8

Métodos de muestreo y teorema del límite central

OBJETIVOS

Al concluir el capítulo, será capaz de:

1. Explicar la razón por la que una muestra es con frecuencia la única forma viable para conocer algo sobre una población.

2. Describir métodos para seleccionar una muestra.

3. Definir y construir una distribución muestral de la media de la muestra.

4. Comprender y explicar el *teorema del límite central.*

5. Aplicar el teorema del límite central para calcular probabilidades de seleccionar posibles medias muestrales de una población específica.

El informe anual de Nike indica que el estadounidense promedio compra 6.5 pares de zapatos deportivos al año. Suponga que la desviación estándar de la población es de 2.1 y que se analizará una muestra de 81 clientes el siguiente año. ¿Cuál es el error estándar de la media en este experimento? (Véase el objetivo 5 y el ejercicio 45.)

Estadística en acción

Con el importante papel que desempeña la estadística inferencial en todas las ramas de la ciencia, es ya una necesidad la disponibilidad de fuentes copiosas de números aleatorios. En 1927 se publicó el primer libro de números aleatorios, con 41 600 dígitos aleatorios, generados por L. Tippett. En 1938, R. A. Fisher y E. Yates publicaron 15 000 dígitos aleatorios, generados con dos barajas. En 1955, RAND Corporation publicó un millón de dígitos aleatorios, generados por pulsos de frecuencia aleatorios de una ruleta electrónica. Para 1970, las aplicaciones del muestreo requerían miles de millones de números aleatorios. Desde entonces se han creado métodos para generar, con ayuda de computadoras, dígitos "casi" aleatorios, por lo que se les llama *seudoaleatorios*. Aún es motivo de debate la pregunta acerca de si un programa de computadora sirve para generar números aleatorios que de verdad sean aleatorios.

Introducción

De los capítulos 2 a 4 se hizo hincapié en las técnicas para describir datos. Con el fin de ilustrar dichas técnicas, se organizaron los precios de 80 vehículos vendidos el mes pasado en Whitner Autoplex en una distribución de frecuencias para calcular las diversas medidas de ubicación y dispersión. Dichas medidas, como la media y la desviación estándar, describen el precio de venta habitual y la dispersión de los precios de venta. En estos capítulos se destacó la descripción de la condición de los datos: se describió algo que ya había sucedido.

El capítulo 5 comienza a establecer el fundamento de la inferencia estadística con el estudio de la probabilidad. Recuerde que, en la inferencia estadística, el objetivo es determinar algo sobre una *población* a partir sólo de una *muestra*. La población es todo el grupo de individuos u objetos en estudio, y la muestra es una parte o subconjunto de dicha población. El capítulo 6 amplía los conceptos de probabilidad al describir tres distribuciones de probabilidad discreta: binomial, hipergeométrica y de Poisson. El capítulo 7 describe la distribución de probabilidad uniforme y la distribución de probabilidad normal. Ambas son distribuciones continuas. Las distribuciones de probabilidad abarcan todos los posibles resultados de un experimento, así como la probabilidad asociada con cada resultado. Mediante las distribuciones de probabilidad se evaluó la probabilidad de que ocurra algo en el futuro.

Este capítulo inicia el estudio del muestreo, herramienta para inferir algo sobre una población. Primero se analizan los métodos para seleccionar una muestra de una población. Después se construye una distribución de la media de la muestra para entender la forma como las medias muestrales tienden a acumularse en torno a la media de la población. Por último, se demuestra que, para cualquier población, la forma de esta distribución de muestreo tiende a seguir la distribución de probabilidad normal.

Métodos de muestreo

Ya se mencionó en el capítulo 1 que el propósito de la estadística inferencial consiste en determinar algo sobre una población a partir de una muestra. Una muestra es una porción o parte de la población de interés. En muchos casos, el muestreo resulta más accesible que el estudio de toda la población. En esta sección se explican las razones principales para muestrear y, enseguida, diversos métodos para elegir una muestra.

Razones para muestrear

Cuando se estudian las características de una población, existen diversas razones prácticas para preferir la selección de porciones o muestras de una población para observar y medir. He aquí algunas razones para muestrear:

1. **Establecer contacto con toda la población requeriría mucho tiempo.** Un candidato para un puesto federal quizá desee determinar las posibilidades que tiene de resultar electo. Una encuesta de muestreo en la que se utiliza el personal y las entrevistas de campo convencionales de una empresa especializada en encuestas tardaría de uno o dos días. Con el mismo personal y los mismos entrevistadores, y laborando siete días a la semana, se requerirían 200 años para ponerse en contacto con toda la población en edad de votar. Aunque fuera posible reunir a un numeroso equipo de encuestadores, quizá no valdría la pena entrar en contacto con todos los votantes.
2. **El costo de estudiar todos los elementos de una población resultaría prohibitivo.** Las organizaciones que realizan encuestas de opinión pública y pruebas entre consumidores, como Gallup Polls y Roper ASW, normalmente entran en contacto con menos de 2 000 de las casi 60 millones de familias en Estados Unidos. Una organización que entrevista a consumidores en panel cobra cerca de $40 000 por enviar muestras por correo y tabular las respuestas con el fin de probar un producto (como un cereal para el desayuno, alimento para gato o algún perfume). La misma prueba del producto con los 60 millones de familias tendría un costo de aproximadamente $1 000 000 000.

3. **Es imposible verificar de manera física todos los elementos de la población.** Algunas poblaciones son infinitas. Sería imposible verificar toda el agua del lago Erie en lo que se refiere a niveles de bacterias, así que se eligen muestras en diversos lugares. Las poblaciones de peces, aves, serpientes o mosquitos son grandes, y se desplazan, nacen y mueren continuamente. En lugar de intentar contar todos los patos que hay en Canadá o todos los peces del lago Pontchartrain, se hacen aproximaciones mediante diversas técnicas: se cuentan todos los patos que hay en un estanque, capturados al azar, se revisan las cestas de los cazadores o se colocan redes en lugares predeterminados en el lago.
4. **Algunas pruebas son de naturaleza destructiva.** Si los catadores de vino de Sutter Home Winery, California, se bebieran todo el vino para evaluar la vendimia, acabarían con la cosecha y no quedaría nada disponible para la venta. En el área de producción industrial: las placas de acero, cables y productos similares deben contar con una resistencia mínima a la tensión. Para cerciorarse de que el producto satisface la norma mínima, el departamento de control de calidad elige una muestra de la producción actual. Cada pieza se somete a tensión hasta que se rompe y se registra el punto de ruptura (medido en libras por pulgada cuadrada). Es obvio que si se sometieran todos los cables o todas las placas a pruebas de resistencia a la tensión no habría productos disponibles para vender u utilizar. Por la misma razón, Kodak selecciona sólo una muestra de película fotográfica y la somete a pruebas para determinar la calidad de todos los rollos que se producen; y sólo unas cuantas semillas se someten a pruebas de germinación en Burpee, antes de la temporada de siembra.

5. **Los resultados de la muestra son adecuados.** Aunque se contara con recursos suficientes, es difícil que la precisión de una muestra de 100% —toda la población— resulte esencial en la mayoría de los problemas. Por ejemplo, el gobierno federal utiliza una muestra de tiendas de comestibles distribuidas en Estados Unidos para determinar el índice mensual de precios de los alimentos. Los precios del pan, frijol, leche y otros productos de primera necesidad se incluyen en el índice. Resulta poco probable que la inclusión de todas las tiendas de comestibles de Estados Unidos influya significativamente en el índice, pues los precios de la leche, el pan y otros productos de primera necesidad no varían más de unos cuantos centavos de una cadena de tiendas a otra.

Muestreo aleatorio simple

El tipo de muestreo más común es el **muestreo aleatorio simple**.

MUESTREO ALEATORIO SIMPLE Muestra seleccionada de manera que cada elemento o individuo de la población tenga las mismas posibilidades de que se le incluya.

Una tabla de números aleatorios es una forma eficiente de seleccionar a los miembros de una muestra.

Para ejemplificar el muestreo aleatorio simple y la selección, suponga que una población consta de 845 empleados de Nitra Industries. Se va a elegir una muestra de 52 empleados de dicha población. Una forma de asegurarse de que todos los empleados de la población tienen las mismas posibilidades de que se les elija consiste en escribir primero el nombre de cada empleado en un papel y depositarlos todos en una caja. Después de mezclarlos , se efectúa la primera selección tomando un papel de la caja sin mirarlo. Se repite este proceso hasta terminar de elegir la muestra de 52 empleados.

Un método más conveniente de seleccionar una muestra aleatoria consiste en utilizar un número de identificación por cada empleado y una **tabla de números aleatorios** como la del apéndice B.6. Como su nombre lo indica, estos números se generaron mediante un proceso aleatorio (en este caso, con una computadora).

La probabilidad de 0, 1, 2, …, 9 es la misma para cada dígito de un número. Por consiguiente, la probabilidad de que se seleccione el empleado 011 es la misma que para los empleados 722 o 382. Al emplear números aleatorios para seleccionar empleados, se elimina la influencia o sesgo del proceso de selección.

En la siguiente ilustración aparece parte de una tabla de números aleatorios. Para seleccionar una muestra de empleados, elija primero un punto de partida en la tabla; cualquier punto sirve. Ahora suponga que el reloj marca las 3:04. Puede observar la tercera columna y enseguida desplazarse hacia abajo hasta el cuarto conjunto de números. El número es 03759. Como sólo hay 845 empleados, utilizará los tres primeros dígitos de un número aleatorio de cinco dígitos. Por tanto, 037 es el número del primer empleado que se convertirá en miembro de la muestra. Otra forma de elegir el punto de partida consiste en cerrar los ojos y señalar un número de la tabla. Para continuar, puede desplazarse en cualquier sentido. Suponga que lo hace hacia la derecha. Los primeros tres dígitos del número a la derecha de 03759 son 447, el número del siguiente empleado seleccionado para integrar la muestra. El siguiente número de tres dígitos a la derecha es 961. Omita 961, pues sólo hay 845 empleados. Continúe hacia la derecha y seleccione al empleado 784; después el 189 y así en lo sucesivo.

Estadística en acción

¿Es discriminación sacar ventaja del físico? Antes de contestar, considere un artículo reciente que apareció en *Personnel Journal*. Estos hallazgos indican que los hombres y mujeres atractivos ganan alrededor de 5% más que los que tienen una apariencia promedio, quienes, a su vez, ganan 5% más que sus compañeros poco agraciados. Esto se aplica tanto en hombres como en mujeres. También es cierto en el caso de gran variedad de ocupaciones, desde la construcción hasta la reparación de automóviles y los empleos de telemarketing, ocupaciones para las que, según se cree, la apariencia no es importante.

50525	57454	28455	68226	34656	38884	39018
72507	53380	53827	42486	54465	71819	91199
34986	74297	00144	38676	89967	98869	39744
68851	27305	03759	44723	96108	78489	18910
06738	62879	03910	17350	49169	03850	18910
11448	10734	05837	24397	10420	16712	94496
		Punto de partida	Segundo empleado		Tercer empleado	Cuarto empleado

La mayoría de los paquetes de software contienen una rutina para seleccionar una muestra aleatoria simple. En el siguiente ejemplo se emplea el sistema Excel para elegir una muestra aleatoria.

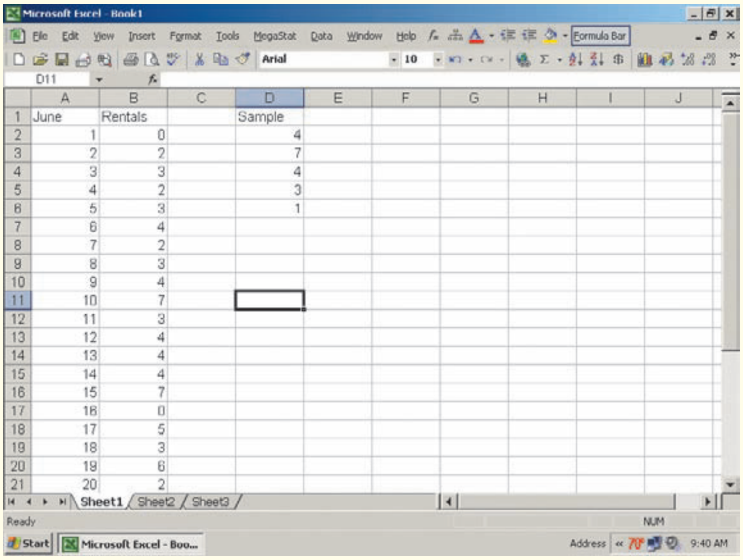
Ejemplo

Jane y Joe Millar administran el Foxtrot Inn, una pensión donde dan alojamiento y desayuno, localizada en Tryon, Carolina del Norte. Se rentan ocho habitaciones en esta pensión. A continuación aparece el número de estas ocho habitaciones rentadas diariamente durante junio de 2006. Utilice Excel para seleccionar una muestra de cinco noches de junio.

Junio	Habitaciones en renta	Junio	Habitaciones en renta	Junio	Habitaciones en renta
1	0	11	3	21	3
2	2	12	4	22	2
3	3	13	4	23	3
4	2	14	4	24	6
5	3	15	7	25	0
6	4	16	0	26	4
7	2	17	5	27	1
8	3	18	3	28	1
9	4	19	6	29	3
10	7	20	2	30	3

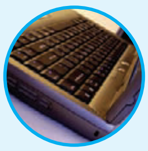
Solución

Excel seleccionará la muestra aleatoria y arrojará los resultados. En la primera fecha muestreada había cuatro habitaciones rentadas de las ocho. En la segunda fecha muestreada de junio, se rentaron siete de las ocho habitaciones. La información aparece en la columna D de la hoja de cálculo de Excel. Los pasos en Excel se incluyen en la sección **Comandos de software**, al final del capítulo. El sistema Excel lleva a cabo el muestreo *con* reemplazo. Esto significa que tal vez el mismo día aparezca más de una vez en una muestra.



Excel

Autoevaluación 8.1



La siguiente lista incluye a los estudiantes que se matricularon en un curso de introducción a la estadística administrativa. Se elige al azar a tres estudiantes, a quienes se formulan varias preguntas relacionadas con el contenido del curso y el método de enseñanza.

a) Se escriben a mano los números 00 a 45 en papeletas y se colocan en un recipiente. Los tres números seleccionados son 31, 7 y 25. ¿Qué estudiantes se van a incluir en la muestra?

b) Ahora utilice la tabla de dígitos aleatorios, apéndice B.6, para seleccionar su propia muestra.

c) ¿Qué haría si localizara el número 59 en la tabla de números aleatorios?

CSPM 264 01 BUSINESS & ECONOMIC STAT
8:00 AM 9:40 AM MW ST 118 LIND D

RANDOM NUMBER	NAME	CLASS RANK	RANDOM NUMBER	NAME	CLASS RANK
00	ANDERSON, RAYMOND	SO	23	MEDLEY, CHERYL ANN	SO
01	ANGER, CHERYL RENEE	SO	24	MITCHELL, GREG R	FR
02	BALL, CLAIRE JEANETTE	FR	25	MOLTER, KRISTI MARIE	SO
03	BERRY, CHRISTOPHER G	FR	26	MULCAHY, STEPHEN ROBERT	SO
04	BOBAK, JAMES PATRICK	SO	27	NICHOLAS, ROBERT CHARLES	JR
05	BRIGHT, M. STARR	JR	28	NICKENS, VIRGINIA	SO
06	CHONTOS, PAUL JOSEPH	SO	29	PENNYWITT, SEAN PATRICK	SO
07	DETLEY, BRIAN HANS	JR	30	POTEAU, KRIS E	JR
08	DUDAS, VIOLA	SO	31	PRICE, MARY LYNETTE	SO
09	DULBS, RICHARD ZALFA	JR	32	RISTAS, JAMES	SR
10	EDINGER, SUSAN KEE	SR	33	SAGER, ANNE MARIE	SO
11	FINK, FRANK JAMES	SR	34	SMILLIE, HEATHER MICHELLE	SO
12	FRANCIS, JAMES P	JR	35	SNYDER, LEISHA KAY	SR
13	GAGHEN, PAMELA LYNN	JR	36	STAHL, MARIA TASHERY	SO
14	GOULD, ROBYN KAY	SO	37	ST. JOHN, AMY J	SO
15	GROSENBACHER, SCOTT ALAN	SO	38	STURDEVANT, RICHARD K	SO
16	HEETFIELD, DIANE MARIE	SO	39	SWETYE, LYNN MICHELE	SO
17	KABAT, JAMES DAVID	JR	40	WALASINSKI, MICHAEL	SO
18	KEMP, LISA ADRIANE	FR	41	WALKER, DIANE ELAINE	SO
19	KILLION, MICHELLE A	SO	42	WARNOCK, JENNIFER MARY	SO
20	KOPERSKI, MARY ELLEN	SO	43	WILLIAMS, WENDY A	SO
21	KOPP, BRIDGETTE ANN	SO	44	YAP, HOCK BAN	SO
22	LEHMANN, KRISTINA MARIE	JR	45	YODER, ARLAN JAY	JR

Estadística en acción

Los métodos de muestreo aleatorio y sin sesgos son muy importantes para realizar inferencias estadísticas válidas. En 1936 se efectuó un sondeo de opinión para predecir el resultado de la carrera presidencial entre Franklin Roosevelt y Alfred Landon. Se enviaron diez millones de papeletas en forma de postales retornables gratuitas a domicilios tomados de directorios telefónicos y registros de automóviles. Se contestó una alta proporción de papeletas, con 59% en favor de Landon y 41% de Roosevelt. El día de la elección, Roosevelt ganó con 61% de los votos. Landon obtuvo 39%. Sin duda, a mediados de la década de 1930, la gente que tenía teléfono y automóvil no era representativa de los votantes estadounidenses.

Muestreo aleatorio sistemático

El procedimiento de muestreo aleatorio simple resulta complicado en algunos estudios. Por ejemplo, suponga que la división de ventas de Computer Graphic, Inc., necesita calcular rápidamente el ingreso medio en dólares por venta del mes pasado. La división encontró que se registraron 2 000 ventas y se almacenaron en cajones de archivo, y se decidió seleccionar 100 recibos para calcular el ingreso medio en dólares. El muestreo aleatorio simple requiere que la numeración de cada recibo antes de utilizar la tabla de números aleatorios para seleccionar los 100 recibos. Dicho proceso de numeración puede tardar mucho tiempo. En su lugar, es posible aplicar el **muestreo aleatorio sistemático**.

MUESTREO ALEATORIO SISTEMÁTICO Se selecciona un punto aleatorio de inicio y posteriormente se elige cada *k*-ésimo miembro de la población.

Primero se calcula *k*, que es el resultado de dividir el tamaño de la población entre el tamaño de la muestra. En el caso de Computers Graphic, Inc., seleccione cada vigésimo recibo (2 000/100) de los cajones del archivo; al hacerlo evita el proceso de numeración. Si *k* no es un número entero, hay que redondearlo.

En la selección del primer recibo emplee el muestreo aleatorio simple. Por ejemplo, seleccionará un número de la tabla de números aleatorios entre 1 y *k*, en este caso, 20. Suponga que el número aleatorio resultó ser 18. Entonces, a partir del recibo 18, se seleccionará cada vigésimo recibo (18, 38, 58, etc.) como muestra.

Antes de aplicar el muestreo aleatorio sistemático, debe observar con cuidado el orden físico de la población. Cuando el orden físico se relaciona con la característica de la población, no debe aplicar el muestreo aleatorio sistemático. Por ejemplo, si los recibos se archivan en orden creciente de ventas, el muestreo aleatorio sistemático no garantiza una muestra aleatoria. Debe aplicar otros métodos de muestreo.

Muestreo aleatorio estratificado

Cuando una población se divide en grupos a partir de ciertas características, se aplica el **muestreo aleatorio estratificado** con el fin de garantizar el hecho de que cada grupo se encuentre representado en la muestra. A los grupos también se les denomina **estratos**. Por ejemplo, los estudiantes universitarios se pueden agrupar en estudiantes de tiempo completo o de medio tiempo, por sexo, masculino o femenino, tradicionales o no tradicionales. Una vez definidos los estratos, se aplica el muestreo aleatorio simple en cada grupo o estrato con el fin de formar la muestra.

MUESTRA ALEATORIA ESTRATIFICADA Una población se divide en subgrupos, denominados *estratos*, y se selecciona al azar una muestra de cada estrato.

Por ejemplo, puede estudiar los gastos en publicidad de las 352 empresas más grandes de Estados Unidos. Suponga que el objetivo del estudio consiste en determinar si las empresas con altos rendimientos sobre el capital (una media de rentabilidad) gastan en publicidad la mayor parte del dinero ganado en ventas que las empresas con un registro de bajo rendimiento o déficit. Para asegurar que la muestra sea una representación imparcial de las 352 empresas, éstas se agrupan de acuerdo con su rendimiento porcentual sobre el capital. La tabla 8.1 incluye los estratos y las frecuencias relativas. Si aplicara el muestreo aleatorio simple, observe que las empresas del tercero y cuarto estratos tienen una probabilidad alta de que se les seleccione (0.87), mientras que las empresas de los demás estratos tienen pocas probabilidades de que se les seleccione (0.13). Podría no seleccionar ninguna de las empresas que aparecen en los estratos 1 o 5 *sencillamente por azar.* No obstante, el muestreo aleatorio estratificado garantizará que por lo menos una empresa de los estratos 1 o 5 aparezca en la muestra. Considere una selección de 50 compañías para llevar a cabo un estudio minucioso. Entonces se seleccionará de forma aleatoria 1 (0.02 × 50) empresa del estrato 1; 5 (0.10 × 50), del estrato 2, etc. En este caso, el número de empresas en cada estrato es proporcional a la frecuencia relativa del estrato en la población. El muestreo estratificado ofrece la ventaja de que, en algunos casos, refleja con mayor fidelidad las características de la población que el muestreo aleatorio simple o el muestreo aleatorio sistemático.

TABLA 8.1 **Número seleccionado para una muestra aleatoria estratificada proporcional**

Estrato	Probabilidad (recuperación de capital)	Número de empresas	Frecuencia relativa	Número muestreado
1	30% y más	8	0.02	1*
2	De 20% a 30%	35	0.10	5*
3	De 10% a 20%	189	0.54	27
4	De 0% a 10%	115	0.33	16
5	Déficit	5	0.01	1
Total		352	1.00	50

*0.02 de 50 = 1, 0.10 de 50 = 5, etcétera.

Muestreo por conglomerados

Otro tipo común de muestreo es el **muestreo por conglomerados**. Éste se emplea a menudo para reducir el costo de muestrear una población dispersa en cierta área geográfica.

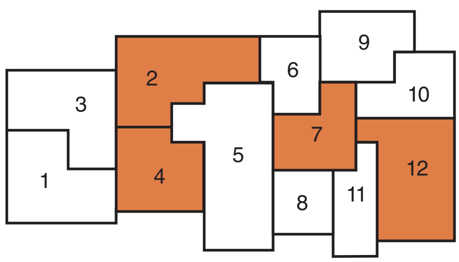
MUESTREO ACUMULADO Una población se divide en conglomerados a partir de los límites naturales geográficos o de otra clase. A continuación se seleccionan los conglomerados al azar y se toma una muestra de forma aleatoria con elementos de cada grupo.

Suponga que desea determinar la opinión de los residentes de algún estado con referencia a las políticas federales y estatales de protección ambiental. Seleccionar una muestra aleatoria de residentes y ponerse en contacto con cada persona requeriría mucho tiempo y resultaría muy costoso. Sería mejor aplicar el muestreo por conglomerados y subdividir el estado en pequeñas unidades: condados o regiones. Con frecuencia, se les conoce como *unidades primarias.*

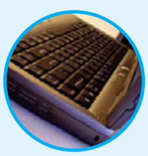
Suponga que dividió el estado en 12 unidades primarias, seleccionó al azar cuatro regiones, 2, 7, 4 y 12, y concentró su atención en estas unidades primarias. Usted puede tomar una muestra aleatoria de los residentes de cada una de estas regiones y entrevistarse con ellos (observe que se trata de una combinación de un muestreo por conglomerados y un muestreo aleatorio simple).

Muchos métodos más de muestreo

El estudio de los métodos de muestreo de las secciones anteriores no incluye todos los métodos de muestreo disponibles para el investigador. Si usted emprendiera un proyecto de investigación importante de marketing, finanzas, contabilidad u otras áreas, necesitaría consultar libros dedicados exclusivamente a la teoría del muestreo y al diseño de muestras.



Autoevaluación 8.2



Consulte la autoevaluación 8.1 y la lista de alumnos de la página 264. Suponga que en un muestreo aleatorio sistemático se elegirá a cada noveno estudiante de la clase. Al principio se elige al azar al cuarto estudiante de la lista. Dicho estudiante es el número 03. Recuerde que los números aleatorios comienzan con 00, entonces, ¿qué estudiantes se elegirán como miembros de la muestra?

Ejercicios

1. La siguiente lista incluye las tiendas de Marco's Pizza en el condado de Lucas. También se indica si la tienda es propiedad de alguna corporación (C) o del administrador (A). Se seleccionará e inspeccionará una muestra de cuatro establecimientos en relación con la conveniencia para el cliente, la seguridad, la higiene y otras características.

Número de identificación	Dirección	Tipo	Número de identificación	Dirección	Tipo
00	2607 Starr Av	C	12	2040 Ottawa River Rd	C
01	309 W Alexis Rd	C	13	2116 N Reynolds Rd	C
02	2652 W Central Av	C	14	3678 Rugby Dr	C
03	630 Dixie Hwy	A	15	1419 South Av	C
04	3510 Dorr St	C	16	1234 W Sylvania Av	C
05	5055 Glendale Av	C	17	4624 Woodville Rd	A
06	3382 Lagrange St	A	18	5155 S Main	A
07	2525 W Laskey Rd	C	19	106 E Airport Hwy	C
08	303 Louisiana Av	C	20	6725 W Central	A
09	149 Main St	C	21	4252 Monroe	C
10	835 S McCord Rd	A	22	2036 Woodville Rd	C
11	3501 Monroe St	A	23	1316 Michigan Av	A

a) Los números aleatorios seleccionados son 08, 18, 11, 02, 41 y 54. ¿Qué tiendas se eligieron?
b) Utilice la tabla de números aleatorios para seleccionar su propia muestra de establecimientos.
c) Una muestra consta de cada séptimo establecimiento. El número 03 es el punto de partida. ¿Qué establecimientos se incluirán en la muestra?
d) Suponga que una muestra consta de tres establecimientos, de los cuales dos son propiedad corporativa y uno del administrador. Seleccione una muestra adecuada.

2. La siguiente lista incluye hospitales localizados en las regiones de Cincinnati (Ohio) y la región norte de Kentucky. También indica si se trata de un hospital general médico o quirúrgico (M/Q), o de especialidades (E). Interesa calcular el promedio de enfermeras que trabaja medio tiempo en los hospitales del área.
a) Se va a seleccionar de forma aleatoria una muestra de cinco hospitales. Los números aleatorios son 09, 16, 00, 49, 54, 12 y 04. ¿Qué hospitales se incluyen en la muestra?
b) Utilice una tabla de números aleatorios para formar su propia muestra de cinco hospitales.

Número de identificación	Nombre	Dirección	Tipo	Número de identificación	Nombre	Dirección	Tipo
00	Bethesda North	10500 Montgomery Cincinnati, Ohio 45242	M/Q	10	Christ Hospital	2139 Auburn Avenue Cincinnati, Ohio 45219	M/Q
01	Ft. Hamilton-Hughes	630 Eaton Avenue Hamilton, Ohio 45013	M/Q	11	Deaconess Hospital	311 Straight Street Cincinnati, Ohio 45219	M/Q
02	Jewish Hospital-Kenwood	4700 East Galbraith Rd. Cincinnati, Ohio 45236	M/Q	12	Good Samaritan Hospital	375 Dixmyth Avenue Cincinnati, Ohio 45220	M/Q
03	Mercy Hospital-Fairfield	3000 Mack Road Fairfield, Ohio 45014	M/Q	13	Jewish Hospital	3200 Burnet Avenue Cincinnati, Ohio 45229	M/Q
04	Mercy Hospital-Hamilton	100 Riverfront Plaza Hamilton, Ohio 45011	M/Q	14	University Hospital	234 Goodman Street Cincinnati, Ohio 45267	M/Q
05	Middletown Regional	105 McKnight Drive Middletown, Ohio 45044	M/Q	15	Providence Hospital	2446 Kipling Avenue Cincinnati, Ohio 45239	M/Q
06	Clermont Mercy Hospital	3000 Hospital Drive Batavia, Ohio 45103	M/Q	16	St. Francis-St. George Hospital	3131 Queen City Avenue Cincinnati, Ohio 45238	M/Q
07	Mercy Hospital-Anderson	7500 State Road Cincinnati, Ohio 45255	M/Q	17	St. Elizabeth Medical Center, North Unit	401 E. 20th Street Covington, Kentucky 41014	M/Q
08	Bethesda Oak Hospital	619 Oak Street Cincinnati, Ohio 45206	M/Q	18	St. Elizabeth Medical Center, South Unit	One Medical Village Edgewood, Kentucky 41017	M/Q
09	Children's Hospital Medical Center	3333 Burnet Avenue Cincinnati, Ohio 45229	M/Q	19	St. Luke's Hospital West	7380 Turfway Drive Florence, Kentucky 41075	M/Q

Número de identificación	Nombre	Dirección	Tipo	Número de identificación	Nombre	Dirección	Tipo
20	St. Luke's Hospital East	85 North Grand Avenue Ft. Thomas, Kentucky 41042	M/Q	25	Drake Center Rehab—Long Term	151 W. Galbraith Road Cincinnati, Ohio 45216	E
21	Care Unit Hospital	3156 Glenmore Avenue Cincinnati, Ohio 45211	E	26	No. Kentucky Rehab Hospital—Short Term	201 Medical Village Edgewood, Kentucky	E
22	Emerson Behavioral Science	2446 Kipling Avenue Cincinnati, Ohio 45239	E	27	Shriners Burns Institute	3229 Burnet Avenue Cincinnati, Ohio 45229	E
23	Pauline Warfield Lewis Center for Psychiatric Treat.	1101 Summit Road Cincinnati, Ohio 45237	E	28	VA Medical Center Cincinnati, Ohio 45220	3200 Vine	E
24	Children's Psychiatric No. Kentucky	502 Farrell Drive Covington, Kentucky 41011	E				

c) Una muestra incluirá cada quinto establecimiento. Se selecciona 02 como punto de partida. ¿Qué hospitales se incluirán en la muestra?

d) Una muestra consta de cuatro hospitales médicos o quirúrgicos y un hospital de especialidades. Seleccione una muestra adecuada.

3. A continuación aparece una lista de los 35 miembros de la Metro Toledo Automobile Dealers Association. Se desea calcular el ingreso medio de los departamentos de servicios de los distribuidores.

Número de identificación	Distribuidor	Número de identificación	Distribuidor	Número de identificación	Distribuidor
00	Dave White Acura	11	Thayer Chevrolet/Toyota	23	Kistler Ford, Inc.
01	Autofair Nissan	12	Spurgeon Chevrolet Motor Sales, Inc.	24	Lexus of Toledo
02	Autofair Toyota-Suzuki	13	Dunn Chevrolet	25	Mathews Ford Oregon, Inc.
03	George Ball's Buick GMC Truck	14	Don Scott Chevrolet-Pontiac	26	Northtowne Chevrolet
04	Yark Automotive Group	15	Dave White Chevrolet Co.	27	Quality Ford Sales, Inc.
05	Bob Schmidt Chevrolet	16	Dick Wilson Pontiac	28	Rouen Chrysler Jeep Eagle
06	Bowling Green Lincoln Mercury Jeep Eagle	17	Doyle Pontiac Buick	29	Saturn of Toledo
		18	Franklin Park Lincoln Mercury	30	Ed Schmidt Pontiac Jeep Eagle
07	Brondes Ford	19	Genoa Motors	31	Southside Lincoln Mercury
08	Brown Honda	20	Great Lakes Ford Nissan	32	Valiton Chrysler
09	Brown Mazda	21	Grogan Towne Chrysler	33	Vin Divers
10	Charlie's Dodge	22	Hatfield Motor Sales	34	Whitman Ford

a) Seleccione una muestra aleatoria de cinco distribuidores. Los números aleatorios son: 05, 20, 59, 21, 31, 28, 49, 38, 66, 08, 29 y 02. ¿Qué distribuidores se van a incluir en la muestra?

b) Utilice la tabla de números aleatorios para seleccionar su propia muestra de cinco distribuidores.

c) Una muestra constará de cada séptimo distribuidor. El número 04 se selecciona como punto de partida. ¿Qué distribuidores se incluyen en la muestra?

4. Enseguida se enumera a los 27 agentes de seguros de Nationwide Insurance en el área metropolitana de Toledo, Ohio. Se desea calcular el promedio de años que han laborado en Nationwide.

Número de identificación	Agente	Número de identificación	Agente	Número de identificación	Agente
00	**Bly Scott** 3332 W Laskey Rd	10	**Heini Bernie** 7110 W Centra	19	**Riker Craig** 2621 N Reynolds Rd
01	**Coyle Mike** 5432 W Central Av	11	**Hinckley Dave** 14 N Holland Sylvania Rd	20	**Schwab Dave** 572 W Dussel Dr
02	**Denker Brett** 7445 Airport Hwy			21	**Seibert John H** 201 S Main
03	**Denker Rollie** 7445 Airport Hwy	12	**Joehlin Bob** 3358 Navarre Av	22	**Smithers Bob** 229 Superior St
04	**Farley Ron** 1837 W Alexis Rd	13	**Keisser David** 3030 W Sylvania Av	23	**Smithers Jerry** 229 Superior St
05	**George Mark** 7247 W Central Av	14	**Keisser Keith** 5902 Sylvania Av	24	**Wright Steve** 105 S Third St
06	**Gibellato Carlo** 6616 Monroe St	15	**Lawrence Grant** 342 W Dussel Dr	25	**Wood Tom** 112 Louisiana Av
07	**Glemser Cathy** 5602 Woodville Rd	16	**Miller Ken** 2427 Woodville Rd	26	**Yoder Scott** 6 Willoughby Av
08	**Green Mike** 4149 Holland Sylvania Rd	17	**O'Donnell Jim** 7247 W Central Av		
09	**Harris Ev** 2026 Albon Rd	18	**Priest Harvey** 5113 N Summit St		

a) Seleccione una muestra aleatoria de cuatro agentes. Los números aleatorios son: 02, 59, 51, 25, 14, 29, 77, 69 y 18. ¿Qué distribuidores se incluirán en la muestra?
b) Utilice la tabla de números aleatorios para seleccionar su propia muestra de cuatro agentes.
c) Una muestra consta de cada séptimo distribuidor. El número 04 se selecciona como punto de partida. ¿Qué agentes se incluirán en la muestra?

“Error” de muestreo

En la sección anterior se estudiaron métodos de muestreo útiles para seleccionar una muestra que constituya una representación imparcial o sin sesgos de la población. Es importante señalar que, en cada método, la selección de cualquier posible muestra de determinado tamaño de una población tiene una posibilidad o probabilidad conocidas. Ésta constituye otra forma de describir un método de muestreo sin sesgo.

Las muestras se emplean para determinar características de la población. Por ejemplo, con la media de una muestra se calcula la media de la población. No obstante, como la muestra forma parte o es una porción representativa de la población, es poco probable que la media de la muestra sea *exactamente igual* a la media poblacional. Asimismo, es poco probable que la desviación estándar de la muestra sea *exactamente igual* a la desviación estándar de la población. Por tanto, puede esperar una diferencia entre un *estadístico de la muestra* y el *parámetro de la población* correspondiente. Esta diferencia recibe el nombre de **error de muestreo**.

ERROR DE MUESTREO Diferencia entre el estadístico de una muestra y el parámetro de la población correspondiente.

El siguiente ejemplo aclara el concepto de error de muestreo.

Ejemplo

Revise el ejemplo anterior de la página 263, en el que estudió el número de habitaciones rentadas en Foxtrot Inn, en Tryon, Carolina del Norte. La población se refiere al número de habitaciones rentadas cada uno de los 30 días de junio de 2006. Determine la media de la población. Utilice Excel u otro software de estadística para seleccionar tres muestras aleatorias de cinco días. Calcule la media de cada muestra y compárela con la media poblacional. ¿Cuál es el error de muestreo en cada caso?

Solución

Durante el mes se rentaron un total de 94 habitaciones. Así, la media de las unidades rentadas por noche es de 3.13. Ésta es la media de la población. Este valor se designa con la letra griega μ.

$$\mu = \frac{\Sigma X}{N} = \frac{0+2+3+\ldots+3}{30} = \frac{94}{30} = 3.13$$

La primera muestra aleatoria de cinco noches dio como resultado el siguiente número de habitaciones rentadas: 4, 7, 4, 3 y 1. La media de esta muestra de cinco noches es de 3.8 habitaciones, que se representa como $\bar{X}_1$. La barra sobre la X recuerda que se trata de una media muestral, y el subíndice 1 indica que se trata de la media de la primera muestra.

$$\bar{X}_1 = \frac{\Sigma X}{n} = \frac{4+7+4+3+1}{5} = \frac{19}{5} = 3.80$$

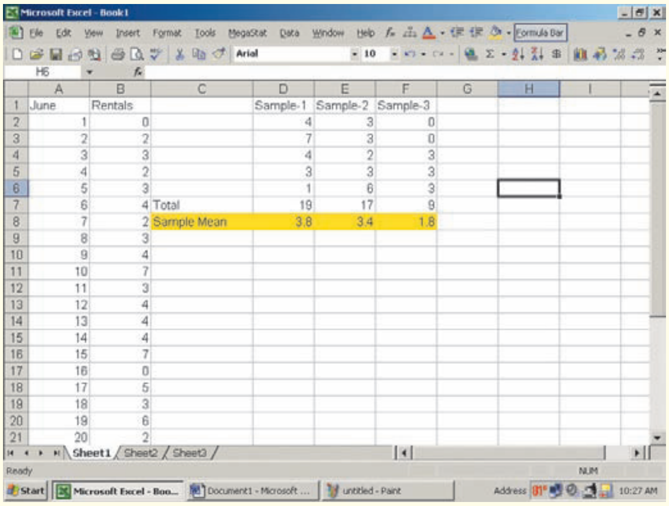
El error de muestreo para la primera muestra es la diferencia entre la media poblacional (3.13) y la media muestral (3.80). De ahí que el error muestral sea ($\bar{X}_1 - \mu = 3.80 - 3.13 = 0.67$). La segunda muestra aleatoria de cinco días de la población de 30 días de junio arrojó el siguiente número de habitaciones rentadas: 3, 3, 2, 3 y 6. La media de estos cinco valores es de 3.4, que se calcula de la siguiente manera:

$$\bar{X}_2 = \frac{\Sigma X}{n} = \frac{3+3+2+3+6}{5} = 3.4$$

El error de muestreo es ($\bar{X}_2 - \mu = 3.4 - 3.13 = 0.27$).

En la tercera muestra aleatoria, la media fue de 1.8, y el error de muestro fue de –1.33.

Cada una de estas diferencias, 0.67, 0.27 y –1.33, representa el error de muestreo cometido al calcular la media de la población. A veces estos errores son valores positivos, lo cual indica que la media muestral sobreexcedió la media poblacional; otras veces son valores negativos, lo cual indica que la media muestral resultó inferior a la media poblacional.



June	Rentals		Sample-1	Sample-2	Sample-3
1	0		4	3	0
2	2		7	3	0
3	3		4	2	3
4	2		3	3	3
5	3		1	6	3
6	4	Total	19	17	9
7	2	Sample Mean	3.8	3.4	1.8
8	3				
9	4				
10	7				
11	3				
12	4				
13	4				
14	4				
15	7				
16	0				
17	5				
18	3				
19	6				
20	2				

En este caso, con una población de 30 valores y muestras de 5 valores, existe una gran cantidad de posibles muestras, 142 506, para ser exactos. Para calcular este valor se aplica la fórmula de las combinaciones 5.10, de la página 168. Cada una de las 142 506 diferentes muestras cuenta con las mismas posibilidades de que se le seleccione. Cada muestra puede tener una media muestral diferente y, por consiguiente, un error de muestreo distinto. El valor del error de muestreo se basa en el valor particular de las 142 506 posibles muestras seleccionadas. Por consiguiente, los errores de muestreo son aleatorios y se presentan al azar. Si determinara la suma de estos errores de muestreo en una gran cantidad de muestras, el resultado se aproximaría mucho a cero. Sucede así porque la media de la muestra constituye un estimador sin sesgo de la media de la población.

Distribución muestral de la media

Ahora que aparece la posibilidad de que se presente un error de muestreo cuando se emplean los resultados del muestreo para aproximar un parámetro poblacional, ¿cómo hacer un pronóstico preciso relacionado con el posible éxito de un nuevo dentífrico u otro producto sobre la única base de los resultados del muestreo? ¿Cómo puede el departamento de control de calidad, de una compañía de producción en serie, enviar un cargamento de microchips a partir de una muestra de 10 chips? ¿Cómo pueden las organizaciones electorales de CNN-*USA Today* o ABC News-*Washington Post* hacer pronósticos precisos sobre la elección presidencial con base en una muestra de 1 200 electores registrados de una población de cerca de 90 millones? Para responder estas preguntas, primero hay que precisar el concepto de *distribución muestral de la media*.

Las medias muestrales varían de muestra en muestra

Las medias muestrales del ejemplo anterior varían de una muestra a la siguiente. La media de la primera muestra de 5 días fue de 3.80 habitaciones, y la media de la segunda muestra fue de 3.40 habitaciones. La media poblacional fue de 3.13 habitaciones. Si organiza las medias de todas las muestras posibles de 5 días en una distribución de probabilidad, el resultado recibe el nombre de **distribución muestral de la media.**

DISTRIBUCIÓN MUESTRAL DE LA MEDIA Distribución de probabilidad de todas las posibles medias de las muestras de un determinado tamaño muestra de la población.

El siguiente ejemplo ilustra la construcción de una distribución muestral de la media.

Ejemplo

Tartus Industries cuenta con siete empleados de producción (a quienes se les considera la población). En la tabla 8.2 se incluyen los ingresos por hora de cada empleado.

TABLA 8.2 Ingresos por hora de empleados de producción en Tartus Industries

Empleado	Ingresos por hora	Empleado	Ingresos por hora
Joe	$7	Jan	$7
Sam	7	Art	8
Sue	8	Ted	9
Bob	8		

1. ¿Cuál es la media de la población?
2. ¿Cuál es la distribución muestral de la media para muestras de tamaño 2?
3. ¿Cuál es la media de la distribución muestral de la media?
4. ¿Qué observaciones es posible hacer sobre la población y la distribución muestral de la media?

Solución

He aquí las respuestas.

1. La media de la población es de $7.71, que se determina de la siguiente manera:

$$\mu = \frac{\Sigma X}{N} = \frac{\$7 + \$7 + \$8 + \$8 + \$7 + \$8 + \$9}{7} = \$7.71$$

Identifique la media de la población por medio de la letra griega μ. En los capítulos 1, 3 y 4 se convino en identificar los parámetros poblacionales con letras griegas.

2. Para obtener la distribución muestral de la media se seleccionó, sin reemplazos de la población, todas las muestras posibles de tamaño 2 y se calcularon las medias de cada muestra. Hay 21 posibles muestras, que se calcularon con la fórmula (5.10) de la página 168.

$$_NC_n = \frac{N!}{n!(N-n)!} = \frac{7!}{2!(7-2)!} = 21$$

Aquí, $N = 7$ es el número de elementos de la población, y $n = 2$, el número de elementos de la muestra.

En la tabla 8.3 se ilustran las 21 medias muestrales de todas las muestras posibles de tamaño 2 que pueden tomarse de la población. Estas 21 muestras se utilizan para construir una distribución de probabilidad, que es la distribución muestral de la media, la cual se resume en la tabla 8.4.

TABLA 8.3 Medias muestrales de todas las posibles muestras de 2 empleados

Muestra	Empleados	Ingresos por hora	Suma	Media	Muestra	Empleados	Ingresos por hora	Suma	Media
1	Joe, Sam	$7, $7	$14	$7.00	12	Sue, Bob	$8, $8	$16	$8.00
2	Joe, Sue	7, 8	15	7.50	13	Sue, Jan	8, 7	15	7.50
3	Joe, Bob	7, 8	15	7.50	14	Sue, Art	8, 8	16	8.00
4	Joe, Jan	7, 7	14	7.00	15	Sue, Ted	8, 9	17	8.50
5	Joe, Art	7, 8	15	7.50	16	Bob, Jan	8, 7	15	7.50
6	Joe, Ted	7, 9	16	8.00	17	Bob, Art	8, 8	16	8.00
7	Sam, Sue	7, 8	15	7.50	18	Bob, Ted	8, 9	17	8.50
8	Sam, Bob	7, 8	15	7.50	19	Jan, Art	7, 8	15	7.50
9	Sam, Jan	7, 7	14	7.00	20	Jan, Ted	7, 9	16	8.00
10	Sam, Art	7, 8	15	7.50	21	Art, Ted	8, 9	17	8.50
11	Sam, Ted	7, 9	16	8.00					

La media de la población es igual a la media de las medias muestrales

TABLA 8.4 Distribución muestral de la media para $n = 2$

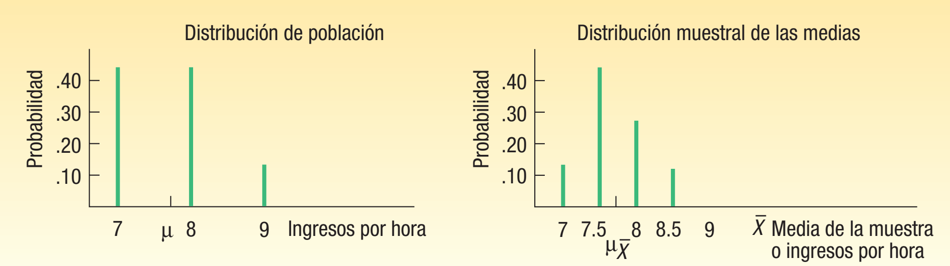
Media muestral	Número de medias	Probabilidad
\$7.00	3	.1429
7.50	9	.4285
8.00	6	.2857
8.50	3	.1429
	21	1.0000

3. La media de la distribución muestral de la media se obtiene al sumar las medias muestrales y dividir la suma entre el número de muestras. La media de todas las medias muestrales se representa mediante $\mu_{\bar{X}}$. La μ recuerda que se trata de un valor poblacional, pues tomó en cuenta todas las muestras posibles. El subíndice $\bar{X}$ indica que se trata de la distribución muestral de la media.

$$\mu_{\bar{X}} = \frac{\text{Suma de todas las medias muestrales}}{\text{Total de muestras}} = \frac{\$7.00 + \$7.50 + \cdots + \$8.50}{21}$$

$$= \frac{\$162}{21} = \$7.71$$

4. Consulte la gráfica 8.1, donde aparecen las dos distribuciones poblacionales y la distribución muestral de la media. Caben las siguientes observaciones:
   a) La media de la distribución muestral de la media (\$7.71) es igual a la media de la población: $\mu = \mu_{\bar{X}}$.
   b) La dispersión de la distribución muestral de las medias es menor que la dispersión de los valores de población. La media de las muestras varía de \$7.00 a \$8.50, mientras que los valores de población varían de \$7.00 a \$9.00. Observe que, conforme se incrementa el tamaño de la muestra, se reduce la dispersión de la distribución muestral de las medias.
   c) La forma de la distribución muestral de la media y la forma de la distribución de frecuencias de los valores de población son diferentes. La distribución muestral de las medias tiende a adoptar más forma de campana y a aproximarse a la distribución de probabilidad normal.



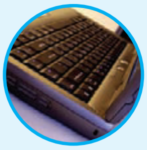
GRÁFICA 8.1 Distribución de los valores de población y distribución muestral de las medias

En resumen, tome todas las posibles muestras aleatorias de una población y calcule un estadístico muestral (la media de los ingresos percibidos) para cada una. Este ejemplo ilustra las importantes relaciones entre la distribución poblacional y la distribución muestral de la media:

1. La media de las medias de las muestras es exactamente igual a la media de la población.
2. La dispersión de la distribución muestral de la media es más estrecha que la distribución poblacional.
3. La distribución muestral de la media suele tener forma de campana y se aproxima a la distribución de probabilidad normal.

Dada una distribución de probabilidad normal o de forma de campana, se aplican los conceptos del capítulo 7 para determinar la probabilidad de seleccionar una muestra con una media muestral específica. En la siguiente sección resalta la importancia del tamaño de una muestra en relación con la distribución muestral de la media.

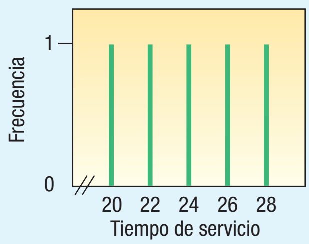
Autoevaluación 8.3



Los tiempos de servicio de los ejecutivos que laboran en Standard Chemicals son los siguientes:

Nombre	Años
Señor Snow	20
Señora Tolson	22
Señor Kraft	26
Señora Irwin	24
Señor Jones	28

a) De acuerdo con la fórmula de las combinaciones, ¿cuántas muestras de tamaño 2 son posibles?
b) Elabore una lista de todas las muestras posibles de 2 ejecutivos de la población y calcule las medias.
c) Organice las medias en una distribución muestral.
d) Compare la media poblacional y la media de las medias de las muestras.
e) Compare la dispersión en la población con la dispersión de la distribución muestral de la media.
f) A continuación se muestra una gráfica con los valores de la población. ¿Tienen los valores de población una distribución normal (en forma de campana)?



g) ¿Comienza la distribución muestral de la media que se calculó en el inciso *c*) a indicar una tendencia a adoptar forma de campana?

Ejercicios

5. Una población consta de los siguientes cuatro valores: 12, 12, 14 y 16.
a) Enumere todas las muestras de tamaño 2 y calcule la media de cada muestra.
b) Calcule la media de la distribución muestral de la media y la media de la población. Compare los dos valores.
c) Compare la dispersión en la población con la de las medias de las muestras.

6. Una población consta de los siguientes cinco valores: 2, 2, 4, 4 y 8.
a) Enumere todas las muestras de tamaño 2 y calcule la media de cada muestra.
b) Calcule la media de la distribución muestral de las medias y la media de la población. Compare los dos valores.
c) Compare la dispersión en la población con la de las medias de las muestras.

7. Una población consta de los siguientes cinco valores: 12, 12, 14, 15 y 20.
a) Enumere todas las muestras de tamaño 3 y calcule la media de cada muestra.
b) Calcule la media de la distribución muestral de las medias y la media de la población. Compare los dos valores.
c) Compare la dispersión en la población con la de las medias de las muestras.

8. Una población consta de los siguientes cinco valores: 0, 0, 1, 3 y 6.
a) Enumere todas las muestras de tamaño 3 y calcule la media de cada muestra.
b) Calcule la media de la distribución muestral de las medias y la media de la población. Compare los dos valores.
c) Compare la dispersión en la población con la de las medias de las muestras.

9. En el despacho de abogados Tybo and Associates, hay seis socios. En la siguiente tabla se incluye el número de casos que en realidad atendió cada socio en los tribunales durante el mes pasado.

Socio	Número de casos
Ruud	3
Wu	6
Sass	3
Flores	3
Wilhelms	0
Schueller	1

a) ¿Cuántas muestras de 3 son posibles?
b) Enumere todas las posibles muestras de 3 y calcule el número medio de casos en cada muestra.
c) Compare la media de la distribución muestral de las medias con la de la media poblacional.
d) En una gráfica similar a la 8.1, compare la dispersión en la población con la de las medias muestrales.

10. Hay cinco vendedores en Mid-Motors Ford. Los cinco representantes de ventas y el número de automóviles que vendieron la semana pasada son los siguientes:

Representantes de ventas	Autos vendidos
Peter Hankish	8
Connie Stallter	6
Juan Lopez	4
Ted Barnes	10
Peggy Chu	6

a) ¿Cuántas muestras de 2 son posibles?
b) Enumere todas las posibles muestras de 2 y calcule la media de casos en cada muestra.
c) Compare la media de la distribución muestral de la media con la de la media poblacional.
d) En una gráfica similar a la 8.1, compare la dispersión en la población con la de la media de la muestra.

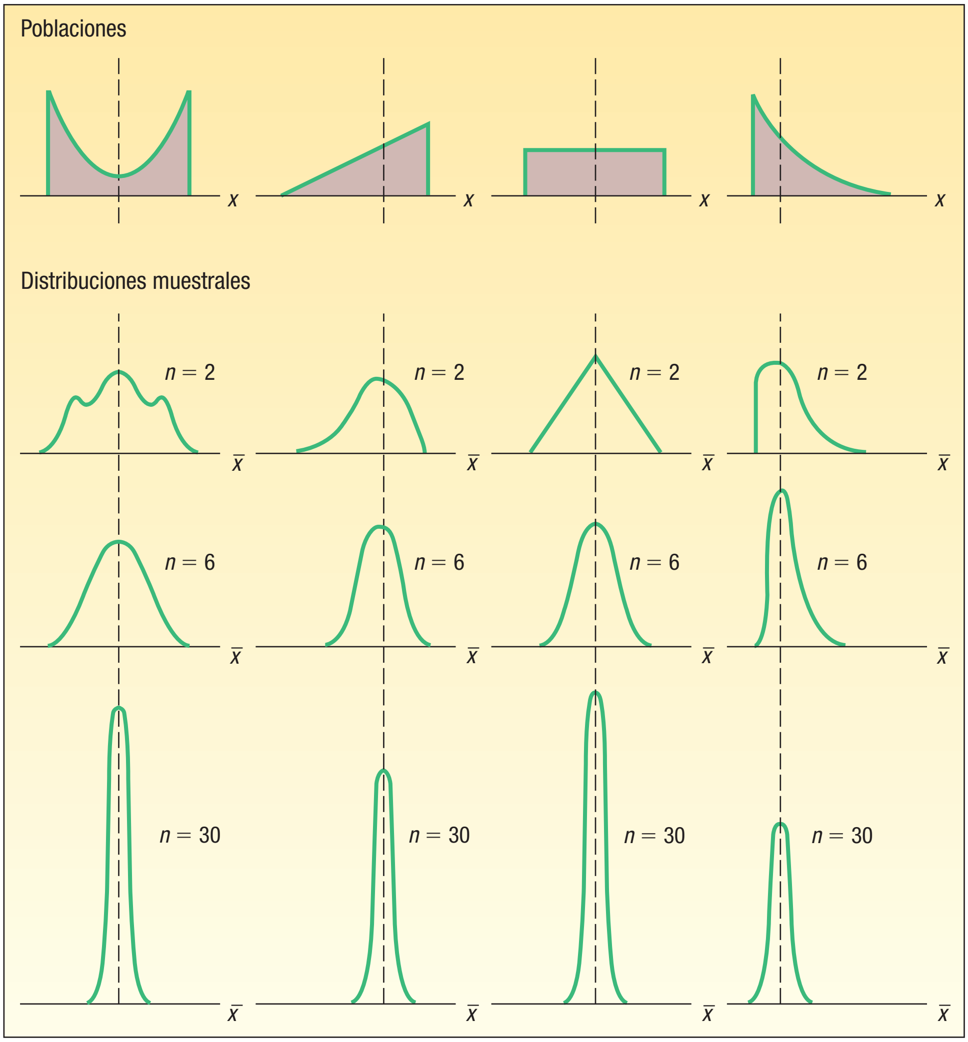
Teorema del límite central

En esta sección se estudia el **teorema del límite central**. Su aplicación a la distribución muestral de medias, en la sección anterior, permite utilizar la distribución de probabilidad normal para crear intervalos de confianza para la media poblacional (que se describe en el capítulo 9) y llevar a cabo pruebas de hipótesis (descritas en el capítulo 10). El teorema del límite central hace hincapié en que, en el caso de muestras aleatorias grandes, la forma de la distribución muestral de la media se aproxima a la distribución de probabilidad normal. La aproximación es más exacta en el caso de muestras grandes que en el de muestras pequeñas. Ésta es una de las conclusiones más útiles de la estadística. Permite razonar sobre la distribución de las medias muestrales sin ninguna información acerca de la forma de la distribución de población de la que se toma la muestra. En otras palabras, el teorema del límite central se cumple en el caso de todas las distribuciones.

En seguida aparece el enunciado formal del teorema del límite central.

TEOREMA DEL LÍMITE CENTRAL Si todas las muestras de un tamaño en particular se seleccionan de cualquier población, la distribución muestral de la media se aproxima a una distribución normal. Esta aproximación mejora con muestras más grandes.

Si la población obedece a una distribución normal, entonces, en el caso de cualquier tamaño de muestra, la distribución muestral de las medias también será de naturaleza normal. Si la distribución poblacional es simétrica (pero no normal), se verá que la forma normal de la distribución muestral de las medias se presenta con muestras tan pequeñas como 10. Por otra parte, si se comienza con una distribución sesgada o con colas gruesas, quizá se requieran muestras de 30 o más para observar la característica de normalidad. Este concepto se resume en la gráfica 8.2 para diversas formas de



GRÁFICA 8.2 Resultados del teorema del límite central para diversas poblaciones

población. Observe la convergencia hacia una distribución normal sin importar la forma de la distribución de población. La mayoría de los especialistas en estadística consideran que una muestra de 30 o mayor es lo bastante grande para aplicar el teorema del límite central.

La idea de que la distribución muestral de las medias de una población que no es normal converge hacia la normalidad se ilustra en las gráficas 8.3, 8.4 y 8.5. En breve se analiza este ejemplo con más detalles, pero la gráfica 8.3 es la gráfica de una distribución de probabilidad discreta con sesgo positivo. Hay varias posibles muestras de 5 que puede seleccionar de esta población. Suponga que selecciona al azar 25 muestras de tamaño 5 cada una y calcula la media de cada muestra. Estos resultados se muestran en la gráfica 8.4. Observe que la forma de la distribución muestral de las medias cambió la forma de la población original aunque sólo seleccionó 25 de las diversas posibles muestras. En otras palabras, eligió 25 muestras al azar de tamaño 5 de una población positivamente sesgada, y encontró que la distribución muestral de las medias cambió en lo que se refiere a la forma de la población. A medida que toma muestras más grandes, es decir, $n = 20$ en lugar de $n = 5$, la distribución muestral de las medias se aproximará a la distribución normal. La gráfica 8.5 muestra los resultados de 25 muestras aleatorias de 20 observaciones cada una tomadas de la misma población. Note la clara tendencia hacia la distribución de probabilidad normal. Ésta es la esencia del teorema del límite central. El siguiente ejemplo pondrá de relieve esta condición.

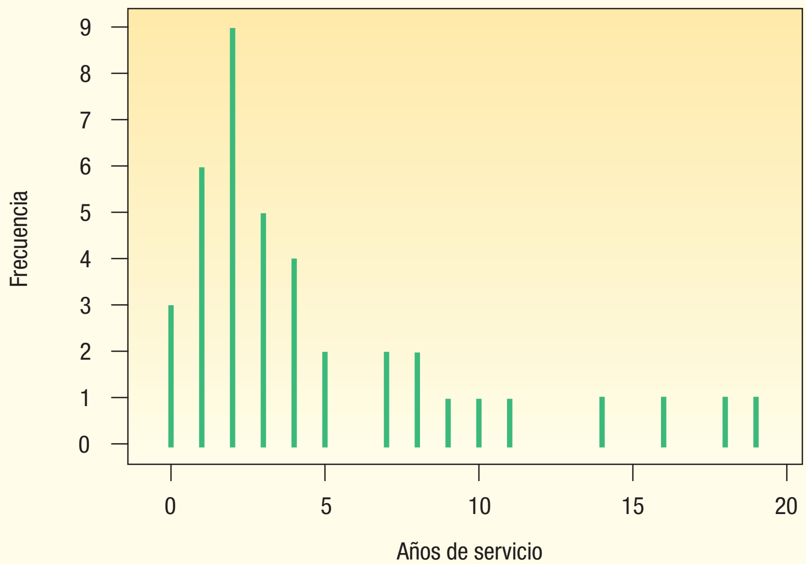
Ejemplo

Ed Spence dio inicio a su negocio de engranes hace 20 años. El negocio creció a lo largo del tiempo y ahora cuenta con 40 empleados. Spence Sprockets, Inc., encara algunas decisiones importantes relacionadas con la atención médica de sus empleados. Antes de tomar una decisión definitiva sobre el programa de atención médica que va a comprar, Ed decide formar un comité de cinco empleados. Se pedirá al comité que estudie el tema del cuidado de la salud y haga alguna recomendación sobre el plan que mejor convenga a los empleados. Ed cree que el punto de vista de los empleados más recientes en relación con el cuidado de la salud difiere de los empleados con más experiencia. Si Ed selecciona al azar este comité, ¿qué puede esperar en términos del promedio de años que llevan con Spence Sprockets los miembros del comité? ¿Cuál es la forma de la distribución de años de experiencia de todos los empleados (la población) en comparación con la forma de la distribución muestral de las medias? Los tiempos de servicio (redondeados al año inmediato) de los 40 empleados que actualmente están en nómina en Spence Sprockers, Inc., son los siguientes:

11	4	18	2	1	2	0	2	2	4
3	4	1	2	2	3	3	19	8	3
7	1	0	2	7	0	4	5	1	14
16	8	9	1	1	2	5	10	2	3

Solución

La gráfica 8.3 muestra la distribución de los años de experiencia de la población de 40 empleados actuales. La distribución de tiempos de servicio tiene un sesgo positivo, pues unos cuantos empleados han laborado en Spence Sprockets por un periodo extenso. En específico, seis empleados han laborado en la compañía 10 años o más. Sin embargo, como el negocio creció, el número de empleados se incrementó en los últimos cinco años. De los 40 empleados, 18 han laborado en la compañía dos años o menos.



GRÁFICA 8.3 Tiempo de servicio en Spence Sprockets, Inc., de los empleados

Considere el primero de los problemas de Ed Spence. A él le gustaría formar un comité de cinco empleados con el objeto de que estudien la cuestión del cuidado de la salud y sugieran el tipo de cobertura de gastos médicos más adecuada para la mayoría de los trabajadores. ¿Cómo elegiría al comité? Si lo selecciona al azar, ¿qué puede esperar respecto del tiempo medio de servicio de quienes forman parte del comité?

Para comenzar, Ed anota el tiempo de servicio de cada uno de los 40 empleados en papeles y los coloca en una gorra de béisbol. Después los revuelve y selecciona al azar cinco de ellos. Los tiempos de servicio de estos cinco empleados son: 1, 9, 0, 19 y 14 años. Por tanto, el tiempo medio de servicio de estos cinco empleados muestreados es de 8.60 años. ¿Cómo se compara este resultado con la media de la población? En este momento, Ed no conoce la media de la población, aunque el número de empleados de la población es de sólo 40, así que decide calcular la media del tiempo de servicio de *todos* sus empleados. Ésta es de 4.8 años, que se determina al sumar los tiempos de servicio de *todos* los empleados y dividir el total entre 40.

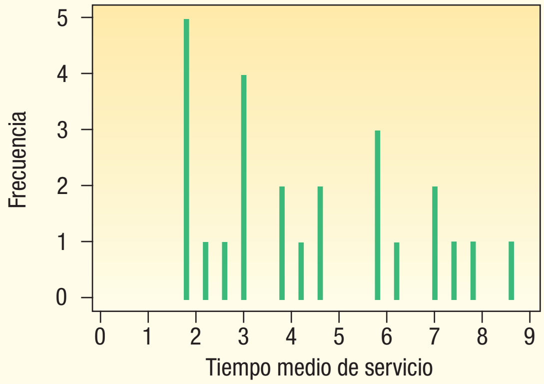
$$\mu = \frac{11+4+18+\cdots+2+3}{40} = 4.80$$

La diferencia entre la media de la muestra ($\overline{X}$) y la media de la población (μ) recibe el nombre de **error de muestreo**. En otras palabras, la diferencia de 3.80 años entre la media poblacional de 4.80 y la media muestral de 8.60 es el error de muestreo. Éste se debe al azar. Por consiguiente, si Ed selecciona a estos cinco empleados para formar el comité, el tiempo medio de servicio de éstos sería mayor que el de la media de la población.

¿Qué sucedería si Ed colocara de nuevo los papeles en la gorra y tomara otra muestra? ¿Esperaría que la media de esta segunda muestra fuera exactamente la misma que la anterior? Suponga que selecciona otra muestra de cinco empleados y encuentra que los tiempos de servicio de esta muestra son de 7, 4, 4, 1 y 3. La media muestral es de 3.80 años. El resultado de seleccionar 25 muestras de cinco empleados cada una se muestra en la tabla 8.5 y en la gráfica 8.4. En realidad hay 658 008 posibles muestras de 5 tomas de la población de 40 empleados, las cuales se determinan con la fórmula de las combinaciones (5.10) con 40 objetos tomados de 5 en 5. Observe la diferencia de forma de las distribuciones poblacional y mues-

TABLA 8.5 Veinticinco muestras aleatorias de cinco empleados

Muestra de identificación			Datos de la muestra			Media muestral
A	1	9	0	19	14	8.6
B	7	4	4	1	3	3.8
C	8	19	8	2	1	7.6
D	4	18	2	0	11	7.0
E	4	2	4	7	18	7.0
F	1	2	0	3	2	1.6
G	2	3	2	0	2	1.8
H	11	2	9	2	4	5.6
I	9	0	4	2	7	4.4
J	1	1	1	11	1	3.0
K	2	0	0	10	2	2.8
L	0	2	3	2	16	4.6
M	2	3	1	1	1	1.6
N	3	7	3	4	3	4.0
O	1	2	3	1	4	2.2
P	19	0	1	3	8	6.2
Q	5	1	7	14	9	7.2
R	5	4	2	3	4	3.6
S	14	5	2	2	5	5.6
T	2	1	1	4	7	3.0
U	3	7	1	2	1	2.8
V	0	1	5	1	2	1.8
W	0	3	19	4	2	5.6
X	4	2	3	4	0	2.6
Y	1	1	2	3	2	1.8



GRÁFICA 8.4 Histograma de tiempos de servicio medios para 25 muestras de cinco empleados

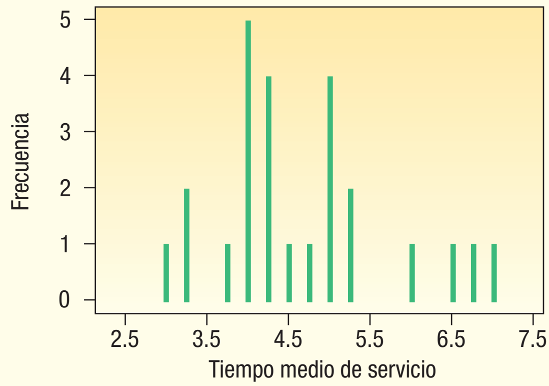
tral de medias. La población de tiempos de servicio de los empleados (gráfica 8.3) tiene un sesgo positivo, y la distribución de estas 25 medias muestrales no refleja el mismo sesgo positivo. También existe una diferencia en el rango de las medias muestrales en comparación con el rango de la población. La población varía de 0 a 19 años, mientras que las medias muestrales varían de 1.6 a 8.6 años.

La tabla 8.6 contiene los resultados de seleccionar 25 muestras de 20 empleados cada una y el cálculo de las medias muestrales. Estas medias muestrales aparecen en la gráfica 8.5. Compare la forma de esta distribución con la población (gráfica 8.3) y con la distribución muestral de medias si la muestra es de $n = 5$ (gráfica 8.4). Observe dos importantes características:

TABLA 8.6 Muestras aleatorias y medias muestrales de 25 muestras de 20 empleados de Spence Sprockets, Inc.

Número de muestra	Datos de la muestra (tiempo de servicio)																				Media muestral
A	3	8	3	0	2	1	2	3	11	5	1	3	4	2	7	1	1	2	4	16	3.95
B	2	3	8	2	1	5	2	0	3	1	0	7	1	4	3	11	4	4	3	1	3.25
C	14	5	0	3	2	14	11	9	2	2	1	2	19	1	0	1	4	2	19	8	5.95
D	9	2	1	1	4	10	0	8	4	3	2	1	0	8	1	14	5	10	1	3	4.35
E	18	1	2	2	4	3	2	8	2	1	0	19	4	19	0	1	4	0	3	14	5.35
F	10	4	4	18	3	3	1	0	0	2	2	4	7	10	2	0	3	4	2	1	4.00
G	5	7	11	8	11	18	1	1	16	2	2	16	2	3	2	16	2	2	2	4	6.55
H	3	0	2	0	5	4	5	3	8	3	2	5	1	1	2	9	8	3	16	5	4.25
I	0	0	18	2	1	7	4	1	3	0	3	2	11	7	2	8	5	1	2	3	4.00
J	2	7	2	4	1	3	3	2	5	10	0	1	1	2	9	3	2	19	3	2	4.05
K	7	4	5	3	3	0	18	2	0	4	2	7	2	7	4	2	10	1	1	2	4.20
L	0	3	10	5	9	2	1	4	1	2	1	8	18	1	4	3	3	2	0	4	4.05
M	4	1	2	1	7	3	9	14	8	19	4	4	1	2	0	3	1	2	1	2	4.40
N	3	16	1	2	4	4	4	2	1	5	2	3	5	3	4	7	16	1	11	1	4.75
O	2	19	2	0	2	2	16	2	3	11	9	2	8	0	8	2	7	3	2	2	5.10
P	2	18	16	5	2	2	19	0	1	2	11	4	2	2	1	4	2	0	4	3	5.00
Q	3	2	3	11	10	1	1	5	19	16	7	10	3	1	1	1	2	2	3	1	5.10
R	2	3	1	2	7	4	3	19	9	2	2	1	1	2	2	2	1	8	0	2	3.65
S	2	14	19	1	19	2	8	4	2	2	14	2	8	16	4	7	2	9	0	7	7.10
T	0	1	3	3	2	2	3	1	1	0	3	2	3	5	2	10	14	4	2	0	3.05
U	1	0	1	2	16	1	1	2	5	1	4	1	2	2	2	2	2	8	9	3	3.25
V	1	9	4	4	2	8	7	1	14	18	1	5	10	11	19	0	3	7	2	11	6.85
W	8	1	9	19	3	19	0	5	2	1	5	3	3	4	1	5	3	1	8	7	5.35
X	4	2	0	3	1	16	1	11	3	3	2	18	2	0	1	5	0	7	2	5	4.30
Y	1	2	1	2	0	2	7	2	4	8	19	2	5	3	3	0	19	2	1	18	5.05

1. La forma de la distribución muestral de las medias es diferente a la de la población. En la gráfica 8.3, la distribución de empleados tiene un sesgo positivo. No obstante, conforme selecciona muestras aleatorias de la población, cambia la forma de la distribución muestral de las medias. A medida que incrementa el tamaño de la muestra, la distribución muestral de las medias se aproxima a la distribución de probabilidad normal. Este hecho se ilustra con el teorema del límite central.



GRÁFICA 8.5 **Histograma del tiempo medio de servicio de 25 muestras de 20 empleados**

2. Hay menos dispersión en la distribución muestral de las medias que en la distribución de la población. En la población, los periodos de servicio variaron de 0 a 19 años. Cuando seleccionó muestras de tamaño 5, las medias de las muestras variaron de 1.6 a 8.6 años, y cuando seleccionó muestras de 20, las medias variaron de 3.05 a 7.10 años.

También puede comparar la media de las medias de la muestra con la media de la población. La media de las 25 muestras de los 20 empleados de la tabla 8.6 es de 4.676 años.

$$\mu_{\bar{X}} = \frac{3.95 + 3.25 + \ldots + 4.30 + 5.05}{25} = 4.676$$

Emplee el símbolo $\mu_{\bar{X}}$ para identificar la media de la distribución muestral de las medias. El subíndice recuerda que la distribución se refiere a la media muestral. Se lee *mu subíndice X barra.* Observe que la media de las medias muestrales, 4.676 años, se encuentra muy próxima a la media de la población de 4.80.

¿Qué concluye de este ejemplo? El teorema del límite central indica que, sin importar la forma de la distribución de población, la distribución muestral de la media se aproximará a la distribución de probabilidad normal. Cuanto mayor sea el número de observaciones en cada muestra, más evidente será la convergencia. El ejemplo de Spence Sprockets, Inc., demuestra el mecanismo del teorema del límite central. Comenzó con una población con sesgo positivo (gráfica 8.3). Después seleccionó 25 muestras aleatorias de 5 observaciones; calculó la media de cada muestra y, por último, organizó las 25 medias de muestra en una gráfica (gráfica 8.4). Observó un cambio en la forma de la distribución muestral de las medias respecto de la propia de la población. El desplazamiento va de una distribución con sesgo positivo a una que tiene la forma de la distribución de probabilidad normal.

Para aclarar más los efectos del teorema del límite central, incremente el número de observaciones en cada muestra de 5 a 20. Seleccione 25 muestras de 20 observaciones cada una y calcule la media de cada muestra. Por último, organice estas medias muestrales en una gráfica (gráfica 8.5). La forma del histograma de la gráfica 8.5 se desplaza claramente hacia la distribución de probabilidad normal.

En el capítulo 6, la gráfica 6.4 muestra diversas distribuciones binomiales con una proporción de *éxitos* de 0.10, lo cual es otra demostración del teorema del límite central. Observe que, conforme n se incrementa de 7 a 12 y de 20 a 40, el perfil de las distribuciones de probabilidad se desplaza para acercarse cada vez más a una distribución de probabilidad normal. La gráfica 8.5 de la página 279 también muestra la convergencia hacia la normalidad conforme n se incrementa. Esto confirma de nuevo el hecho de que, conforme se incluyen más observaciones de la muestra de cualquier distribución poblacional, la forma de la distribución muestral de las medias se aproximará cada vez más a la distribución normal.

El teorema del límite central mismo (lea de nuevo la definición de la página 274) no dice nada sobre la dispersión de la distribución muestral de medias ni sobre la comparación entre la media de la distribución muestral de medias y la media de la población. Sin embargo, en el ejemplo de Spence Sprockets hay menor dispersión en la distribución de la media muestral que en la distribución de población, lo que indica la diferencia en el rango de la población y en el rango de las medias muestrales. Observe que la media de las medias de las muestras se encuentra cerca de la media de la población. Se puede demostrar que la media de la distribución muestral es la media poblacional, es decir, que $\mu_{\bar{X}} = \mu$, y si la desviación estándar de la población es σ, la desviación estándar de las medias muestrales es $\sigma/\sqrt{n}$, en la que n es el número de observaciones de cada muestra. Entonces, $\sigma/\sqrt{n}$ es el **error estándar de la media**. En realidad, el nombre completo es *desviación estándar de la distribución muestral de medias.*

ERROR ESTÁNDAR DE LA MEDIA $$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad [8.1]$$

Esta sección permite importantes conclusiones.

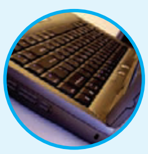
1. La media de la distribución muestral de las medias será *exactamente* igual a la media poblacional si selecciona todas las muestras posibles del mismo tamaño de una población dada. Es decir,

$$\mu = \mu_{\bar{X}}$$

Aunque no seleccione todas las muestras, es de esperar que la media de la distribución muestral de medias se aproxime a la media poblacional.

2. Habrá menos dispersión en la distribución muestral de las medias que en la población. Si la desviación estándar de la población es σ, la desviación estándar de la distribución muestral de medias es $\sigma/\sqrt{n}$. Note que, cuando se incrementa el tamaño de la muestra, disminuye el error estándar de la media.

Autoevaluación 8.4



Repase los datos de Spence Sprockets, Inc., de la página 276. Seleccione al azar 10 muestras de 5 empleados cada una. Utilice los métodos descritos en el capítulo y la tabla de números aleatorios (apéndice B.6) para determinar los empleados por incluir en la muestra. Calcule la media de cada muestra y trace una gráfica de las medias muestrales en una gráfica similar a la gráfica 8.3. ¿Cuál es la media de las 10 medias muestrales?

Ejercicios

11. El apéndice B.6 es una tabla de números aleatorios. De ahí que cada dígito de 0 a 9 tenga la misma probabilidad de presentarse.

a) Trace una gráfica que muestre la distribución de la población. ¿Cuál es la media de la población?

b) A continuación aparecen los 10 primeros renglones de cinco dígitos del apéndice B.6. Suponga que se trata de 10 muestras aleatorias de cinco valores cada una. Determine la media de cada muestra y trace una gráfica similar a la gráfica 8.3. Compare la media de la distribución muestral de las medias con la media poblacional.

0	2	7	1	1
9	4	8	7	3
5	4	9	2	1
7	7	6	4	0
6	1	5	4	5
1	7	1	4	7
1	3	7	4	8
8	7	4	5	5
0	8	9	9	9
7	8	8	0	4

12. Scrapper Elevator Company tiene 20 representantes de ventas, que distribuyen su producto en Estados Unidos y Canadá. La cantidad de unidades vendidas el mes pasado por cada representante se incluye a continuación. Suponga que estas cifras representan los valores la población.

2	3	2	3	3	4	2	4	3	2	2	7	3	4	5	3	3	3	3	5

a) Trace una gráfica que muestre la distribución de población.
b) Calcule la media de la población.
c) Seleccione cinco muestras aleatorias de 5 cada una. Calcule la media de cada muestra. Utilice los métodos descritos en el capítulo y en el apéndice B.6 para determinar los elementos que deben incluirse en la muestra.
d) Compare la media de la distribución muestral de medias con la media poblacional. ¿Esperaría que los dos valores fueran aproximadamente iguales?
e) Trace un histograma de las medias muestrales. ¿Nota alguna diferencia en la forma de la distribución muestral de las medias en comparación con la forma de la distribución de población?

13. Considere que todas las monedas (un centavo, 25 centavos, etc.) que tenga en el bolsillo o monedero constituyen una población. Elabore una tabla de frecuencias, comience por el año en curso y cuente de manera regresiva, para registrar la antigüedad (en años) de las monedas. Por ejemplo, si el año en curso es 2006, una moneda que tiene impreso el año 2004 tiene dos años de antigüedad.
a) Trace un histograma u otro tipo de gráfica que muestre la distribución de población.
b) Seleccione de manera aleatoria cinco monedas y registre la antigüedad media de las monedas seleccionadas. Repita el proceso 20 veces. Ahora trace un histograma u otro tipo de gráfica que muestre la distribución muestral de las medias.
c) Compare las formas de los dos histogramas.

14. Considere los dígitos de los números telefónicos en una página seleccionada al azar del directorio telefónico local como una población. Elabore una tabla de frecuencias con el último dígito de 30 números telefónicos seleccionados al azar. Por ejemplo, si el número telefónico es 5-55-97-04, registre un 4.
a) Trace un histograma u otro tipo de gráfica que muestre la distribución de población. Con la distribución uniforme, calcule la media de la población y la desviación estándar de la población.
b) Registre, asimismo, la media de la muestra de los últimos cuatro dígitos (97-04 daría una media de 5). Ahora elabore un histograma u otro tipo de gráfica que muestre la distribución muestral de las medias.
c) Compare la forma de los dos histogramas.

Uso de la distribución muestral de las medias

El análisis anterior reviste importancia, pues la mayoría de las decisiones tomadas en los negocios tiene como fundamento los resultados de un muestreo. He aquí algunos ejemplos.

1. Arm and Hammer Company desea cerciorarse de que su detergente para lavandería contiene realmente 100 onzas líquidas, como indica la etiqueta. Los registros de

los procesos de llenado indican que la cantidad media por recipiente es de 100 onzas líquidas y que la desviación estándar es de 2 onzas líquidas. A las diez de la mañana el técnico de calidad realiza la verificación de 40 recipientes y encuentra que la cantidad media por recipiente es de 99.8 onzas líquidas. ¿Debe interrumpir el proceso de llenado, o el error de muestreo es razonable?

2. A.C. Nielsen Company proporciona información a las empresas que se anuncian en televisión. Las investigaciones anteriores indican que, en promedio, los adultos estadounidenses ven televisión 6.0 horas al día. La desviación estándar es de 1.5 horas. Para una muestra de 50 adultos que viven en el área de Greater de Boston, ¿sería razonable seleccionar al azar una muestra y encontrar que en promedio ven un promedio de 6.5 horas al día?
3. Haughton Elevator Company pretende formular especificaciones relacionadas con el número de personas que pueden desplazarse en un elevador nuevo de gran capacidad. Suponga que el peso medio de un adulto es de 160 libras, y que la desviación estándar es de 15 libras. Ahora bien, la distribución de pesos no sigue una distribución de probabilidad normal. Tiene un sesgo positivo. ¿Cuál es la probabilidad de que, en una muestra de 30 adultos, el peso medio sea de 170 o más libras?

En cada una de estas situaciones hay una población de la cual existe determinada información. Se toma una muestra de esta población y se quiere saber si el error de muestreo, es decir, la diferencia entre el parámetro de población y la muestra estadística, se debe al azar.

De acuerdo con los conceptos analizados en la sección anterior, es posible calcular la probabilidad de que la media de una muestra se encuentre dentro de cierto margen. La distribución de muestreo seguirá la distribución de probabilidad normal con dos condiciones:

1. Cuando se sabe que las muestras se toman de poblaciones regidas por la distribución normal. En este caso, el tamaño de la muestra no constituye un factor.
2. Cuando se desconoce la forma de la distribución de población o se sabe que no es normal, pero la muestra contiene por lo menos 30 observaciones. En este caso, el teorema del límite central garantiza que la distribución muestral de las medias sigue una distribución normal.

Aplique la fórmula (7.5) del capítulo anterior para convertir cualquier distribución normal en una distribución normal estándar. A este hecho también se le denomina valor *z*. Así, se emplea la tabla estándar normal del apéndice B.1 para determinar la probabilidad de seleccionar una observación que caerá dentro de un intervalo específico. La fórmula para determinar un valor *z* es:

$$z = \frac{X - \mu}{\sigma}$$

En esta fórmula, X es el valor de la variable aleatoria; μ es la media de la población y σ es la desviación estándar de la población.

Sin embargo, la mayor parte de las decisiones de negocios se refiere a una muestra, no a una sola observación. Así, lo importante es la distribución de $\overline{X}$, la media muestral, en lugar de X, el valor de una observación. Éste es el primer cambio en la fórmula (7.5). El segundo consiste en emplear el error estándar de la media de n observaciones en lugar de la desviación estándar de la población. Es decir, se usa $\sigma/\sqrt{n}$ en el denominador en vez de σ. Por consiguiente, para determinar la probabilidad de una media muestral con rango especificado, primero aplique la fórmula para determinar el valor z correspondiente. Después consulte el apéndice B.1 para localizar la probabilidad.

CÁLCULO DEL VALOR *z* DE $\overline{X}$ CUANDO SE CONOCE LA DESVIACIÓN ESTÁNDAR DE LA POBLACIÓN

$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \qquad \textbf{[8.2]}$$

El siguiente ejemplo muestra la aplicación.

Ejemplo

El departamento de control de calidad de Cola, Inc., conserva registros sobre la cantidad de bebida de cola en su botella gigante. La cantidad real de bebida en cada botella es de primordial importancia, pero varía en una mínima cantidad de botella en botella. Cola, Inc., no desea llenar botellas con menos líquido del debido, pues tendría problemas en lo que se refiere a la confiabilidad de la etiqueta. Por otra parte, no puede colocar líquido de más en las botellas porque regalaría bebida, lo cual reduciría sus utilidades. Los registros indican que la cantidad de bebida de cola tiene una distribución de probabilidad normal. La cantidad media por botella es de 31.2 onzas, y la desviación estándar de la población, de 0.4 onzas. Hoy, a las 8 de la mañana, el técnico de calidad seleccionó al azar 16 botellas de la línea de llenado. La cantidad media de bebida en las botellas es de 31.38 onzas. ¿Es un resultado poco probable? ¿Es probable que el proceso permita colocar demasiada bebida en las botellas? En otras palabras, ¿es poco común el error de muestreo de 0.18 onzas?

Solución

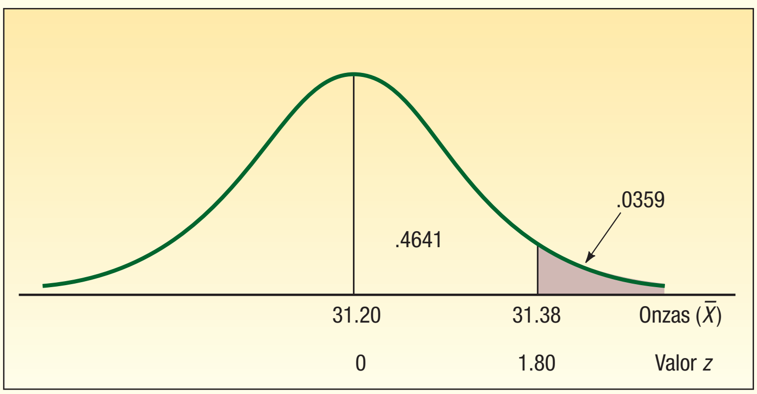
Utilice los resultados de la sección anterior para determinar la probabilidad de seleccionar una muestra de 16 (n) botellas de una población normal con una media de 31.2 (μ) onzas y una desviación estándar de la población de 0.4 (σ) onzas, y encontrar que la media muestral es de 31.38($\bar{X}$). Aplique la fórmula (8.2) para determinar el valor de z.

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{31.38 - 31.20}{0.4/\sqrt{16}} = 1.80$$

El numerador de esta ecuación, $\bar{X} - \mu = 31.38 - 31.20 = .18$, es el error muestral. El denominador, $\sigma/\sqrt{n} = 0.4/\sqrt{16} = 0.1$, es el error estándar de la distribución muestral de la media. Así, los valores z expresan el error muestral en unidades estándar; en otras palabras, el error estándar.

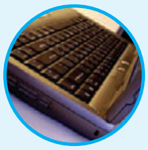
Después, calcule la probabilidad de un valor z mayor que 1.80. En el apéndice B.1 localice la probabilidad correspondiente a un valor z de 1.80. Este valor es de 0.4641. La probabilidad de un valor z mayor que 1.80 es de 0.0359, que se calcula con la resta 0.5000 – 0.4641.

¿Qué concluye? No es probable —menos de 4% de probabilidad— que seleccione una muestra de 16 observaciones de una población normal con una media de 31.2 onzas y una desviación estándar poblacional de 0.4 onzas, y determine que la media de la muestra es igual o mayor que 31.38 onzas. La conclusión es que en el proceso se vierte demasiada bebida de cola en las botellas. El técnico de control de calidad debe entrevistarse con el supervisor de producción para sugerir la reducción de la cantidad de bebida en cada botella. La información se resume en la gráfica 8.6.



GRÁFICA 8.6 Distribución muestral de la cantidad media de bebida de cola en una botella gigante

Autoevaluación 8.5



Consulte la información relativa a Cola, Inc. Suponga que el técnico de control de calidad seleccionó una muestra de 16 botellas gigantes con un promedio de 31.08 onzas. ¿Qué concluye sobre el proceso de llenado?

Ejercicios

15. Una población normal tiene una media de 60 y una desviación estándar de 12. Usted selecciona una muestra aleatoria de 9. Calcule la probabilidad de que la media muestral:
- ***a)*** Sea mayor que 63.
- ***b)*** Sea menor que 56.
- ***c)*** Se encuentre entre 56 y 63.

16. Una población normal posee una media de 75 y una desviación estándar de 5. Usted selecciona una muestra de 40. Calcule la probabilidad de que la media muestral:
- ***a)*** Sea menor que 74.
- ***b)*** Se encuentre entre 74 y 76.
- ***c)*** Se encuentre entre 76 y 77.
- ***d)*** Sea mayor que 77.

17. En el sur de California, la renta de un departamento con una recámara tiene una distribución normal con una media de $2 200 mensuales y una desviación estándar de $250 mensuales. La distribución del costo mensual no se rige por la distribución normal. De hecho, tiene un sesgo positivo. ¿Cuál es la probabilidad de seleccionar una muestra de 50 departamentos de una recámara y hallar que la media es de por lo menos $1 950 mensuales?

18. De acuerdo con un estudio del Internal Revenue Service, los contribuyentes tardan 330 minutos en promedio en preparar, copiar y archivar en un medio electrónico la forma fiscal 1040. Esta distribución de tiempos se rige por una distribución normal, y la desviación estándar es de 80 minutos. Un organismo de control selecciona una muestra aleatoria de 40 consumidores.
- ***a)*** ¿Cuál es el error estándar de la media de este ejemplo?
- ***b)*** ¿Cuál es la probabilidad de que la media de la muestra sea mayor que 320 minutos?
- ***c)*** ¿Cuál es la probabilidad de que la media de la muestra se encuentre entre 320 y 350 minutos?
- ***d)*** ¿Cuál es la probabilidad de que la media de la muestra sea superior que 350 minutos?

Resumen del capítulo

I. Hay muchas razones para realizar el muestreo de una población.
- **A.** Los resultados de una muestra permiten calcular adecuadamente el valor del parámetro poblacional, con lo cual se ahorra tiempo y dinero.
- **B.** Entrar en contacto con todos los miembros de la población consume demasiado tiempo.
- **C.** Resulta imposible verificar y localizar a todos los miembros de la población.
- **D.** El costo de estudiar a todos los elementos de la población resulta prohibitivo.
- **E.** En una prueba con frecuencia se destruye el elemento de la muestra y no se puede regresar a la población.

II. En una muestra sin sesgo, todos los miembros de la población tienen una posibilidad de ser seleccionados para la muestra. Existen diversos métodos de muestreo de probabilidad.
- **A.** En una muestra aleatoria simple, todos los miembros de la población tienen la misma posibilidad de ser seleccionados para la muestra.
- **B.** En una muestra sistemática, se selecciona un punto de partida aleatorio y después se selecciona cada *k*-ésimo elemento subsiguiente de la población para formar la muestra.
- **C.** En una muestra estratificada, la población se divide en varios grupos, a los que se denomina *estratos*, y enseguida se selecciona una muestra aleatoria de cada estrato.
- **D.** En el muestreo por conglomerados, la población se divide en unidades primarias; después se toman las muestras de las unidades primarias.

III. El error de muestreo es la diferencia entre un parámetro poblacional y un estadístico de la muestra.

IV. La distribución muestral de las medias es una distribución de probabilidad de todas las posibles medias muestrales del mismo tamaño de muestra.

A. Para un tamaño de muestra dado, la media de todas las posibles medias muestrales tomadas de una población es igual a la media de la población.

B. Existe una menor variación en la distribución de las medias muestrales que en la distribución de la población.

C. El error estándar de la media mide la variación de la distribución muestral de las medias. El error estándar se calcula de la siguiente manera:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \qquad \textbf{[8.1]}$$

D. Si la población se rige por una distribución normal, la distribución muestral de las medias también se regirá por la distribución normal para muestras de cualquier tamaño. Suponga que conoce la desviación estándar de la población. Para determinar la probabilidad de que una media muestral caiga dentro de determinada región, se aplica la fórmula

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \qquad \textbf{[8.2]}$$

Clave de pronunciación

SÍMBOLO	SIGNIFICADO	PRONUNCIACIÓN
$\mu_{\bar{X}}$	Media de la distribución muestral de las medias	*mu subíndice X barra*
$\sigma_{\bar{X}}$	Error estándar de la población de las medias de las muestras	*sigma subíndice X barra*

Ejercicios del capítulo

19. Las tiendas de venta al menudeo en el centro comercial de North Towne Square son las siguientes:

00	Elder-Beerman	09	Lion Store	18	County Seat
01	Sears	10	Bootleggers	19	Kid Mart
02	Deb Shop	11	Formal Man	20	Lerner
03	Frederick's of Hollywood	12	Leather Ltd.	21	Coach House Gifts
04	Petries	13	B Dalton Bookseller	22	Spencer Gifts
05	Easy Dreams	14	Pat's Hallmark	23	CPI Photo Finish
06	Summit Stationers	15	Things Remembered	24	Regis Hairstylists
07	E. B. Brown Opticians	16	Pearle Vision Express		
08	Kay-Bee Toy & Hobby	17	Dollar Tree		

a) Si selecciona los números aleatorios 11, 65, 86, 62, 06, 10, 12, 77 y 04, ¿con qué tiendas es necesario ponerse en contacto para realizar una encuesta?

b) Seleccione una muestra aleatoria de cuatro tiendas. Utilice el apéndice B.6.

c) Debe aplicar un procedimiento de muestreo sistemático. Es necesario ponerse en contacto con la primera tienda y a continuación con cada tercer establecimiento. ¿Con qué tiendas entrará en contacto?

20. Medical Mutual Insurance investiga el costo de una visita de rutina a consultorios de médicos familiares en el área de Rochester, Nueva York. La siguiente constituye una lista de médicos familiares de la región. Se seleccionará a los médicos de forma aleatoria y se establecerá comunicación con ellos para conocer el monto de sus honorarios. Los 39 médicos se codificaron del 00 al 38. También se indica si cuentan con consultorio propio (P), si tienen un socio (S) o si tiene un consultorio en grupo (G).

Número	Médico	Tipo de consultorio	Número	Médico	Tipo de consultorio
00	R. E. Scherbarth, M.D.	P	20	Gregory Yost, M.D.	S
01	Crystal R. Goveia, M.D.	S	21	J. Christian Zona, M.D.	S
02	Mark D. Hillard, M.D.	S	22	Larry Johnson, M.D.	S
03	Jeanine S. Huttner, M.D.	S	23	Sanford Kimmel, M.D.	S
04	Francis Aona, M.D.	S	24	Harry Mayhew, M.D.	P
05	Janet Arrowsmith, M.D.	S	25	Leroy Rodgers, M.D.	P
06	David DeFrance, M.D.	P	26	Thomas Tafelski, M.D.	P
07	Judith Furlong, M.D.	P	27	Mark Zilkoski, M.D.	G
08	Leslie Jackson, M.D.	G	28	Ken Bertka, M.D.	G
09	Paul Langenkamp, M.D.	P	29	Mark DeMichiei, M.D.	G
10	Philip Lepkowski, M.D.	P	30	John Eggert, M.D.	S
11	Wendy Martin, M.D.	P	31	Jeanne Fiorito, M.D.	S
12	Denny Mauricio, M.D.	S	32	Michael Fitzpatrick, M.D.	S
13	Hasmukh Parmar, M.D.	S	33	Charles Holt, D.O.	S
14	Ricardo Pena, M.D.	S	34	Richard Koby, M.D.	S
15	David Reames, M.D.	S	35	John Meier, M.D.	S
16	Ronald Reynolds, M.D.	G	36	Douglas Smucker, M.D.	P
17	Mark Steinmetz, M.D.	G	37	David Weldy, M.D.	S
18	Geza Torok, M.D.	P	38	Cheryl Zaborowski, M.D.	S
19	Mark Young, M.D.	S			

a) Los números aleatorios que se obtuvieron del apéndice B.6 son 31, 94, 43, 36, 03, 24, 17 y 09. ¿Con qué médicos se debe establecer comunicación?
b) Seleccione una muestra aleatoria con los números aleatorios del apéndice B.6.
c) Una muestra incluirá a cada quinto médico. El número 04 se selecciona como punto de partida. ¿Con qué médicos se debe establecer contacto?
d) Una muestra constará de dos médicos con consultorio propio (P), dos que tienen socios (S) y uno con consultorio en grupo (G). Seleccione la muestra correspondiente. Explique su procedimiento.

21. ¿Qué es el error de muestreo? ¿Puede ser cero el valor de una muestra? De ser cero, ¿qué significaría?

22. Señale las razones del muestreo. Proporcione un ejemplo de cada una.

23. El fabricante de eMachines, que manufactura una computadora económica, recién concluyó el diseño de un nuevo modelo de computadora portátil. A los altos ejecutivos de eMachines les gustaría obtener ayuda para poner precio a la nueva computadora portátil. Se solicitaron los servicios de empresas de investigación de mercados y se les pidió que prepararan una estrategia de precios. Marketing-Gets-Results probó las nuevas computadoras portátiles de eMachines con 50 consumidores elegidos al azar, quienes indicaron que tenían planes de adquirir la computadora el año entrante. La segunda empresa de investigación de mercados, llamada Marketing-Reaps-Profits, probó en el mercado la nueva computadora portátil de eMachines con 200 actuales propietarios de una computadora portátil. ¿Cuál de las pruebas de las empresas de investigación de mercados resulta la más útil? Explique las razones.

24. Responda las siguientes preguntas en uno o dos enunciados bien construidos.
a) ¿Qué sucede con el error estándar de la media si aumenta el tamaño de la muestra?
b) ¿Qué sucede con la distribución muestral de las medias si aumenta el tamaño de la muestra?
c) Cuando se utiliza la distribución de las medias muestrales para aproximar la media poblacional, ¿cuál es el beneficio de utilizar tamaños muestrales más grandes?

25. Hay 25 moteles en Goshen, Indiana. El número de habitaciones en cada motel es el siguiente:

90	72	75	60	75	72	84	72	88	74	105	115	68	74	80	64	104	82	48	58	60	80	48	58	100

a) De acuerdo con la tabla de números aleatorios (apéndice B.6), seleccione una muestra aleatoria de cinco moteles de esta población.
b) Obtenga una muestra sistemática seleccionando un punto de partida aleatorio entre los primeros cinco moteles y después haga una selección cada quinto motel.
c) Suponga que los últimos cinco moteles son *de tarifas rebajadas*. Describa la forma en que seleccionaría una muestra aleatoria de tres moteles normales y dos de tarifas rebajadas.

26. Como parte de su programa de servicio al cliente, United Airlines seleccionó de forma aleatoria a 10 pasajeros del vuelo de hoy que parte de Chicago a Tampa a las nueve de la mañana. A cada pasajero de la muestra se le hará una entrevista a fondo en relación con las instalaciones, servicios, alimentos, etc., en los aeropuertos. Para identificar la muestra, a cada pasajero se le proporcionó un número al abordar la nave. Los números comenzaron por 001 y terminaron en 250.

a) Seleccione al azar 10 números con ayuda del apéndice B.6.
b) La muestra de 10 pudo seleccionarse con una muestra sistemática. Elija el primer número con ayuda del apéndice B.6 y, después, mencione los números con los que se entrevistará.
c) Evalúe ambos métodos señalando las ventajas y posibles desventajas.
d) ¿De qué otra forma se puede seleccionar una muestra aleatoria de los 250 pasajeros?

27. Suponga que el profesor de estadística le aplicó seis exámenes durante el semestre. Usted obtuvo las siguientes calificaciones (porcentaje corregido): 79, 64, 84, 82, 92 y 77. En lugar de promediar las seis calificaciones, el profesor le indicó que escogería dos al azar y calcularía el porcentaje final con base en dos porcentajes.

a) ¿Cuántas muestras de dos calificaciones se pueden tomar?
b) Enumere todas las posibles muestras de tamaño dos y calcule la media de cada una.
c) Calcule la media de las medias de la muestra y compárela con la media de la población.
d) Si usted fuera estudiante, ¿le gustaría este sistema? ¿Sería diferente el resultado si se eliminara la calificación más baja? Redacte un breve informe.

28. En la oficina del First National Bank, ubicada en el centro de la ciudad, hay cinco cajeros automáticos. La semana pasada cada uno de los cajeros incurrió en el siguiente número de errores: 2, 3, 5, 3 y 5.

a) ¿Cuántas muestras de dos cajeros se pueden seleccionar?
b) Escriba todas las posibles muestras de tamaño 2 y calcule la media de cada una.
c) Calcule la media de las medias de las muestras y compárela con la media de la población.

29. El departamento de control de calidad tiene como empleados a cinco técnicos en el turno matutino. A continuación aparece el número de veces que cada técnico indicó al supervisor de producción que interrumpiera el proceso durante la última semana.

Técnico	Interrupciones
Taylor	4
Hurley	3
Gupta	5
Rousche	3
Huang	2

a) ¿Cuántas muestras de dos técnicos se forman con esta población?
b) Enumere todas las muestras de dos observaciones que se pueden tomar y calcule la media de cada muestra.
c) Compare la media de las medias de las muestras con la media de la población.
d) Compare la forma de la distribución de la población con la forma de la distribución muestral de las medias.

30. The Appliance Center cuenta con seis representantes de ventas en su sucursal del norte de Jacksonville. A continuación aparece el número de refrigeradores vendidos por cada representante el último mes.

Representante de ventas	Refrigeradores vendidos
Zina Craft	54
Woon Junge	50
Ernie DeBrul	52
Jan Niles	48
Molly Camp	50
Rachel Myak	52

a) ¿Cuántas muestras de tamaño 2 se pueden tomar?
b) Seleccione todas las muestras posibles de tamaño 2 y calcule la cantidad media de refrigeradores vendidos.
c) Organice las medias de las muestras en una distribución de frecuencias.
d) ¿Cuál es la media de la población? ¿Cuál es la media de las medias de la muestra?
e) ¿Cuál es la forma de la distribución de población?
f) ¿Cuál es la forma de la distribución muestral de la media?

31. Mattel Corporation produce autos de control remoto que funcionan con baterías AA. La vida media de las baterías para este producto es de 35.0 horas. La distribución de las vidas de las baterías se aproxima a una distribución de probabilidad normal con una desviación estándar de 5.5 horas. Como parte de su programa, Sony prueba muestras de 25 baterías.
- ***a)*** ¿Qué se puede decir sobre la forma de la distribución muestral de la media?
- ***b)*** ¿Cuál es el error estándar de la distribución muestral de la media?
- ***c)*** ¿Qué proporción de las muestras tendrá una media de vida útil de más de 36 horas?
- ***d)*** ¿Qué proporción de la muestra tendrá una media de vida útil mayor que 34.5 horas?
- ***e)*** ¿Qué proporción de la muestra tendrá una media de vida útil entre 34.5 y 36 horas?

32. CRA CDs, Inc., desea que las extensiones medias de los "cortes" de un CD sean de 135 segundos (2 minutos y 15 segundos). Esto permitirá a los disc jockeys contar con tiempo de sobra para comerciales entre cada segmento de 10 minutos. Suponga que la distribución de la extensión de los cortes sigue una distribución normal con una desviación estándar de la población de 8 segundos, y también que selecciona una muestra de 16 cortes de varios CD vendidos por CRA CDs, Inc.
- ***a)*** ¿Qué puede decir sobre la forma de la distribución muestral de la media?
- ***b)*** ¿Cuál es el error estándar de la media?
- ***c)*** ¿Qué porcentaje de las medias muestrales será superior a 140 segundos?
- ***d)*** ¿Qué porcentaje de las medias muestrales será superior a 128 segundos?
- ***e)*** ¿Qué porcentaje de las medias muestrales será superior a 128 segundos e inferior a 140?

33. Estudios recientes indican que la mujer común de 50 años de edad gasta $350 anuales en productos de cuidado personal. La distribución de las sumas que se gastan se rige por una distribución normal con una desviación estándar de $45 anuales. Se selecciona una muestra aleatoria de 40 mujeres. La cantidad media que gasta dicha muestra es de $335. ¿Cuál es la probabilidad de hallar una media muestral igual o superior a la de la población indicada?

34. La información del American Institute of Insurance indica que la cantidad media de seguros de vida por familia en Estados Unidos asciende a $110 000. Esta distribución sigue la distribución normal con una desviación estándar de $40 000.
- ***a)*** Si selecciona una muestra aleatoria de 50 familias, ¿cuál es el error estándar de la media?
- ***b)*** ¿Cuál es la forma que se espera que tenga la distribución muestral de la media?
- ***c)*** ¿Cuál es la probabilidad de seleccionar una muestra con una media de por lo menos $112 000?
- ***d)*** ¿Cuál es la probabilidad de seleccionar una muestra con una media de más de $100 000?
- ***e)*** Determine la probabilidad de seleccionar una muestra con una media de más de $100 000 e inferior a $112 000.

35. La edad media a la que los hombres se casan en Estados Unidos por primera vez se rige por la distribución normal con una media de 24.8 años. La desviación estándar de la distribución es de 2.5 años. En el caso de una muestra aleatoria de 60 hombres, ¿cuál es la probabilidad de que la edad a la que se casaran por primera vez sea menor de 25.1 años?

36. Un estudio reciente llevado a cabo por la Greater Los Angeles Taxi Drivers Association mostró que la tarifa media por servicio de Hermosa Beach al aeropuerto internacional de Los Ángeles es de $18.00, y la desviación estándar, de $3.50. Seleccione una muestra de 15 tarifas.
- ***a)*** ¿Cuál es la probabilidad de que la media de la muestra se encuentre entre $17.00 y $20.00?
- ***b)*** ¿Qué debe suponer para llevar a cabo el cálculo anterior?

37. Crosset Trucking Company afirma que el peso medio de sus camiones cuando se encuentran completamente cargados es de 6 000 libras, y la desviación estándar, de 150 libras. Suponga que la población se rige por la distribución normal. Se seleccionan al azar 40 camiones y se pesan. ¿Dentro de qué límites se presentará 95% de las medias de la muestra?

38. La cantidad media de abarrotes que compra cada cliente en Churchill Grocery Store es de $23.50, con una desviación estándar de $5.00. Suponga que la distribución de cantidades compradas sigue la distribución normal. En el caso de una muestra de 50 clientes, conteste las siguientes preguntas.
- ***a)*** ¿Cuál es la probabilidad de que la media de la muestra sea de por lo menos $25.00?
- ***b)*** ¿Cuál es la probabilidad de que la media de la muestra sea superior a $22.50 e inferior a $25.00?
- ***c)*** ¿Dentro de qué límites se presentará 90% de las medias muestrales?

39. La calificación media SAT para estudiantes atletas de la División I es de 947, con una desviación estándar de 205. Si selecciona una muestra aleatoria de 60 estudiantes, ¿cuál es la probabilidad de que la media se encuentre por debajo de 900?

40. Suponga que lanza un dado dos veces.
- ***a)*** ¿Cuántas muestras se pueden seleccionar?
- ***b)*** Enumere cada una de las posibles muestras y calcule la media.
- ***c)*** En una gráfica similar a la 8.1, compare la distribución de las medias muestrales con la distribución de la población.
- ***d)*** Calcule la media y la desviación estándar de cada distribución y compárelas.

41. La siguiente tabla contiene los ingresos personales per cápita de cada uno de los 50 estados en 2004.

Número	Estado	2004	Número	Estado	2004
0	Alabama	$27 795	25	Montana	$26 857
1	Alaska	34 454	26	Nebraska	31 339
2	Arizona	28 442	27	Nevada	33 405
3	Arkansas	25 725	28	New Hampshire	37 040
4	California	35 019	29	New Jersey	41 332
5	Colorado	36 063	30	New Mexico	26 191
6	Connecticut	45 398	31	New York	38 228
7	Delaware	35 861	32	North Carolina	29 246
8	Florida	31 455	33	North Dakota	31 398
9	Georgia	30 051	34	Ohio	31 322
10	Hawaii	32 160	35	Oklahoma	28 089
11	Idaho	27 098	36	Oregon	29 971
12	Illinois	34 351	37	Pennsylvania	33 348
13	Indiana	30 094	38	Rhode Island	33 733
14	Iowa	30 560	39	South Carolina	27 172
15	Kansas	30 811	40	South Dakota	30 856
16	Kentucky	27 709	41	Tennessee	30 005
17	Louisiana	27 581	42	Texas	30 222
18	Maine	30 566	43	Utah	26 606
19	Maryland	39 247	44	Vermont	32 770
20	Massachusetts	41 801	45	Virginia	35 477
21	Michigan	31 954	46	Washington	35 299
22	Minnesota	35 861	47	West Virginia	25 872
23	Mississippi	24 650	48	Wisconsin	32 157
24	Missouri	30 608	49	Wyoming	34 306

a) Usted pretende seleccionar una muestra de ocho elementos de la lista. Los números aleatorios seleccionados son 45, 15, 81, 09, 39, 43, 90, 26, 06, 45, 01 y 42. ¿Qué estados se incluyen en la muestra?

b) Usted desea utilizar una muestra sistemática de cada sexto elemento y elige el dígito 02 como punto de partida. ¿Qué estados se incluyen?

42. Human Resource Consulting (HRC) lleva a cabo un sondeo con una muestra de 60 empresas con el fin de estudiar los costos del cuidado de la salud del cliente. Uno de los elementos que se estudia es el deducible anual que deben pagar los empleados. La Bureau of Labor estatal informa que la media de esta distribución es de $502, con una desviación estándar de $100.

a) Calcule el error estándar de la media muestral para HRC.

b) ¿Cuál es la probabilidad de que HRC encuentre una media muestral entre $477 y $527?

c) Calcule la probabilidad de que la media muestral oscile entre $492 y $512.

d) ¿Cuál es la probabilidad de que la media muestral sea superior a $550?

43. La década pasada, el número medio de miembros de la Information Systems Security Association, que tenían experiencia en ataques por negación de servicios cada año es de 510, con una desviación estándar de 14.28 ataques. Suponga que nada cambia en este ambiente.

a) ¿Cuál es la probabilidad de que este grupo sufra un promedio de más de 600 ataques los próximos 10 años?

b) Calcule la probabilidad de que experimenten un promedio de entre 500 y 600 ataques durante los próximos 10 años.

c) ¿Cuál es la probabilidad de que experimenten un promedio de menos de 500 ataques durante los próximos 10 años?

44. El Oil Price Information Center informa que el precio medio por galón de gasolina normal es de $3.26, con una desviación estándar de población de $0.18. Suponga que se selecciona una muestra aleatoria de 40 estaciones de gasolina, cuyo costo medio de gasolina normal se calcula.

a) ¿Cuál es el error estándar de la media de este experimento?

b) ¿Cuál es la probabilidad de que la media de la muestra oscile entre $3.24 y $3.28?

c) ¿Cuál es la probabilidad de que la diferencia entre la media muestral y la media poblacional sea inferior a 0.01?

d) ¿Cuál es la probabilidad de que la media de la muestra sea superior a $3.34?