

ANÁLISIS DEL ESTUDIO DE CASO

PROBLEMA DE LAS CARACTERÍSTICAS DEL BARRIO

Al principio de este capítulo se sugirió el uso de los datos de la Oficina de Censos sobre estadísticas de manzana para estimar la proporción de residentes con una edad mayor o igual a 65 años en un área de 40 manzanas. Las $n = 5$ manzanas fueron muestreadas aleatoriamente de las 40 y se obtuvieron los siguientes datos:

Número de residentes, m_i	Personas con 65 años o más, a_i	$\hat{p}m_i$	$a_i - \hat{p}m_i$	$(a_i - \hat{p}m_i)^2$
90	15	21.60	-6.60	43.5600
32	8	7.68	0.32	0.1024
47	14	11.28	2.72	7.3984
25	9	6.00	3.00	9.0000
16	4	3.84	0.16	0.0256
210	50			60.0864

$$\hat{p} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n m_i} = \frac{50}{210} = 0.24$$

Así que la mejor estimación de la proporción de personas con edad igual o mayor a 65 años es 0.24.

El límite para el error de estimación es

$$\begin{aligned} 2\sqrt{V(\hat{p})} &= 2\sqrt{\left(\frac{N-n}{Nnm^2}\right)\left(\frac{1}{n-1}\right)\sum_{i=1}^n (a_i - pm_i)^2} \\ &= 2\sqrt{\left[\frac{35}{(40)(5)(42)^2}\right]\left(\frac{1}{4}\right)(60.0864)} \\ &= 0.08 \end{aligned}$$

Entonces la estimación de la proporción verdadera para el área de 40 manzanas es 0.24 ± 0.08 o bien 0.16 a 0.32. Tenemos confianza en que más del 16% de los residentes tiene una edad igual o mayor a 65 años.

EJERCICIOS

- 8.1 Una experimentadora que trabaja en un área urbana desea estimar el valor promedio de una variable altamente correlacionada con el renta. Ella piensa que debe usar muestreo por conglomerados, con manzanas como conglomerados y adultos dentro de manzanas como elementos.

Explique por qué se debería o no usar muestreo por conglomerados en cada una de las siguientes situaciones.

- La mayoría de los adultos en ciertas manzanas son blancos y la mayoría son no blancos en otras manzanas.
- La proporción de no blancos es la misma en cada bloque y no está cercana a 1 o a 0.
- La proporción de no blancos difiere de manzana a manzana en la manera que se podría esperar si los conglomerados fueran hechos asignando aleatoriamente los adultos de la población a los conglomerados.

8.2.

Un fabricante de sierras de cinta quiere estimar el costo de reparación promedio mensual para las sierras que ha vendido a ciertas industrias. El fabricante no puede obtener un costo de reparación para cada sierra, pero puede obtener la cantidad total gastada en reparación y el número de sierras que tiene cada industria. Entonces decide usar muestreo por conglomerados, con cada industria como un conglomerado. El fabricante selecciona una muestra irrestricta aleatoria de $n = 20$ de $N = 96$ industrias a las que da servicio. Los datos sobre costo total de reparaciones por industria y el número de sierras por industria se presentan en la tabla anexa. Estime el costo promedio de reparación por sierra para el mes pasado, y establezca un límite para el error de estimación.

Industria	Número de sierras	Costo total de reparación para el mes pasado (en dólares)	Industria	Número de sierras	Costo total de reparación para el mes pasado (en dólares)
1	3	50	11	8	140
2	7	110	12	6	130
3	11	230	13	3	70
4	9	140	14	2	50
5	2	60	15	1	40
6	12	280	16	4	60
7	14	240	17	12	280
8	3	45	18	6	150
9	5	60	19	5	110
10	9	230	20	8	120

- Para los datos en el Ejercicio 8.2, estime la cantidad total gastada por las 96 industrias en la reparación de sierras. Establezca un límite para el error de estimación.
- Después de verificar sus registros de ventas, el fabricante del Ejercicio 8.2 se percató de que ha vendido un total de 710 sierras a esas industrias. Usando esta información adicional, estime la cantidad total gastada en reparación de sierras por estas industrias, y establezca un límite para el error de estimación.
- El mismo fabricante (Ejercicio 8.2) quiere estimar el costo de reparación promedio por sierra para el mes siguiente. ¿Cuántos conglomerados debe seleccionar en la muestra si quiere que el límite para el error de estimación sea menor que \$2.00?
- Un politólogo desarrolla una prueba para medir el grado de conocimiento sobre acontecimientos actuales. Él quiere estimar la calificación promedio que obtendrán en su prueba todos los estudiantes de una escuela preparatoria. La administración de la escuela no le permitirá seleccionar aleatoriamente a los estudiantes fuera de clases, pero sí interrumpir un pequeño número de clases con el propósito de aplicar la prueba a cada miembro de la clase. Entonces el experimentador selecciona al azar 25 clases de un total de 108 a una hora determinada. Se aplica la prueba a cada miembro de las clases muestreadas, con los resultados que se presentan en la tabla anexa.

Estime la calificación promedio que sería obtenida para esta prueba por todos los estudiantes en la escuela. Establezca un límite para el error de estimación.

Clase	Número de estudiantes	Calificación total	Clase	Número de estudiantes	Calificación total
1	31	1590	14	40	1980
2	29	1510	15	38	1990
3	25	1490	16	28	1420
4	35	1610	17	17	900
5	15	800	18	22	1080
6	31	1720	19	41	2010
7	22	1310	20	32	1740
8	27	1427	21	35	1750
9	25	1290	22	19	890
10	19	860	23	29	1470
11	30	1620	24	18	910
12	18	710	25	31	1740
13	21	1140			

- 8.7 El politólogo del Ejercicio 8.6 quiere estimar la calificación promedio en la prueba para una escuela preparatoria similar. Él quiere que el límite para el error de estimación sea menor que 2 puntos. ¿Cuántas clases debe tomar en la muestra? Supóngase que la escuela tiene 100 clases durante cada hora en este periodo escolar.
- 8.8 Una industria está considerando la revisión de su política de jubilación y quiere estimar la proporción de empleados que apoyan la nueva política. La industria consiste de 87 plantas separadas localizadas en todo Estados Unidos. Ya que los resultados deben ser obtenidos rápidamente y con poco dinero, la industria decide usar muestreo por conglomerados; con cada planta como un conglomerado. Se selecciona una muestra irrestricta aleatoria de 15 plantas y se obtienen las opiniones de los empleados en estas plantas a través de un cuestionario. Los resultados se presentan en la tabla anexa. Estime la proporción de empleados en la industria que apoyan la nueva política de jubilación y establezca un límite para el error de estimación.

Planta	Número de empleados	Número de empleados que apoyan la nueva política	Planta	Número de empleados	Número de empleados que apoyan la nueva política
1	51	42	9	73	54
2	62	53	10	61	45
3	49	40	11	58	51
4	73	45	12	52	29
5	101	63	13	65	46
6	48	31	14	49	37
7	65	38	15	55	42
8	49	30			

- 8.9 La industria del Ejercicio 8.8 modificó su política de jubilación después de obtener los resultados de la encuesta. Ahora se quiere estimar la proporción de empleados a favor de la política modificada. ¿Cuántas plantas deben ser muestreadas para tener un límite de 0.08 para el error de estimación? Use los datos del Ejercicio 8.8 para aproximar los resultados de la nueva encuesta.

- 8.10 Se diseña una encuesta económica para estimar la cantidad promedio gastada en servicios para el hogar en una ciudad. Ya que no se encuentra disponible una lista de hogares, se usa muestreo por conglomerados, con divisiones (barrios) formando los conglomerados. Se selecciona una muestra aleatoria de 20 barrios de la ciudad de un total de 60. Los entrevistadores obtienen el costo de los servicios de cada hogar dentro de los barrios seleccionados; los costos totales se muestran en la tabla anexa. Estime la cantidad promedio de gastos en servicios por hogar en la ciudad y establezca un límite para el error de estimación.

Barrio muestreado	Número de hogares	Cantidad total gastada en servicios	Barrio muestreado	Número de hogares	Cantidad total gastada en servicios
1	55	\$2210	11	73	\$2930
2	60	2390	12	64	2470
3	63	2430	13	69	2830
4	58	2380	14	58	2370
5	71	2760	15	63	2390
6	78	3110	16	75	2870
7	69	2780	17	78	3210
8	58	2370	18	51	2430
9	52	1990	19	67	2730
10	71	2810	20	70	2880

- 8.11 En la encuesta del Ejercicio 8.10 se desconoce el número de hogares en la ciudad. Estime la cantidad total gastada en servicios por todos los hogares de la ciudad y establezca un límite para el error de estimación.
- 8.12 La encuesta económica del Ejercicio 8.10 se va a llevar a cabo en una ciudad vecina de estructura similar. El objetivo es estimar la cantidad total gastada en servicios por los hogares de la ciudad, con un límite de \$ 5000 para el error de estimación. Use los datos del Ejercicio 8.10 para encontrar el número aproximado de conglomerados que se necesitan para obtener este límite.
- 8.13 Un inspector quiere estimar el peso promedio de llenado para cajas de cereal empacadas en una fábrica. El cereal está en paquetes que contienen 12 cajas cada uno. El inspector selecciona aleatoriamente 5 y mide el peso de llenado de cada caja en los paquetes muestreados, con los resultados (en onzas) que se muestran en la tabla acompañante. Estime el peso promedio de llenado para las cajas empacadas por esta fábrica, y establezca un límite para el error de estimación. Suponga que el número total de cajas empacadas por la fábrica es lo suficientemente grande para que no se tome en cuenta la corrección por población finita.

Paquete	Onzas de llenado											
1	16.1	15.9	16.1	16.2	15.9	15.7	16.1	16.2	16.0	15.9	15.8	16.0
2	15.9	16.2	15.8	16.0	16.3	16.1	15.8	15.9	16.0	16.1	16.1	15.9
3	16.2	16.0	16.2	16.3	15.8	16.0	15.9	16.0	16.1	16.0	15.9	16.1
4	15.9	16.1	16.2	16.1	16.1	16.3	15.9	16.1	15.9	15.9	16.0	16.0
5	16.0	15.8	16.3	15.7	16.1	15.9	16.0	16.1	15.8	16.0	16.1	15.9

- 8.14 Un periódico quiere estimar la proporción de votantes que apoyan a cierto candidato, candidato A, en una elección estatal. Ya que la selección y entrevista de una muestra irrestricta aleatoria de votantes registrados es muy costosa, se utiliza muestreo por conglomerados, con distritos como conglomerados. Se selecciona una muestra irrestricta aleatoria de 50 distritos de un total de 497 que tiene el estado. El periódico quiere hacer la estimación el día de la elección, pero antes de que se haya hecho la cuenta final de los votos. Es por eso que los reporteros son enviados a los lugares de votación de cada distrito en la muestra, para obtener la información pertinente directamente de los votantes. Los resultados se muestran en la tabla acompañante. Estime la proporción de votantes que apoyan al candidato A, y establezca un límite para el error de estimación.

Número de votantes	Número que vota por A	Número de votantes	Número que vota por A	Número de votantes	Número que vota por A
1290	680	1893	1143	843	321
1170	631	1942	1187	1066	487
840	475	971	542	1171	596
1620	935	1143	973	1213	782
1381	472	2041	1541	1741	980
1492	820	2530	1679	983	693
1785	933	1567	982	1865	1033
2010	1171	1493	863	1888	987
974	542	1271	742	1947	872
832	457	1873	1010	2021	1093
1247	983	2142	1092	2001	1461
1896	1462	2380	1242	1493	1301
1943	873	1693	973	1783	1167
798	372	1661	652	1461	932
1020	621	1555	523	1237	481
1141	642	1492	831	1843	999
1820	975	1957	932		

- 23-57 13044 29604 1322 24556 112
- 8.15 El periódico del Ejercicio 8.14 quiere realizar una encuesta similar durante la siguiente elección. ¿Qué tan grande se necesitará la muestra para estimar la proporción de votantes que favorecen un candidato similar, con un límite de 0.05 para el error de estimación?
- 8.16 Un guardabosques desea estimar la altura promedio de los árboles en una plantación. La plantación se divide en parcelas de un cuarto de acre. Se selecciona una muestra irrestricta aleatoria de 20 parcelas de un total de 386 parcelas en la plantación. Se miden todos los árboles en las parcelas muestreadas, con los resultados que se muestran en la tabla anexa. Estime la altura promedio de los árboles en la plantación y establezca un límite para el error de estimación. (Sugerencia: el total para el conglomerado i se puede encontrar tomando m_i veces el promedio del conglomerado.)

Número de árboles	Altura promedio (en pies)	Número de árboles	Altura promedio (en pies)
42	6.2	60	6.3
51	5.8	52	6.7
49	6.7	61	5.9
55	4.9	49	6.1

47	5.2	57	6.0
58	6.9	63	4.9
43	4.3	45	5.3
59	5.2	46	6.7
48	5.7	62	6.1
41	6.1	58	7.0

- 8.17 Para reafirmar la seguridad, una compañía de taxis quiere estimar la proporción de llantas inseguras en sus 175 taxis. (No considere las llantas de refacción.) La selección de una muestra aleatoria de llantas es impráctica, así que se usa muestreo por conglomerados, con cada taxi como un conglomerado. Una muestra irrestricta aleatoria de 25 taxis nos da los siguientes números de llantas inseguras por taxi:

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

- Estime la proporción de llantas inseguras que se están usando en la compañía de taxis, y establezca un límite para el error de estimación.
- 8.18 Los comercios solicitan frecuentemente a los contadores la realización de inventarios. Ya que un inventario completo es costoso, a través del muestreo se pueden realizar inventarios cada cuatro meses. Supóngase que una empresa abastecedora de artículos de plomería desea un inventario para muchos artículos pequeños en existencia. La obtención de una muestra aleatoria de artículos es muy difícil. Sin embargo, los artículos se encuentran dispuestos en anaqueles, y la selección de una muestra aleatoria de anaqueles es relativamente fácil, considerando a cada anaquel como un conglomerado de artículos. Una muestra de 10 anaqueles de un total de 48 dio los resultados que se muestran en la tabla siguiente. Estime la cantidad total de dólares de los artículos en los anaqueles y establezca un límite para el error de estimación.

Conglomerado	Número de artículos, m_i	Cantidad total de dólares, y_i
1	42	83
2	27	62
3	38	45
4	63	112
5	72	96
6	12	58
7	24	75
8	14	58
9	32	67
10	41	80

- 8.19 Una empresa especializada en la fabricación y venta de ropa de descanso tiene 80 almacenes en Florida y 140 en California. Con cada estado como un estrato, la empresa desea estimar el tiempo promedio de ausencia por enfermedad por empleado durante el año pasado. Cada almacén puede ser considerado como un conglomerado de empleados, y se puede determinar de los registros el tiempo total de ausencia por enfermedad para cada almacén. Muestras irrestrictas aleatorias de 8 almacenes de Florida y 10 almacenes de California nos dan los resultados que se muestran en la tabla acompañante (m_i denota el número de empleados y y_i denota el total de días de ausencia por enfermedad para el i -ésimo almacén). Estime la cantidad promedio de ausencia por enfermedad por empleado, y calcule un estimador de la varianza de su estimador.

Florida		California	
m_i	y_i	m_i	y_i
12	40	16	51
20	52	8	32
8	30	4	11
14	36	3	10
24	71	12	33
15	48	17	39
10	39	24	61
6	21	30	37
		21	40
		9	41

- 8.20 Las estadísticas de manzana reportan el número de unidades habitacionales, el número de residentes y el número total de cuartos dentro de las unidades habitacionales para una muestra aleatoria de ocho manzanas seleccionadas de una gran ciudad. (Suponga que el número de manzanas en la ciudad es muy grande.) Los datos se presentan en la tabla acompañante.

Manzana	Número de unidades habitacionales	Número de residentes	Número de cuartos
1	12	40	58
2	14	39	72
3	3	12	26
4	20	52	98
5	12	37	74
6	8	33	57
7	10	41	76
8	6	14	48

- (a) Estime el número promedio de residentes por unidad habitacional y establezca un límite para el error de estimación.
 (b) Estime el número promedio de cuartos por residente y establezca un límite para el error de estimación.

- 8.21 Cierta tipo de tableros de circuitos fabricados para su instalación en computadoras tiene 12 microcircuitos por tablero. Durante la inspección de control de calidad de 10 de esos tableros, el número de microcircuitos defectuosos por tablero fue como sigue:

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

Estime la proporción de microcircuitos defectuosos en la población de la cual se extrajo la muestra y establezca un límite para el error de estimación.

- 8.22 Considere la situación del Ejercicio 8.21. Suponga que la muestra utilizada proviene de un embarque de 50 de tales tableros. Estime el número total de microcircuitos defectuosos en este embarque y establezca un límite para el error de estimación.

- 8.23 Una empresa grande tiene sus inventarios de equipo listados separadamente por departamento. De los 15 departamentos en la empresa, se van a muestrear aleatoriamente 5, por un auditor que

va a verificar que todo el equipo esté identificado y localizado apropiadamente. La proporción de artículos del equipo que no estén identificados propiamente es de interés al auditor. Los datos se dan en la tabla siguiente. Estime la proporción de artículos del equipo en la empresa que no están identificados propiamente y establezca un límite para el error de estimación.

Departamento	Número de artículos del equipo	Número de artículos identificados inapropiadamente
1	15	2
2	27	3
3	9	1
4	31	1
5	16	2

- 8.24 Suponga que para la empresa del Ejercicio 8.23, los 15 departamentos tienen el número de artículos del equipo que se da en la tabla acompañante. Seleccione una muestra de 3 departamentos, con probabilidades proporcionales al número de artículos del equipo.

Departamento	Número de artículos	Departamento	Número de artículos
1	12	9	31
2	9	10	26
3	27	11	22
4	40	12	19
5	35	13	16
6	15	14	33
7	18	15	6
8	10		

- 8.25 Suponga que los tres departamentos seleccionados en el Ejercicio 8.24 tienen cada uno dos artículos del equipo identificados inapropiadamente. Estime el número total de artículos impropriadamente identificados en la empresa y establezca un límite para el error de estimación.
- 8.26 Un gran embarque de mariscos congelados es empaquetado en cajas, conteniendo cada uno veinticuatro paquetes de 5 libras. Hay cien cajas en el embarque. Un inspector del gobierno determina el peso total (en libras) de mariscos dañados para cada una de cinco cajas muestreadas. Los datos son como sigue:

9, 6, 3, 10, 2

Estime el peso total de mariscos dañados en el embarque y establezca un límite para el error de estimación.

- 8.27 Usando los datos del Ejercicio 8.26, estime la cantidad promedio de mariscos dañados por paquete de 5 libras y establezca un límite para el error de estimación.
- 8.28 Un profesor desea muestrear a los estudiantes residentes de una universidad. Las unidades muestrales pueden ser convenientemente usadas como conglomerados de estudiantes.

lecciones de unidades habitacionales (dormitorios para estudiantes de primer año, casas de fraternidad, y así sucesivamente) pueden ser usadas como estratos. Analice los méritos de muestreo por conglomerados contra muestreo aleatorio estratificado, si el objetivo es estimar la proporción de estudiantes que favorecen a cierto candidato en los siguientes tipos de elecciones.

(a) Una elección de dirigentes estudiantiles.

(b) Una elección del presidente del país.

- 8.29 ¿En qué condiciones el muestreo por conglomerados produce un límite más pequeño para el error de estimación de una media que el muestreo irrestricto aleatorio?
- 8.30 Sin considerar los costos de muestreo, ¿qué criterio usaría usted para seleccionar conglomerados apropiados en un problema de muestreo por conglomerados?

EXPERIENCIAS CON DATOS REALES

- 8.1 En la Tabla 3 del Apéndice se muestra el ingreso por persona en Estados Unidos (durante 1977). Se presentan también valores para la población de 1980. Tratando a cada estado como un conglomerado de personas, seleccione una muestra aleatoria de estados y estime el ingreso personal total para Estados Unidos. Establezca un límite para el error de estimación.
- 8.2 Trate de realizar un estudio económico, tal vez considerando los hogares en cierta área geográfica fija (tal vez unas cuantas manzanas de la ciudad) como conglomerados de personas. Seleccione una muestra de n hogares y, después de obtener el permiso para la entrevista, registre la cantidad total semanal que se gasta en alimentos por todos los individuos en el hogar, y el número de individuos. Entonces estime la cantidad promedio gastada en alimentos por persona entre los hogares de esta población. Aun si todo el dinero es realmente gastado por una persona (digamos la madre), la cantidad total es la misma que se hubiera registrado si cada individuo hubiera comprado su propia alimentación. Entonces, se dispone del total para el conglomerado, aun cuando puede no contarse con las observaciones por elemento.