

# 8 Métodos de muestreo y el teorema del límite central

## METAS

Al terminar este capítulo, usted podrá:

- 1 Explicar por qué una muestra a menudo es la única forma de saber algo acerca de una población.
- 2 Describir métodos para seleccionar una muestra.
- 3 Definir y construir una distribución muestral de la media de la muestra.
- 4 Explicar el *teorema del límite central*.
- 5 Utilizar el teorema del límite central para el cálculo de probabilidades en una distribución muestral de medias.



En la oficina del centro de la ciudad del First National City Bank hay cinco cajeros automáticos. ¿Cuántas muestras diferentes de dos cajeros son posibles? (Vea la Meta 3 y el Ejercicio 28.)



### Estadística en acción

Con el importante papel que desempeña la estadística inferencial en todas las ramas de la ciencia, la disponibilidad de fuentes de números aleatorios se ha convertido en una necesidad. El primer libro de números aleatorios, contiene 41 600 dígitos aleatorios generados por L. Tippett, se publicó en 1927. En 1938, R. A. Fisher y F. Yates publicaron 15 000 números aleatorios generados con el uso de dos barajas. En 1955, RAND Corporation publicó un millón de dígitos aleatorios, generados por los pulsos de frecuencia al azar de una ruleta electrónica. Para 1970, las aplicaciones de muestreo requerían de miles de millones de números aleatorios. Desde entonces, se han desarrollado métodos para generar, mediante el uso de una computadora, dígitos que son "casi" aleatorios y, por tanto, se conoce como *seudoaleatorios*. La pregunta sobre si un programa se puede utilizar para generar números que realmente sean aleatorios sigue siendo objeto de discusión.

## Introducción

Los Capítulos 1 al 4 enfatizaron las técnicas para describir datos. A fin de ilustrar estas técnicas, organizamos los precios de los 80 vehículos que se vendieron durante el mes pasado en Whitner Autoplex en una distribución de frecuencia y calculamos diversas medidas de ubicación y dispersión. Medidas como la media y la desviación estándar describen el precio típico de venta y la dispersión en los precios de venta. En estos capítulos el énfasis se centró en describir la condición de la información. Es decir, describimos algo que ya sucedió.

El Capítulo 5 empieza por definir los fundamentos de la inferencia estadística con el estudio de la probabilidad. Recuerde que nuestro objetivo en la inferencia estadística es determinar algo acerca de una *población* con base sólo en una *muestra*. La población es la totalidad del grupo de individuos u objetos que se someten a consideración, y la muestra es una parte o un subconjunto representativo de esa población. El Capítulo 6 amplía los conceptos de probabilidad al describir tres distribuciones de probabilidad discreta: la binomial, la hipergeométrica y la de Poisson. El Capítulo 7 describe la distribución de probabilidad uniforme y la distribución de probabilidad normal. Estas dos son distribuciones continuas. Las distribuciones de probabilidad abarcan todos los resultados posibles de un experimento y la probabilidad relacionada con cada resultado. Utilizamos las distribuciones de probabilidad para evaluar la posibilidad de que algo ocurra en el futuro.

En este capítulo empezamos a estudiar el muestreo. Una muestra es una herramienta para inferir algo acerca de una población. Iniciamos el presente capítulo con un análisis de los métodos de selección de una muestra de una población. Luego, construimos una distribución muestral de medias para entender la manera en que las medias de la muestra suelen hacinarse alrededor de la media de la población. Por último, demostramos que para cualquier población la forma de esta distribución de muestreo tiende a seguir la distribución de probabilidad normal.

## Métodos de muestreo

En el Capítulo 1 dijimos que el propósito de la estadística inferencial es averiguar algo acerca de una población con base en una muestra. Una muestra es una porción o parte representativa de la población de interés. En muchos casos, el muestreo es más viable que estudiar a la población en su totalidad. En esta sección demostramos los motivos principales para realizar un muestreo, y después diversos métodos para seleccionar una muestra.

## Razones del muestreo

Cuando se estudian las características de una población, existen muchas razones de tipo práctico por las que preferimos seleccionar muestras de una población para observar y medir. Algunas de las razones del muestreo son las siguientes.

1. **Establecer contacto con la totalidad de la población requeriría de demasiado tiempo.** Es posible que un candidato a algún puesto federal de elección popular quiera determinar las posibilidades que tiene de resultar electo. Una encuesta de muestreo utilizando el personal y las entrevistas de campo convencionales de una empresa especializada en encuestas tardaría sólo uno o dos días. Utilizando el mismo personal y los mismos entrevistadores y trabajando siete días a la semana, ¿se necesitarían casi 200 años para ponerse en contacto con la totalidad de la población en edad de votar! Aunque fuera posible reunir a un numeroso equipo de encuestadores, quizá no valdría la pena contactar a todos los votantes.
2. **El costo de estudiar todos los elementos en una población puede resultar prohibitivo.** Las organizaciones que realizan encuestas de opinión pública y pruebas entre los consumidores, como Gallup Polls y Roper ASW, por lo regular establecen contacto con menos de 2 000 de las casi 60 millones de familias que viven en Estados Unidos. Una organización tipo panel de consumidores cobra cerca de 40 000 dólares por enviar muestras por correo y tabular las respuestas a fin de probar un producto (como cereal para el desayuno, alimento para gato o algún perfume). La misma prueba de producto utilizando la totalidad de las 60 millones de familias costaría alrededor de 1 000 millones de dólares.

3. **La imposibilidad física de verificar todos los elementos de la población.** Las poblaciones de peces, aves, serpientes, mosquitos y similares son muy numerosas y se desplazan, nacen y mueren en forma constante. En lugar de intentar contar todos los patos que existen en Canadá o todos los peces que habitan en el lago Erie, hacemos cálculo empleando diversas técnicas, como contar todos los patos que viven en un estanque elegido al azar, revisando las cestas de los cazadores o colocando redes en lugares predeterminados en el lago.
4. **La naturaleza destructiva de algunas pruebas.** Si los catadores de vino de Sutter Home Winery, en California, se tomaran todo el vino para evaluar la vendimia, consumirían la totalidad de la cosecha y no quedaría nada disponible para vender. En el área de la producción industrial, las placas de acero, los cables y productos similares deben tener cierta resistencia mínima a la tensión. Para asegurar que el producto cumple con las especificaciones mínimas, el Departamento de Aseguramiento de la Calidad selecciona una muestra de la producción actual. Cada pieza se estira hasta que se rompe, y se registra el punto de fractura (por lo regular en libras por pulgada cuadrada). Es evidente que si todo el cable o todas las placas se sometieran a pruebas de resistencia a la tensión no habría productos disponibles para la venta o para utilizarlos. Por el mismo motivo, Kodak sólo selecciona 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 dispusiera de recursos suficientes, dudamos que la precisión adicional de una muestra del 100% (es decir, estudiar a toda la población) sea de importancia fundamental en la mayor parte de los problemas. El gobierno federal de Estados Unidos, por ejemplo, utiliza una muestra de supermercados repartidos por todo el país para determinar el índice mensual de precios de alimentos. En el índice se incluyen los precios del pan, los frijoles, la leche y otros productos alimenticios. Es poco probable que la inclusión de todos los supermercados de Estados Unidos afecte el índice en términos significativos, porque los precios de la leche, el pan y otros alimentos no suelen variar más que unos cuantos centavos de una cadena de tiendas a otra.



Al seleccionar una muestra, los investigadores o analistas deben prestar mucha atención en que la muestra sea una representación imparcial de la población. En otras palabras, la muestra debe estar libre de influencias o sesgos. En el Capítulo 1, un ejemplo del abuso de la estadística es la selección intencional de los dentistas para reportar que “dos de cada tres dentistas entrevistados indicaban que recomendarían la pasta dentífrica marca X a sus pacientes”. Es evidente que la gente puede seleccionar una muestra que respalde sus propias influencias o sesgos. El aspecto ético de la estadística exige siempre que las muestras estén libres de influencias o sesgos y que los informes acerca de los resultados se elaboren en forma objetiva. A continuación, varios métodos de muestreo ilustran cómo seleccionar una muestra imparcial y libre de influencias de una población.

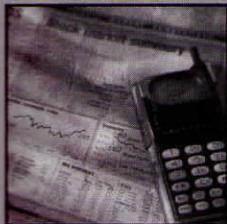
## Muestreo aleatorio simple

El tipo de muestreo que más se utiliza es el **muestreo aleatorio simple**.

**MUESTRA ALEATORIA SIMPLE** Muestra seleccionada de modo que cada elemento o persona en la población tiene la misma oportunidad de resultar seleccionado.

Para ilustrar el muestreo aleatorio simple y la selección, supongamos que una población consiste en 845 empleados de Nitra Industries. Se va a seleccionar una muestra de 50 empleados de esa población. Una forma de asegurar que todos los empleados en la población tienen la misma posibilidad de ser electos consiste en escribir el nombre de cada empleado en un trozo de papel pequeño y depositar todos los trozos en una caja. Una vez que

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



### Estadística en acción

¿Será discriminación sacar ventaja de su físico? Antes de responder, considere un artículo reciente en *Personnel Journal*. Estos hallazgos indican que los hombres y mujeres atractivos ganan alrededor de 5% más que aquellos que tienen una apariencia promedio, quienes a su vez ganan 5% más que sus contrapartes francamente feos. Esto se aplica tanto para hombres como para mujeres. Lo anterior también es cierto en el caso de gran variedad de ocupaciones, desde los trabajadores de la construcción hasta la reparación automotriz y los empleos de telemarketing: ocupaciones para las que, según se cree, la apariencia no es importante.

los pedazos de papel se han mezclado bien, se hace la primera selección extrayendo una papeleta de la caja sin verla. Este proceso se repite hasta terminar de elegir el tamaño de la muestra de 52.

Un método más conveniente para seleccionar una muestra aleatoria simple consiste en utilizar el número de identificación de cada uno de los empleados y una **tabla de números aleatorios** como la que aparece en el Apéndice E. Como su nombre lo dice, estos números se generaron a través de un proceso aleatorio (en este caso por medio de una computadora). La probabilidad de 0, 1, 2, ..., 9 es la misma para cada dígito de un número. Por consecuencia, la probabilidad de que se seleccione el empleado 011 es la misma que la del empleado 722 o el 382. Al utilizar números aleatorios para seleccionar empleados, se elimina la influencia o sesgo del proceso de selección.

En la ilustración siguiente aparece una porción de una tabla de números aleatorios. Para seleccionar una muestra de empleados, elija primero un punto de partida en la tabla; cualquier punto es adecuado. Suponga que la hora que marca nuestro reloj es 3:04. Puede observar la tercera columna y luego desplazarse hacia abajo hasta el cuarto conjunto de números. El número es 03759. Como sólo hay 845 empleados, utilizaremos los tres primeros dígitos de un número aleatorio de cinco dígitos. Por tanto, 037 es el número del primer empleado que será miembro de la muestra. Otra manera de elegir el punto de partida consiste en cerrar los ojos y señalar un número en la tabla. Para seguir escogiendo empleados, podría 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 empleado seleccionado en segundo término para integrar la muestra. El siguiente número de tres dígitos a la derecha es 961. Omite 961 porque sólo hay 845 empleados. Continúe hacia la derecha y seleccione el empleado 784, después 189 y así sucesivamente.

5 0 5 2 5	5 7 4 5 4	2 8 4 5 5	6 8 2 2 6	3 4 6 5 6	3 8 8 8 4	3 9 0 1 8
7 2 5 0 7	5 3 3 8 0	5 3 8 2 7	4 2 4 8 6	5 4 4 6 5	7 1 8 1 9	9 1 1 9 9
3 4 9 8 6	7 4 2 9 7	0 0 1 4 4	3 8 6 7 6	8 9 9 6 7	9 8 8 6 9	3 9 7 4 4
6 8 8 5 1	2 7 3 0 5	0 3 7 5 9	4 4 7 2 3	9 6 1 0 8	7 8 4 8 9	1 8 9 1 0
0 6 7 3 8	6 2 8 7 9	0 3 9 1 0	1 7 3 5 0	4 9 1 6 9	0 3 8 5 0	1 8 9 1 0
1 1 4 4 8	1 0 7 3 4	0 5 9 3 7	2 4 3 9 7	1 0 4 2 0	1 6 7 1 2	9 4 4 9 6
		↓	↓		↓	↓
		Punto de partida	Segundo empleado		Tercer empleado	cuarto empleado

### Autoevaluación 8-1



La siguiente lista de alumnos presenta los estudiantes que se inscribieron a una clase de inducción a la estadística administrativa. Se van a seleccionar tres estudiantes al azar a los que se les van a hacer varias preguntas acerca del contenido del curso y el método de enseñanza.

- Los números 00 a 45 se escriben a mano 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?
- Ahora utilice la tabla de dígitos aleatorios, Apéndice E, para seleccionar su propia muestra.
- ¿Qué haría si encontrara el número 59 en la tabla de números aleatorios? (Ver tabla de la página siguiente.)

## Muestreo aleatorio sistemático

El procedimiento de muestreo aleatorio simple puede resultar complicado en algunas situaciones de investigación. Suponga, por ejemplo, que la división de ventas de Computers Printers Unlimited necesita estimar rápidamente el ingreso medio en dólares por venta durante el mes pasado. Encontraron que se habían registrado 2 000 recibos de ventas y se habían almacenado en cajones de un archivero, y decidieron seleccionar 100 recibos para calcular el ingreso medio en dólares. El muestreo aleatorio simple requiere que se numere cada recibo antes de utilizar la tabla de números aleatorios para seleccionar los 100 recibos. El proceso de numeración puede tardar mucho. En lugar de ello, es posible utilizar el **muestreo aleatorio sistemático**.

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

**MUESTRA ALEATORIA SISTEMÁTICA** Se selecciona un punto de inicio aleatorio y después se elige cada  $k$  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. Para Computers Printers Unlimited, seleccionaríamos cada 20 recibos (2 000/100) de los cajones del archivero; al hacerlo evitamos el proceso de numerar. Si  $k$  no es un número completo, hay que redondearlo.

En la selección del primer recibo se utiliza el muestreo aleatorio simple. Por ejemplo, se seleccionará un número de una tabla de números aleatorios entre 1 y  $k$ , o sea 20. Digamos que el número aleatorio resultó ser 18; entonces, empezando por el recibo 18, cada 20 recibos (18, 38, 58, etc.) se seleccionará como la muestra.

Antes de utilizar el muestreo aleatorio sistemático, debemos observar con todo cuidado el orden físico de la población. Cuando el orden físico se relaciona con la característica de la población, entonces no se debe utilizar el muestreo aleatorio sistemático. Por ejemplo, si los recibos que se mencionan en el ejemplo se archivaran en orden de incremento en ventas, el muestreo aleatorio sistemático no garantizaría una muestra aleatoria. Debemos utilizar otros métodos de muestreo.

## Muestreo aleatorio estratificado

Cuando una población es susceptible de dividirse en grupos con base en algunas características, se puede utilizar el **muestreo aleatorio estratificado** para garantizar que cada grupo esté representado en la muestra. A los grupos también se les llama **estratos**. Los estudiantes universitarios, por ejemplo, se pueden agrupar como de tiempo completo o medio tiempo, de sexo masculino o femenino, regulares o no regulares. Una vez que se han definido los estratos, podemos aplicar el muestreo aleatorio simple en cada grupo o estrato para formar la muestra.

**MUESTRA ALEATORIA ESTRATIFICADA** Una población se divide en subgrupos, llamados estratos, y selecciona al azar un elemento de cada estrato.

Por ejemplo, podríamos estudiar los gastos en publicidad de las 352 empresas más grandes de Estados Unidos. Suponga que el objetivo del estudio es determinar si las compañías con altos rendimientos sobre el capital (una medida de la productividad) gastan una mayor parte del dinero ganado en las ventas en propaganda que aquellas empresas que registran un bajo rendimiento o un déficit. Para asegurarse de que la muestra es una representación imparcial de las 352 empresas, las compañías se agrupan de acuerdo con su rendimiento porcentual sobre el capital. La Tabla 8-1 incluye los estratos y las frecuencias relativas. Si se utilizara el muestreo aleatorio simple, observe que las compañías que se encuentran en los estratos 3º y 4º tienen una alta probabilidad de ser seleccionadas (una probabilidad de 0.87), mientras que aquellas que aparecen en los otros estratos tienen pocas probabilidades de resultar seleccionadas (probabilidad de 0.13). Podríamos no elegir ninguna de las empresas que aparecen en los estratos 1 o 5 simplemente por casualidad. Sin embargo, el muestreo aleatorio estratificado garantizará que por lo menos una empresa de los estratos 1 y 5 estén representadas en la muestra. Digamos que se seleccionan 50 compañías para efectuar un estudio minucioso. Entonces se seleccionará 1 ( $0.02 \times 50$ ) empresa del estrato 1 en forma aleatoria, 5 ( $0.10 \times 50$ ) del estrato 2 y así sucesivamente. En este caso, el número de empresas de 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 precisión 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 en la muestra
1	30 por ciento y más	8	0.02	1*
2	20 arriba de 30 por ciento	35	0.10	5*
3	10 arriba de 20 por ciento	189	0.54	27
4	0 arriba de 10 por ciento	115	0.33	16
5	Déficit	5	0.01	1
Total		352	1.00	50

\* $0.02$  de  $50 = 1$ ,  $0.10$  de  $50 = 5$ , etc.

## Muestreo por conglomerados

Otro tipo de muestreo común es el **muestreo por conglomerados**. Éste se utiliza a menudo a fin de reducir el costo del muestreo de una población dispersa en una región geográfica extensa.

**MUESTREO POR CONGLOMERADOS.** Una población se divide en grupos o conglomerados utilizando los límites naturales geográficos o de otros tipos. Luego, los grupos se seleccionan al azar y se recopila una muestra al elegir en forma aleatoria elementos de cada grupo.

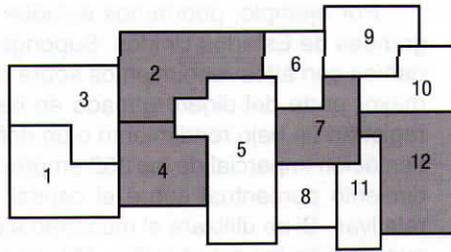
Suponga que desea determinar las opiniones de los residentes en algún estado en particular acerca de las políticas federales y estatales de protección del ambiente. Seleccionar una muestra aleatoria de residentes en el estado y contactar personalmente a cada persona requeriría demasiado tiempo y resultaría muy costoso. En lugar de ello, podría utilizar el muestreo por conglomerado subdividiendo el estado en unidades pequeñas, ya sea condados o regiones, que con frecuencia se conocen como *unidades primarias*.

Suponga que dividió el estado en 12 unidades primarias, luego seleccionó al azar cuatro regiones, 2, 7, 4 y 12, y posteriormente centrar sus esfuerzos en estas unidades primarias. Podría tomar una muestra aleatoria de los residentes en cada una de estas regiones y entrevistarlos. (Observe que ésta es una combinación de muestreo por aglomerado y muestreo aleatorio simple.)



### Estadística en acción

Los métodos de muestreo aleatorio y sin sesgo son muy importantes para hacer inferencias estadísticas válidas. En 1936 se realizó una votación para pronosticar el resultado de la competencia presidencial entre Franklin Roosevelt y Alfred Landon. Se enviaron diez millones de papeletas en forma de tarjetas postales con porte pagado a direcciones tomadas de directorios telefónicos y de registros automovilísticos. Una alta proporción de las papeletas se regresaron, con 59% en favor de Landon y 41% en favor de Roosevelt. El día de las elecciones, Roosevelt ganó con 61% de los votos, y Landon obtuvo 39%. ¡Es evidente que, a mediados de la década de 1930, la gente que tenía teléfono y conducía un automóvil no representaba a los votantes estadounidenses!



Muchos otros métodos de muestreo.

El análisis de los métodos de muestreo en las secciones anteriores no incluyó todos los métodos de muestreo que tiene a su disposición un investigador. Si usted participara en un importante proyecto de investigación sobre mercadotecnia, finanzas, contabilidad u otras áreas, necesitaría consultar libros dedicados a la teoría del muestreo y al diseño de muestras.

### Autoevaluación 8-2



Consulte la Autoevaluación 8-1 (página 253) y la lista de alumnos en la página 254. Suponga que un muestreo aleatorio sistemático va a seleccionar a cada noveno estudiante inscrito en el curso. En un principio, se eligió al azar al cuarto alumno de la lista, y se le da el número 03. Sin olvidar que los números aleatorios empiezan con el 00, ¿qué estudiantes se van a elegir como miembros de la muestra?

## Ejercicios

- La siguiente es una lista de las tiendas de Marco's Pizza en el condado de Lucas. También indica si la tienda es de propiedad corporativa (C) o propiedad del administrador (A). Se va a seleccionar y a inspeccionar una muestra de cuatro establecimientos en cuanto a la conveniencia para el cliente, seguridad, 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

- Los números aleatorios seleccionados son 08, 18, 11, 54, 02, 41 y 54. ¿Qué tiendas se seleccionaron?
  - Utilice la tabla de números aleatorios para seleccionar su propia muestra de establecimientos.
  - Una muestra debe consistir de cada 7<sup>o</sup> establecimiento. El número 03 es el punto de partida. ¿Qué establecimientos se incluirán en la muestra?
  - Suponga que una muestra debe consistir de tres establecimientos, de los cuales dos son de propiedad corporativa y uno es propiedad del administrador. Seleccione una muestra apropiada.
- La siguiente es una lista de hospitales en las regiones de Cincinnati (Ohio) y de la región norte de Kentucky. También indica si el hospital es general médico/quirúrgico (M/Q) o de especialidades (E). Nos interesa calcular el número promedio de enfermeras que trabajan medio tiempo y tiempo completo en los hospitales del área.
    - Se va a seleccionar una muestra de cinco hospitales en forma aleatoria. Los números aleatorios son 09, 16, 00, 49, 54, 12 y 04. ¿Qué hospitales se incluyen en la muestra?

Número de identificación	Nombre	Dirección	Tipo	Número de identificación	Nombre	Dirección	Tipo
00	Bethesda North	10500 Montgomery Cincinnati, Ohio 45242	M/Q	15	Providence Hospital	2446 Kipling Avenue Cincinnati, Ohio 45239	M/Q
01	Ft. Hamilton-Hughes	630 Eaton Avenue Hamilton, Ohio 45013	M/Q	16	St. Francis-St. George Hospital	3131 Queen City Avenue Cincinnati, Ohio 45238	M/Q
02	Jewish Hospital-Kenwood	4700 East Galbraith Rd. Cincinnati, Ohio 45236	M/Q	17	St. Elizabeth Medical Center, North Unit	401 E. 20th Street Covington, Kentucky 41014	M/Q
03	Mercy Hospital-Fairfield	3000 Mack Road Fairfield, Ohio 45014	M/Q	18	St. Elizabeth Medical Center, South Unit	One Medical Village Edgewood, Kentucky 41017	M/Q
04	Mercy Hospital-Hamilton	100 Riverfront Plaza Hamilton, Ohio 45011	M/Q	19	St. Luke's Hospital West	7380 Turfway Dr. Florence, Kentucky 41075	M/Q
05	Middletown Regional	105 McKnight Drive Middletown, Ohio 45044	M/Q	20	St. Luke's Hospital East	85 North Grand Avenue Ft. Thomas, Kentucky 41042	M/Q
06	Clermont Mercy Hospital	3000 Hospital Dr. Batavia, Ohio 45103	M/Q	21	Care Unit Hospital Cinti.	3156 Glenmore Avenue Cincinnati, Ohio 45211	E
07	Mercy Hospital-Anderson	7500 State Road Cincinnati, Ohio 45255	M/Q	22	Emerson Behavioral Science	2446 Kipling Avenue Cincinnati, Ohio 45239	E
08	Bethesda Oak Hospital	619 Oak Street Cincinnati, Ohio 45206	M/Q	23	Pauline Warfield Lewis Center for Psychiatric Treat.	1101 Summit Rd. Cincinnati, Ohio 45237	E
09	Children's Hospital Medical Center	3333 Burnet Avenue Cincinnati, Ohio 45229	M/Q	24	Children's Psychiatric No. Kentucky	502 Farrell Drive Covington, Kentucky 41011	E
10	Christ Hospital	2139 Auburn Avenue Cincinnati, Ohio 45219	M/Q	25	Drake Center Rehab—Long Term	151 W. Galbraith Road Cincinnati, Ohio 45216	E
11	Deaconess Hospital	311 Straight Street Cincinnati, Ohio 45219	M/Q	26	No. Kentucky Rehab Hospital—Short Term	201 Medical Village Edgewood, Kentucky	E
12	Good Samaritan Hospital	375 Dixmyth Avenue Cincinnati, Ohio 45220	M/Q	27	Shriners Burns Institute	3229 Burnet Avenue Cincinnati, Ohio 45229	E
13	Jewish Hospital	3200 Burnet Avenue Cincinnati, Ohio 45229	M/Q	28	VA Medical Center	3200 Vine Cincinnati, Ohio 45220	E
14	University Hospital	234 Goodman Street Cincinnati, Ohio 45267	M/Q				

- b. Utilice una tabla de números aleatorios para desarrollar su propia muestra de cinco hospitales.
  - c. Una muestra debe incluir cada 5<sup>o</sup> establecimiento. Seleccionamos 02 como el punto de partida. ¿Qué hospitales se incluirán en la muestra?
  - d. Una muestra debe consistir en cuatro hospitales médico/quirúrgicos y un hospital de especialidades. Seleccione una muestra adecuada.
3. A continuación, presentamos una lista de los 35 miembros de la Metro Toledo Automobile Dealers Association. Quisiéramos estimar 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	12	Spurgeon Chevrolet Motor Sales, Inc.	24	Lexus of Toledo
01	Autofair Nissan	13	Dunn Chevrolet	25	Mathews Ford Oregon, Inc.
02	Autofair Toyota-Suzuki	14	Don Scott Chevrolet-Pontiac-Geo, Inc.	26	Northtowne Chevrolet-Geo
03	George Ball's Buick GMC Truck	15	Dave White Chevrolet Co.	27	Quality Ford Sales, Inc.
04	Yark Automotive Group	16	Dick Wilson Pontiac	28	Rouen Chrysler Jeep Eagle
05	Bob Schmidt Chevrolet	17	Doyle Pontiac Buick	29	Saturn of Toledo
06	Bowling Green Lincoln Mercury Jeep Eagle	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
11	Thayer Chevrolet Geo Toyota	23	Kistler Ford, Inc.		

- Queremos seleccionar 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?
  - Utilice la tabla de números aleatorios para seleccionar su propia muestra de cinco distribuidores.
  - Una muestra debe consistir en cada 7º distribuidor. Se selecciona como punto de partida el número 04. ¿Qué distribuidores se incluyen en la muestra?
4. A continuación se enumeran los 27 agentes de seguros de Nationwide Insurance en el área metropolitana de Toledo, Ohio. Queremos calcular el número medio de años que han trabajado en Nationwide.

Número de identificación	Agente	Número de identificación	Agente	Número de identificación	Agente
00	<b>Bly Scott</b> 3332 W Laskey Rd	09	<b>Harris Ev</b> 2026 Albon Rd	18	<b>Priest Harvey</b> 5113 N Summit St
01	<b>Coyle Mike</b> 5432 W Central Av	10	<b>Heini Bernie</b> 7110 W Central	19	<b>Riker Craig</b> 2621 N Reynolds Rd
02	<b>Denker Brett</b> 7445 Airport Hwy	11	<b>Hinckley Dave</b> 14 N Holland Sylvania Rd	20	<b>Schwab Dave</b> 572 W Dussel Dr
03	<b>Denker Rollie</b> 7445 Airport Hwy	12	<b>Joehlin Bob</b> 3358 Navarre Av	21	<b>Seibert John H</b> 201 S Main
04	<b>Farley Ron</b> 1837 W Alexis Rd	13	<b>Keisser David</b> 3030 W Sylvania Av	22	<b>Smithers Bob</b> 229 Superior St
05	<b>George Mark</b> 7247 W Central Av	14	<b>Keisser Keith</b> 5902 Sylvania Av	23	<b>Smithers Jerry</b> 229 Superior St
06	<b>Gibellato Carlo</b> 6616 Monroe St 3521 Navarre Av	15	<b>Lawrence Grant</b> 342 W Dussel Dr	24	<b>Wright Steve</b> 105 S Third St
07	<b>Glemser Cathy</b> 5602 Woodville Rd	16	<b>Miller Ken</b> 2427 Woodville Rd	25	<b>Wood Tom</b> 112 Louisiana Av
08	<b>Green Mike</b> 4149 Holland Sylvania Rd	17	<b>O'Donnell Jim</b> 7247 W Central Av	26	<b>Yoder Scott</b> 6 Willoughby Av

- Deseamos seleccionar una muestra aleatoria de cuatro agentes. Los números aleatorios son: 02, 59, 51, 25, 145, 29, 77, 69 y 18. ¿Qué distribuidores se van a incluir en la muestra?
- Utilice la tabla de números aleatorios para seleccionar su propia muestra de cuatro agentes.
- Una muestra consiste de cada 7º distribuidor. El número 04 se selecciona como punto de partida. ¿Qué agentes se van a incluir en la muestra?

## “Error” de muestreo

En la sección anterior se analizaron métodos de muestreo que pueden utilizarse para seleccionar una muestra que es una representación imparcial o sin sesgos de la población. En cada método es importante señalar que la selección de cualquier muestra posible de un tamaño específico de una población tiene una posibilidad o probabilidad conocida. Ésta es otra forma de describir un método de muestreo sin sesgo.

Las muestras se utilizan para estimar características de la población. Por ejemplo, la media de una muestra se utiliza para estimar la media de la población. Sin embargo, como la muestra es una parte o porción representativa de la población, es poco probable que la media de la muestra sea *exactamente igual* a la media de la población. De manera análoga, es poco probable que la desviación estándar de la muestra sea *exactamente igual* a la desviación estándar de la población. Por tanto, podemos esperar una diferencia entre el *estadístico de una muestra* y el *parámetro de la población* correspondiente. A la diferencia se le llama **error de muestreo**.

**ERROR DE MUESTREO** La diferencia entre un estadístico de la muestra y el parámetro de la población correspondiente.

Tome la población formada por cinco empleados de Kiser Industries. La semana pasada, la producción de cada empleado fue 97, 103, 96, 99 y 105 unidades. Suponga que seleccionamos una muestra de dos empleados y que la producción de los dos fue 97 y 105, respectivamente. La media de esta muestra es 101, que se encuentra por medio de  $(97 + 105)/2$ . Otra muestra de dos empleados dio como resultado producciones de 103 y 96; por tanto, la media de esta muestra es 99.5. No obstante, la media de *todas* las producciones (la media de la población) es 100, que se calcula así

$$\mu = \frac{\sum X}{N} = \frac{97 + 103 + 96 + 99 + 105}{5} = 100$$

El error de muestreo para la primera muestra es 1.0, calculado  $\bar{X} - \mu = 101 - 100$ . El error de muestreo para la segunda muestra es  $-0.5$ , calculado  $\bar{X} - \mu = 99.5 - 100$ . Cada una de estas diferencias, 1.0 y  $-0.5$ , es el error de muestreo que se cometió al estimar la media de la población con base en la media de la muestra. Estos errores de muestreo se deben a la casualidad.

En este ejemplo, cada una de las 10 muestras posibles de tamaño 2 [véase la fórmula (5-10)], correspondiente a la regla de la combinación, que aparece en la página 168) tiene la misma posibilidad de ser seleccionada. Cada muestra puede tener una media de la muestra distinta y un error de muestreo diferente. El valor del error de muestreo se basa en la selección aleatoria de una muestra. Por tanto, los errores de muestreo son aleatorios y ocurren por casualidad.

Ahora que hemos descubierto la posibilidad de que ocurra un error de muestreo cuando los resultados de un muestreo se utilizan para estimar un parámetro de una población, ¿cómo podemos hacer una proyección precisa acerca del posible éxito de una pasta dentífrica o de otro producto recién desarrollado con base en los resultados de la muestra únicamente? ¿Cómo puede el departamento de aseguramiento de la calidad de una empresa de producción en serie liberar un embarque basándose sólo en una muestra de 10 chips? ¿Cómo pueden las organizaciones que levantan encuestas de *CNN/USA Today* o *ABC News/Washington Post* hacer un pronóstico exacto acerca de una elección presidencial con base en una muestra de 2 000 votantes registrados de una población de casi 90 millones de votantes? Para responder estas preguntas, primero desarrollamos una *distribución muestral de medias*.

## Distribución muestral de medias

El resultado de producción en el ejemplo anterior demostró que las medias para muestras de un tamaño específico varían de una muestra a otra. La producción media de la primera muestra de dos empleados fue 101, y la media de la segunda muestra fue 99.5. Es probable que una tercera muestra dé como resultado una media distinta. La media de la población fue 100. Si organizamos las medias de todas las muestras posibles de dos empleados en una distribución de probabilidad, obtendremos la distribución muestral de medias.

**DISTRIBUCIÓN MUESTRAL DE MEDIAS** Una distribución de probabilidad de todas las medias posibles de las muestras de un determinado tamaño de muestra de la población.

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

## EJEMPLO

Tartus Industries cuenta con siete empleados de producción (se considera como la población). En la Tabla 8-2 se proporcionan los ingresos por hora de cada empleado.

TABLA 8-2 Ingresos por hora de producción de Tartus Industries

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

1. ¿Cuál es la media de la población?
2. ¿Cuál es la distribución muestral de medias para muestras de tamaño 2?