres, una altura de entrada que se ajuste a los hombres también servirá para las mujeres.
   b. 1, pero en realidad una cantidad muy pequeña menor que 1 (Tabla: 0.9999)
   c. Debido a que la altura media de 60 hombres es inferior a 72 pulgadas, no se sigue que los 60 hombres individuales tengan alturas inferiores a 72 pulgadas. Al determinar la idoneidad de la altura de la puerta para los hombres, la media de 60 alturas es irrelevante, pero las alturas de los hombres individuales son relevantes.
8. a. No. Un histograma está lejos de tener forma de campana. Una gráfica cuantilar normal revela un patrón de puntos que está lejos de ser un patrón en línea recta.
   b. No. El tamaño de muestra de $n = 13$ no satisface la condición de $n > 30$, y los valores no parecen ser de una población que tiene una distribución normal.
9. Si se usa la aproximación normal: 0.2286 (Tabla: 0.2296). Tecnología que usa binomial: 0.2278. La ocurrencia de 787 plantas descendientes con tallos largos no es significativamente baja porque la probabilidad de 787 o menos plantas con tallos largos no es pequeña. Los resultados son consistentes con la proporción declarada por Mendel de 3/4.
10. a. 1.49% (Tabla: 1.50%) b. 69.4 pulg.

Capítulo 6: Ejercicios de repaso acumulados

1. a. \$6.02 millones o \$6,020,000 b. \$2.80 millones o \$ 2,800,000
   c. \$7.47 millones o \$7,470,000 d. 55.73 (millones de dólares)^2
   e. $z = 2.33$ f. De razón g. Discreto
2. a. \$1.30 millones, \$2.80 millones, \$7.10 millones
   b. 

   c. La muestra no parece ser de una población que tiene una distribución normal.
3. No, la distribución no parece ser una distribución normal.



4. a. $\overline{B}$ es el evento de seleccionar a alguien que no tiene ojos azules.
   b. 0.65 c. 0.0429
   d. Si se usa la aproximación normal: 0.1727 (Tabla: 0.1736); Tecnología que usa binomial: 0.1724.
   e. No
5. a. 0.7881 b. 0.9968 (Tabla: 0.9967) c. 10.4 pulg. d. 0.0228

Capítulo 7 Respuestas

Sección 7-1

1. El nivel de confianza, como el 95%, no se proporcionó.
3. $\hat{p} = 0.14$ es la proporción muestral; $\hat{q} = 0.86$ (encontrada a partir de la evaluación de $1 - \hat{p}$); $n = 1000$ es el tamaño de muestra; $E = 0.04$ es el margen de error; p es la proporción poblacional, que se desconoce. El valor de α es 0.05.
5. 1.645 7. 2.81
9. 0.130 ± 0.087 11. $0.0169 < p < 0.143$
13. a. $\hat{p} = 0.0912$ b. $E = 0.0297$
    c. $0.0615 < p < 0.121$
    d. Tenemos una confianza del 95% de que el intervalo de 0.0615 a 0.121 realmente contiene el verdadero valor de la proporción poblacional de los pedidos de McDonald's que no son precisos.
15. a. $\hat{p} = 0.143$
    b. $E = 0.00815$
    c. $0.135 < p < 0.152$
    d. Tenemos un 90% de confianza de que el intervalo de 0.135 a 0.152 realmente contiene el valor verdadero de la proporción poblacional de las encuestas devueltas.
17. $0.462 < p < 0.529$. Debido a que 0.512 está contenido dentro del intervalo de confianza, no hay pruebas sólidas contra 0.512 como el valor de la proporción de niños en todos los nacimientos.
19. a. $11.6\% < p < 22.4\%$
    b. Debido a que los dos intervalos de confianza se superponen, es posible que Burger King y Wendy's tengan el mismo índice de pedidos que no son precisos. Ninguno de los restaurantes parece tener una tasa de precisión significativamente mejor en los pedidos.
21. a. 0.5 b. 0.439 c. $0.363 < p < 0.516$
    d. Si los terapeutas táctiles realmente tienen la capacidad de seleccionar la mano correcta detectando un campo de energía, su índice de éxito sería significativamente mayor que 0.5, pero la tasa de éxito de la muestra de 0.439 y el intervalo de confianza sugieren que no

tienen la capacidad de seleccionar la mano correcta detectando un campo de energía.

23. a. $0.0276\% < p < 0.0366\%$ (si se usa $x = 135$: $0.0276\% < p < 0.0367\%$).
    b. No, porque 0.0340% está incluido en el intervalo de confianza.
25. Grupo placebo: $0.697\% < p < 4.49\%$. Grupo de tratamiento: $0.288\% < p < 1.57\%$. Debido a que los dos intervalos de confianza se superponen, no parece haber una diferencia significativa entre las tasas de reacciones alérgicas. Las reacciones alérgicas no parecen ser una preocupación para los usuarios de Lipitor.
27. Cuidado sostenido: $77.6\% < p < 88.1\%$ (si se usa $x = 164$). Atención estándar: $56.1\% < p < 69.5\%$ (si se usa $x = 125$). Los dos intervalos de confianza no se superponen. Parece que la tasa de éxito es mayor con la atención sostenida.
29. $\hat{p} = 18/34$, o 0.529. IC: $36.2\% < p < 69.7\%$. Una mayor estatura no parece ser una ventaja para los candidatos presidenciales. Si una mayor estatura fuera una ventaja, los candidatos más altos deberían ganar sustancialmente más del 50% de las elecciones, pero el intervalo de confianza muestra que el porcentaje de elecciones ganadas por candidatos más altos es probable que esté entre 36.2% y 69.7%.
31. a. 1844 (Tabla: 1842) b. 664 c. No cambian
33. a. 385 b. 369
    c. No, el tamaño de muestra no cambia mucho.
35. a. 1537 b. 1449
37. a. 752 b. 405
    c. No. Una muestra de las personas que usted conoce es una muestra de conveniencia, no una muestra aleatoria simple, por lo que es muy posible que los resultados no sean representativos de la población.
39. 1238 (Tabla: 1237)
41. a. El requisito de al menos 5 éxitos y al menos 5 fallas no se cumple, por lo que no se puede usar la distribución normal.
    b. 0.075

Sección 7-2

1. a. 13.05 Mbps $< \mu <$ 22.15 Mbps
   b. La mejor estimación puntual de μ es 17.60 Mbps. El margen de error es $E = 4.55$ Mbps.
   c. Debido a que el tamaño de muestra de 50 es mayor que 30, podemos considerar que la media de la muestra proviene de una población con una distribución normal.
3. Tenemos un 95% de confianza en que los límites de 13.05 Mbps y 22.15 Mbps contienen el valor real de la media de la población de todas las velocidades de datos de Verizon en los aeropuertos.
5. No se aplica la distribución normal ni la t.
7. $z_{\alpha/2} = 2.576$ (Tabla: 2.575)
9. 29.4 hg $< \mu <$ 31.4 hg. No, los resultados no difieren mucho.
11. 98.08°F $< \mu <$ 98.32°F. Debido a que el intervalo de confianza no contiene 98.6°F, parece que la temperatura corporal media no es 98.6°F, como se cree comúnmente.
13. 71.4 min $< \mu <$ 126.4 min. El intervalo de confianza incluye la media de 102.8 minutos que se midió antes del tratamiento, por lo que la media podría ser la misma después del tratamiento. Este resultado sugiere que el tratamiento con zopiclona no tiene un efecto significativo.
15. $1.8 < \mu < 3.4$. Los números dados son simplemente sustitutos de los cuatro nombres básicos de ADN, por lo que los números no miden ni cuentan nada, y se encuentran en el nivel nominal de medición. El intervalo de confianza no tiene uso práctico.
17. $5.0 < \mu < 9.0$. Los resultados no nos dicen nada sobre la población de mujeres adultas.
19. 0.284 ppm $< \mu <$ 1.153 ppm. Si se usa la guía de la FDA, el intervalo de confianza sugiere que podría haber demasiado mercurio en los peces porque es posible que la media sea mayor a 1 ppm. Además, uno de los valores muestrales excede la directriz de la FDA de 1 ppm, por lo que al menos algunos de los peces tienen demasiado mercurio.
21. 143.3 millones de dólares $< \mu <$ 200.7 millones de dólares. Debido a que las cantidades son de las diez celebridades más ricas, el intervalo de confianza no nos dice nada acerca de la población de todas las celebridades. Los datos no parecen ser de una población distribuida normalmente, por lo que el intervalo de confianza puede no ser una buena estimación de la media poblacional.
23. $3.67 < \mu < 4.17$. Debido a que todos los estudiantes estaban en la Universidad de Texas en Austin, el intervalo de confianza no nos dice nada acerca de la población de estudiantes universitarios en Texas.
25. Mujeres: 72.0 lpm $< \mu <$ 76.1 lpm. Hombres: 67.8 lpm $< \mu <$ 71.4 lpm. Las mujeres adultas parecen tener una frecuencia media de pulso que es más alta que la frecuencia media del pulso de los hombres adultos. Es bueno saberlo.
27. McDonald's: 161.4 seg $< \mu <$ 197.2 seg. Burger King: 139.1 seg $< \mu <$ 167.5 seg (Tabla: 139.2 seg $< \mu <$ 167.4 seg). No parece haber una diferencia significativa entre los tiempos medios del servicio de la cena en McDonald's y Burger King.
29. El tamaño de muestra es 94, y parece ser muy práctico.
31. El tamaño de muestra requerido es 38,415 (Tabla: 38,416). La muestra parece ser demasiado grande para ser práctica.
33. 4814 (Tabla: 4815). Sí, el supuesto parece razonable.
35. a. 425 b. 212
    c. El resultado del inciso (a) es sustancialmente mayor que el resultado del inciso (b). Es probable que el resultado del inciso (b) sea mejor porque usa s en lugar de la σ estimada que se obtuvo a partir de la regla práctica del rango.
37. 29.4 hg $< \mu <$ 31.4 hg
39. 0.8462 g $< \mu <$ 0.8668 g; 0.8474 g $< \mu <$ 0.8656 g; el segundo intervalo de confianza es más estrecho, lo que indica que tenemos una estimación más precisa cuando se selecciona la muestra relativamente grande sin reemplazo de una población finita relativamente pequeña.

Sección 7-3

1. 95.0 cm$^3 < \sigma <$ 182.5 cm^3. Tenemos un 95% de confianza de que los límites de 95.0 cm^3 y 182.5 cm^3 contienen el verdadero valor de la desviación estándar de los volúmenes cerebrales.
3. La gráfica de puntos no parece representar datos muestrales de una población normalmente distribuida. El tamaño de muestra grande no justifica el tratamiento de los valores como de una población con distribución normal. Debido a que el requisito de normalidad no se cumple, la estimación del intervalo de confianza de σ no se debe elaborar utilizando los métodos de esta sección.
5. gl = 24. $\chi_L^2 = 12.401$ y $\chi_R^2 = 39.364$. IC: 0.19 mg $< \sigma <$ 0.33 mg.
7. gl = 146. $\chi_L^2 = 105.741$ (Tabla: 67.328) y $\chi_R^2 = 193.761$ (Tabla: 140.169). IC: $56.8 < \sigma < 76.8$ (Tabla: $66.7 < \sigma < 96.3$).
9. 0.55°F $< \sigma <$ 0.72°F (Tabla: 0.56°F $< \sigma <$ 0.74°F)

11. 29.6 min $< \sigma <$ 71.6 min. No, el intervalo de confianza no indica si el tratamiento es efectivo.
13. $1.6 < \sigma < 3.8$
15. 2.9 mi/h $< \sigma <$ 6.9 mi/h. Debido a que las condiciones del tránsito varían considerablemente en diferentes momentos durante el día, el intervalo de confianza es una estimación de la desviación estándar de la población de velocidades a las 3:30 en un día laborable, no en otras.
17. a. Evaluaciones de cursos: $0.46 < \sigma < 0.62$ (Tabla: $0.47 < \sigma < 0.62$)
    b. Evaluaciones de profesores: $0.49 < \sigma < 0.66$
    c. Las cantidades de variación son casi iguales.
19. 19,205 es demasiado grande. No hay 19,205 profesores de estadística en la población, e incluso si los hubiera, el tamaño de muestra es demasiado grande para ser práctico.
21. El tamaño de muestra es 48. No, con muchos ingresos muy bajos y algunos ingresos altos, es probable que la distribución se desvíe hacia la derecha y no satisfaga el requisito de una distribución normal.
23. Si se usa $z_{\alpha/2} = 2.575829303$: $\chi^2_L = 109.980$ y $\chi^2_R = 199.655$ (Tabla: $\chi^2_L = 109.993$ y $\chi^2_R = 199.638$). Los valores de la aproximación dada son bastante cercanos a los valores críticos reales.

Sección 7-4

1. Sin reemplazo, cada muestra sería idéntica a la muestra original, por lo que las proporciones o medias, o desviaciones estándar, o varianzas serían todas iguales, y no habría un "intervalo" de confianza.
3. Los incisos b, d, e no son posibles muestras bootstrap.
5. $0.000 < p < 0.500$
7. a. 0.1 kg $< \mu <$ 8.6 kg b. 1.9 kg $< \sigma <$ 6.3 kg
9. Las respuestas varían, pero a continuación se muestran respuestas típicas.
    a. -0.8 kg $< \mu <$ 7.8 kg
    b. 1.2 kg $< \sigma <$ 7.0 kg
11. Las respuestas varían, pero a continuación se muestran respuestas típicas.
    a. $5.3 < \mu < 8.5$. Esto no es dramáticamente diferente de $5.0 < \mu < 9.0$.
    b. $1.2 < \sigma < 2.9$. Esto no es dramáticamente diferente de $1.6 < \sigma < 3.8$.
13. Las respuestas varían, pero a continuación se presenta un resultado típico: $0.0608 < p < 0.123$. Esto es bastante cercano al intervalo de confianza de $0.0615 < p < 0.121$ encontrado en el ejercicio 13 de la sección 7-1.
15. Las respuestas varían, pero a continuación se presenta un resultado típico: $0.135 < p < 0.152$. El resultado es esencialmente el mismo que el intervalo de confianza de $0.135 < p < 0.152$ encontrado en el ejercicio 15 de la sección 7-1.
17. Las respuestas varían, pero a continuación se presenta un resultado típico: $3.69 < \mu < 4.15$. Este resultado está muy cerca del intervalo de confianza $3.67 < \mu < 4.17$ encontrado en el ejercicio 23 de la sección 7-2.
19. a. Las respuestas varían, pero a continuación se presenta un resultado típico: 233.6 seg $< \mu <$ 245.1 seg.
    b. 234.4 seg $< \mu <$ 246.0 seg
    c. El resultado del método bootstrap es razonablemente cercano al resultado encontrado usando los métodos de la sección 7-2.
21. a. Las respuestas varían, pero a continuación se presenta un resultado típico: $2.5 < \sigma < 3.3$.
    b. $2.4 < \sigma < 3.7$
    c. El intervalo de confianza del método bootstrap no es muy diferente del intervalo de confianza encontrado usando los métodos de la sección 7-3. Debido a que un histograma o una gráfica cuantilar normal indica que la muestra parece provenir de una población que no tiene una distribución normal, el intervalo de confianza bootstrap de $2.5 < \sigma < 3.3$ sería una mejor estimación de σ.
23. Las respuestas varían, pero a continuación se presenta un resultado típico usando 10,000 muestras bootstrap: $2.5 < \sigma < 3.3$. Este resultado está muy cerca del intervalo de confianza de $2.4 < \sigma < 3.3$ encontrado usando 1000 muestras bootstrap. En este caso, aumentar el número de muestras bootstrap de 1000 a 10,000 no tiene mucho efecto en el intervalo de confianza.

Capítulo 7: Examen rápido

1. 0.720
2. Tenemos un 95% de confianza de que los límites de 0.692 y 0.748 contienen el verdadero valor de la proporción de adultos en la población que dicen que la ley se inclina por las celebridades.
3. $z = 2.576$ (Tabla: 2.575)
4. $36.9\% < p < 43.1\%$
5. 601 6. 136
7. Existe un requisito poco estricto de que los valores muestrales provengan de una población distribuida normalmente.
8. Los grados de libertad son el número de valores muestrales que pueden variar después de que se hayan impuesto restricciones en todos los valores. Para los datos muestrales descritos en el ejercicio 7, gl = 11.
9. $t = 2.201$
10. No, el uso de la distribución χ^2 tiene un requisito bastante estricto de que los datos deben ser de una distribución normal. El método bootstrap podría usarse para encontrar una estimación del intervalo de confianza del 95% de σ.

Capítulo 7: Ejercicios de repaso

1. $37.9\% < p < 42.1\%$. Debido a que tenemos un 95% de confianza de que los límites del 37.9% y del 42.1% contienen el porcentaje verdadero para la población de adultos, podemos decir con seguridad que menos del 50% de los adultos prefieren recibir sus noticias en línea.
2. 423
3. a. 2.926
    b. $2.749 < \mu < 3.102$
    c. Tenemos un 95% de confianza de que los límites de 2.749 y 3.102 contienen el valor de la media poblacional μ.
4. 94
5. a. Distribución *t* de Student
    b. Distribución normal
    c. Ninguna de las tres distribuciones es apropiada, pero se puede elaborar un intervalo de confianza mediante el uso de métodos *bootstrap*.
    d. χ^2 (distribución ji cuadrada)
    e. Distribución normal
6. a. 1068
    b. 340
    c. 1068
7. -22.1 seg $< \mu <$ 308.1 seg
8. 184.0 seg $< \mu <$ 441.1 seg
9. Las respuestas varían, pero a continuación se presenta un resultado típico: 7.1 seg $< \mu <$ 293.7 seg.
10 a. $0.0113 < p < 0.0287$
    b. Las respuestas varían, pero a continuación se presenta un resultado típico: $0.0120 < p < 0.0290$.
    c. Los intervalos de confianza son bastante cercanos.

Capítulo 7: Ejercicios de repaso acumulado

1. $\bar{x} = -3.6$ min, mediana $= -20.0$ min, $s = 39.9$ min, rango $= 149.0$ min. Estos resultados son estadísticos.

2. Los valores significativamente bajos son −83.4 min o menos, y los valores significativamente altos son 76.2 min o más. Debido a que 103 min excede 76.2 min, el tiempo de retraso de llegada de 103 min es significativamente alto.
3. Nivel de razón de la medición; datos continuos.
4. $-28.9 \text{ min} < \mu < 21.7 \text{ min}$
5. a. 0.2553 (Tabla: 0.2546)
   b. 15.5 min (Tabla: 15.4 min)
6. 143 vuelos
7. $77.1\% < p < 83.5\%$
8. Las gráficas sugieren que la población tiene una distribución sesgada hacia la derecha en lugar de ser normal. El histograma muestra que algunos tiempos de salida pueden ser muy largos y pueden ocurrir con mucho tráfico, pero poco o nada de tráfico no pueden hacer que el tiempo de salida sea muy bajo. Hay un tiempo mínimo requerido, independientemente de las condiciones del tráfico. La elaboración de una estimación del intervalo de confianza de una desviación estándar de población tiene un requisito bastante estricto de que los datos muestrales provengan de una población distribuida normalmente, y las gráficas muestran que este requisito estricto de normalidad no se cumple.

Capítulo 8 Respuestas

Sección 8-1

1. El rechazo del reclamo sobre la aspirina es más serio porque es un medicamento utilizado para tratamientos médicos. La dosis incorrecta de aspirina podría causar reacciones adversas más graves que una dosis incorrecta de vitamina C. Sería conveniente usar un nivel de significancia menor para probar la hipótesis sobre la aspirina.
3. a. H_0: $\mu = 174.1$ cm   b. H_1: $\mu \neq 174.1$ cm
   c. Rechazar la hipótesis nula o no rechazar la hipótesis nula.
   d. No. En este caso, la afirmación original se convierte en la hipótesis nula. Para la hipótesis de que la altura media de los hombres es igual a 174.1 cm, podemos rechazar esa hipótesis o no rechazarla, pero no podemos afirmar que hay suficiente evidencia para *respaldar* la hipótesis.
5. a. $p > 0.5$ b. H_0: $p = 0.5$; H_1: $p > 0.5$
7. a. $\mu = 69$ lpm   b. H_0: $\mu = 69$ lpm; H_1: $\mu \neq 69$ bpm
9. Hay suficiente evidencia para respaldar la hipótesis de que la mayoría de los adultos borrarían toda su información personal en línea si pudieran.
11. No hay pruebas suficientes para justificar el rechazo de la hipótesis de que la frecuencia media del pulso (en latidos por minuto) de los hombres adultos es de 69 lpm.
13. $z = 4.28$ (o $z = 4.25$ si se usa $x = 333$)
15. $t = 0.657$
17. a. De cola derecha b. Valor $P = 0.1587$
    c. No se puede rechazar H_0
19. a. Dos colas   b. Valor $P = 0.0444$ c. Se rechaza H_0.
21. a. $z = 1.645$ b. No se puede rechazar H_0
23. a $z = \pm 1.96$ b. Se rechaza H_0
25. a. No se puede rechazar H_0
    b. No hay evidencia suficiente para respaldar la hipótesis de que más de 58% de los adultos borrarían toda su información personal en línea si pudieran.
27. a. Se rechaza H_0
    b. Existe suficiente evidencia para justificar el rechazo de la hipótesis de que la frecuencia media del pulso (en latidos por minuto) de los hombres adultos es de 72 lpm.
29. Error tipo I: en realidad $p = 0.1$, pero rechazamos la hipótesis de que $p = 0.1$. Error tipo II: en realidad $p \neq 0.1$, pero no rechazamos la hipótesis de que $p = 0.1$.
31. Error tipo I: en realidad $p = 0.87$, pero respaldamos la hipótesis de que $p > 0.87$. Error tipo II: en realidad $p > 0.87$, pero no podemos respaldar esa conclusión.
33. La potencia de 0.96 muestra que hay un 96% de posibilidad de rechazar la hipótesis nula de $p = 0.08$ cuando la proporción verdadera es realmente 0.18. Es decir, si la proporción de usuarios de Chantix que experimentan dolor abdominal es en realidad de 0.18, entonces hay un 96% de posibilidades de respaldar la hipótesis de que la proporción de usuarios de Chantix que experimentan dolor abdominal es mayor a 0.08.
35. 617

Sección 8-2

1. a. 270   b. $\hat{p} = 0.53$
3. El método basado en un intervalo de confianza no es equivalente al método del valor P y al método del valor crítico.
5. a. Cola izquierda b. $z = -4.46$ c. Valor P: 0.000004
   d. H_0: $p = 0.10$. Se rechaza la hipótesis nula.
   e. Hay pruebas suficientes para respaldar la hipótesis de que menos del 10% de los sujetos tratados experimentan dolores de cabeza.
7. a. Dos colas   b. $z = -1.69$ c. Valor P: 0.091
   d. H_0: $p = 0.92$. No se puede rechazar la hipótesis nula.
   e. No hay pruebas suficientes para justificar el rechazo de la hipótesis de que 92% de los adultos poseen teléfonos celulares.
9. H_0: $p = 0.10$. H_1: $p \neq 0.10$. Estadístico de prueba: $z = -0.56$. Valor $P = 0.5751$ (Tabla: 0.5754). Valores críticos: $z = \pm 1.96$. No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la hipótesis de que la tasa de pedidos inexactos es igual al 10%. Dado que 10% de las órdenes son inexactas, parece que McDonald's debería trabajar para reducir esa tasa.
11. H_0: $p = 0.5$. H_1: $p \neq 0.5$. Estadístico de prueba: $z = 2.69$. Valor $P = 0.0071$ (Tabla: 0.0072). Valores críticos: $z = \pm 2.576$ (Tabla: ± 2.575). Se rechaza H_0. Hay evidencia suficiente para rechazar la hipótesis de que la proporción de los que están a favor es igual a 0.5. El resultado sugiere que el político está equivocado al afirmar que las respuestas son conjeturas al azar equivalentes a lanzar una moneda.
13. H_0: $p = 0.20$. H_1: $p > 0.20$. Estadístico de prueba: $z = 1.10$. Valor $P = 0.1367$ (Tabla: 0.1357). Valor crítico: $z = 1.645$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que más de 20% de los usuarios de OxyContin desarrollan náuseas. Sin embargo, con $\hat{p} = 0.229$, vemos que un gran porcentaje de usuarios de OxyContin experimentan náuseas, por lo que la tasa parece ser muy alta.
15. H_0: $p = 0.15$. H_1: $p < 0.15$. Estadístico de prueba: $z = -1.31$. Valor $P = 0.0956$ (Tabla: 0.0951). Valor crítico: $z = -2.33$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que la tasa de devolución es inferior al 15%.
17. H_0: $p = 0.512$. H_1: $p \neq 0.512$. Estadístico de prueba: $z = -0.98$. Valor $P = 0.3286$ (Tabla: 0.3270). Valores críticos: $z = \pm 1.96$. No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la hipótesis de que 51.2% de los bebés recién nacidos son varones. Los resultados no *respaldan* la creencia de que 51.2% de los recién nacidos son varones; los resultados simplemente muestran que no hay pruebas sólidas contra la tasa del 51.2%.

19. H_0: $p = 0.80$. H_1: $p < 0.80$. Estadístico de prueba: $z = -1.11$. Valor $P = 0.1332$ (Tabla: 0.1335). Valor crítico: $z = -1.645$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que los resultados del polígrafo son correctos menos de 80% del tiempo. Sin embargo, con base en la proporción muestral de resultados correctos en 75.5% de los 98 casos, los resultados del polígrafo no parecen tener el alto grado de confiabilidad que justificaría el uso de resultados del polígrafo en la corte, por lo que los resultados de esa prueba deberían prohibirse como evidencia en los juicios.
21. H_0: $p = 0.5$. H_1: $p \neq 0.5$. Estadístico de prueba: $z = -2.03$. Valor $P = 0.0422$ (Tabla: 0.0424). Valores críticos: $z = \pm 1.645$. Se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la hipótesis de que los terapeutas táctiles usan un método equivalente a las suposiciones al azar. Sin embargo, su tasa de éxito de 123/280, o 43.9%, indica que tuvieron un peor desempeño que las suposiciones al azar, por lo que no parecen ser efectivos.
23. H_0: $p = 0.000340$. H_1: $p \neq 0.000340$. Estadístico de prueba: $z = -0.66$. Valor $P = 0.5122$ (Tabla: 0.5092). Valores críticos: $z = \pm 2.81$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que la tasa es diferente de 0.0340%. Los usuarios de teléfonos celulares no deberían preocuparse por el cáncer cerebral o del sistema nervioso.
25. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 1.00$. Valor P: 0.1587. Valor crítico: $z = 1.645$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que la probabilidad de que un equipo de la NFC gane el Súper Bowl es mayor a un medio.
27. H_0: $p = 0.5$. H_1: $p \neq 0.5$. Estadístico de prueba: $z = 2.05$. Valor $P = 0.0402$ (Tabla: 0.0404). Valores críticos: $z = \pm 1.96$. Se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la hipótesis de que el lanzamiento de la moneda es justo en el sentido de que ninguno de los equipos tiene una ventaja al ganarlo. La regla del lanzamiento de una moneda no parece ser justa. Esto ayuda a explicar por qué se cambiaron las reglas del tiempo extra.
29. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 40.91$ (si se usa $\hat{p} = 0.64$) o $z = 40.90$ (si se usa $x = 13{,}661$). Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 2.33$. Se rechaza H_0. Hay suficiente evidencia para respaldar la hipótesis de que la mayoría de las personas cree que existe el monstruo de Loch Ness. Debido a que la muestra es una muestra de respuesta voluntaria, la conclusión sobre la población podría no ser válida.
31. H_0: $p = 0.791$. H_1: $p < 0.791$. Estadístico de prueba: $z = -29.09$ (si se usa $\hat{p} = 0.39$) o $z = -29.11$ (si se usa $x = 339$). Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = -2.33$. Se rechaza H_0. Existe suficiente evidencia para respaldar la hipótesis de que el porcentaje de estadounidenses seleccionados de ascendencia mexicana es inferior al 79.1%, por lo que el proceso de selección del jurado parece ser parcial.
33. Los valores de P concuerdan razonablemente bien con el tamaño de muestra grande de $n = 1009$. La aproximación normal a la distribución binomial funciona mejor a medida que aumenta el tamaño de la muestra. Entradas de aproximación normales: 0.0114, 0.0012, 0.0054. Entradas exactas: 0.0215, 0.0034, 0.0059. Exacto con corrección de continuidad simple: 0.0117, 0.0018, 0.0054.
35. a. 0.7219 (Tabla: 0.7224) b. 0.2781 (Tabla: 0.2776)
    c. La potencia de 0.7219 muestra que hay una posibilidad razonablemente buena de tomar la decisión correcta de rechazar la hipótesis nula falsa. Sería mejor si la potencia fuera aún mayor, por ejemplo mayor a 0.8 o 0.9.

Sección 8-3

1. Los requisitos son (1) la muestra debe ser una muestra aleatoria simple, y (2) una o ambas de las siguientes condiciones deben cumplirse: La población se distribuye normalmente o $n > 30$. No se proporciona suficiente información para determinar si la muestra es una muestra aleatoria simple. Como el tamaño de la muestra no es superior a 30, debemos verificar la normalidad, pero el valor de 583 segundos parece ser un valor atípico, y una gráfica cuantilar normal o un histograma sugieren que la muestra no parece proceder de una población distribuida normalmente. Los requisitos no se satisfacen.
3. Una prueba t es una prueba de hipótesis que utiliza la distribución t de Student, como el método de prueba de una hipótesis sobre la media poblacional tal como se presenta en esta sección. La letra t se usa en referencia a la distribución t de Student, que se usa en una prueba t. Los métodos de prueba z requieren un valor conocido de σ, pero sería muy raro realizar una prueba de hipótesis para una hipótesis sobre un valor desconocido de μ mientras conocemos de alguna manera el valor de σ.
5. Valor $P = 0.1301$ (Tabla: $0.10 <$ Valor $P < 0.20$)
7. Valor $P = 0.2379$ (Tabla: valor $P > 0.20$)
9. H_0: $\mu = 4.00$ Mbps. H_1: $\mu < 4.00$ Mbps. Estadístico de prueba: $t = -0.366$. Valor $P = 0.3579$. Valor crítico suponiendo un nivel de significancia de 0.05: $t = -1.677$ (Tabla: -1.676 aproximadamente). No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que las velocidades de datos de Sprint en los aeropuertos provienen de una población que tiene una media inferior a 4.00 Mbps.
11. H_0: $\mu = 0$ min. H_1: $\mu \neq 0$ min. Estadístico de prueba: $t = -8.720$. Valor P: 0.0000. Valores críticos: $t = \pm 1.970$. Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la hipótesis de que el error de predicción promedio es igual a cero. Las predicciones no parecen ser muy precisas.
13. H_0: $\mu = 4.00$. H_1: $\mu \neq 4.00$. Estadístico de prueba: $t = -1.638$. Valor $P = 0.1049$ (Tabla: >0.10). Valores críticos: $t = \pm 1.986$ (Tabla: ± 1.987 aproximadamente). No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la hipótesis de que la población de las evaluaciones de los cursos de los estudiantes tiene una media igual a 4.00.
15. H_0: $\mu = 0$. H_1: $\mu > 0$. Estadístico de prueba: $t = 0.133$. Valor $P = 0.4472$ (Tabla: >0.10). Valor crítico: $t = 1.677$ (Tabla: 1.676 aproximadamente). No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que con el tratamiento del ajo, el cambio medio en el colesterol es mayor que 0. No hay evidencia suficiente para respaldar la hipótesis de que el tratamiento con ajo no es eficaz para reducir los niveles del colesterol LDL.
17. H_0: $\mu = 0$ lb. H_1: $\mu > 0$ lb. Estadístico de prueba: $t = 3.872$. Valor $P = 0.0002$ (Tabla: <0.005). Valor crítico: $t = 2.426$. Se rechaza H_0. Hay suficiente evidencia para respaldar la hipótesis de que la pérdida de peso promedio es mayor que 0. Aunque la dieta parece tener significancia estadística, no parece tener significancia práctica, porque la pérdida de peso promedio de sólo 3.0 lb no parece valer el esfuerzo y el costo involucrados.
19. H_0: $\mu = 12.00$ oz. H_1: $\mu \neq 12.00$ oz. Estadístico de prueba: $t = 10.364$. Valor $P = 0.0000$ (Tabla: <0.01). Valores críticos: $t = \pm 2.030$. Se rechaza H_0. Hay suficiente evidencia para justificar el rechazo de la hipótesis de que el volumen medio es igual a 12.00 oz. Debido a que la media parece ser mayor que 12.00 oz, los consumidores no están siendo engañados porque están obteniendo un poco más de 12.00 oz.

21. Los datos muestrales cumplen el requisito relajado de tener una distribución normal. H_0: $\mu = 14$ μg/g. H_1: $\mu < 14$ μg/g. Estadístico de prueba: $t = -1.444$. Valor $P = 0.0913$ (Tabla: >0.05). Valor crítico: $t = -1.833$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que la concentración media de plomo para todos estos medicamentos es inferior a 14 μg/g.
23. H_0: $\mu = 1000$ hic. H_1: $\mu < 1000$ hic. Estadístico de prueba: $t = -2.661$. Valor $P = 0.0224$ (Tabla: el valor P está entre 0.01 y 0.025). Valor crítico: $t = -3.365$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que la media de la población es inferior a 1000 hic. No hay pruebas sólidas de que la media sea inferior a 1000 hic, y uno de los asientos elevadores tiene una medida de 1210 hic, que no satisface el requisito especificado de ser inferior a 1000 hic.
25. H_0: $\mu = 75$ lpm. H_1: $\mu < 75$ lpm. Estadístico de prueba: $t = -0.927$. Valor $P = 0.1777$ (Tabla: >0.10). Valor crítico: $t = -1.655$ (Tabla: -1.660 aproximadamente). No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que la frecuencia media del pulso de las mujeres adultas es inferior a 75 lpm.
27. H_0: $\mu = 90$ mm Hg. H_1: $\mu < 90$ mm Hg. Estadístico de prueba: $t = -21.435$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = -1.655$ (Tabla: -1.660 aproximadamente). Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la población de mujeres adultas tiene un nivel de presión arterial diastólica promedio menor de 90 mm Hg. La conclusión aborda la media de una población, no individuos, por lo que no podemos concluir que no hay mujeres adultas en la muestra con hipertensión.
29. Estadístico de prueba: $z = -1.64$. Valor $P = 0.1015$ (Tabla: 0.1010). Valor crítico: $z = -1.645$. Las hipótesis nula y alternativa son las mismas y las conclusiones son iguales entre sí. Los resultados no se ven afectados en gran medida por conocer σ.
31. La puntuación t crítica calculada es 1.648, que es igual al valor de 1.648 encontrado mediante el uso de tecnología. La aproximación parece funcionar bastante bien.

Sección 8-4

1. La muestra debe ser una muestra aleatoria simple y provenir de una población distribuida normalmente. El requisito de normalidad para una prueba de hipótesis de una afirmación sobre una desviación estándar es mucho más estricto, lo que significa que la distribución de la población debe ser mucho más cercana a una distribución normal.
3. a. Se rechaza H_0.
   b. Se rechaza la afirmación de que el nuevo proceso de llenado da como resultado volúmenes con la misma desviación estándar de 0.115 oz.
   c. Parece que con el nuevo proceso de llenado, la variación entre los volúmenes ha aumentado, por lo que los volúmenes no son tan consistentes. El nuevo proceso de llenado parece ser inferior al proceso de llenado original.
5. H_0: $\sigma = 10$ lpm. H_1: $\sigma \neq 10$ lpm. Estadístico de prueba: $\chi^2 = 195.172$. Valor $P = 0.0208$. Se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la afirmación de que las frecuencias de pulso de los hombres tienen una desviación estándar de 10 latidos por minuto. Usar la regla práctica del rango con el rango normal de 60 a 100 latidos por minuto no es muy bueno para estimar σ en este caso.
7. H_0: $\sigma = 2.08$°F. H_1: $\sigma < 2.08$°F. Estadístico de prueba: $\chi^2 = 9.329$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 74.252$ (Tabla: 70.065 aproximadamente). Se rechaza H_0. Hay suficiente evidencia para respaldar la afirmación de que las temperaturas corporales tienen una desviación estándar menor a 2.08°F. Es muy poco probable que la conclusión en la prueba de hipótesis del Ejemplo 5 de la sección 8-3 cambie, debido a una desviación estándar de una muestra diferente.
9. H_0: $\sigma = 27.8$ lb. H_1: $\sigma \neq 27.8$ lb. Estadístico de prueba: $\chi^2 = 8.505$. Valor $P = 0.0384$ (Tabla: >0.05). Valores críticos de χ^2: 8.907, 32.852. Se rechaza H_0. Existe evidencia suficiente para garantizar el rechazo de la afirmación de que las latas con un espesor de 0.0109 pulgadas tienen cargas axiales con la misma desviación estándar que las cargas axiales de las latas que tienen un grosor de 0.0111 pulgadas. El espesor de las latas parece afectar la variación de las cargas axiales.
11. H_0: $\sigma = 0.15$ oz. H_1: $\sigma > 0.15$ oz. Estadístico de prueba: $\chi^2 = 33.396$. Valor $P = 0.1509$ (Tabla: >0.10). Valor crítico: $\chi^2 = 38.885$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que la máquina dispensa cantidades con una desviación estándar mayor que la desviación estándar de 0.15 oz. especificada en el diseño de la máquina.
13. H_0: $\sigma = 32.2$ pies. H_1: $\sigma > 32.2$ pies. Estadístico de prueba: $\chi^2 = 29.176$. Valor P: 0.0021. Valor crítico: $\chi^2 = 19.675$. Se rechaza H_0. Hay evidencia suficiente para respaldar la afirmación de que el nuevo método de producción tiene errores con una desviación estándar mayor a 32.2 pies. La variación parece ser mayor que en el pasado, por lo que el nuevo método parece ser peor porque habrá más altímetros que tendrán errores más grandes. La compañía debe tomar medidas inmediatas para reducir la variación.
15. H_0: $\sigma = 55.93$ seg. H_1: $\sigma \neq 55.93$ seg. Estadístico de prueba: $\chi^2 = 20.726$. Valor $P = 0.0084$ (Tabla: <0.01). Valores críticos de χ^2: 0.989, 20.278. Se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la afirmación de que los tiempos de servicio en McDonald's tienen la misma variación que los tiempos de servicio en Wendy's. Los tiempos de servicio durante la cena parecen variar más en McDonald's que en Wendy's. Dada la composición similar de los menús, McDonald's debería considerar métodos para reducir la variación.
17. H_0: $\sigma = 55.93$ seg. H_1: $\sigma \neq 55.93$ seg. Estadístico de prueba: $\chi^2 = 62.049$. Valor $P = 0.1996$ (Tabla: >0.10). Valores críticos de χ^2: 27.249, 78.231 (Tabla: 27.991, 79.490 aproximadamente). No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los tiempos de servicio en McDonald's tienen la misma variación que los tiempos de servicio en Wendy's. Los tiempos de servicio durante la comida parecen tener la misma variación en McDonald's y Wendy's. No se recomienda realizar ninguna acción.
19. χ^2 crítica $= 81.540$ (o 81.494 si se usa $z = 2.326348$ encontrada a partir de la tecnología), que está cerca del valor de 82.292 obtenido de Statdisk y Minitab.

Capítulo 8: Examen rápido

1. a. distribución t b. Distribución normal
   c. Distribución ji cuadrada
2. a. Dos colas b. Cola izquierda c. Cola derecha
3. a. H_0: $p = 0.5$. H_1: $p > 0.5$. b. $z = 1.39$
   c. No se puede rechazar H_0
   d. No hay evidencia suficiente para respaldar la afirmación de que la mayoría de los usuarios de Internet de entre 18 y 29 años usan Instagram.
4. 0.10 5. Verdadero 6. Falso 7. Falso
8. No. Todos los valores críticos de χ^2 son siempre positivos.
9. La prueba t requiere que la muestra proceda de una población distribuida normalmente, y la prueba es sólida en el sentido de que la

prueba funciona razonablemente bien si la desviación de la normalidad no es demasiado extrema. La prueba χ^2 (ji cuadrada) no es sólida frente a una desviación de la normalidad, lo que significa que la prueba no funciona bien si la población tiene una distribución que está lejos de ser normal.

Capítulo 8: Ejercicios de repaso

1. a. Falso b. Verdadero c. Falso d. Falso e. Falso
2. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 6.09$. Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 2.33$. Se rechaza H_0. Hay suficiente evidencia para respaldar la hipótesis de que el método de selección de boletas favorece a los Demócratas.
3. H_0: $\mu = 30$ años. H_1: $\mu > 30$ años. Estadístico de prueba: $t = 5.029$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = 2.370$ (Tabla: 2.368 aproximadamente). Se rechaza H_0. Hay suficientes pruebas para respaldar la hipótesis de que la edad media de las actrices cuando ganan el Óscar es mayor de 30 años.
4. H_0: $\mu = 5.4$ millones de células por microlitro. H_1: $\mu < 5.4$ millones de células por microlitro. Estadístico de prueba: $t = -5.873$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = -2.426$. Se rechaza H_0. Existe evidencia suficiente para respaldar la hipótesis de que la muestra proviene de una población con una media inferior a 5.4 millones de células por microlitro. La prueba trata de la distribución de las medias muestrales, no de los valores individuales, por lo que el resultado no sugiere que cada uno de los 40 hombres tenga un conteo de glóbulos rojos por debajo de 5.4 millones de células por microlitro.
5. H_0: $p = 0.43$. H_1: $p \neq 0.43$. Estadístico de prueba: $z = 3.70$. Valor P: 0.0002. Valores críticos: $z = \pm 1.96$. Se rechaza H_0. Hay suficiente evidencia para justificar el rechazo de la hipótesis de que el porcentaje que cree que votaron por el candidato ganador es igual al 43%. Parece haber una discrepancia sustancial entre cómo las personas dijeron que votaron y cómo realmente lo hicieron.
6. H_0: $\mu = 20.16$. H_1: $\mu < 20.16$. Estadístico de prueba: $t = -3.732$. Valor $P = 0.0023$ (Tabla: <0.005). Valor crítico: $t = -2.821$. Se rechaza H_0. Existen pruebas suficientes para respaldar la hipótesis de que la población de ganadores recientes tiene un IMC medio inferior a 20.16. Los ganadores recientes parecen ser significativamente más pequeños que los de los años veinte y treinta.
7. H_0: $\sigma = 1.34$. H_1: $\sigma \neq 1.34$. Estadístico de prueba: $\chi_2 = 7.053$. Valor $P = 0.7368$ (Tabla: >0.20). Valores críticos de χ^2: 1.735, 23.589. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la hipótesis de que los ganadores recientes tienen valores de IMC con una variación diferente a la de las décadas de 1920 y 1930.
8. a. Un error de tipo I es el error de rechazar una hipótesis nula cuando en realidad es verdadera. Un error de tipo II es el error de no rechazar una hipótesis nula cuando en realidad es falsa.
   b. Error tipo I: en realidad, el IMC medio es igual a 20.16, pero respaldamos la hipótesis de que el IMC medio es menor que 20.16. Error tipo II: en realidad, el IMC medio es inferior a 20.16, pero no logramos respaldarlo.

Capítulo 8: Ejercicios de repaso acumulado

1. a. 37.1 muertes b. 36.0 muertes c. 9.8 muertes
   d. 96.8 muertes^2 e. 28.0 muertes
   f. El patrón de los datos a través del tiempo no es revelado por los estadísticos. Una gráfica de series de tiempo sería muy útil para comprender el patrón a lo largo del tiempo.
2. a. De razón b. Discreto c. Cuantitativo
   d. No. Los datos son de años recientes y consecutivos, por lo que no se seleccionan al azar.
3. 29.1 muertes $< \mu <$ 45.0 muertes. Tenemos un 99% de confianza de que los límites de 29.1 muertes y 45.0 muertes contienen el valor de la media poblacional.
4. H_0: $\mu = 72.6$ muertes. H_1: $\mu < 72.6$ muertes. Estadístico de prueba: $t = -13.509$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = -2.650$. Se rechaza H_0. Hay suficientes pruebas para respaldar la hipótesis de que el número medio de muertes anuales por rayos es ahora menor que el promedio de 72.6 muertes de los años ochenta. Posibles factores: Cambio en la población de las áreas rurales a las urbanas; mejor protección contra rayos y conexión a tierra en líneas eléctricas y de cable y telefónicas; mejor tratamiento médico de las personas afectadas por un rayo; menos personas usan teléfonos conectados a cables; mejores predicciones meteorológicas.
5. Debido a que la escala vertical comienza en 50 y no en 0, la diferencia entre el número de hombres y el número de mujeres es exagerada, por lo que la gráfica es engañosa al crear la falsa impresión de que los hombres representan casi todas las muertes por rayos. Una comparación del número de muertes muestra que el número de muertes de hombres es aproximadamente 4 veces el número de muertes femeninas, pero el gráfico hace que parezca que el número de muertes de hombres es alrededor de 25 veces el número de muertes femeninas.
6. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 10.45$. Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 2.33$. Se rechaza H_0. Hay evidencia suficiente para respaldar la hipótesis de que la proporción de muertes de hombres es mayor de 1/2. Más hombres están involucrados en ciertas actividades al aire libre como la construcción, la pesca y el golf.
7. $0.763 < p < 0.854$. Debido a que el intervalo de confianza total es mayor a 0.5, no parece factible que los hombres y las mujeres tengan las mismas posibilidades de morir por la caída de un rayo.
8. a. 0.512 b. 0.008 c. 0.992 d. 0.205
   e. $\mu = 40.0$ hombres; $\sigma = 2.8$ hombres
   f. Sí. Usando la regla práctica del rango, los valores significativamente altos son $\mu + 2\sigma$ o mayores. Con $\mu + 2\sigma = 45.6$, los valores superiores a 45.6 son significativamente altos, por lo que 46 sería un número significativamente alto de víctimas masculinas en un grupo de 50.

Respuestas del capítulo 9

Sección 9-1

1. Las muestras son muestras aleatorias simples que son independientes. Para cada uno de los dos grupos, el número de éxitos es de al menos 5 y el número de fracasos es de al menos 5. (Dependiendo de lo que llamemos un éxito, los cuatro números son 33, 115, 201, 196 y 200, 630 y todos esos números son al menos 5). Se satisfacen los requisitos.
3. a. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$.
   b. Existen pruebas suficientes para respaldar la afirmación de que la tasa de poliomielitis es menor para los niños que recibieron la vacuna Salk que para los niños que recibieron un placebo. La vacuna Salk parece ser efectiva.
5. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 12.82$. Valor P: 0.0000. Valor crítico: $z = 2.33$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que los guantes de vinilo tienen una mayor tasa de fuga de virus que los guantes de látex.

7. a. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = -0.95$. Valor $P = 0.8280$ (Tabla: 0.8289). Valor crítico: $z = 1.645$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que los propietarios de automóviles violan las leyes de placas de matrícula a una tasa mayor que los propietarios de camiones comerciales.
   b. IC del 90%: $-0.0510 < p_1 - p_2 < 0.0148$. Debido a que los límites del intervalo de confianza contienen 0, no hay una diferencia significativa entre las dos proporciones. No hay evidencia suficiente para respaldar la afirmación de que los propietarios de automóviles violan las leyes de placas de matrícula a un ritmo mayor que los propietarios de camiones comerciales.
9. a. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 2.64$. Valor P: 0.0041. Valor crítico: $z = 2.33$. Se rechaza H_0. Existe suficiente evidencia para respaldar la afirmación de que la tasa de éxito para dejar de fumar es mayor con el programa de cuidado sostenido.
   b. IC del 98%: $0.0135 < p_1 - p_2 < 0.200$ (Tabla: $0.0134 < p_1 - p_2 < 0.200$). Debido a que los límites del intervalo de confianza no contienen 0, existe una diferencia significativa entre las dos proporciones. Debido a que el intervalo consta sólo de números positivos, parece que la tasa de éxito para el programa de atención sostenida es mayor que la tasa de éxito para el programa de atención estándar.
   c. Según las muestras, las tasas de éxito de los programas son del 25.8% (atención sostenida) y del 15.1% (atención estándar). Esa diferencia parece ser sustancial, por lo que la diferencia entre los programas parece tener una significancia práctica.
11. a. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 6.44$. Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 2.33$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la proporción de personas mayores de 55 años que sueñan en blanco y negro es mayor que la proporción de menores de 25 años.
   b. IC del 98%: $0.117 < p_1 - p_2 < 0.240$. Debido a que los límites del intervalo de confianza no incluyen 0, parece que las dos proporciones no son iguales. Debido a que los límites del intervalo de confianza incluyen sólo valores positivos, parece que la proporción de personas mayores de 55 años que sueñan en blanco y negro es mayor que la proporción de menores de 25 años.
   c. Los resultados sugieren que la proporción de personas mayores de 55 años que sueñan en blanco y negro es mayor que la proporción de menores de 25 años, pero los resultados no se pueden utilizar para verificar la causa de esa diferencia.
13. a. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 6.11$. Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 1.645$. Se rechaza H_0. Hay pruebas suficientes para respaldar la afirmación de que la tasa de mortalidad es más alta para quienes no usan cinturones de seguridad.
   b. IC del 90%: $0.00559 < p_1 - p_2 < 0.0123$. Debido a que los límites del intervalo de confianza no incluyen 0, parece que las dos tasas de mortalidad no son iguales. Debido a que los límites del intervalo de confianza incluyen sólo valores positivos, parece que la tasa de mortalidad es mayor para quienes no usan cinturones de seguridad.
   c. Los resultados sugieren que el uso de cinturones de seguridad se asocia con tasas de mortalidad inferiores a las asociadas con el uso de los cinturones de seguridad.
15. a. H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$. Estadístico de prueba: $z = 0.57$. Valor P: 0.5720 (Tabla: 0.5686). Valores críticos: $z = \pm 1.96$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que el tratamiento de la equinácea tiene un efecto.
   b. IC del 95%: $-0.0798 < p_1 - p_2 < 0.149$. Debido a que los límites del intervalo de confianza contienen 0, no hay una diferencia significativa entre las dos proporciones. No hay evidencia suficiente para respaldar la afirmación de que el tratamiento de la equinácea tiene algún efecto. c. La equinácea no parece tener un efecto significativo en la tasa de infección. Debido a que no parece tener un efecto, no debería recomendarse.
17. a. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = -7.94$. Valor P: 0.0000 (Tabla: 0.0001). Valor crítico: $z = -2.33$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la tasa de diestros que prefieren usar su oreja izquierda para los teléfonos celulares es menor que la tasa de diestros que prefieren usar su oreja derecha para los teléfonos celulares.
   b. IC del 98%: $-0.266 < p < -0.126$. Debido a que los límites del intervalo de confianza no contienen 0, existe una diferencia significativa entre las dos proporciones. Como el intervalo consta únicamente de números negativos, parece que la afirmación se sostiene. La diferencia entre las poblaciones parece tener una significancia práctica.
19. a. H_0: $p_1 = p_2$. H_1: $p_1 > p_2$. Estadístico de prueba: $z = 9.97$. Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 2.33$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la tasa de curación con el tratamiento con oxígeno es más alta que la tasa de curación para los que reciben un placebo. Parece que el tratamiento con oxígeno es efectivo.
   b. IC del 98%: $0.467 < p_1 - p_2 < 0.687$. Debido a que los límites del intervalo de confianza no incluyen 0, parece que las dos tasas de curación no son iguales. Debido a que los límites del intervalo de confianza incluyen sólo valores positivos, parece que la tasa de curación con el tratamiento con oxígeno es más alta que la tasa de curación para los que recibieron un placebo. Parece que el tratamiento con oxígeno es efectivo.
   c. Los resultados sugieren que el tratamiento con oxígeno es efectivo para curar los dolores de cabeza en racimo.
21. a. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = -1.17$. Valor $P = 0.1214$ (Tabla: 0.1210). Valor crítico: $z = -2.33$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que la tasa de zurdera entre los hombres es menor que la de las mujeres.
   b. IC del 98%: $-0.0848 < p_1 - p_2 < 0.0264$ (Tabla: $-0.0849 < p_1 - p_2 < 0.0265$). Debido a que los límites del intervalo de confianza incluyen 0, no parece haber una diferencia significativa entre la tasa de zurdera entre los hombres y la tasa entre las mujeres. No hay pruebas suficientes para apoyar la afirmación de que la tasa de zurdera entre los hombres es menor que la de las mujeres.
   c. La tasa de zurdera entre los hombres no parece ser menor que la tasa de zurdera entre las mujeres.
23. Las muestras deberían incluir 2135 hombres y 2135 mujeres.
25. a. $0.0227 < p_1 - p_2 < 0.217$; debido a que los límites del intervalo de confianza no contienen 0, parece que $p_1 = p_2$ puede rechazarse.
   b. $0.491 < p_1 < 0.629$; $0.371 < p_2 < 0.509$; debido a que los intervalos de confianza se superponen, parece que $p_1 = p_2$ no se puede rechazar.
   c. H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$. Estadístico de prueba: $z = 2.40$. Valor P: 0.0164. Valores críticos: $z = \pm 1.96$. Se rechaza H_0. Hay suficiente evidencia para rechazar $p_1 = p_2$.
   d. Se rechaza $p_1 = p_2$. Método menos efectivo: usar la superposición entre intervalos de confianza individuales.

Sección 9-2

1. Sólo el inciso (c) describe muestras independientes.
3. a. Sí b. Sí c. 98%
5. a. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$. Estadístico de prueba: $t = -22.092$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = -1.672$ (Tabla: -1.690). Se rechaza H_0. Hay suficiente evidencia para respaldar la afirmación de que el contenido de las latas de Diet Coke tiene pesos con una media que es menor que la media para la Coca Cola normal.
   b. IC del 90%: $-0.03445 \text{ lb} < (\mu_1 - \mu_2) < -0.02961 \text{ lb}$ (Tabla: $-0.03448 \text{ lb} < (\mu_1 - \mu_2) < -0.02958 \text{ lb}$).
   c. El contenido en latas de Diet Coke parece pesar menos, probablemente debido al azúcar presente en la Coca Cola normal pero no en la Diet Coke.
7. a. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$. Estadístico de prueba: $t = -2.979$. Valor $P = 0.0021$ (Tabla: <0.005). Valor crítico: $t = -2.392$ (Tabla: -2.441). Se rechaza H_0. Existe suficiente evidencia para respaldar la afirmación de que el azul mejora el rendimiento en una tarea creativa.
   b. IC del 98%: $-1.05 < \mu_1 - \mu_2 < -0.11$ (Tabla: $-1.06 < \mu_1 - \mu_2 < -0.10$). El intervalo de confianza consiste únicamente en números negativos y no incluye 0, por lo que la puntuación promedio de creatividad con el fondo rojo parece ser menor que la puntuación promedio de creatividad con el fondo azul. Parece que el azul mejora el rendimiento en una tarea creativa.
9. a. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. Estadístico de prueba: $t = 0.132$. Valor $P = 0.4480$ (Tabla: > 0.10). Valor crítico: $t = 1.691$ (Tabla: 1.729). No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que los imanes son efectivos para reducir el dolor.
   b. IC del 90%: $-0.59 < \mu_1 - \mu_2 < 0.69$ (Tabla: $-0.61 < \mu_1 - \mu_2 < 0.71$).
   c. Los imanes no parecen ser efectivos para tratar el dolor de espalda. Es válido argumentar que los imanes podrían parecer efectivos si los tamaños de muestra fueran más grandes.
11. a. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 0.674$. Valor $P = 0.5015$ (Tabla: > 0.20). Valores críticos: $t = \pm 1.979$ (Tabla: ± 1.994). No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las mujeres y los hombres tienen el mismo IMC medio.
   b. IC del 95%: $-1.39 < (\mu_1 - \mu_2) < 2.83$ (Tabla: $-1.41 < (\mu_1 - \mu_2) < 2.85$). Debido a que el intervalo de confianza incluye 0, no hay evidencia suficiente para justificar el rechazo de la afirmación de que las dos muestras provienen de poblaciones con la misma media.
   c. Según los datos muestrales disponibles, parece que los hombres y las mujeres tienen el mismo IMC medio, pero sólo podemos concluir que no hay evidencia suficiente para decir que son diferentes.
13. a. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = -2.025$. Valor $P = 0.0460$ (Tabla: <0.05). Valores críticos: $t = \pm 1.988$ (Tabla: ± 2.023). Se rechaza H_0. Existe evidencia suficiente para justificar el rechazo de la afirmación de que las dos muestras provienen de poblaciones con la misma media.
   b. IC del 95%: $-0.44 < (\mu_1 - \mu_2) < 0.00$. Debido a que el intervalo de confianza sólo incluye números negativos y no incluye 0, existe evidencia suficiente para justificar el rechazo de la afirmación de que las dos muestras provienen de poblaciones con la misma media.
   c. Sí. Con las muestras más pequeñas de tamaño 12 y 15, no hubo evidencia suficiente para justificar el rechazo de la hipótesis nula, pero hay pruebas suficientes con muestras más grandes.
15. a. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. Estadístico de prueba: $t = 32.771$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = 1.667$ (Tabla: 1.685). Se rechaza H_0. Hay suficiente evidencia para respaldar la afirmación de que las monedas de ¢25 anteriores a 1964 tienen un peso medio que es mayor que el peso medio de las monedas de ¢25 posteriores a 1964.
   b. IC del 90%: $0.52522 \text{ lb} < (\mu_1 - \mu_2) < 0.58152$ (Tabla: $0.52492 \text{ lb} < (\mu_1 - \mu_2) < 0.58182$)
   c. Sí. Las máquinas expendedoras no se ven afectadas porque las monedas de ¢25 anteriores a 1964 están fuera de circulación.
17. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = -0.315$. Valor $P = 0.7576$ (Tabla: > 0.20). Suponiendo un nivel de significancia de 0.05, los valores críticos son $t = \pm 2.159$ (Tabla: ± 2.262). No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las profesoras y los profesores tienen las mismas puntuaciones medias de evaluación. No parece haber una diferencia entre las puntuaciones de evaluación de los profesores y las profesoras.
19. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = -2.385$. Valor $P = 0.0244$ (Tabla: <0.05). Suponiendo un nivel de significancia de 0.05, los valores críticos son $t = \pm 2.052$ (Tabla: ± 2.201). La conclusión depende de la elección del nivel de significancia. Hay una diferencia significativa entre las dos medias poblacionales en el nivel de significancia de 0.05, pero no en el nivel de significancia de 0.01.
21. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$. Estadístico de prueba: $t = -0.132$. Valor $P = 0.4477$ (Tabla: > 0.10). Valor crítico: $t = -1.669$ (Tabla: -1.688). No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que los hombres hablan menos que las mujeres.
23. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$. Estadístico de prueba: $t = -3.450$. Valor $P = 0.0003$ (Tabla: <0.005). Suponiendo un nivel de significancia de 0.05, el valor crítico es $t = -1.649$ (Tabla: -1.653). Se rechaza H_0. Hay evidencia suficiente para respaldar la afirmación de que, al nacer, las niñas tienen un peso medio menor que los niños.
25. Con el agrupamiento, los gl aumentan dramáticamente a 97, pero el estadístico de prueba disminuye de 2.282 a 1.705 (porque la desviación estándar estimada aumenta de 2.620268 a 3.507614), el valor P aumenta a 0.0457, y el intervalo de confianza del 90% se hace más amplio. Con la agrupación, estos resultados no muestran mayor significancia.
27. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = 15.322$. Valor $P = 0.0000$ (Tabla: <0.01). Valores críticos: $t = \pm 2.080$. Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la afirmación de que las dos poblaciones tienen la misma media.

Sección 9-3

1. Sólo los incisos (a) y (c) son verdaderos.
3. Los resultados serán los mismos.
5. a. H_0: $\mu_d = 0$ años. H_1: $\mu_d < 0$ años. Estadístico de prueba: $t = -2.609$. Valor $P = 0.0142$ (Tabla: <0.025). Valor crítico: $t = -1.833$. Se rechaza H_0. Hay suficiente evidencia para respaldar la afirmación de que para la población de las edades de las mejores actrices y mejores actores, las diferencias tienen una media inferior a 0. Hay suficiente evidencia para concluir que las mejores actrices son en general más jóvenes que los mejores actores.

b. IC del 90%: -16.5 años $< \mu_d < -2.9$ años. El intervalo de confianza consiste sólo en números negativos y no incluye 0.

7. a. H_0: $\mu_d = 0$°F. H_1: μ_d: $\neq 0$°F. Estadístico de prueba: $t = -7.499$. Valor $P = 0.0003$ (Tabla: <0.01). Valores críticos: $t = \pm 2.447$. Se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la afirmación de que no hay diferencia entre las temperaturas corporales medidas a las 8 AM y a las 12 AM. Parece haber una diferencia.
   b. -1.97°F $< \mu_d < -1.00$ °F. El intervalo de confianza consiste en números negativos solamente y no incluye 0.
9. H_0: $\mu_d = 0$ pulg. H_1: $\mu_d \neq 0$ pulg. Estadístico de prueba: $t = -1.379$. Valor $P = 0.2013$ (Tabla: >0.20). Valores críticos: $t = \pm 2.262$. No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la afirmación de que no hay diferencia en las estaturas entre las madres y sus primeras hijas.
11. H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = 0.793$. Valor $P = 0.4509$ (Tabla: >0.20). Valores críticos: $t = \pm 2.306$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que existe una diferencia entre las calificaciones de los atributos femeninos y las calificaciones de los atributos masculinos.
13. -6.5 admisiones $< \mu_d < -0.2$ admisiones. Debido a que el intervalo de confianza no incluye 0 admisiones, parece que hay suficiente evidencia para justificar el rechazo de la afirmación de que cuando el día 13 de un mes cae en viernes, el número de ingresos hospitalarios por choques automovilísticos no se ve afectado. Las admisiones al hospital parecen verse afectadas.
15. $0.69 < \mu_d < 5.56$. Debido a que los límites del intervalo de confianza no contienen 0 y consisten sólo en valores positivos, parece que las mediciones "previas" son mayores que las mediciones "posteriores", por lo que la hipnosis parece ser efectiva para reducir el dolor.
17. a. H_0: $\mu_d = 0$ años. H_1: $\mu_d < 0$ años. Estadístico de prueba: $t = -5.185$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = -1.663$ (Tabla: -1.662). Se rechaza H_0. Hay evidencia suficiente para respaldar la afirmación de que las actrices generalmente son más jóvenes que los actores.
    b. IC del 90%: -10.4 años $< \mu_d < -5.4$ años. El intervalo de confianza consiste sólo en números negativos y no incluye 0.
19. a. H_0: $\mu_d = 0$°F. H_1: $\mu_d \neq 0$°F. Estadístico de prueba: $t = -8.485$. Valor $P = 0.0000$ (Tabla: <0.01). Valores críticos: $t = \pm 1.996$ (Tabla: ± 1.994). Se rechaza H_0. Existe evidencia suficiente para garantizar el rechazo de la afirmación de que no hay diferencia entre las temperaturas corporales medidas a las 8 AM y a las 12 AM. Parece haber una diferencia.
    b. IC del 95%: -1.05°F $< \mu_d < -0.65$°F. El intervalo de confianza consiste en números negativos solamente y no incluye 0.
21. H_0: $\mu_d = 0$ pulg. H_1: $\mu_d \neq 0$ pulg. Estadístico de prueba: $t = -4.090$. Valor $P = 0.0001$ (Tabla: <0.01). Valores críticos: $t = \pm 1.978$ (Tabla: ± 1.984 aproximadamente). Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la afirmación de que no hay diferencias en las estaturas entre las madres y sus primeras hijas.
23. H_0: $\mu_d = 0$. H_1: $\mu_d \neq 0$. Estadístico de prueba: $t = 0.191$. Valor $P = 0.8485$ (Tabla: >0.20). Valores críticos: ± 1.972. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que existe una diferencia entre las calificaciones de los atributos femeninos y las calificaciones de los atributos masculinos.
25. Para las temperaturas en grados Fahrenheit y las temperaturas en grados Celsius, el estadístico de prueba ($t = 0.124$) es el mismo, el valor P de 0.9023 es el mismo, los valores críticos ($t = \pm 2.028$) son los mismos, y las conclusiones son las mismas, por lo que los resultados de la prueba de hipótesis son iguales en ambos casos. Los intervalos de confianza son -0.25°F $< \mu_d < 0.28$°F y -0.14°C $< \mu_d < 0.16$°C. Los límites del intervalo de confianza de -0.14°C y 0.16°C tienen valores numéricos que son 5/9 de los valores numéricos de -0.25°F y 0.28°F.

Sección 9-4

1. a. No. b. No.
   c. Las dos muestras tienen desviaciones estándar (o varianzas) que tienen un valor muy cercano.
   d. Asimétrico a la derecha
3. No. A diferencia de algunas otras pruebas que requieren que las muestras sean de poblaciones distribuidas normalmente o que las muestras deben tener más de 30 valores, la prueba F requiere que las muestras sean de poblaciones normalmente distribuidas, independientemente del tamaño de las muestras.
5. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 2.3706$. Valor P: 0.0129. Valor crítico superior F: 1.9678 (Tabla: El valor crítico superior F está entre 1.8752 y 2.0739). Se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la afirmación de que las puntuaciones de las tareas creativas tienen la misma variación con un fondo rojo y un fondo azul.
7. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 > \sigma_2$. Estadístico de prueba: $F = 9.3364$. Valor P: 0.0000. El valor crítico F es 2.0842 (Tabla: El valor crítico F está entre 2.0540 y 2.0960). Se rechaza H_0. Hay evidencia suficiente para respaldar la afirmación de que el grupo de tratamiento tiene errores que varían más que los errores del grupo placebo.
9. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 2.9265$. Valor P: 0.0020. El valor crítico superior F es 1.9611 (Tabla: El valor crítico superior F está entre 1.8752 y 2.0739). Se rechaza H_0. Existe evidencia suficiente para justificar el rechazo de la afirmación de que la variación es la misma para ambos tipos de Coca Cola.
11. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 > \sigma_2$. Estadístico de prueba: $F = 2.1267$. Valor P: 0.0543. El valor crítico F es 2.1682 (Tabla: El valor crítico F está entre 2.1555 y 2.2341). No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que quienes reciben un tratamiento simulado tienen reducciones de dolor que varían más que las reducciones de dolor para los tratados con imanes.
13. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 4.1750$. Valor P: 0.0447. Valor crítico: 4.0260. Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la afirmación de que las profesoras y los profesores tienen puntuaciones de evaluación con la misma variación.
15. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 2.3095$. Valor P: 0.1635. Suponiendo un nivel de significancia de 0.05, el valor crítico es 3.3044 (Tabla: El valor crítico está entre 3.2261 y 3.3299). No se puede rechazar H_0. No hay evidencia suficiente para respaldar una afirmación de que la variación de los tiempos entre las erupciones ha cambiado.
17. $c_1 = 3$, $c_2 = 0$, el valor crítico es 7.4569. No se puede rechazar H_0. No hay evidencia suficiente para respaldar una afirmación de que las dos poblaciones de puntuaciones tienen diferentes cantidades de variación.
19. $F_L = 0.4103$; $F_R = 2.7006$.

Capítulo 9: Examen rápido

1. H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$.
2. $x_1 = 258$, $x_2 = 282$, $\hat{p}_1 = 258/1121 = 0.230$, $\hat{p}_2 = 282/1084 = 0.260$, $\bar{p} = 0.245$.
3. 0.1015 (Tabla: 0.1010)
4. a. $-0.0659 < p_1 - p_2 < 0.00591$
   b. El intervalo de confianza incluye el valor de 0, por lo que es posible que las dos proporciones sean iguales. No hay una diferencia significativa.
5. No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la afirmación de que, para las personas que conocían la declaración, la proporción de mujeres es igual a la proporción de hombres.
6. Verdadero 7. Falso
8. Debido a que los datos consisten en pares relacionados, son dependientes.
9. H_0: $\mu_d = 0$ mm Hg. H_1: $\mu_d \neq 0$ mm Hg.
10. a. $t = \dfrac{\bar{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}}$

    b. $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

    c. $z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1} + \dfrac{\bar{p}\bar{q}}{n_2}}}$

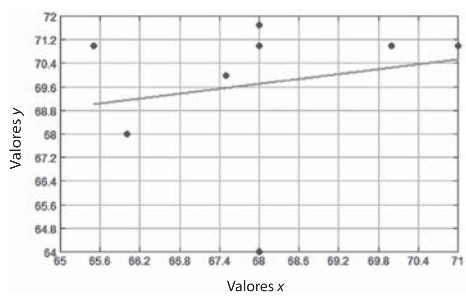
    d. $F = \dfrac{s_1^2}{s_2^2}$

Capítulo 9: Ejercicios de repaso

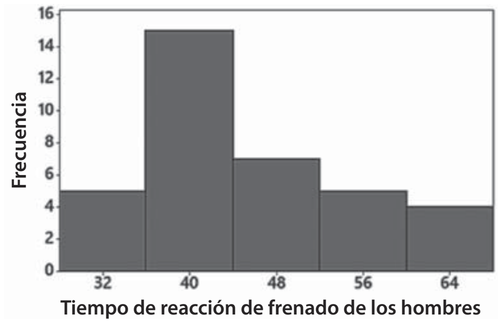
1. H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. Estadístico de prueba: $z = -3.49$. Valor P: 0.0002. Valor crítico: $z = -1.645$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que es menos probable que el dinero en una denominación grande se gaste en relación con una cantidad equivalente en denominaciones más pequeñas.
2. IC del 90%: $-0.528 < p_1 - p_2 < -0.206$. Los límites del intervalo de confianza no contienen 0, por lo que parece que hay una diferencia significativa entre las dos proporciones. Debido a que el intervalo de confianza consiste sólo en valores negativos, parece que p_1 es menor que p_2, por lo que parece que es menos probable que el dinero en una denominación grande se gaste en relación con una cantidad equivalente en denominaciones más pequeñas.
3. -25.33 cm $< (\mu_1 - \mu_2) < -7.51$ cm (Tabla: -25.70 cm $< (\mu_1 - \mu_2) < -7.14$ cm). Con un 95% de confianza, concluimos que la estatura media de las mujeres es menor que la estatura media de los hombres en una cantidad que está entre 7.51 cm y 25.33 cm (Tabla: 7.14 cm y 25.70 cm).
4. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$. Estadístico de prueba: $t = -4.001$. Valor $P = 0.0008$ (Tabla: <0.005). Valor crítico: $t = -2.666$ (Tabla: -2.821). Se rechaza H_0. Hay suficiente evidencia para apoyar la afirmación de que las mujeres tienen estaturas con una media que es menor que la estatura media de los hombres. Pero usted ya lo sabía.
5. H_0: $\mu_d = 0$. H_1: $\mu_d > 0$. Estadístico de prueba: $t = 6.371$. Valor $P = 0.0000$. (Tabla: <0.005). Valor crítico: $t = 2.718$. Se rechaza H_0. Hay evidencia suficiente para respaldar la afirmación de que el captopril es eficaz para reducir la presión arterial sistólica.
6. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. Estadístico de prueba: $t = 2.879$. Valor $P = 0.0026$ (Tabla: <0.005). Valor crítico: $t = 2.376$ (Tabla: 2.429). Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la cantidad media de detalles recordados es menor para el grupo de estrés. Parece que "el estrés disminuye la cantidad recordada", pero no debemos concluir que el estrés es la causa de la disminución.
7. H_0: $\mu_d = 0$. H_1: $\mu_d > 0$. Estadístico de prueba: $t = 14.061$. Valor $P = 0.0000$ (Tabla: <0.005). Valor crítico: $t = 3.365$. Se rechaza H_0. Existe suficiente evidencia para respaldar la afirmación de que los vuelos programados con 1 día de antelación cuestan más que los vuelos programados con 30 días de anticipación. Ahorre dinero programando vuelos con 30 días de anticipación.
8. H_0: $\sigma_1 = \sigma_2$. H_1: $\sigma_1 \neq \sigma_2$. Estadístico de prueba: $F = 4.9933$. Valor P: 0.0252. Valor superior crítico F: 4.0260. Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la afirmación de que las mujeres y los hombres tienen estaturas con la misma variación.

Capítulo 9: Ejercicios de repaso acumulado

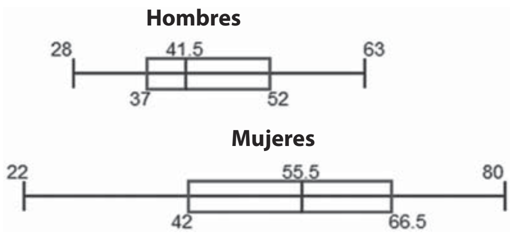
1. a. Debido a que los datos muestrales se corresponden con cada columna que consta de las estaturas de la misma familia, los datos son dependientes.
   b. $\bar{x} = 69.7$ pulg; mediana $= 71.0$ pulg; rango $= 7.7$ pulg; $s = 2.6$ pulg; $s^2 = 6.6$ pulg^2
   c. De razón
   d. Continuo
2. No parece haber una correlación o asociación entre las estaturas de los padres y las estaturas de sus hijos.



3. 67.6 pulg. $< \mu <$ 71.9 pulg. Tenemos un 95% de confianza de que los límites de 67.6 pulg. y 71.9 pulg. Contienen realmente el valor real de la estatura media de todos los hijos adultos.
4. H_0: $\mu_d = 0$ pulg. H_1: $\mu_d \neq 0$ pulg. Estadístico de prueba: $t = -1.712$. Valor $P = 0.1326$ (Tabla: >0.10). Valores críticos: $t = \pm 2.365$. No se puede rechazar H_0. No hay evidencia suficiente para garantizar el rechazo de la afirmación de que las diferencias entre las estaturas de los padres y sus hijos tienen una media de 0. No parece haber una diferencia entre las estaturas de los padres y sus hijos.
5. Debido a que los puntos se encuentran razonablemente cerca de un patrón en línea recta, y no hay otro patrón que no sea un patrón en línea recta y no hay valores atípicos, los datos muestrales parecen provenir de una población con una distribución normal.
6. La forma del histograma indica que los datos muestrales parecen provenir de una población con una distribución que es aproximadamente normal. (La respuesta continúa en la página siguiente)



7. Debido a que los puntos están razonablemente cerca de un patrón en línea recta y no hay otro patrón que no sea un patrón en línea recta, parece que los tiempos de reacción de frenado de las mujeres provienen de una población con una distribución normal.
8. Debido a que los diagramas de caja y bigote se superponen, no parece haber una diferencia significativa entre los tiempos de reacción de frenado de hombres y mujeres, pero los tiempos de reacción de frenado para los hombres parecen ser generalmente más bajos que los tiempos de reacción de frenado de las mujeres.

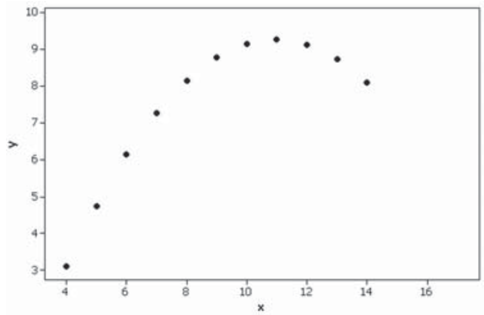


9. H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Estadístico de prueba: $t = -3.259$. Valor $P = 0.0019$ (Tabla: <0.005). Valores críticos: $t = \pm 2.664$ (Tabla: ± 2.724). Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la afirmación de que los hombres y las mujeres tienen el mismo tiempo de reacción de frenado medio. Los hombres parecen tener tiempos de reacción más bajos.
10 a. Hombres: $40.1 < \mu < 48.7$. Mujeres: $47.2 < \mu < 61.4$. Los intervalos de confianza se superponen, por lo que no parece haber una diferencia significativa entre los tiempos medios de reacción de frenado de hombres y mujeres.
   b. $-18.0 < \mu_1 - \mu_2 < -1.8$ (Tabla: $-18.2 < \mu_1 - \mu_2 < -1.6$). Debido a que el intervalo de confianza consiste en números negativos y no incluye 0, parece haber una diferencia significativa entre los tiempos medios de reacción de frenado de hombres y mujeres.
   c. Los resultados del inciso (b) son mejores.

Capítulo 10 Respuestas

Sección 10-1

1. a. r es un dato estadístico que representa el valor del coeficiente de correlación lineal calculado a partir de los datos muestrales emparejados, y ρ es un parámetro que representa el valor del coeficiente de correlación lineal que se calcularía utilizando todos los datos emparejados en la población de todos los estudiantes de estadística.
   b. El valor de r se estima en 0, porque es probable que no haya una correlación entre la temperatura corporal y la circunferencia de la cabeza.
   c. El valor de r no cambia si las temperaturas corporales se convierten a grados Fahrenheit.
3. No. Una correlación entre dos variables indica que de alguna manera están asociadas, pero esa asociación no implica necesariamente que una de las variables tenga un efecto directo sobre la otra variable. La correlación no implica causalidad.
5. Sí. $r = 0.963$. Valor $P = 0.000$. Valores críticos: ± 0.268 (Tabla: ± 0.279 aproximadamente). Hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre el peso de los osos y el tamaño de sus pechos. Es más fácil medir el tamaño del pecho de un oso que su peso, lo que requeriría levantar al oso sobre una báscula. Parece que el tamaño del pecho podría usarse para predecir el peso.
7. No. $r = 0.117$. Valor $P > 0.05$. Valores críticos: ± 0.250 (Tabla: ± 0.254 aproximadamente). No hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre los pesos del papel y el vidrio desechados.
9. a.



   b. $r = 0.816$. Valor $P = 0.002$ (Tabla: <0.01). Valores críticos: $r = \pm 0.602$ suponiendo un nivel de significancia de 0.05. Hay suficiente evidencia para apoyar el reclamo de una correlación lineal entre las dos variables.
   c. El diagrama de dispersión revela un patrón distinto que no es un patrón en línea recta.
11. a. La respuesta varía Debido a que parece haber un patrón ascendente, es razonable pensar que existe una correlación lineal.
   b. $r = 0.906$. Valor $P = 0.000$ (Tabla: <0.01). Valores críticos: $r = \pm 0.632$ (para un nivel de significancia de 0.05). Hay suficiente evidencia para respaldar la afirmación de una correlación lineal.
   c. $r = 0$. Valor $P = 1.000$ (Tabla: >0.05). Valores críticos: $r = \pm 0.666$ (para un nivel de significancia de 0.05). No hay suficiente evidencia para respaldar la afirmación de una correlación lineal.
   d. El efecto de un solo par de valores puede ser muy importante y puede cambiar la conclusión.
13. $r = 0.799$. Valor $P = 0.056$ (Tabla: >0.05). Valores críticos: $r = \pm 0.811$. No hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre los usuarios de Internet y los premios Nobel.
15. $r = 0.992$. Valor $P = 0.000$ (Tabla: <0.01). Valores críticos: $r = \pm 0.666$. Hay suficiente evidencia para respaldar la afirmación de que existe una correlación lineal significativa entre el costo de una rebanada de pizza y la tarifa del metro.
17. $r = 0.591$. Valor $P = 0.294$ (Tabla: >0.05). Valores críticos: $r = \pm 0.878$. No hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre la longitud de la huella del zapato y la estatura de los hombres. Los resultados dados no sugieren que la policía pueda usar la longitud de un zapato para calcular la estatura de un hombre.

19. $r = -0.959$. Valor $P = 0.010$. Valores críticos: $r = \pm 0.878$. Existen suficientes pruebas para respaldar la afirmación de que existe una correlación lineal entre el peso de las importaciones de limón de México y las tasas de mortalidad automovilística de Estados Unidos. Los resultados no sugieren ninguna relación de causa-efecto entre las dos variables.
21. $r = -0.288$. Valor $P = 0.365$ (Tabla: >0.05). Valores críticos: $r = \pm 0.576$. No hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal significativa entre las edades de las mejores actrices y los mejores actores. Debido a que las mejores actrices y los mejores actores suelen aparecer en diferentes películas, no deberíamos esperar que haya una correlación entre sus edades en el momento en que ganaron los premios.
23. $r = 0.948$. Valor $P = 0.004$ (Tabla: <0.01). Valores críticos: $r = \pm 0.811$. Hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre el ancho global de una foca en una fotografía y su peso.
25. $r = 0.828$. Valor $P = 0.042$ (Tabla: <0.05). Valores críticos: $r = \pm 0.811$. Existe suficiente evidencia para respaldar la afirmación de que existe una correlación lineal entre los montos de las cuentas y los montos de las propinas. Si todos tuvieran que dejar el mismo porcentaje, r sería 1.
27. $r = 1.000$. Valor $P = 0.000$. Valores críticos: $r = \pm 0.707$. Hay suficiente evidencia para respaldar la afirmación de que existe una correlación lineal entre los diámetros y las circunferencias. El diagrama de dispersión confirma una asociación lineal.
29. $r = 0.702$. Valor $P = 0.000$ (Tabla: <0.01). Valores críticos: ± 0.413 (Tabla: los valores críticos están entre ± 0.396 y ± 0.444). Existen pruebas suficientes para respaldar la afirmación de que existe una correlación lineal entre los usuarios de Internet y los premios Nobel.
31. $r = 0.594$. Valor $P = 0.007$ (Tabla: <0.01). Valores críticos: $r = \pm 0.456$. Hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre la longitud de la huella del zapato y la estatura de los hombres. Los resultados dados sugieren que la policía puede usar la longitud de un zapato para estimar la estatura de un hombre.
33. $r = 0.319$. Valor $P = 0.017$ (Tabla: <0.05). Valores críticos: ± 0.263 (Tabla: $r = \pm 0.254$ aproximadamente). Hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre el número de palabras habladas por hombres y mujeres que están en relaciones de pareja.
35. a. 0.911 b. 0.787 c. 0.9999 (más grande)
    d. 0.976 e. -0.948

Sección 10-2

1. a. $\hat{y} = -368 + 130x$
   b. $\hat{y}$ representa el valor predicho del precio.
3. a. Un residuo es un valor de $y - \hat{y}$, que es la diferencia entre un valor observado de y y un valor predicho de y.
   b. La línea de regresión tiene la propiedad de que la suma de los cuadrados de los residuos es la suma más baja posible.
5. Sin una correlación lineal significativa, el mejor valor predicho es $\bar{y} = 5.9$.
7. Con una correlación lineal significativa, el mejor valor predicho es 92.0 kg.
9. $\hat{y} = 3.00 + 0.500x$. Los datos tienen un patrón que no es una línea recta.
11. a. $\hat{y} = 0.264 + 0.906x$
    b. $\hat{y} = 2 + 0x$ (o $\hat{y} = 2$)
    c. Los resultados son muy diferentes, lo que indica que un punto puede afectar drásticamente la ecuación de regresión.
13. $\hat{y} = -8.44 + 0.203x$. El mejor valor predicho: $\bar{y} = 5.1$ por cada 10 millones de personas. El mejor valor predicho no es para nada cercano al valor real.
15. $\hat{y} = -0.0111 + 1.01x$. El mejor valor predicho: \$3.02. No es probable que se implemente la mejor tarifa de metro prevista de \$3.02 porque no es un valor conveniente, como \$3.00 o \$3.25.
17. $\hat{y} = 125 + 1.73x$. El mejor valor predicho: $\bar{y} = 177$ cm. Debido a que el mejor valor predicho es la altura media, no sería útil que la policía tratara de obtener una descripción del hombre.
19. $\hat{y} = 16.5 - 0.00282x$. El mejor valor predicho: 15.1 muertes por 100,000 habitantes. El sentido común sugiere que la predicción no tiene mucho sentido.
21. $\hat{y} = 51.6 - 0.165x$. El mejor valor predicho: $\bar{y} = 45$ años, que no es cercano a la edad real de 33 años.
23. $\hat{y} = -157 + 40.2x$. El mejor valor predicho: -76.6 kg. La predicción es un peso negativo que no puede ser correcto. El ancho general de 2 cm está fuera del alcance de los anchos muestrales, por lo que la extrapolación podría estar equivocada en una cantidad considerable. Claramente, el peso negativo predicho no tiene sentido.
25. $\hat{y} = -0.347 + 0.149x$. El mejor valor predicho: \$14.55. Regla de propina: multiplique la cuenta por 0.149 (o 14.9%) y reste 35 centavos. Una regla más aproximada pero más fácil es esta: deje una propina del 15%.
27. $\hat{y} = -0.00396 + 3.14x$. La mejor circunferencia pronosticada: 4.7 cm. Aunque el diámetro de 1.50 cm está fuera del alcance de los diámetros de la muestra, el valor predicho produce la circunferencia real.
29. $\hat{y} = -23.2 + 0.456x$. El mejor valor predicho: 12.9 por cada 10 millones de personas. El mejor valor pronosticado no se acerca al índice real de premio Nobel en Japón de 1.5 por cada 10 millones de personas.
31. $\hat{y} = 93.5 + 2.85x$. El mejor valor predicho: 183 cm. Aunque hay una correlación lineal, con $r = 0.594$, vemos que no es muy fuerte, por lo que una estimación de la estatura de un hombre podría estar equivocada por una cantidad considerable.
33. $\hat{y} = 13{,}400 + 0.302x$. El mejor valor predicho: 18,200 palabras.
35. a. 823.64
    b. Si se usa $\hat{y} = -3 + 2.5x$, la suma de los cuadrados de los residuos es 827.45, que es mayor que 823.64, la suma de los cuadrados de los residuos para la línea de regresión.

Sección 10-3

1. El valor de $s_e = 16.27555$ cm es el error estándar de estimación, que es una medida de las diferencias entre los pesos observados y los pesos predichos a partir de la ecuación de regresión. Es una medida de la variación de los puntos muestrales sobre la línea de regresión.
3. El coeficiente de determinación es $r^2 = 0.155$. Sabemos que el 15.5% de la variación en el peso se explica por la correlación lineal entre la estatura y el peso, y el 84.5% de la variación en el peso se explica por otros factores y/o por la variación aleatoria.
5. $r^2 = 0.764$. El 76.4% de la variación de la temperatura se explica por la correlación lineal entre los chirridos y la temperatura, y el 23.6% de la variación de la temperatura se explica por otros factores y/o por la variación aleatoria.
7. $r^2 = 0.783$. El 78.3% de la variación en el tamaño de la cintura se explica por la correlación lineal entre el peso y el tamaño de la cintura, y el 21.7% de la variación en el tamaño de la cintura se explica por otros factores y/o por la variación aleatoria.

9. $r = 0.850$. Valores críticos: $r = \pm 0.404$ (Tabla: $r = \pm 0.396$ aproximadamente), suponiendo un nivel de significancia de 0.05. Existe evidencia suficiente para respaldar la afirmación de una correlación lineal entre los barcos registrados y las muertes de manatíes.
11. 70.5 manatíes
13. 42.7 manatíes $< y <$ 98.3 manatíes
15. 65.1 manatíes $< y <$ 106.8 manatíes
17. a. 10,626.59
 b. 68.83577
 c. 38.0°F $< y <$ 60.4°F
19. a. 352.7278
 b. 109.3722
 c. 71.09°F $< y <$ 88.71°F
21. $17.1 < y < 26.0$ (los valores son premios Nobel por cada 10 millones de personas)

Sección 10-4

1. La variable de respuesta es el peso y las variables de predicción son la longitud y el tamaño del pecho.
3. El R^2 no ajustado aumenta (o permanece igual) a medida que se incluyen más variables, pero el R^2 ajustado se ajusta de acuerdo con el número de variables y el tamaño de muestra. El R^2 no ajustado sugiere incorrectamente que la mejor ecuación de regresión múltiple se obtiene al incluir todas las variables disponibles, pero al tomar en cuenta el tamaño de muestra y el número de variables de predicción, el R^2 ajustado es mucho más útil para descartar variables que no deberían incluirse.
5. Hijo = 18.0 + 0.504 Padre + 0.277 Madre
7. El valor P menor que 0.0001 es bajo, pero los valores de R^2 (0.3649) y R^2 ajustado (0.3552) no son altos. Aunque la ecuación de regresión múltiple se ajusta mejor a los datos muestrales, no es una buena opción, por lo que no debe usarse para predecir la estatura de un hijo en función de las estaturas de su padre y de su madre.
9. CARR (consumo de combustible en carretera) porque tiene la mejor combinación de pequeño valor P (0.000) y R^2 ajustado (0.920).
11. CIUDAD = −3.15 + 0.819 CARR. Esa ecuación tiene un valor P bajo de 0.000 y su valor de R^2 ajustado de 0.920 no es mucho menor que los valores de 0.928 y 0.935 que usan dos variables de predicción, por lo que en este caso es mejor usar la variable de predicción en lugar de dos.
13. La mejor ecuación de regresión es $\hat{y} = 0.127 + 0.0878x_1 - 0.0250x_2$, donde x_1 representa alquitrán y x_2 representa monóxido de carbono. Es mejor porque tiene el valor R^2 ajustado más alto de 0.927 y el valor P más bajo de 0.000. Es una buena ecuación de regresión para predecir el contenido de nicotina porque tiene un alto valor de R^2 ajustado y un valor P bajo.
15. La mejor ecuación de regresión es $\hat{y} = 109 - 0.00670x_1$, donde x_1 representa el volumen. Es mejor porque tiene el valor R^2 ajustado más alto de -0.0513 y el valor P más bajo de 0.791. Las tres ecuaciones de regresión tienen valores ajustados de R^2 que están muy cerca de 0, por lo que ninguno de ellos es bueno para predecir el coeficiente intelectual. No parece que las personas con cerebros más grandes tengan puntuaciones de IQ más altas.
17. Para H_0: $\beta_1 = 0$, el estadístico de prueba es $t = 10.814$, el valor P es menor que 0.0001; por lo tanto, se rechaza H_0 y se concluye que se debe mantener el coeficiente de regresión de $\beta_1 = 0.769$. Para H_0: $\beta_2 = 0$, el estadístico de prueba es $t = 29.856$, el valor P es menor que 0.0001; por lo tanto, se rechaza H_0 y se concluye que debe mantenerse el coeficiente de regresión de $\beta_2 = 1.01$. Parece que la ecuación de regresión debe incluir ambas variables independientes de altura y circunferencia de la cintura.
19. $\hat{y} = 3.06 + 82.4x_1 + 2.91x_2$, donde x_1 representa el sexo y x_2 representa la edad. Mujer: 61 lb; Hombre: 144 lb. El sexo del oso parece tener un efecto sobre su peso. La ecuación de regresión indica que el peso predicho de un oso macho es aproximadamente 82 lb más que el peso predicho de un oso hembra con otras características iguales.

Sección 10-5

1. $y = x^2$; cuadrático; $R^2 = 1$
3. El 25.5% de la variación en los puntos del Súper Bowl se puede explicar por el modelo cuadrático que relaciona la variable del año y la variable de los puntos anotados. Debido a que el modelo explica un porcentaje tan pequeño de la variación, el modelo no es muy útil.
5. Cuadrático: $d = -4.88t^2 + 0.0214t + 300$
7. Exponencial: $y = 1000(1.0122^x)$
9. Exponencial: $y = 10(2^x)$
11. Logarítmico: $y = 3.22 + 0.293 \ln x$
13. Potencia: $y = 2107.9x^{0.615}$. (El resultado se basa en 1990 codificado como 1). El valor proyectado para 2014 es 15,271 (utilizando coeficientes redondeados: 15,261), que es considerablemente menor que el valor real de 18,054.
15. Cuadrático: $y = 0.700x^2 - 3.41x + 299$. Valor predicho: 563. La década de 2090-2099 está demasiado lejos del alcance de los datos disponibles, por lo que el valor predicho es cuestionable.
17. 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: $y = (1.36558)(1.42774)^x$, donde 1971 está codificado como 1.
 c. La ley de Moore parece funcionar razonablemente bien. Con $R^2 = 0.990$, el modelo parece ser muy bueno.

Capítulo 10: Examen rápido

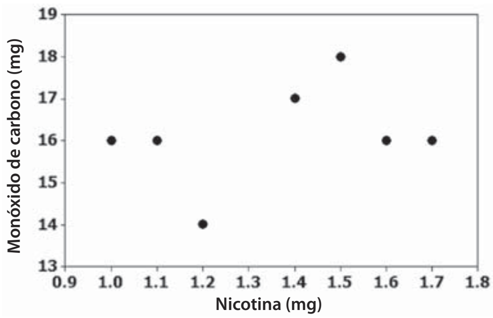
1. Se concluye que no hay evidencia suficiente para respaldar la afirmación de una correlación lineal entre la inscripción y los robos.
2. Ninguno de los valores dados cambia cuando se modifican las variables.
3. No. El valor de r no cambia si todos los valores de una de las variables se multiplican por la misma constante.
4. Como r debe estar entre -1 y 1 inclusive, el valor de 1.500 es el resultado de un error en los cálculos.
5. El mejor número predicho de robos es 92.6, que es el promedio de los cinco conteos muestrales de robos.
6. El mejor número predicho de robos sería 123.3, que se encuentra al sustituir 50 por x en la ecuación de regresión.
7. $r^2 = 0.249$
8. Falso.
9. Falso.
10. $r = -1$

Capítulo 10: Ejercicios de repaso

1. a. $r = 0.962$. Valor $P = 0.000$ (Tabla: <0.01). Valores críticos: $r = \pm 0.707$ (suponiendo un nivel de significancia de 0.05). Existe evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre la cantidad de alquitrán y la cantidad de nicotina.

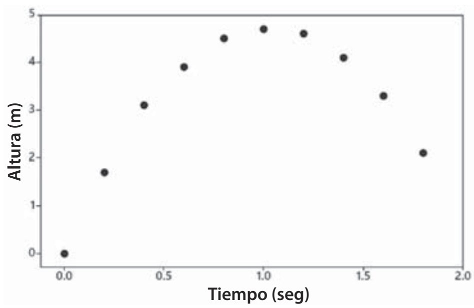
b. 92.5%
c. $\hat{y} = -0.758 + 0.0920x$
d. El valor predicho es de 1.358 mg o 1.4 mg redondeado, que es cercano a la cantidad real de 1.3 mg.

2. a. La gráfica de dispersión muestra un patrón en el que la nicotina y el CO aumentan de izquierda a derecha, pero es un patrón muy débil y los puntos no están muy cerca de un patrón en línea recta, por lo que parece que no hay suficiente evidencia muestral para respaldar la afirmación de una correlación lineal entre las cantidades de nicotina y de monóxido de carbono.



b. $r = 0.329$. Valor $P = 0.427$ (Tabla: >0.05). Valores críticos: $r = \pm 0.707$ (suponiendo un nivel de significancia de 0.05). No hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre la cantidad de nicotina y la cantidad de monóxido de carbono.
c. $\hat{y} = 14.2 + 1.42x$
d. El valor predicho es $\bar{y} = 16.1$ mg, que está cerca de la cantidad real de 15 mg.

3. $r = 0.450$. Valor $P = 0.192$ (Tabla: >0.05). Valores críticos: $r = \pm 0.632$ (suponiendo un nivel de significancia de 0.05). No hay evidencia suficiente para respaldar la afirmación de que existe una correlación lineal entre el tiempo y la altura. Aunque no existe una correlación *lineal* entre el tiempo y la altura, el diagrama de dispersión muestra un patrón muy distinto que revela que el tiempo y la altura están asociados por alguna función que no es lineal.



4. a. NICOTINA = −0.443 + 0.0968 ALQUITRÁN − 0.0262 CO, o $\hat{y} = -0.443 + 0.0968x_1 - 0.0262x_2$.
b. $R^2 = 0.936$; R^2 ajustado = 0.910; Valor $P = 0.001$.
c. Con valores altos de R^2 y R^2 ajustado y un valor P pequeño de 0.001, parece que la ecuación de regresión puede usarse para predecir la cantidad de nicotina, dadas las cantidades de alquitrán y monóxido de carbono.
d. El valor predicho es de 1.39 mg o 1.4 mg redondeado, que está cerca del valor real de 1.3 mg de nicotina.

Capítulo 10: Ejercicios de repaso acumulado

1. a. $\bar{x} = 35.91$, mediana = 36.10, rango = 76.40, $s = 31.45$, $s^2 = 989.10$.
b. Datos cuantitativos
c. De razón

2. $r = 0.731$. Valor $P = 0.039$ (Tabla: <0.05) Valores críticos: $r = \pm 0.707$. Existe evidencia suficiente para respaldar la afirmación de una correlación lineal entre los valores del DJIA y los números de manchas solares. Debido a que sería razonable pensar que no existe una correlación entre las acciones y el número de manchas solares, el resultado no es el esperado. Aunque parece haber una correlación lineal, un inversionista razonable sería inteligente al ignorar los números de manchas solares cuando invierte en acciones.

3. El número más alto de manchas solares es 79.3, que se convierte en $z = 1.38$. El número más alto de manchas solares no es significativamente alto porque su puntuación z de 1.38 muestra que está dentro de 2 desviaciones estándar de la media.

4. H_0: $\mu = 49.7$. H_1: $\mu \neq 49,7$. Estadístico de prueba: $t = -1.240$. Valor $P = 0.255$ (Tabla: >0.20). Valores críticos: $t = \pm 2.365$. No se puede rechazar H_0. No hay evidencia suficiente para justificar el rechazo de la afirmación de que la muestra proviene de una población con una media igual a 49.7.

5. $9.62 < \mu < 62.21$. Tenemos un 95% de confianza de que los límites de intervalo de 9.62 y 62.21 contienen el valor real de la media de la población de números de manchas solares.

6. H_0: $p = 0.10$. H_1: $p < 0.10$. Estadístico de prueba: $z = -5.25$ (usando $x = 191$) o $z = -5.23$ (usando $\hat{p} = 0.07$). Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = -1.645$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que menos del 10% de las detenciones de tráfico policiales son atribuibles al uso indebido del teléfono celular.

7. $\bar{x} = 35.2$ años, $s = 19.7$ años, $s^2 = 389.6$ años^2

8. a. 40.13%
b. 21.5 años (Tabla: 21.6 años)
c. 0.1056
d. 0 + o 0.0000 (a partir de $0.4013^{25} = 0.0000$). La audiencia de una película y presentación particular no es una muestra aleatoria simple. Algunas películas y presentaciones atraen a un público muy joven.

Capítulo 11 Respuestas

Sección 11-1

1. a. Los valores observados están representados por O y los valores esperados por E.
b. Para el dígito inicial de 2, $O = 62$ y $E = (317)(0.176) = 55.792$.
c. Para el dígito inicial de 2, $(O - E)^2/E = 0.691$.

3. Hay evidencia suficiente para justificar el rechazo de la afirmación de que los dígitos iniciales tienen una distribución que se ajusta bien a la ley de Benford.

5. Valor $P = 0.516$ (Tabla: >0.10). Valor crítico: $\chi^2 = 16.919$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los resultados observados concuerdan con las frecuencias esperadas. La máquina tragamonedas parece estar funcionando como se esperaba.

7. Estadístico de prueba: $\chi^2 = 5.860$. Valor $P = 0.320$ (Tabla: >0.10). Valor crítico: $\chi^2 = 11.071$. No hay evidencia suficiente para respaldar la afirmación de que los resultados no son igualmente probables. Los resultados parecen ser igualmente probables, por lo que el dado cargado no parece comportarse de manera diferente a un dado legal.

9. Estadístico de prueba: $\chi^2 = 11.161$. Valor $P = 0.011$ (Tabla: <0.025). Valor crítico: $\chi^2 = 7.815$. Hay suficiente evidencia para respaldar la afirmación de que los resultados contradicen la teoría de Mendel.
11. Estadístico de prueba: $\chi^2 = 29.814$. Valor $P = 0.000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 16.812$. Hay evidencia suficiente para justificar el rechazo de la afirmación de que los diferentes días de la semana tienen las mismas frecuencias de llamadas policiales. El número más grande de llamadas parece ser en viernes y en sábado, y estos son días de fin de semana con más fiesta y bebida en exceso.
13. Estadístico de prueba: $\chi^2 = 13.855$. Valor $P = 0.128$ (Tabla: >0.10). Valor crítico: $\chi^2 = 16.919$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que la probabilidad de ganar es la misma para las diferentes posiciones de inicio. Con base en estos resultados, la posición de inicio no se debe considerar al apostar en la carrera del Kentucky Derby.
15. Estadístico de prueba: $\chi^2 = 8.882$. Valor $P = 0.031$ (Tabla: <0.05). Valor crítico: $\chi^2 = 7.815$. Existe evidencia suficiente para justificar el rechazo de la afirmación de que el número real de juegos se ajusta a la distribución indicada por las proporciones listadas en la tabla dada.
17. Estadístico de prueba: $\chi^2 = 9.500$. Valor $P = 0.147$ (Tabla: >0.10). Valor crítico: $\chi^2 = 16.812$. No hay evidencia suficiente para respaldar la afirmación de que los nacimientos no ocurren en los siete diferentes días de la semana con la misma frecuencia.
19. Estadístico de prueba: $\chi^2 = 6.682$. Valor $P = 0.245$ (Tabla: >0.10). Valor crítico: $\chi^2 = 11.071$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que la distribución del color es como se declara.
21. Estadístico de prueba: $\chi^2 = 3650.251$. Valor $P = 0.000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 20.090$. Hay suficiente evidencia para justificar el rechazo de la afirmación de que los dígitos iniciales provienen de una población con una distribución que se ajusta a la ley de Benford. Parece que los cheques son el resultado de un fraude (aunque los resultados no pueden confirmar que el fraude sea la causa de la discrepancia entre los resultados observados y los resultados esperados).
23. Estadístico de prueba: $\chi^2 = 1.762$. Valor $P = 0.988$ (Tabla: >0.10). Valor crítico: $\chi^2 = 15.507$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los dígitos iniciales provienen de una población con una distribución que se ajusta a la ley de Benford. Las entradas de impuestos parecen ser legítimas.
25. a. 26, 46, 49, 26
    b. 0.2023, 0.3171, 0.3046, 0.1761 (Tabla: 0.2033, 0.3166, 0.3039, 0.1762)
    c. 29.7381, 46.6137, 44.7762, 25.8867 (Tabla: 29.8851, 46.5402, 44.6733, 25.9014)
    d. Estadístico de prueba: $\chi^2 = 0.877$ (Utilizando las probabilidades de la tabla: 0.931). Valor $P = 0.831$ (Tabla: >0.10). Valor crítico: $\chi^2 = 11.345$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las alturas se seleccionaron al azar de una población distribuida normalmente. La prueba sugiere que no podemos descartar la posibilidad de que los datos provengan de una población distribuida normalmente.

Sección 11-2

1. a. $E = 4.173$.
   b. Debido a que la frecuencia esperada de una celda es menor que 5, los requisitos para la prueba de hipótesis no se cumplen.
3. Estadístico de prueba: $\chi^2 = 64.517$. Valor P: 0.000. Se rechaza la hipótesis nula de independencia entre la destreza y la preferencia del oído del teléfono celular.
5. Estadístico de prueba: $\chi^2 = 25.571$. Valor $P = 0.000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 3.841$. Existe suficiente evidencia para justificar el rechazo de la afirmación de que si un sujeto miente es independiente de la indicación de la prueba del polígrafo. Los resultados sugieren que los polígrafos son efectivos para distinguir entre verdades y mentiras, pero hay muchos falsos positivos y falsos negativos, por lo que no son muy confiables.
7. Estadístico de prueba: $\chi^2 = 576.224$. Valor $P = 0.000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 3.841$. Existe suficiente evidencia para justificar el rechazo de la afirmación de independencia entre enviar mensajes de texto mientras se conduce o manejar cuando se bebe alcohol. Esos dos comportamientos riesgosos parecen estar relacionados de alguna manera.
9. Estadístico de prueba: $\chi^2 = 12.162$. Valor $P = 0.001$ (Tabla: <0.005). Valor crítico: $\chi^2 = 3.841$. Hay suficiente evidencia para justificar el rechazo de la afirmación de que si los estudiantes compraron goma de mascar o guardaron el dinero es independiente de si recibieron cuatro monedas de ¢25 o un billete de \$1. Parece que hay un efecto de denominación.
11. Estadístico de prueba: $\chi^2 = 0.064$. Valor $P = 0.801$ (Tabla: >0.10). Valor crítico: $\chi^2 = 3.841$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que el sexo del tenista es independiente de si la decisión se anula. Ni los hombres ni las mujeres parecen ser mejores en lo desafíos a las decisiones arbitrales.
13. Estadístico de prueba: $\chi^2 = 14.589$. Valor $P = 0.0056$ (Tabla: <0.01). Valor crítico: $\chi^2 = 9.488$. Hay evidencia suficiente para justificar el rechazo de la afirmación de que la dirección de la patada es independiente de la dirección del movimiento del portero. Los resultados no respaldan la teoría de que debido a que las patadas son tan rápidas, los porteros no tienen tiempo para reaccionar. Parece que los porteros pueden elegir su dirección en función de la dirección de la patada.
15. Estadístico de prueba: $\chi^2 = 2.925$. Valor $P = 0.232$ (Tabla: >0.10). Valor crítico: $\chi^2 = 5.991$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que pescar un resfriado es independiente del grupo de tratamiento. Los resultados sugieren que la equinácea no es efectiva para prevenir los resfriados.
17. Estadístico de prueba: $\chi^2 = 20.271$. Valor $P = 0.0011$ (Tabla: <0.005). Valor crítico: $\chi^2 = 15.086$. Existe evidencia suficiente para justificar el rechazo de la afirmación de que la cooperación del sujeto es independiente de la categoría de edad. El grupo de edad de 60 y más parece ser particularmente poco cooperativo.
19. Estadístico de prueba: $\chi^2 = 50.446$. Valor $P = 0.000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 5.991$. Hay suficiente evidencia para justificar el rechazo de la afirmación de independencia entre el estado y si un automóvil tiene placas delanteras y traseras. No parece que las leyes de las placas se sigan con las mismas proporciones en los tres estados.
21. Estadístico de prueba: $\chi^2 = 12.1619258$ y $z = 3.487395274$, de modo que $z^2 = \chi^2$. Valores críticos: $\chi^2 = 3.841$ y $z = \pm 1.96$, entonces $z^2 = \chi^2$ (aproximadamente).

Capítulo 11: Examen rápido

1. H_0: $p_0 = p_1 = \cdots = p_9$. H_1: Al menos una de las probabilidades es diferente de las demás.
2. $O = 27$ y $E = 30$.
3. De cola derecha.
4. gl = 9

5. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los últimos dígitos son igualmente probables. Debido a que las estaturas reportadas probablemente incluirían más dígitos de 0 y 5, parece que las alturas se midieron en lugar de reportarse. (Además, la mayoría de los residentes de Estados Unidos tendrían dificultades para informar sus estaturas en centímetros, porque Estados Unidos, Liberia y Myanmar son los únicos países que continúan utilizando el sistema de medición imperial).
6. H_0: Sobrevivir al hundimiento es independiente de si la persona es un hombre, una mujer, un niño o una niña.
   H_1: Sobrevivir al hundimiento y si la persona es un hombre, una mujer, un niño o una niña están relacionados de alguna manera.
7. Distribución ji cuadrada.
8. De cola derecha.
9. gl $= 3$
10. Hay evidencia suficiente para justificar el rechazo de la afirmación de que sobrevivir al hundimiento es independiente de si la persona es un hombre, una mujer, un niño o una niña. La mayoría de las mujeres sobrevivieron, el 45% de los niños sobrevivieron, y la mayoría de las niñas sobrevivieron, pero sólo alrededor del 20% de los hombres sobrevivieron, por lo que parece que la regla fue seguida bastante bien.

Capítulo 11: Ejercicios de repaso

1. Estadístico de prueba: $\chi^2 = 787.018$. Valor P: 0.000. Valor crítico: $\chi^2 = 16.812$. Existe evidencia suficiente para justificar el rechazo de la afirmación de que las muertes en auto ocurren en los diferentes días de la semana con la misma frecuencia. Debido a que las personas generalmente tienen más tiempo libre los fines de semana y se bebe más durante esos días, los días viernes, sábado y domingo parecen tener desproporcionadamente más muertes.
2. Estadístico de prueba: $\chi^2 = 0.751$. Valor $P = 0.386$ (Tabla: >0.10). Valor crítico: $\chi^2 = 3.841$. No hay evidencia suficiente para justificar el rechazo de la afirmación de independencia entre el tipo de llenado y las condiciones de salud adversas. Los rellenos que contienen mercurio no parecen afectar las condiciones de salud.
3. Estadístico de prueba: $\chi^2 = 5.624$. Valor $P = 0.467$ (Tabla: >0.10). Valor crítico: $\chi^2 = 12.592$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las eliminaciones reales concuerdan con los números esperados. Los primeros cantantes parecen estar en desventaja porque 20 de ellos fueron eliminados en comparación con el valor esperado de 12.9 eliminaciones, pero ese resultado no parece ser *significativamente* alto.
4. Estadístico de prueba: $\chi^2 = 0.773$. Valor $P = 0.856$ (Tabla: >0.10). Valor crítico: $\chi^2 = 11.345$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que obtener una infección es independiente del tratamiento. El tratamiento con atorvastatina (Lipitor) no parece tener un efecto sobre las infecciones.
5. Estadístico de prueba: $\chi^2 = 269.147$. Valor $P = 0.000$ (Tabla: <0.005). Valor crítico: $\chi^2 = 24.725$. Hay evidencia suficiente para justificar el rechazo de la afirmación de que las muertes relacionadas con el clima ocurren en los diferentes meses con la misma frecuencia. Los meses de mayo, junio y julio parecen tener desproporcionadamente más muertes relacionadas con el clima, y eso probablemente se deba al hecho de que las vacaciones y las actividades al aire libre son mucho mayores durante esos meses.

Capítulo 11: Ejercicios de repaso acumulado

1. H_0: $p = 0.5$. H_1: $p \neq 0.5$. Estadístico de prueba: $z = 8.96$. Valor P: 0.0000 (Tabla: 0.0002). Valores críticos: $z = \pm 1.96$. Se rechaza H_0. Hay evidencia suficiente para justificar el rechazo de la afirmación de que, entre las personas que mueren en eventos relacionados con el clima, el porcentaje de hombres es igual a 50%. Una posible explicación es que más hombres participan en algunas actividades al aire libre, como golf, pesca y paseos en bote.
2. a. Existe la posibilidad de que los resultados se vean afectados porque el patrocinador de la encuesta produce chocolate y, por lo tanto, tiene interés en los resultados.
   b. 1452
3. $82.8\% < p < 87.2\%$. Tenemos un 99% de confianza de que los límites del 82.8 y 87.2% contienen el valor del porcentaje verdadero de la población de mujeres que dicen que el chocolate las hace más felices.
4. H_0: $p = 0.80$. H_1: $p > 0.80$. Estadístico de prueba: $z = 5.18$. Valor $P = 0.0000$ (Tabla: 0.0001). Valor crítico: $z = 2.33$. Se rechaza H_0. Hay evidencia suficiente para respaldar la afirmación de que cuando se les pregunta, más del 80% de las mujeres dicen que el chocolate las hace más felices.
5. Estadístico de prueba: $\chi^2 = 3.409$. Valor $P = 0.0648$ (Tabla: >0.05). Valor crítico: $\chi^2 = 3.841$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que la forma del regalo de 100 yuanes es independiente de si el dinero se gastó. No hay suficiente evidencia para apoyar la afirmación de un efecto de denominación. Las mujeres en China no parecen verse afectadas por si 100 yuanes tienen la forma de un solo billete o varios billetes más pequeños.
6. a. $128/150 = 0.853$ b. $143/150 = 0.953$
   c. 0.727 (no 0.728)
7. $r = -0.283$. Valor $P = 0.539$. Valores críticos: ± 0.754 (suponiendo un nivel de significancia de 0.05). No hay evidencia suficiente para respaldar la afirmación de una correlación lineal entre los costos de reparación de choques frontales completos y colisiones traseras completas.
8. a. 630 mm
   b. 14.48% (Tabla: 14.46%). Ese porcentaje es demasiado alto, porque muchas mujeres no serían atendidas.
   c. 0.7599 (Tabla: 0.7611). Grupos de 16 mujeres no ocupan el asiento del conductor ni la cabina; debido a que mujeres *individuales* ocupan el asiento del conductor/cabina, este resultado no tiene ningún efecto en el diseño.

Capítulo 12 Respuestas

Sección 12-1

1. a. Los tiempos de demora en la llegada se clasifican según la característica del número de vuelo.
   b. La terminología de *análisis de varianza* se refiere al método utilizado para evaluar la igualdad de las tres medias poblacionales. Ese método se basa en dos estimaciones diferentes de una varianza poblacional común.
3. El estadístico de prueba es $F = 1.334$, y se aplica la distribución F.
5. Estadístico de prueba: $F = 0.39$. Valor P: 0.677. No se puede rechazar H_0: $\mu_1 = \mu_2 = \mu_3$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las tres categorías de niveles de plomo en la sangre tienen la misma puntuación verbal promedio. La exposición al plomo no parece tener un efecto en las puntuaciones del IQ verbal.

7. Estadístico de prueba: $F = 5.5963$. Valor P: 0.0045. Se rechaza H_0: $\mu_1 = \mu_2 = \mu_3$. Hay evidencia suficiente para justificar el rechazo de la afirmación de que las tres muestras provienen de poblaciones con la misma media. Parece que al menos uno de los tiempos de servicio promedio es diferente de los demás.
9. Estadístico de prueba: $F = 7.9338$. Valor P: 0.0005. Se rechaza H_0: $\mu_1 = \mu_2 = \mu_3$. Existe suficiente evidencia para justificar el rechazo de la afirmación de que las mujeres de los tres grupos de edad tienen la misma frecuencia media de pulso. Parece que los pulsos de las mujeres se ven afectados por el grupo de edad.
11. Estadístico de prueba: $F = 27.2488$. Valor P: 0.000. Se rechaza H_0: $\mu_1 = \mu_2 = \mu_3$. Existe suficiente evidencia para justificar el rechazo de la afirmación de que las tres diferentes millas tienen el mismo tiempo medio. Estos datos sugieren que la tercera milla parece tomar más tiempo, y una explicación razonable es que ésta tiene una pendiente.
13. Estadístico de prueba: $F = 2.3163$. Valor P: 0.123. No se puede rechazar H_0: $\mu_1 = \mu_2 = \mu_3$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los tres vuelos diferentes tienen el mismo tiempo medio de demora en la salida. Los tiempos de demora en la salida del vuelo 1 tienen muy poca variación, y las salidas del vuelo 1 parecen ser puntuales o ligeramente tempranas. Los tiempos de demora en la salida del vuelo 21 parecen tener una variación considerable. Con varianzas de 2.5 min^2, 709.8 min^2 y 2525.4 min^2, el requisito ANOVA de la misma varianza parece ser violado incluso para este requisito relajado.
15. Estadístico de prueba: $F = 28.1666$. Valor P: 0.000. Se rechaza H_0: $\mu_1 = \mu_2 = \mu_3$. Existe suficiente evidencia para justificar el rechazo de la afirmación de que los tres tipos diferentes de galletas Chips Ahoy tienen la misma cantidad media de chispas de chocolate. Las galletas bajas en grasa tienen una media de 19.6 chispas de chocolate, que es un poco más que la media de 19.1 chispas de chocolate para las galletas masticables, por lo que la reducción de grasa no parece ser resultado de incluir menos chispas de chocolate. Tal vez el contenido de grasa en las chispas de chocolate es diferente y/o el contenido de grasa en el material de la galleta es diferente.
17. Los resultados de la prueba de Tukey muestran diferentes valores P, pero no son dramáticamente diferentes. Los resultados de Tukey sugieren las mismas conclusiones que la prueba de Bonferroni.

Sección 12-2

1. Las frecuencias del pulso se categorizan usando dos factores diferentes de (1) rango de edad y (2) género.
3. a. Se produce una interacción entre dos factores o variables si el efecto de uno de los factores cambia para diferentes categorías del otro factor.
   b. Si hay un efecto de interacción, no deberíamos proceder con pruebas individuales de los efectos del factor fila y el factor columna. Si hay una interacción, no deberíamos considerar los efectos de un factor sin considerar los efectos del otro.
   c. Debido a que las líneas están lejos de ser paralelas, los dos sexos tienen efectos muy diferentes para los diversos grupos de edad, por lo que parece haber una interacción entre el género y el grupo de edad.
5. Para la interacción, el estadístico de prueba es $F = 9.58$ y el valor de P es 0.0003, por lo que hay suficiente evidencia para justificar el rechazo de la hipótesis nula de que no hay efecto de interacción. Debido a que parece haber una interacción entre el grupo de edad y el sexo, no deberíamos proceder con una prueba para un efecto del grupo de edad y una prueba para un efecto de género. Parece que la interacción entre el grupo de edad y el género tiene un efecto en los pulsos. (Recuerde, estos resultados se basan en datos fabricados utilizados en una de las celdas, por lo que esta conclusión no necesariamente corresponde con los datos reales).
7. Para la interacción, el estadístico de prueba es $F = 1.7970$ y el valor P es 0.1756, por lo que no hay evidencia suficiente para concluir que haya un efecto de interacción. Para la variable de fila del rango de edad, el estadístico de prueba es $F = 2.0403$ y el valor P es 0.1399, por lo que no hay evidencia suficiente para concluir que el grupo de edad tiene un efecto sobre la altura. Para la variable de columna de género, el estadístico de prueba es $F = 43.4607$ y el valor P es menor que 0.0001, por lo que hay suficiente evidencia para apoyar la afirmación de que el género tiene un efecto sobre la altura.
9. Para la interacción, el estadístico de prueba es $F = 1.1653$ y el valor de P es 0.3289, por lo que no hay un efecto de interacción significativo. Para el género, el estadístico de prueba es $F = 1.6864$ y el valor P es 0.2064, por lo que no hay un efecto significativo del género. Para la edad, el estadístico de prueba es $F = 5.0998$ y el valor P es 0.0143, por lo que hay un efecto significativo de la edad.
11. a. Los estadísticos de prueba y los valores P no cambian.
    b. Los estadísticos de prueba y los valores P no cambian.
    c. Los estadísticos de prueba y los valores P no cambian.
    d. Un valor atípico puede afectar y cambiar dramáticamente los estadísticos de prueba y los valores P.

Capítulo 12: Examen rápido del capítulo

1. H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$. Debido a que el valor P mostrado de 0.000 es pequeño, se rechaza H_0. Existe suficiente evidencia para justificar el rechazo de la afirmación de que las cuatro muestras tienen el mismo peso medio.
2. No. Parece que los pesos medios de Diet Coke y Diet Pepsi son más bajos que los pesos medios de Coca-Cola regular y Pepsi regular, pero el método de análisis de varianza no justifica la conclusión de que cualquier media en particular es significativamente diferente de las otras.
3. De cola derecha.
4. Estadístico de prueba: $F = 503.06$. Los estadísticos de prueba más grandes dan como resultado valores P más pequeños.
5. Las cuatro muestras se clasifican utilizando un solo factor: el tipo de bebida de cola (Coca-Cola regular, Diet Coke, Pepsi regular, Diet Pepsi).
6. El análisis de varianza de un factor se usa para probar una hipótesis nula de que tres o más muestras provienen de poblaciones con las mismas medias.
7. Con el análisis de varianza de un factor, los datos de las diferentes muestras se clasifican utilizando un solo factor, pero con el análisis de varianza de dos factores, los datos muestrales se clasifican en diferentes celdas determinadas por dos factores diferentes.
8. No se puede rechazar la hipótesis nula de no interacción. No parece haber un efecto debido a una interacción entre el sexo y la especialidad.
9. No hay evidencia suficiente para respaldar una afirmación de que las estimaciones de longitud se ven afectadas por el sexo del sujeto.
10. No hay evidencia suficiente para respaldar una afirmación de que las estimaciones de longitud se ven afectadas por la especialidad del sujeto.

Capítulo 12: Ejercicios de repaso

1. Estadístico de prueba: $F = 2.7347$. Valor P: 0.0829. No se puede rechazar H_0: $\mu_1 = \mu_2 = \mu_3$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los hombres en los diferentes grupos de edad otorgan calificaciones de atributos con la misma media. La edad no parece ser un factor en las calificaciones de los atributos masculinos.

2. Estadístico de prueba: $F = 9.4695$. Valor P: 0.0006. Se rechaza H_0: $\mu_1 = \mu_2 = \mu_3$. Existe suficiente evidencia para justificar el rechazo de la afirmación de que los tres libros tienen la misma puntuación media de Flesch Reading Ease. Los datos sugieren que los libros parecen tener puntuaciones medias que no son todas iguales, por lo que los autores no parecen tener el mismo nivel de legibilidad.
3. Para la interacción, el estadístico de prueba es $F = 1.7171$ y el valor P es 0.1940, por lo que no hay evidencia suficiente para justificar el rechazo de ningún efecto de interacción. No parece haber una interacción entre el fémur y el tamaño del automóvil. Para la variable de fila del fémur, el estadístico de prueba es $F = 1.3896$ y el valor P es 0.2462, por lo que no hay evidencia suficiente para concluir que si el fémur es derecho o izquierdo tiene un efecto sobre la carga. Para la variable de columna del tamaño del automóvil, el estadístico de prueba es $F = 2.2296$ y el valor P es 0.1222, por lo que no hay evidencia suficiente para justificar el rechazo de la afirmación de que el tamaño del automóvil no tiene efecto. Parece que las cargas de la prueba de colisión no se ven afectadas por una interacción entre el fémur y el tamaño del automóvil, no se ven afectadas por el fémur y no se ven afectadas por el tamaño del automóvil.
4. Para la interacción, el estadístico de prueba es $F = 0.4784$ y el valor P es 0.7513, por lo que no hay evidencia suficiente para concluir que haya un efecto de interacción. Para la variable de fila del rango de edad de las mujeres, el estadístico de prueba es $F = 0.3149$ y el valor P es 0.7318, por lo que no hay evidencia suficiente para concluir que el rango de edad de las mujeres tiene un efecto sobre las calificaciones. Para la variable de columna del grupo de edad de los hombres, el estadístico de prueba es $F = 1.1939$ y el valor P es 0.3148, por lo que no hay evidencia suficiente para concluir que el grupo de edad de los hombres tiene un efecto sobre las calificaciones.

Capítulo 12: Ejercicios de repaso acumulado

1. a. 2.0 min, 9.9 min, 33.4 min
   b. 10.6 min, 26.6 min, 50.3 min
   c. 112.0 min^2, 709.8 min^2, 2525.4 min^2
   d. El tiempo de demora en la salida de 142 min es un valor atípico.
   e. De razón
2. Estadístico de prueba: $t = -1.728$. Valor $P = 0.1241$ (Tabla: >0.10). Valores críticos que suponen un nivel de significancia de 0.05: $t = \pm 2.326$ (Tabla: ± 2.365). No se puede rechazar H_0: $\mu_1 = \mu_2$. No hay evidencia suficiente para respaldar la afirmación de que existe una diferencia entre los tiempos de demora en la salida para los dos vuelos.
3. Debido a que el patrón de puntos está lejos de ser un patrón en línea recta, los tiempos de demora en la salida para el Vuelo 19 no parecen ser de una población con una distribución normal.
4. $-6.8 \text{ min} < \mu < 10.8 \text{ min}$. Tenemos un 95% de confianza de que los límites de -6.8 min y 10.8 min contienen el valor de la media poblacional para todos los retrasos en la salida del Vuelo 3.
5. a. H_0: $\mu_1 = \mu_2 = \mu_3$
   b. Debido a que el valor P de 0.1729 es mayor que el nivel de significancia de 0.05, no se rechaza la hipótesis nula de igualdad de las medias. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las tres medias son iguales. Las tres poblaciones no parecen tener medias que sean significativamente diferentes.
6. a. 0.5563 (Tabla: 0.5552)
   b. 0.3434 (Tabla: 0.3446)
   c. 1/256 o 0.00391
   d. 5.591 g
7. a. 200
   b. $0.175 < p < 0.225$
   c. Sí. El intervalo de confianza indica que tenemos una confianza del 95% de que la proporción real de la población está dentro de los límites de 0.175 y 0.225, y 1/4 no está incluido dentro de ese rango.
8. a. Debido a que la escala vertical comienza en 15 en lugar de 0, la gráfica es engañosa al exagerar las diferencias entre las frecuencias.
   b. No. Una distribución normal tiene aproximadamente una forma de campana, pero el histograma dado está lejos de poseer esa forma. Debido a que se supone que los dígitos son igualmente probables, el histograma debe ser plano y todas las barras tienen aproximadamente la misma altura.
   c. Las frecuencias son 19, 21, 22, 21, 18, 23, 16, 16, 22, 22.
   d. Estadístico de prueba: $\chi^2 = 3.000$. Valor $P = 0.964$ (Tabla: >0.95). Valor crítico: $\chi_2 = 16.919$ (suponiendo un nivel de significancia de 0.05). No hay evidencia suficiente para justificar el rechazo de la afirmación de que los dígitos se seleccionan de una población en la que todos los dígitos son igualmente probables. No parece haber un problema con la lotería.

Respuestas del capítulo 13

Sección 13-2

1. El único requisito para los pares relacionados es que constituyan una muestra aleatoria simple. No hay ningún requisito de una distribución normal o cualquier otra distribución específica. La prueba del signo es "sin distribución" en el sentido de que no requiere una distribución normal o cualquier otra distribución específica.
3. H_0: No hay diferencia entre las poblaciones de los pesos de septiembre y los pesos de abril correspondientes. H_1: Hay una diferencia entre las poblaciones de los pesos de septiembre y los pesos de abril correspondientes. Los datos muestrales no contradicen H_1 porque el número de signos positivos (2) y signos negativos (7) no son exactamente iguales.
5. El estadístico de prueba de $x = 3$ no es menor o igual que el valor crítico de 1 (de la Tabla A-7). No hay evidencia suficiente para justificar el rechazo de la afirmación de que no hay diferencia. No existe evidencia suficiente para respaldar la afirmación de que existe una diferencia entre las calificaciones de atributos femeninos y las calificaciones de atributos masculinos.
7. El estadístico de prueba de $z = -1.61$ da como resultado un valor P de 0.1074, y no cae en la región crítica limitada por $z = -1.96$ y 1.96. No hay evidencia suficiente para justificar el rechazo de la afirmación de ninguna diferencia. No hay evidencia suficiente para respaldar la afirmación de que existe una diferencia entre las calificaciones de atributos femeninos y las calificaciones de atributos masculinos.
9. El estadístico de prueba de $z = -2.66$ da como resultado un valor P de 0.0078, y está en la región crítica limitada por $z = -2.575$ y 2.575. Existe evidencia suficiente para justificar el rechazo de la afirmación de que no hay diferencia entre las proporciones de quienes se oponen y quienes están a favor.
11. El estadístico de prueba de $z = -0.24$ da como resultado un valor P de 0.8103, y no se encuentra en la región crítica limitada por $z = -1.96$ y 1.96. No hay evidencia suficiente para rechazar la afirmación de que los niños y las niñas son igualmente probables.

13. El estadístico de prueba de $z = -14.93$ da como resultado un valor P de 0.0000, y se encuentra en la región crítica limitada por $z = -2.575$ y 2.575. Hay suficiente evidencia para justificar el rechazo de la afirmación de que la mediana es igual a 2.00.
15. El estadístico de prueba de $z = -2.37$ da como resultado un valor P de 0.0178 y no se encuentra en la región crítica limitada por $z = -2.575$ y 2.575. No hay evidencia suficiente para justificar el rechazo de la afirmación de que la mediana es igual a 5.670 g. Las monedas de ¢25 parecen estar acuñados de acuerdo con las especificaciones.
17. Segundo método: El estadístico de prueba de $z = -4.29$ da como resultado un valor P de 0.0000 y está en la región crítica limitada por $z = -1.645$, por lo que las conclusiones son las mismas que en el ejemplo 4. Tercer método: El estadístico de prueba de $z = -2.82$ da como resultado un valor P de 0.0048 y está en la región crítica limitada por $z = -1.645$, por lo que las conclusiones son las mismas que en el ejemplo 4. Los diferentes métodos pueden conducir a resultados muy diferentes; vea los estadísticos de prueba de -4.21, -4.29 y -2.82. Las conclusiones son las mismas en este caso, pero podrían ser diferentes en otras situaciones.

Sección 13-3

1. a. Los únicos requisitos son que los pares relacionados sean una muestra aleatoria simple y la población de diferencias sea aproximadamente simétrica.
   b. No hay ningún requisito de una distribución normal o cualquier otra distribución específica.
   c. La prueba de rangos con signo de Wilcoxon es “sin distribución” en el sentido de que no requiere una distribución normal ni ninguna otra distribución específica.
3. La prueba del signo usa sólo los signos de las diferencias, pero la prueba de rangos con signo de Wilcoxon usa rangos que se ven afectados por las magnitudes de las diferencias.
5. Estadístico de prueba: $T = 16.5$. Valor crítico: $T = 6$. No se puede rechazar la hipótesis nula de que la población de diferencias tiene una mediana de 0. No hay evidencia suficiente que respalde la afirmación de que existe una diferencia entre las calificaciones de atributos femeninos y las calificaciones de atributos masculinos.
7. Convierta $T = 8323.5$ al estadístico de prueba $z = -0.63$. Valor P: 0.5287. Valores críticos: $z = \pm 1.96$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que no hay diferencia. No hay evidencia suficiente para respaldar la afirmación de que existe una diferencia entre las calificaciones de atributos femeninos y las calificaciones de atributos masculinos.
9. Convierta $T = 18{,}014$ al estadístico de prueba $z = -16.92$. Valor P: 0.0000. Valores críticos: $z = \pm 2.575$. Hay suficiente evidencia para justificar el rechazo de la afirmación de que la mediana es igual a 2.00.
11. Convierta $T = 196$ al estadístico de prueba $z = -2.88$. Valor P: 0.0040. Valores críticos: $z = \pm 2.575$. Existe suficiente evidencia para justificar el rechazo de la afirmación de que la mediana es igual a 5.670 g. Las monedas de ¢25 no parecen estar acuñadas de acuerdo con las especificaciones.
13. a. 0 y 31,375
    b. 15,687.5
    c. 30,141
    d. $\dfrac{n(n+1)}{2} - k$

Sección 13-4

1. Sí. Las dos muestras son independientes. Las evaluaciones de las profesoras y los profesores no se relacionan de ninguna manera. Las muestras son muestras aleatorias simples. Cada muestra tiene más de 10 valores.
3. H_0: Las evaluaciones de las profesoras y los profesores tienen la misma mediana. Hay tres diferentes hipótesis alternativas posibles:
   H_1: Las evaluaciones de las profesoras y los profesores tienen diferentes medianas.
   H_1: Las evaluaciones de las profesoras tienen una mediana mayor que la mediana de las evaluaciones de los profesores.
   H_1: Las evaluaciones de las profesoras tienen una mediana inferior a la mediana de las evaluaciones de los profesores.
5. $R_1 = 163$, $R_2 = 188$, $\mu_R = 189$, $\sigma_R = 19.4422$, estadístico de prueba: $z = -1.34$. Valor P: 0.1802. Valores críticos: $z = \pm 1.96$. No se puede rechazar la hipótesis nula de que las poblaciones tienen la misma mediana. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las calificaciones de evaluación de las profesoras tienen la misma mediana que las calificaciones de evaluación de los profesores.
7. $R_1 = 253.5$, $R_2 = 124.5$, $\mu_R = 182$, $\sigma_R = 20.607$, estadístico de prueba: $z = 3.47$. Valor P: 0.0005. Valores críticos: $z = \pm 1.96$. Se rechaza la hipótesis nula de que las poblaciones tienen la misma mediana. Existe suficiente evidencia para rechazar la afirmación de que para los sujetos tratados con 20 mg de Lipitor y los sujetos tratados con 80 mg de Lipitor, los cambios en el colesterol LDL tienen la misma mediana. Parece que la cantidad de la dosis tiene un efecto sobre el cambio en el colesterol LDL.
9. $R_1 = 1615.5$, $R_2 = 2755.5$, $\mu_R = 1880$, $\sigma_R = 128.8669$, estadístico de prueba: $z = -2.05$. Valor P: 0.0404. Valores críticos: $z = \pm 1.96$. Se rechaza la hipótesis nula de que las poblaciones tienen la misma mediana. Existe evidencia suficiente para justificar el rechazo de la afirmación de que las calificaciones de evaluación de las profesoras tienen la misma mediana que las calificaciones de evaluación de los profesores.
11. $R_1 = 501$, $R_2 = 445$, $\mu_R = 484$, $\sigma_R = 41.15823$, estadístico de prueba: $z = 0.41$. Valor P: 0.3409. Valor crítico: $z = 1.645$. No se puede rechazar la hipótesis nula de que las poblaciones tienen la misma mediana. No hay evidencia suficiente para respaldar la afirmación de que los sujetos con niveles de plomo medios tienen una mediana más alta de las puntuaciones de IQ completas que los sujetos con niveles altos de plomo. Según estos datos, no parece que el nivel de plomo afecte a los puntuaciones de IQ completas.
13. Si se usa $U = 98.5$, se obtiene $z = 0.41$. El estadístico de prueba es el mismo valor con signo opuesto.

Sección 13-5

1. $R_1 = 164.5$, $R_2 = 150$, $R_3 = 150.5$
3. $n_1 = 10$, $n_2 = 10$, $n_3 = 10$ y $N = 30$
5. Estadístico de prueba: $H = 0.1748$. Valor crítico: $\chi^2 = 5.991$. (Tecnología: Valor $P = 0.916$). No se puede rechazar la hipótesis nula de medianas iguales. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las mujeres de diferentes grupos de edad otorgan calificaciones de atributos con la misma mediana.

7. Estadístico de prueba: $H = 4.9054$. Valor crítico: $\chi^2 = 5.991$. (Tecnología: Valor $P = 0.086$). No se puede rechazar la hipótesis nula de medianas iguales. Los datos no sugieren que los autos más grandes sean más seguros.
9. Estadístico de prueba: $H = 11.4704$. Valor crítico: $\chi^2 = 5.991$. (Tecnología: Valor $P = 0.003$). Se rechaza la hipótesis nula de medianas iguales. Parece que los tres restaurantes tienen tiempos de servicio en auto con diferentes medianas a la hora de la cena.
11. Estadístico de prueba: $H = 2.5999$. Valor crítico: $\chi^2 = 7.815$. (Tecnología: Valor $P = 0.458$). No se puede rechazar la hipótesis nula de medianas iguales. Parece que los cuatro hospitales tienen pesos al nacer con la misma mediana.
13. Los valores de t son 2, 2, 2, 2 y 4, por lo que los valores de T son 6, 6, 6, 6 y 60 y $\Sigma T = 84$. Si se usa $\Sigma T = 84$ y $N = 19$, el valor corregido de H es 0.703, que no es sustancialmente diferente del valor de 0.694 encontrado en el ejemplo 1. En este caso, la gran cantidad de empates no parece tener un efecto dramático en el estadístico de prueba H.

Sección 13-6

1. Los métodos de la sección 10-2 no deben usarse para predicciones. La ecuación de regresión se basa en una correlación lineal entre las dos variables, pero los métodos de esta sección no requieren una relación lineal. Los métodos de esta sección podrían sugerir que existe una correlación con los datos pareados, asociados por alguna relación no lineal, por lo que la ecuación de regresión no sería un modelo adecuado para hacer predicciones.
3. r representa el coeficiente de correlación lineal calculado a partir de datos muestrales pareados; ρ representa el parámetro del coeficiente de correlación lineal calculado a partir de una población de datos pareados; r_s denota el coeficiente de correlación de rango calculado a partir de datos muestrales pareados; ρ_s representa el coeficiente de correlación de rango calculado a partir de una población de datos pareados. El subíndice s se usa para que el coeficiente de correlación de rango se pueda distinguir del coeficiente de correlación lineal r. El subíndice no representa la desviación estándar s. Se utiliza en reconocimiento a Charles Spearman, quien introdujo el método de correlación de rango.
5. $r_s = 1$. Los valores críticos son -0.886 y 0.886. Se rechaza la hipótesis nula de $\rho_s = 0$. Hay suficiente evidencia para respaldar una afirmación de correlación entre distancia y tiempo.
7. $r_s = 0.888$. Valores críticos: -0.618, 0.618. Se rechaza la hipótesis nula de $\rho_s = 0$. Hay suficiente evidencia para respaldar la afirmación de una correlación entre el consumo de chocolate y la tasa de premios Nobel. No tiene sentido pensar que existe una relación causa efecto, por lo que la correlación podría ser el resultado de una coincidencia u otros factores que afectan las variables de la misma manera.
9. $r_s = 1.000$. Valores críticos: -0.700, 0.700. Se rechaza la hipótesis nula de $\rho_s = 0$. Hay suficiente evidencia para respaldar la afirmación de una correlación entre el costo de una rebanada de pizza y la tarifa del metro.
11. $r_s = 1$. Valores críticos: -0.886, 0.886. Se rechaza la hipótesis nula de $\rho_s = 0$. Hay suficiente evidencia para concluir que existe una correlación entre el ancho global de las focas en fotografías y el peso de las focas.
13. $r_s = 0.902$. Valores críticos: -0.415, 0.415. Se rechaza la hipótesis nula de $\rho_s = 0$. Hay suficiente evidencia para respaldar la afirmación de una correlación entre el consumo de chocolate y la tasa de premios Nobel. No tiene sentido pensar que existe una relación causa-efecto, por lo que la correlación podría ser el resultado de una coincidencia u otros factores que afectan las variables de la misma manera.
15. $r_s = 0.360$. Valores críticos: -0.159, 0.159. Se rechaza la hipótesis nula de $\rho_s = 0$. Hay suficiente evidencia para concluir que existe una correlación entre los niveles de presión arterial sistólica y diastólica en los hombres.
17. -0.159 y 0.159. (Utilice $t = 1.975799$ de la tecnología o utilice la interpolación en la tabla A-3 con 151 grados de libertad, por lo que el valor crítico de t es aproximadamente la mitad entre 1.984 y 1.972, que es 1.978). Los valores críticos son los mismos que los encontrados usando la fórmula 13-1.

Sección 13-7

1. No. La prueba de rachas se puede usar para determinar si la secuencia de partidos políticos no es aleatoria, pero la prueba de rachas no muestra si la proporción de republicanos es significativamente mayor que la proporción de demócratas.
3. Los valores críticos son 8 y 19. Como $G = 16$ no es menor o igual a 8 ni $G = 16$ es mayor o igual que 18, no se puede rechazar la aleatoriedad. Parece que la secuencia de partidos políticos es aleatoria.
5. $\bar{x} = 157.7$ muertes. $n_1 = 11$, $n_2 = 9$, $G = 12$, valores críticos: 6, 16. No se puede rechazar la aleatoriedad. No hay evidencia suficiente para justificar el rechazo de la afirmación de que hay aleatoriedad por encima y por debajo de la media. No parece haber una tendencia
7. $n_1 = 20$, $n_2 = 10$, $G = 16$, valores críticos: 9, 20. No se puede rechazar la aleatoriedad. No hay evidencia suficiente para rechazar la afirmación de que las fechas anteriores y posteriores al 1 de julio se seleccionan al azar.
9. $n_1 = 26$, $n_2 = 23$, $G = 20$, $\mu_G = 25.40816$, $\sigma_G = 3.450091$. Estadístico de prueba: $z = -1.57$. Valor P: 0.1164. Valores críticos: $z = \pm 1.96$. No se rechaza la aleatoriedad. No hay suficiente evidencia para rechazar la aleatoriedad. La prueba de rachas no prueba desproporcionadamente más apariciones de una de las dos categorías, por lo que la prueba de rachas no sugiere que ninguna de las dos conferencias sea superior.
11. La mediana es 2895.5, $n_1 = 25$, $n_2 = 25$, $G = 2$, $\mu_G = 26$, $\sigma_G = 3.49927$. Estadístico de prueba: $z = -6.86$. Valor P: 0.0000. Valores críticos: $z = \pm 1.96$. Se rechaza la aleatoriedad. La secuencia no parece ser aleatoria cuando se consideran los valores por encima y por debajo de la mediana. Parece haber una tendencia al alza, por lo que el mercado de valores parece ser una inversión rentable a largo plazo.
13. b. Las 84 secuencias producen los siguientes resultados: 2 secuencias tienen 2 rachas, 7 secuencias tienen 3 rachas, 20 secuencias tienen 4 rachas, 25 secuencias tienen 5 rachas, 20 secuencias tienen 6 rachas, y 10 secuencias tienen 7 rachas.
    c. Con P(2 rachas) = 2/84, P(3 rachas) = 7/84, P(4 rachas) = 20/84, P (5 rachas) = 25/84, P(6 rachas) = 20/84 y P(7 rachas) = 10/84, cada uno de los valores G de 3, 4, 5, 6, 7 puede ocurrir fácilmente por casualidad, mientras que $G = 2$ es improbable porque P(2 rachas) es menor que 0.025. Por lo tanto, el valor crítico más bajo de G es 2, y no hay ningún valor crítico superior que pueda igualarse o excederse.
    d. El valor crítico de $G = 2$ concuerda con la tabla A-10. La tabla indica 8 como el valor crítico superior, pero es imposible obtener 8 rachas usando los elementos dados.

Capítulo 13: Examen rápido

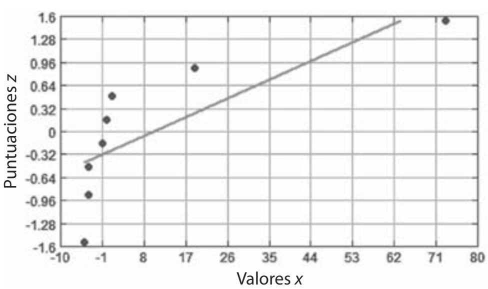
1. 1, 3, 3, 5, 3
2. La clasificación de eficiencia de 0.91 indica que, con todos los demás factores iguales, la correlación de rangos requiere 100 pares de observaciones muestrales para lograr los mismos resultados que 91 pares de observaciones con la prueba paramétrica de correlación lineal, suponiendo que los requisitos más estrictos para usar la correlación lineal se cumplen.
3. a. Prueba sin distribución
   b. El término “prueba sin distribución” sugiere correctamente que la prueba no requiere que una población deba tener una distribución particular, como una distribución normal. El término “prueba no paramétrica” sugiere incorrectamente que la prueba no se basa en un parámetro, pero algunas pruebas no paramétricas se basan en la mediana, que es un parámetro; el término “prueba sin distribución” es mejor porque no hace esa sugerencia incorrecta.
4. Se debe usar la correlación de rangos. La prueba de correlación de rangos se usa para investigar si existe una correlación entre la longitud del pie y la estatura.
5. No, los valores P casi siempre son diferentes, y las conclusiones pueden ser iguales o no.
6. La correlación de rangos se puede usar en una variedad más amplia de circunstancias que la correlación lineal. La correlación de rangos no requiere una distribución normal para ninguna población. La correlación de rangos se puede usar para detectar algunas (no todas) las relaciones que no son lineales.
7. Debido a que sólo hay dos rachas, todos los valores por debajo de la media ocurren al comienzo y todos los valores por encima de la media ocurren al final, o viceversa. Esto indica la presencia de una tendencia ascendente (o descendente).
8. a. Falso
   b. Falso
9. Debido a que la prueba del signo sólo utiliza los signos de las diferencias mientras que la prueba de los rangos con signo de Wilcoxon usa rangos de las diferencias, la prueba de los rangos con signo de Wilcoxon usa más información sobre los datos y tiende a arrojar conclusiones que reflejan mejor la verdadera naturaleza de los datos.
10. Prueba de Kruskal-Wallis

Capítulo 13: Ejercicios de repaso

1. $r_s = 0.400$. Valores críticos: -0.700, 0.700. No se puede rechazar la hipótesis nula de $\rho_s = 0$. No hay evidencia suficiente para respaldar la afirmación de una correlación entre el estrés laboral y el ingreso anual. Según los datos proporcionados, no parece que los trabajos con más estrés tengan salarios más altos.
2. Estadístico de prueba: $H = 2.5288$. (Tecnología: valor $P = 0.2824$). Valor crítico: $\chi^2 = 5.991$. No se puede rechazar la hipótesis nula de medianas iguales. Parece que los tiempos de longevidad después de la ceremonia de investidura para los presidentes de Estados Unidos, papas y monarcas británicos tienen la misma mediana.
3. El estadístico de prueba de $z = -1.43$ da como resultado un valor P de 0.1527 y no es menor o igual que el valor crítico de $z = -1.96$. No se puede rechazar la hipótesis nula de $p = 0.5$. No hay evidencia suficiente para justificar el rechazo de la afirmación de que en cada Serie Mundial, el equipo de la Liga Americana tiene una probabilidad de 0.5 de ganar.
4. $n_1 = 16$, $n_2 = 14$, $G = 11$, valores críticos: 10, 22. No se puede rechazar la aleatoriedad. No hay evidencia suficiente para justificar el rechazo de la afirmación de que los dígitos pares e impares se producen en orden aleatorio. La lotería parece estar funcionando como debería.
5. El estadístico de prueba de $x = 3$ es menor o igual que el valor crítico de 5 (a partir de la tabla A-7). Existe evidencia suficiente para justificar el rechazo de la afirmación de que la muestra proviene de una población con una mediana igual a 5.
6. El estadístico de prueba $T = 21$ es menor o igual que el valor crítico de 59. Existe evidencia suficiente para justificar el rechazo de la afirmación de que la muestra proviene de una población con una mediana igual a 5.
7. $R_1 = 204.5$, $R_2 = 230.5$, $\mu_R = 255$, $\sigma_R = 22.58318$, estadístico de prueba: $z = -2.24$. Tecnología: valor P: 0.025. Valores críticos: $z = \pm 1.96$. Se rechaza la hipótesis nula de que las poblaciones tienen la misma mediana. Hay suficientes pruebas para justificar el rechazo de la afirmación de que las erupciones recientes y erupciones pasadas tienen el mismo intervalo de tiempo medio entre las erupciones. La conclusión sí cambia con un nivel de significancia de 0.01.
8. El estadístico de prueba de $x = 0$ es menor o igual que el valor crítico de 0. Hay suficiente evidencia para rechazar la afirmación de que no hay diferencia. Parece que hay una diferencia de costo entre los vuelos programados con 1 día de anticipación y los programados con 30 días de antelación. Debido a que todos los vuelos programados con 30 días de anticipación cuestan menos que los programados con 1 día de antelación, parece aconsejable programar los vuelos con 30 días de anticipación.
9. El estadístico de prueba de $T = 0$ es menor o igual que el valor crítico de 4. Hay suficiente evidencia para rechazar la afirmación de que las diferencias entre las tarifas para vuelos programados con 1 día de anticipación y los programados con 30 días de antelación tienen una mediana igual a 0. Debido a que todos los vuelos programados con 30 días de anticipación cuestan menos que los programados con 1 día de antelación, parece aconsejable programar los vuelos con 30 días de anticipación.
10. $r_s = 0.714$. Valores críticos: ± 0.738. No se puede rechazar la hipótesis nula de $\rho_s = 0$. No hay evidencia suficiente para respaldar la afirmación de que existe una correlación entre las clasificaciones de los estudiantes y las clasificaciones de la revista. Al clasificar las universidades, los estudiantes y la revista no parecen estar de acuerdo.

Capítulo 13: Ejercicios de repaso acumulado

1. Vuelo 1: $\bar{x} = -1.3$ min, la mediana es -2.0 min, $s = 1.6$ min.
   Vuelo 19: $\bar{x} = 9.9$ min, la mediana es -0.5 min, $s = 26.6$ min.
   Vuelo 21: $\bar{x} = 33.4$ min, la mediana es 15.5 min, $s = 50.3$ min.
   Las medias parecen ser muy diferentes, y el vuelo 21 tiene los mayores tiempos de demora en la salida. Las medianas parecen ser muy diferentes, y el vuelo 21 tiene los mayores tiempos de demora en la salida. Las desviaciones estándar parecen ser muy diferentes, y el vuelo 21 tiene la mayor cantidad de variación. El vuelo 21 parece ser el vuelo menos predecible porque tiene la variación más alta y parece tener los mayores tiempos de demora en la salida.
2. La gráfica cuantilar normal sugiere que los tiempos de demora en la salida para el Vuelo 19 no se distribuyen normalmente.

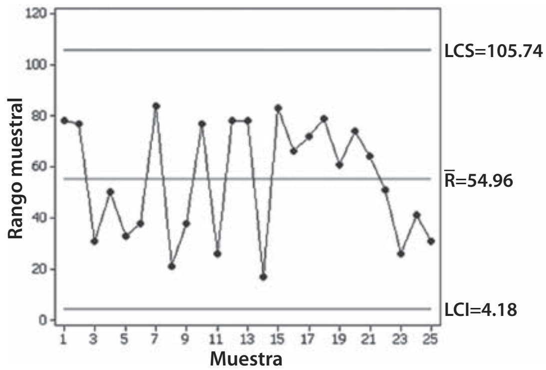


3. Estadístico de prueba de Kruskal-Wallis: $H = 3.2600$. Tecnología: Valor $P = 0.1959$. Valor crítico: $\chi^2 = 5.991$ (suponiendo un nivel de significancia de 0.05). No se puede rechazar la hipótesis nula de medianas iguales. No hay evidencia suficiente para justificar el rechazo de la afirmación de que las tres muestras provienen de poblaciones con el mismo tiempo de demora medio en la salida.
4. $3.1\% < p < 4.7\%$. Tenemos un 95% de confianza en que los límites de 3.1 y 4.7% realmente contienen el porcentaje verdadero de la población de trabajadores que dan positivo en las pruebas de drogas.
5. H_0: $p = 0.03$. H_1: $p > 0.03$. Estadístico de prueba: $z = 2.36$. Valor P: 0.0092 (Tabla: 0.0091). Valor crítico: $z = 1.645$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la tasa de resultados positivos en las pruebas de drogas entre los trabajadores de Estados Unidos es mayor al 3.0%.
6. La media muestral es 54.8 años. $n_1 = 19$, $n_2 = 19$, y el número de ejecuciones es $G = 18$. Los valores críticos son 13 y 27. No se puede rechazar la hipótesis nula de aleatoriedad. No hay evidencia suficiente para garantizar el rechazo de la afirmación de que la secuencia de edades es aleatoria en relación con los valores superiores e inferiores a la media. Los resultados no sugieren que haya una tendencia al alza o una tendencia a la baja.
7. 2401
8. Hay un número relativamente pequeño de jugadores con sueldos que son sustancialmente grandes, por lo que la media se ve fuertemente afectada por esos valores, lo que resulta en un gran valor de la media, pero la mediana no se ve afectada por el pequeño número de salarios muy grandes.
9. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 1.36$. Valor P: 0.0865 (Tabla: 0.0869). Valor crítico: $z = 1.645$. No se puede rechazar H_0. No hay evidencia suficiente para respaldar la afirmación de que la mayoría de la población no teme a las alturas en los rascacielos. Debido a que los propios encuestados eligieron responder, la muestra es de respuesta voluntaria, no una muestra aleatoria, por lo que los resultados podrían no ser válidos.
10. Debe haber un error, porque las tasas de 13.7 y 10.6% no son posibles con muestras de tamaño 100.

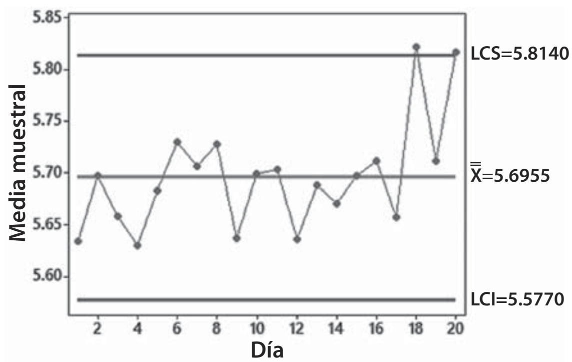
Respuestas del capítulo 14

Sección 14-1

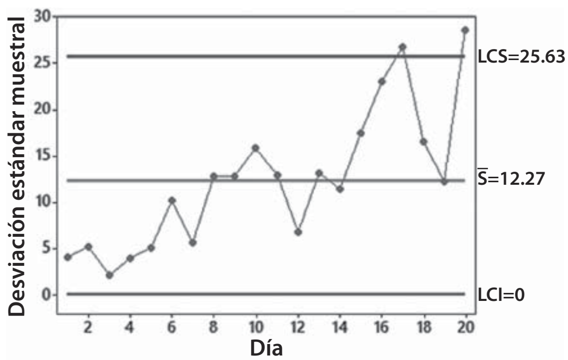
1. No. Si sabemos que el proceso está bajo control estadístico, sabemos que ninguno de los tres criterios fuera de control se satisface, pero no sabemos nada sobre si se cumplen las especificaciones o los requisitos. Es posible estar bajo control estadístico mientras se fabrican altímetros con errores que son demasiado grandes para satisfacer los requisitos de la FAA.
3. La media está fuera de control estadístico. Las elevaciones han disminuido sustancialmente en los últimos años, por lo que el Lago Mead se está volviendo menos profundo. Las disminuciones han sido significativas (y están teniendo un impacto dramático en las poblaciones afectadas).
5. $\bar{x} = 267.11$ lb, $\bar{R} = 54.96$ lb. Para la gráfica R: LCI = 4.18 lb y LCS = 105.74 lb. Para la gráfica $\bar{x}$: LCI = 244.08 lb y LCS = 290.14 lb
7. El gráfico R no cumple con ninguno de los tres criterios fuera de control, por lo que la variación del proceso parece estar bajo control estadístico. (La respuesta continua en la columna de la derecha)



9. $\bar{x} = 5.6955$ g, $\bar{R} = 0.2054$ g. Para la gráfica R: LCI = 0.0000 g y LCS = 0.4342 g. Para la gráfica $\bar{x}$: LCI = 5.5770 g y LCS = 5.8140 g.
11. Hay puntos que se encuentran más allá del límite de control superior, por lo que la media del proceso parece estar fuera de control estadístico.



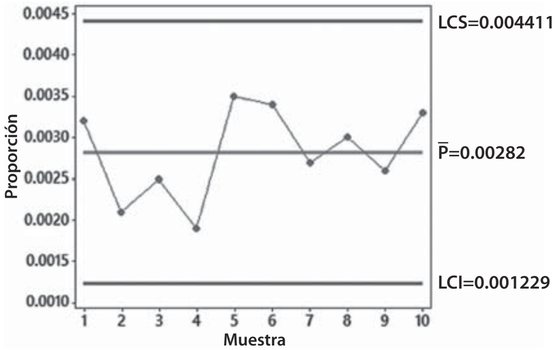
13. Excepto por los valores en la escala vertical, la gráfica s es casi idéntica a la gráfica R que se muestra en el ejemplo 3.



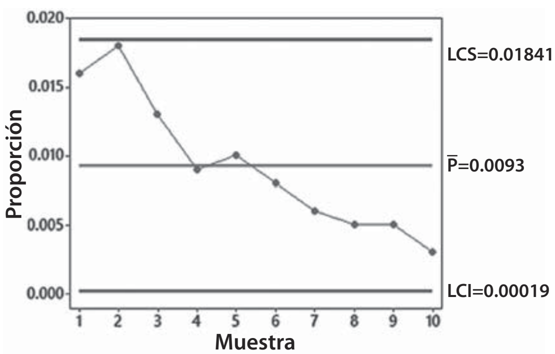
Sección 14-2

1. No, el proceso parece estar fuera de control estadístico. Hay una tendencia a la baja y hay al menos ocho puntos consecutivos que se encuentran por encima de la línea central. Debido a que las proporciones de defectos están disminuyendo, el proceso de fabricación no se está deteriorando; está mejorando.
3. Debido a que el valor de -0.00325 es negativo y la proporción real de defectos no puede ser menor que 0, deberíamos reemplazar ese valor por 0.

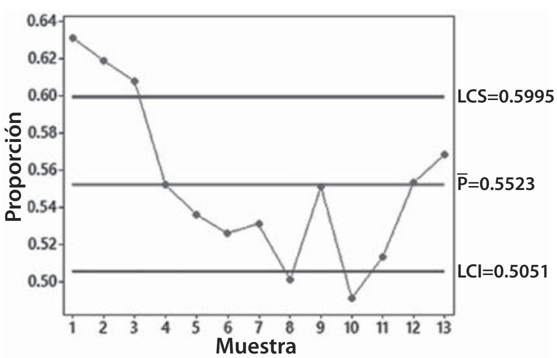
5. El proceso parece estar bajo control estadístico. (Teniendo en cuenta un cambio hacia arriba, note que el primero y el último puntos son aproximadamente iguales).



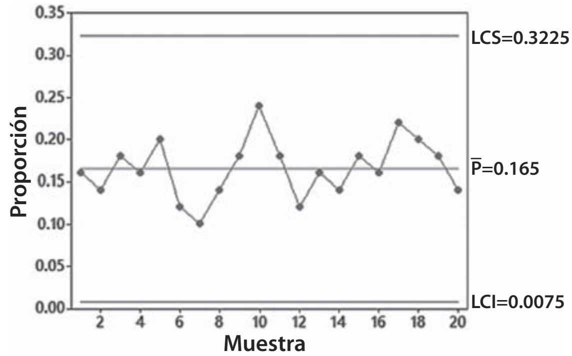
7. El proceso parece estar fuera de control estadístico debido a una tendencia a la baja, pero el número de defectos parece estar disminuyendo, por lo que el proceso está mejorando. Las causas de la disminución del número de defectos deben identificarse para poder continuar.



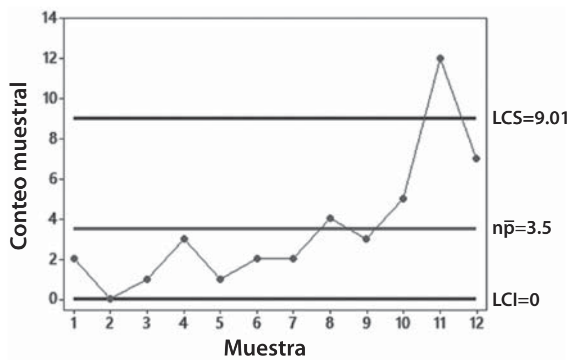
9. El proceso está fuera de control estadístico porque hay puntos que se encuentran más allá del límite de control superior y hay puntos que se encuentran más allá del límite de control inferior. Además, hay ocho puntos consecutivos que se encuentran debajo de la línea central. El porcentaje de votantes está aumentando en las recientes elecciones presidenciales, y debería ser mucho más alto que cualquiera de las tasas mostradas.



11. Aunque el proceso está bajo control estadístico, las proporciones de defectos son sustancialmente altas, por lo que se deben tomar medidas correctivas inmediatas para reducir dramáticamente las proporciones de defectos.



13. Excepto por una escala vertical diferente, la gráfica básica de control es idéntica a la dada para el ejemplo 1.



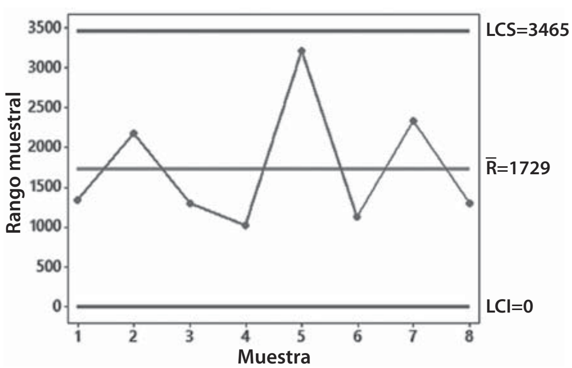
Capítulo 14: Examen rápido

1. Los datos de proceso son datos dispuestos de acuerdo con una secuencia de tiempo. Son medidas de una característica de bienes o servicios que resultan de alguna combinación de equipo, personas, materiales, métodos y condiciones.
2. La variación aleatoria se debe al azar, pero la variación asignable resulta de causas que se pueden identificar, como maquinaria defectuosa o empleados no capacitados.
3. Hay un patrón, tendencia o ciclo que obviamente no es aleatorio. Hay un punto situado fuera de la región entre los límites de control superior e inferior. Hay al menos ocho puntos consecutivos por encima o por debajo de la línea central.
4. Una gráfica *R* usa rangos para monitorear la variación, pero una gráfica *x* usa medias muestrales para monitorear el centro (promedio) de un proceso.
5. No. La gráfica *R* tiene al menos ocho puntos consecutivos ubicados debajo de la línea central, hay al menos ocho puntos consecutivos que se encuentran por encima de la línea central, hay puntos que se encuentran más allá de los límites de control superior e inferior, y hay un patrón que muestra que los rangos han saltado en valor para las muestras más recientes. ¡Vaya lío!

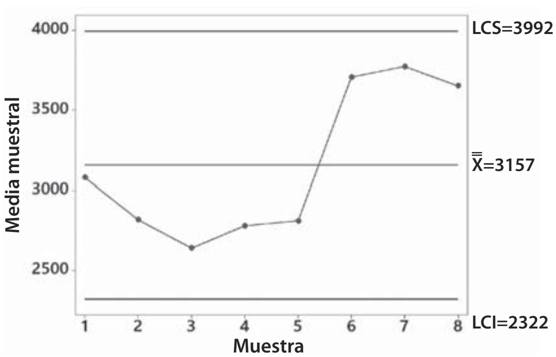
6. $\overline{R} = 67.0$ pies. En general, se encuentra un valor de $\overline{R}$ al determinar primero el rango para los valores dentro de cada subgrupo individual; la media de esos rangos es el valor de $\overline{R}$.
7. No. La gráfica $\overline{x}$ tiene un punto que está más allá del límite superior de control, y hay al menos ocho puntos consecutivos debajo de la línea central.
8. $\overline{\overline{x}} = -2.24$ pies. En general, se encuentra un valor de $\overline{\overline{x}}$ al determinar primero la media de los valores dentro de cada subgrupo individual; la media de esas medias de subgrupo es el valor de $\overline{\overline{x}}$.
9. No. Las gráficas de control se pueden usar para determinar si la media y la variación están bajo control estadístico, pero no revelan nada sobre las especificaciones o requisitos.
10. Debido a que existe una tendencia a la baja, el proceso está fuera del control estadístico, pero la tasa de defectos está disminuyendo, por lo que debemos investigar e identificar la causa de esa tendencia para poder continuar.

Capítulo 14: Ejercicios de repaso

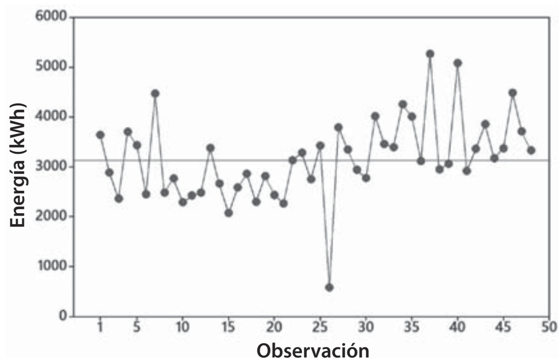
1. $\overline{\overline{x}} = 3157$ kWh, $\overline{R} = 1729$ kWh. Gráfica R: LCI $= 0$ kWh, LCS $= 3465$ kWh. Gráfica : LCI $= 2322$ kWh, LCS $= 3992$ kWh.
2. La variación del proceso está bajo control estadístico.



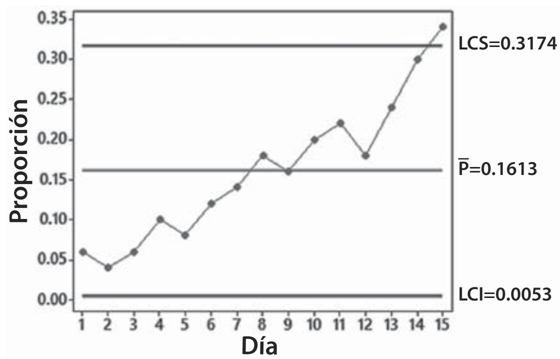
3. Parece haber un cambio en los valores medios, por lo que la media del proceso está fuera de control estadístico.



4. Parece haber una ligera tendencia al alza. Hay un punto que parece ser excepcionalmente bajo. (La compañía eléctrica del autor cometió un error al registrar y reportar el consumo de energía para ese período de tiempo).



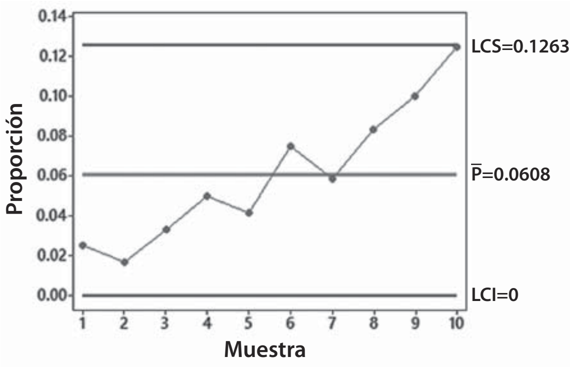
5. Debido a que existe una clara tendencia ascendente y hay un punto más allá del límite de control superior, el proceso está fuera de control estadístico. Debido a que los tiempos de pedido están aumentando claramente, se deben tomar medidas correctivas inmediatas.



Capítulo 14: Ejercicios de repaso acumulado

1. $0.528 < p < 0.572$. Debido a que el rango completo de valores en el intervalo de confianza incluye valores que son todos mayores que 0.5, parece que la mayoría de los adultos se informan más a menudo de los síntomas médicos a través de Internet que con su médico.
2. H_0: $p = 0.5$. H_1: $p > 0.5$. Estadístico de prueba: $z = 4.48$. Valor P: 0.0000 (Tabla: 0.0001). Valor crítico: $z = 1.645$. Se rechaza H_0. Existe evidencia suficiente para respaldar la afirmación de que la mayoría de los adultos se informan más a menudo sobre los síntomas médicos a través de Internet que con su médico.
3. La gráfica es engañosa. La escala vertical comienza con una frecuencia de 800 en lugar de 0, por lo que la diferencia entre las respuestas “sí” y “no” es muy exagerada.
4. a. 0.166 b. 0.909
5. $r = 0.356$. Valor P: 0.313 (Tabla >0.05). Valores críticos: $r = \pm 0.632$. No hay evidencia suficiente para respaldar la afirmación de una correlación lineal entre el DJIA y el número de manchas solares. Debido a que no esperamos ninguna relación entre el número de manchas solares y el comportamiento de las acciones, este resultado no es sorprendente.

6. $\hat{y} = 9772 + 79.2x$. El mejor valor predicho del DJIA en el año 2004 es $\overline{y} = 13{,}423.6$, y ese valor no está cerca del valor real de 2004 de 10,855.
7. a. 99.89% (Tabla: 99.88%)
   b. 0.1587
8. Existe un patrón de tendencia ascendente, por lo que el proceso está fuera de control estadístico.



9. $\overline{x} = 7.3$, mediana $= 6.5$, $s = 4.2$. Estos estadísticos no transmiten información sobre el patrón cambiante de los datos a lo largo del tiempo.
10. Estadístico de prueba: $\chi^2 = 42.557$. Valor P: 0.000 (Tabla <0.005). Valor crítico: $\chi^2 = 3.841$. Hay evidencia suficiente para justificar el rechazo de la afirmación de que la sentencia es independiente de la apelación. Los resultados alientan las apelaciones para los acusados culpables.

CRÉDITOS

Fotografías

Capítulo 1

P1, GongTo/Shutterstock; **P1**, Stokkete/Shutterstock; **P5**, Gary Blakeley/Shutterstock; **P6**, USBFCO/Shutterstock; **P7**, Wavebreakmedia/Shutterstock; **P8**, Donskarpo/Shutterstock; **P14**, 18percentgrey/Shutterstock; **P15**, Khamidulin Sergey/Shutterstock; **P15**, Suppakij1017/Shutterstock; **P16**, Chris DeRidder/Shutterstock; **P17**, Ollyy/Shutterstock; **P18**, Allstar Picture Library/Alamy Stock Photo; **P20**, Lucky Business/Shutterstock; **P26**, Andersen Ross/Stockbyte/Getty Images; **P27**, Pixsooz/Shutterstock; **P27**, Triff/Shutterstock; **P29**, Auremar/Shutterstock; **P30**, Fujji/Shutterstock; **P31**, Dotshock/Shutterstock

Capítulo 2

P40, Ministr-84/Shutterstock; **P43**, Gallofoto/Shutterstock; **P46**, Image Source; **P53**, Valua Vitaly/Shutterstock; **P58**, Stockbyte/Getty Images; **P72**, Dmitriy Eremenkov/Shutterstock

Capítulo 3

P80, 06photo/Shutterstock; **P83**, Wavebreakmedia/Shutterstock; **P84**, Image Point Fr/Shutterstock; **P87**, Stefan Schurr/Shutterstock; **P98**, Kitch Bain/Shutterstock; **P99**, Alfredo Ragazzoni/Shutterstock; **P100**, Sozaijiten; **P102**, Ariwasabi/Shutterstock; **P115**, Trinacria Photo/Shutterstock; **P117**, Poleze/Shutterstock

Capítulo 4

P131, Sculpies/Shutterstock; **P135**, Photomatz/Shutterstock; **P136**, Dansin/E+/Getty Images; **P136**, Pakhnyushcha/Shutterstock; **P136**, Worker/Shutterstock; **P137**, PeJo/Shutterstock; **P138**, Ssuaphotos/Shutterstock; **P139**, Bochkarev Photography/Shutterstock; **P141**, Nomad_Soul/Shutterstock; **P142**, Dgbomb/Shutterstock; **P152**, Ssuaphotos/Shutterstock; **P153**, Jim Barber/Shutterstock; **P154**, Chepe Nicoli/Shutterstock; **P160**, Brian sullivan/E+/Getty Images; **P161**, Africa Studio/Shutterstock; **P162**, Alexander Raths/Shutterstock; **P164**, Philip Gangler/Shutterstock; **P165**, Stephen VanHorn/Shutterstock; **P169**, Bioraven/Shutterstock; **P170**, Serg Shalimoff/Shutterstock; **P172**, Naki Kouyioumtzis/Pearson Education, Inc.; **P173**, Zentilia/123RF

Capítulo 5

P184, Fuse/Getty Images; **P189**, JHDT Stock Images LLC/Shutterstock; **P190**, Laborant/Shutterstock; **P201**, B Calkins/Shutterstock; **P205**, Vitalinka/Shutterstock; **P225**, Alfred Eisenstaedt/The LIFE Picture Collection/Getty Images; **P225**, General Motors Cancer Research Foundation

Capítulo 6

P226, Stacy L. Pearsall/Aurora/Getty Images; **P230**, Rui Santos/123RF; **P237**, Elena Stepanova/Shutterstock; **P238**, Elena Yakusheva/Shutterstock; **P269**, Ciurea Adrian/Shutterstock

Capítulo 7

P297, Syda Productions/Shutterstock; **P300**, Mike Flippo/Shutterstock; **P302**, Stocksnapper/Shutterstock; **P304**, Mama_Mia/Shutterstock; **P305**, Roland Ijdema/Shutterstock; **P308**, Juan Camilo Bernal/Shutterstock; **P319**, Eric Isselée/Fotolia; **P320**, Robin W/Shutterstock; **P321**, Gary Blakeley/Shutterstock; **P322**, Rafael Ramirez Lee/Shutterstock; **P343**, Andresr/Shutterstock

Capítulo 8

P356, Tiero/Fotolia; **P364**, Azuzl/Shutterstock; **P367**, Viktoria/Shutterstock; **P370**, David H.Seymour/Shutterstock; **P376**, Suravid/Shutterstock; **P379**, Alexey Burmakin/123RF; **P381**, Pete Saloutos/Shutterstock; **P392**, Zentilia/Shutterstock

Capítulo 9

P414, Dorothy Alexander/Alamy Stock Photo; **P418**, Bikeriderlondon/Shutterstock; **P420**, Sergei Telegin/Shutterstock; **P421**, Alex Staroseltsev/Shutterstock; **P429**, GJS/Shutterstock; **P432**, Amy Walters/Shutterstock; **P433**, David Lee/Shutterstock; **P435**, Christy Thompson/Shutterstock; **P444**, D7INAMI7S/Shutterstock; **P445**, Andrey Arkusha/Shutterstock; **P446**, Pressmaster/Shutterstock

Capítulo 10

P468, Greatandlittle/123RF; **P475**, John Roman Images/Shutterstock; **P477**, Alexander Raths/Shutterstock; **P480**, Africa Studio/Shutterstock; **P493**, Zffoto/Shutterstock; **P495**, Doglikehorse/Shutterstock; **P517**, David Madison/Photodisc/Getty Images; **P523**, Leah-Anne Thompson/Shutterstock

Capítulo 11

P533, Pedro Miguel Sousa/Shutterstock; **P539**, AlexKalashnikov/Shutterstock; **P541**, Serhiy Shullye/Shutterstock; **P549**, Tatiana Popova/Shutterstock; **P551**, Lenetstan/Shutterstock; **P552**, Dundanim/Shutterstock

Capítulo 12

P566, Bikeriderlondon/Shutterstock

Capítulo 13

P597, 97/E+/Getty Images; **P636**, Sinisa Botas/Shutterstock; **P642**, Dotshock/Shutterstock; **P644**, Luna Vandoorne/Shutterstock

Capítulo 14

P654, Maravic/Vetta/Getty Images; **P657**, Oleg Zabielin/Shutterstock; **P658**, Blaz Kure/Shutterstock; **P660**, Michael Rosskothen/Shutterstock; **P661**, Moreno Soppelsa/Shutterstock; **P662**, Age Fotostock/SuperStock; **P663**, Blend Images/Shutterstock; **P669**, Tupungato/Shutterstock

Capítulo 15

P680, Umberto Shtanzman/Shutterstock

PORTADA

Laura A. Watt/Getty Images; Robert Essel NYC/Getty Images

Texto

Capítulo 1

P4, Ryan y Arnold, *American Journal of Public Health*, vol. 101, núm. 10; **P8**, *From Lies, Damn Lies, And Statistics: The Manipulation Of Public Opinion In America by 1976*. Publicado por Dell Publishing, © 1976; **P9**, con base en datos de Centers for Disease Control and Prevention y Department of Energy; **P10**, *Journal of Nutrition* (vol. 130, núm. 8); **P11**, del National Center for Health Statistics; **P14**, Basado en datos de "Texting While Driving and Other Risky Motor Vehicle Behaviors Among US High School Students," de Olsen, Shults, Eaton, *Pediatrics*, vol. 131, núm. 6; **P23**, basado en datos de KRC Research; **P32**, basado en datos de *Journal of Great Lakes Research*; **P33**, basado en datos de "A Randomized Trial Comparing Acupuncture, Simulated Acupuncture, and Usual Care for Chronic Low Back Pain", de Cherkin *et al*., *Archives of Internal Medicine*, vol. 169, núm. 9; **P33,** basado en datos de Pfizer, Inc.; **P36**, basado en datos de *Consumer Reports magazine*

Capítulo 2

P44-45, basado en datos de the Centers for Disease Control; **varias p**áginas, pantallas de Minitab. Cortesía de Minitab Corporation; **P59-60**, basado en datos de "Misconduct Accounts for the Majority of Retracted Scientific Publications", de Fang, Steen, Casadevall, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, núm. 3; **P66**, basado en datos de TE Connectivity; **P66**, basado en datos de *The Ladders*; **varias páginas**, pantallas de Stat-Crunch. Usadas con autorización de StatCrunch; **varias páginas**, Excel 2016, Windows 10, Microsoft Corporation

Capítulo 3

P63-64, basado en datos de *U.S. News & World Report*; **P93**, basado en datos de *U.S. News & World Report*; **P112**, basado en datos de *California Health Interview Survey*; **P125**, basado datos del estudio antropométrico de Gordon, Clauser, *et al*.; **P126**, basado en datos de "An Unexpected Rise in Strontium-90 in U.S. Deciduous Teeth in the1990s", de Mangano *et al*., *Science of the Total Environment*

Capítulo 4

P139, basado en datos de "Texting While Driving …", de Olsen, Shults, Eaton, *Pediatrics*, vol. 131, núm. 6; **P144**, basado en datos de the Genetics & IVF Institute; **P146**, Basado en datos de "Are Women Carrying 'Basketballs'…", de Perry, DiPietro, Constigan, Birth, vol. 26, núm. 3; **P146**, basado en datos de "Mortality Reduction with Air Bag and Seat Belt Use in Head-on Passenger Car Collisions", de Crandall, Olson, Sklar, *American Journal of Epidemiology*, vol. 153, núm. 3; **P146**, basado en datos de *Journal of the National Cancer Institute*; **P148**, basado en datos del estudio antropométrico de Gordon, Churchill, *et al*.; **P150**, basado en datos de "Prevalence and Comorbidity of Nocturnal Wandering in the U.S. General Population" de Ohayon *et al*., *Neurology*, vol. 78, núm. 20; **P156**, basado en datos de la guía del 5% para cálculos engorrosos; **P156**, datos de un estudio sobre el servicio de comida rápida en automóvil; **P157**, basado en datos de diversas fuentes, incluyendo lifehacker.com; **P158**, basado en datos de Arshad Mansoor, Vicepresidente en jefe del Electric Power Research Institute; **P158**, basado en datos de "Pacemaker and ICD Generator Malfunctions", de Maisel *et al*., *Journal of the American Medical Association*, vol. 295, núm. 16; **P158**, basado en datos de "Association Between Helicopter *vs* Ground Emergency Medical Services and Survival for Adults with Major Trauma", de Galvagno *et al*., *Journal of the American Medical Association*, vol. 307, núm. 15; **P161**, basado en datos de *USA Today*; **P164**, basado en *Probabilistic Reasoning in Clinical Medicine* de David Eddy, Cambridge University Press; **P167**, basado en datos de "The Denomination Effect" de Priya Raghubir y Joydeep Srivastava, *Journal of Consumer Research*, vol. 36; **P168**, basado en datos de diversas fuentes, incluyendo lifehacker.com; **P168**, basado en datos de Arshad Mansoor, Vicepresidente en jefe del Electric Power Research Institute; **P178**, basado en datos de *USA Today*; **P178**, los datos provienen de "Mortality Reduction with Air Bag and Seat Belt Use in Head-on Passenger Car Collisions" de Crandall Olson & Sklar, *American Journal of Epidemiology*, vol. 153, núm. 3; **P179**, basado en datos de Centers for Disease Control and Prevention; **P182**, basado en datos de "Carbon Monoxide Test Can Be Used to Identify Smoker", de Patrice Wendling, *Internal Medicine News*, vol. 40., núm. 1, y Centers for Disease Control and Prevention

Capítulo 5

P188, basado en datos de un estudio de Adecco; **P196**, basado en datos de National Institutes of Health; **P196**, based en un estudio de mensajería instantánea de Microsoft; **P196**, basado en un estudio de TE Connectivity; **P197**, Basado en datos de "Hemispheric Dominance and Cell Phone Use", de Seidman *et al*., *JAMA Otolaryngology—Head & Neck Surgery*, vol. 139, núm. 5; **P198**, basado en datos de "Prevalence and Comorbidity of Nocturnal Wandering In the U.S. Adult General Population", de Ohayon *et al*., *Neurology*, vol. 78, núm. 20; **P199**, basado en datos de U.S. Department of Health and Human Services; **P200**, basado en los resultados de un estudio de Pew Research Center; **varias páginas**, ©Triola Stats, todos los derechos reservados; **P208**, basado en un estudio de Pitney Bowes; **P209**, basado en datos de un estudio de Harris Interactive; **P210**, basado en datos de un estudio de LG Smartphone; **P212**, basado en datos de National Institutes of Health; **P212**, basado en los resultados de una muestra de los reactivos; **P213**, basado en datos de ICR Survey Research Group; **P219**, basado en datos de National Highway Traffic Safety Administration; **P220**, basado en datos de Department of Transportation; **P220-221**, basado en resultados de un estudio de AARP Bulletin; **P221**, basado en un estudio de Coca-Cola; **P221**, basado en datos de National Institutes of Health; **P221**, basado en datos de U.S. National Center for Health Statistics; **P223**, HSS; **P223**, basado en datos del artículo de investigación de IBM "Passenger-Based Predictive Modeling of Airline No-Show Rates", de Lawrence, Hong y Cherrier

Capítulo 6

P230, de *Statisticians and the EPA: What's the Connection Issue -371* de Barry D. Nussbaum. Publicado por la American Statistical Association, © mayo de 2008; **P251**, basado en datos del estudio antropométrico de Gordon, Churchill, *et al*.; **P252**, basado en datos de "Ethological Study of Facial Behavior in Nonparanoid and Paranoid Schizophrenic Patients", de Pittman, Olk, Orr y Singh, *Psychiatry*, vol. 144, núm. 1; **P252**, basado en datos de National Health Survey; **P252**, basado en datos de Department of Transportation; **P270**, basado en datos de investigadores de la University of Maryland; **P273**, basado en datos de National Health and Nutrition Examination Survey; **P274**, basado en datos del estudio antropométrico de Gordon, Churchill, *et al*.; **P275**, basado en datos de "Neurobehavioral Outcomes of School-age Children Born Extremely Low Birth Weight or Very Preterm" de Anderson *et al*., *Journal of the American Medical Association*, vol. 289, núm. 24; **P289**, basado en PPG Industries; **P290**, basado en "Prevalence and Comorbidity of Nocturnal Wandering in the U.S. Adult General Population", de Ohayon *et al*., *Neurology*, vol. 78, núm. 20; **P290**, basado en datos de ICR Research Group; **P291**, basado en datos de un artículo de investigación de IBM por Lawrence, Hong y Cherrier; **P295**, basado en un estudio del Dr. P. Soria en la Indiana University; **P295**, basado en datos del U.S. Army Anthropometry Survey (ANSUR)

Capítulo 7

P312, basado en datos de *QSR magazine*; **P312**, basado en datos de Bristol-Myers Squibb Co.; **P312**, basado en datos de the Physicians Insurers Association of America; **P312**, basado en datos de *QSR magazine*; **P312**, basado en datos de Purdue Pharma L.P.; **P313**, basado en datos de "A Close Look at Therapeutic Touch," *Journal of the American Medical Association*, vol. 279, núm. 13; **P313**, basado en datos de "Use of Prescription and Over-the-Counter Medications and Dietary Supplements Among Older Adults in the United States", de Qato *et al*., *Journal of the American Medical Association*, vol. 300, núm. 24; **P313**, basado en datos de ICR Research Group; **P313**, los datos provienen de la *Journal of the National Cancer Institute*; **P314**, basado en datos de "Sustained Care Intervention and Postdischarge Smoking Cessation Among Hospitalized Adults", de Rigotti *et al*., *Journal of the American Medical Association*, vol. 312, núm. 7; **P315**, basado en datos de NetMarketShare; **P315**, basado en un estudio de 3M Privacy Filters; **P315**, basado en un reporte de Spil Games; **P329**, basado en datos de "Cognitive Behavioral Therapy *vs* Zopiclone for Treatment of Chronic Primary Insomnia in Older Adults", de Sivertsen *et al*., *Journal of the American Medical Association*, vol. 295, núm. 24; **P329**, basado en datos de "Effect of Raw Garlic *vs* Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults with Moderate Hypercholesterolemia", de Gardner *et al*., *Archives of Internal Medicine*, vol. 167; **P329**, basado en datos de Food and Drug Administration; **P330**, basado en datos de National Center for Education Statistics; **P340**, basado en datos de "Cognitive Behavioral Therapy *vs* Zopiclone for Treatment of Chronic Primary Insomnia in Older Adults", de Sivertsen *et al*., *Journal of the American Medical Association*, vol. 295, núm. 24; **P340**, basado en datos de "Effect of Raw Garlic *vs* Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults with Moderate Hypercholesterolemia", de Gardner *et al*., *Archives of Internal Medicine*, vol. 167); **P340**, basado en datos de SigAlert; **P348**, basado en datos del Environmental Working Group; **P348**, basado en datos de *QSR magazine*; **P348**, basado en datos de Bristol-Myers Squibb Co.; **P349**, basado en datos de the Physicians Insurers Association of America; **P350**, basado en datos de un estudio de Rasmussen Reports

Capítulo 8

P373, basado en datos de Pfizer, Inc.; **P383**, basado en datos de Pfizer; **P383**, basado en datos de *QSR magazine*; **P383**, basado en datos de Bristol-Myers Squibb Co.; **P384**, basado en datos de Purdue Pharma L.P.; **P384**, basado en datos de the Physicians Insurers Association of America; **P384**, basado en datos de experimentos realizados por Charles R. Honts de la Boise State University y Gordon H. Barland del Department of Defense Polygraph Institute; **P385**, basado en datos de "A Close Look at Therapeutic Touch", *Journal of the American Medical Association*, vol. 279, núm. 13; **P385**, basado en datos de the *Journal of the National Cancer Institute* según se reportó en *USA Today*; **P385**, basado en datos de "UPC Scanner Pricing Systems: Are They Accurate?" de Goodstein, *Journal of Marketing*, vol. 58; **P385**, basado en datos de "High-Dose Nicotine Patch Therapy", de Dale *et al*., *Journal of the American Medical Association*, vol. 274, núm. 17; **P385**, basado en datos de "Holidays, Birthdays, and Postponement of Cancer Death", de Young y Hade, *Journal of the American Medical Association*, vol. 292, núm. 24; **P386**, basado en datos de "Sustained Care Intervention and Postdischarge Smoking Cessation Among Hospitalized Adults", de Rigotti *et al*., *Journal of the American Medical Association*, vol. 312, núm. 7; **P386**, basado en datos de "Use of Prescription and Over-the-Counter Medications and Dietary Supplements Among Older Adults in the United States," de Qato *et al*., *Journal of the American Medical Association*, vol. 300, núm. 24; **P390**, creado mediante el uso del software JMP®. Copyright © SAS Institute Inc. SAS y todos los demás productos o servicios con el nombre de SAS Institute Inc. Son marcas registradas o marcas de SAS Institute Inc., Cary, NC, USA; **P396**, basado en datos de "Content and Ratings of Teen-Rated Video Games", de Haninger y Thompson, *Journal of the American Medical Association*, vol. 291, núm. 7; **P397**, basado en datos de "Effect of Raw Garlic *vs* Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults with Moderate Hypercholesterolemia", de Gardner *et al*., *Archives of Internal Medicine*, vol. 167; **P397**, basado en datos de "Comparison of the Atkins, Ornish, Weight Watchers, and Zone Diets for Weight Loss and Heart Disease Reduction", de Dansinger *et al*., *Journal of the American Medical Association*, vol. 293, núm. 1; **P398**, basado en datos de "Cognitive Behavioral Therapy *vs* Zopiclone for Treatment of Chronic Primary Insomnia in Older Adults", de Sivertsen *et al*., *Journal of the American Medical Association*, vol. 295, núm. 24; **P398**, basado en datos de "Lead, Mercury, and Arsenic in US and Indian Manufactured Ayurvedic Medicines Sold via the Internet", de Saper *et al*., *Journal of the American Medical Association*, vol. 300, núm. 8; **P408**, basado en datos de payscale.com; **P409**, basado en datos de ICR Survey Research Group

Capítulo 9

P424, basado en datos de *QSR magazine*; **P424**, basado en datos de "Sustained Care Intervention and Postdischarge Smoking Cessation Among Hospitalized Adults", de Rigotti *et al*., *Journal of the American Medical Association*, vol. 312, núm. 7; **P425**, basado en datos de "Do We Dream in Color?" de Eva Murzyn, *Consciousness and Cognition*, vol. 17, núm. 4; **P425**, basado en datos de Purdue Pharma L.P.; **P425**, basado en datos de "Who Wants Airbags?" de Meyer y Finney, *Chance*, vol. 18, núm. 2; **P425**, basado en datos de "Survival from In-Hospital Cardiac Arrest During Nights and Weekends," de Peberdy *et al*., *Journal of the American Medical Association*, vol. 299, núm. 7; **P426**, basado en datos de "An Evaluation of *Echinacea angustifolia* in Experimental Rhinovirus Infections," de Turner *et al*., *New England Journal of Medicine*, vol. 353, núm. 4; **P426**, basado en datos de "Sustainability of Reductions in Malaria Transmission and Infant Mortality in Western Kenya with Use of Insecticide-Treated Bednets", de Lindblade *et al*., *Journal of the American Medical Association*, vol. 291, núm. 21; **P426**, basado en datos de "Hemispheric Dominance and Cell Phone Use", de Seidman *et al*., *JAMA Otolaryngology—Head & Neck Surgery*, vol. 139, núm. 5; **P426**, basado en datos de "The Denomination Effect", de Raghubir y Srivastava, *Journal of Consumer Research*, vol. 36; **P426**, basado en datos de "High-Flow Oxygen for Treatment of Cluster Headache", de Cohen, Burns y Goadsby, *Journal of the American Medical Association*, vol. 302, núm. 22; **P427**, basado en datos de "Final Report on the Aspirin Component of the Ongoing Physicians' Health Study", *New England Journal of Medicine*, vol. 321:129-135; **P427**, basado en datos de "The Left-Handed: Their Sinister History", de Elaine Fowler Costas, *Education Resources Information Center*, artículo 399519; **P427**, basado en datos de "Association Between Helicopter *vs* Ground Emergency Medical Services and Survival for Adults With Major Trauma", de Galvagno *et al*., *Journal of the American Medical Association*, vol. 307, núm. 15; **P439**, basado en 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; **P440**, basado en datos 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; **P441**, basado en datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4; **P442**, basado 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; **P451**, basado en datos de "Is Friday the 13th Bad for Your Health?" de Scanlon *et al*., *British Medical Journal*, vol. 307, según se lista en el registro de información de los recursos en línea de los conjuntos de datos; **P451**, basado en "An Analysis of Factors That Contribute to the Efficacy of Hypnotic Analgesia", de Price y Barber, *Journal of Abnormal Psychology*, vol. 96, núm. 1; **P459**, basado 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; **P460**, basado en 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; **P460**, basado en datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4; **P461**, basado en datos de un sondeo de Harris; **P462**, basado en datos de "Consistency of Blood Pressure Differences Between the Left and Right Arms", de Eguchi *et al*., *Archives of Internal Medicine*, vol. 167; **P463**, basado en datos de "Essential Hypertension: Effect of an Oral Inhibitor of Angiotensin-Converting Enzyme", de MacGregor *et al*., *British Medical Journal*, vol. 2; **P463**, basado en datos de "Eyewitness Memory of Police Trainees for Realistic Role Plays", de Yuille *et al*., *Journal of Applied Psychology*, vol. 79, núm. 6

Capítulo 10

P486, basado en datos de "The Trouble with QSAR (or How I Learned to Stop Worrying and Embrace Fallacy)" de Stephen Johnson, *Journal of Chemical Information and Modeling*, vol. 48, núm. 1; **P487**, basado en datos de *The Song of Insects* de George W. Pierce, Harvard University Press; **P487**, basado en "Mass Estimation of Weddell Seals Using Techniques of Photogrammetry", de R. Garrott de la Montana State University; **P488**, basado en "Mass Estimation of Weddell

Seals Using Techniques of Photogrammetry", de R. Garrott de la Montana State University; **P491**, creado mediante el uso del software JMP®. Copyright ©SAS Institute Inc. SAS y todos los demás productos y servicios con el nombre de SAS Institute Inc. son marcas registradas o marcas del SAS Institute Inc., Cary, NC, USA; **P510-511** basado en datos de la *Poughkeepsie Journal*; **P530**, basado en datos de la Motion Picture Association of America; **P531**, basado en datos de Janssen Pharmaceutical Products, L.P.; **varias páginas**, reproducción cortesía de International Business Machines Corporation, © SPSS, Inc., e IBM Company; **varias páginas**, cortesía de XLSTAT™. Usado con autorización; **varias páginas**, pantallas de Texas Instruments. Cortesía de Texas Instruments

Capítulo 11

P537, basado en datos del California Department of Public Health; **P543-544** basado en datos de "Participation in Cancer Clinical Trials", de Murthy, Krumholz y Gross, *Journal of the American Medical Association*, vol. 291, núm. 22; **P544**, basado en datos de "Efficacy of Hip Protector to Prevent Hip Fracture in Nursing Home Residents", de Kiel *et al*., *Journal of the American Medical Association*, vol. 298, núm. 4; **P549**, basado en datos de "Surgery Unfounded for Tarsal Navicular Stress Fracture", de Bruce Jancin, *Internal Medicine News*, vol. 42, núm. 14; **P556**, basado en datos de "Splinting *vs* Surgery in the Treatment of Carpal Tunnel Syndrome", de Gerritsen *et al*., *Journal of the American Medical Association*, vol. 288, núm. 10; **P557**, basado en datos de "Texting While Driving and Other Risky Motor Vehicle Behaviors Among U.S. High School Students" de O'Malley, Shults e Eaton, *Pediatrics*, vol 131, núm. 6; **P557**, basado en "The Denomination Effect", de Priya Raghubir y Joydeep Srivastava, *Journal of Consumer Research*, vol. 36; **P558**, basado en datos de "Action Bias Among Elite Soccer Goalkeepers: The Case of Penalty Kicks", de Bar-Eli *et al*., *Journal of Economic Psychology*, vol. 28, núm. 5; **P558**, basado en datos de "What Kinds of People Do Not Use Seat Belts?" de Helsing y Comstock, *American Journal of Public Health*, vol. 67, núm. 11; **P558**, basado en datos de "An Evaluation of Echinacea Angustifolia in Experimental Rhinovirus Infections", de Turner *et al*., *New England Journal of Medicine*, vol. 353, núm. 4; **P559**, basado en datos de "Motorcycle Rider Conspicuity and Crash Related Injury: Case-Control Study", de Wells *et al*., *BMJ USA*, vol. 4; **P559**, basado en datos de "I Hear You Knocking But You Can't Come In", de Fitzgerald y Fuller, *Sociological Methods and Research*, vol. 11, núm. 1; **P559**, basado en datos de "Predicting Professional Sports Game Outcomes from Intermediate Game Scores", de Copper, DeNeve, y Mosteller, *Chance*, vol. 5, núm. 3-4; **P561**, basado en datos del Insurance Institute for Highway Safety; **P561**, basado en datos de "Neuropsychological and Renal Effects of Dental Amalgam in Children", de Bellinger *et al*., *Journal of the American Medical Association*, vol. 295, núm. 15; **P562**, basado en "The Denomination Effect" de Priya Raghubir y Joydeep Srivastava, *Journal of Consumer Research*, vol. 36; **P562**, basado en datos de Parke-Davis; **P563**, basado en datos del Insurance Institute for Highway Safety; **P563**, basado en datos del estudio antropométrico de Gordon, Churchill, *et al*.

Capítulo 12

P566, creado mediante el uso del software JMP®. Copyright ©SAS Institute Inc. SAS y todos los demás productos y servicios con el nombre de SAS Institute Inc. Son marcas registradas o marcas del SAS Institute Inc., Cary, NC, USA; **P566-567**, basado en datos de "Neuropsychological Dysfunction in Children with Chronic Low-Level Lead Absorption", de P. J. Landrigan, R. H. Whitworth, R. W. Baloh, N. W. Staehling, W. F. Barthel y B. F. Rosenblum, *Lancet*, vol. 1, artículo 7909

Capítulo 13

P604, basado en datos de Genetics & IVF Institute; **P611**, basado en datos de Physicians Insurers Association of America; **P624**, basado en datos de "An Unexpected Rise in Strontium-90 in U.S. Deciduous Teeth in the 1990s", de Mangano *et al*., *Science of the Total Environment*; **P624**, basado en datos de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, vol. 52, núm. 4; **P635**, basado en datos de *Consumer Reports*; **P639**, basado en "Mass Estimation of Weddell Seals Using Techniques of Photogrammetry", de R. Garrott de la Montana State University; **P640**, basado en datos de *The Song of Insects* de George W. Pierce, Harvard University Press; **P648**, basado en datos de "Job Rated Stress Score" de CareerCast.com; **P650**, basado en datos de Quest Diagnostics

Capítulo 14

P665, basado en datos del U.S. Department of the Interior; **P674**, basado en un estudio de MerckManuals.com; **P675**, basado en datos de "Does It Pay to Plead Guilty? Differential Sentencing and the Functioning of the Criminal Courts", de Brereton y Casper, *Law and Society Review*, vol. 16, núm. 1; **P660**, *Fuente:* Adaptado de *ASTM Manual on the Presentation of Data and Control Chart Analysis*, © 1976 ASTM, pp. 134-136. Reproducido con autorización de la American Society for Testing and Materials

Apéndice A

P2, Mosteller, *Probability with Statistical Applications*, 2a. Ed., © 1970. Reimpreso y reproducido electrónicamente con autorización de Pearson Education, Inc. Upper Saddle River, New Jersey; **P6**, *Handbook of Statistical Tables*, Addison-Wesley Pub. Co.; **P7-8**, basado en datos de Maxine Merrington y Catherine M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution," *Biometrika 33* (1943): 80-84; **P13**, basado en datos de *Some Rapid Approximate Statistical Procedures*, Copyright ©1949, 1964 Lederle Laboratories Division of American Cyanamid Company; **P15**, a partir de *Tables for Testing Randomness of Grouping in a Sequence of Alternatives* de Frieda S. Swed y C. Eisenhart en *The Annals of Mathematical Statistics*, vol. 14, núm. 1, pp. 66-87. Copyright © Institute of Mathematical Statistics

Apéndice B

P2, Rohren, Brenda, "Estimation of Stature from Foot and Shoe Length: Applications in Forensic Science". Copyright © 2006. Reproducido con autorización del autor; **P4**, basado en datos de *Biostatistical Analysis*, 4a. edición © 1999, de Jerrold Zar, Prentice Hall, Inc., Upper Saddle River, New Jersey, y "Distribution of Sums of Squares of Rank Differences to Small Numbers with Individuals", *The Annals of Mathematical Statistics*, vol. 9, núm. 2; **P4**, Mario, F. Triola; **P5**, Mario, F. Triola; **P6**, Mario, F. Triola; **P6**, Mario, F. Triola; **P6**, Mario, F. Triola; **P8**, Mario, F. Triola; **P9**, basado en datos de Andrew Gelman, Jennifer Hill, 2007, "Replication data for Data Analysis Using Regression Multilevel/Hierarchical Models", http://hdl.handle.net/1902.1/10285; **P9**, con base en datos replicados de *Data Analysis Using Regression and Multilevel/Hierarchical Models*, de Andrew Gelman y Jennifer Hill, Cambridge University Press; **P12**, Mario, F. Triola; **P14**, Basado en datos de CNN

ÍNDICE

La estadística es una ciencia relacionada con todos los aspectos de la vida. Se encuentra en encuestas, estudios clínicos, negocios, análisis de datos y hasta en robótica.

Esta relación es la que el autor, Mario F. Triola, considera el eje principal para presentar en este libro el estudio de la estadística con aplicaciones a una cantidad sin precedente de casos reales actuales.

El texto se encuentra organizado en capítulos individuales que se pueden revisar de forma consecutiva o, si lo prefiere el profesor, pueden adaptarse a un orden especial para cualquier curso.

Todos los capítulos tienen conjuntos de datos, ejemplos y ejercicios suficientes e interesantes para cubrir todos los temas de estadística de un primer curso; además, abordan temas de interés tanto de conocimiento general como de la tecnología más moderna.

Entre las novedades sobresalientes de la presente edición destacan las siguientes:

- Más del 90% de los datos que se utilizan en los ejercicios y problemas son reales y están actualizados.
- Objetivos y problemas al inicio de cada capítulo como motivación para el estudio del tema presentado.
- Énfasis en la interpretación de resultados y ejemplos trabajados.
- Notas de interés al margen del contenido.
- Resúmenes de conceptos y procedimientos, y revisión de los temas en cada capítulo.
- Atractivas secciones, como *Su turno*, para que el estudiante practique lo aprendido resolviendo ejercicios.
- Una sección denominada *Centro de tecnología* para que el lector utilice la tecnología como herramienta de apoyo en el análisis estadístico.
- Gran cantidad de apoyos extra, como el software STATDISK proporcionado por el autor en su sitio web.
- Conjuntos de datos y ejercicios de aplicación en Excel, Minitab, JMP, SPSS, STATDISK, Statcrunch y las calculadoras TI-83/84 Plus.

Para más información sobre este libro y los recursos adicionales visite:
www.pearsonenespañol.com/triola

www.pearsonenespañol.com

ISBN 978-607-32-4378-0

90000

9 786073 243780