11 LAS MUESTRAS ESTADÍSTICAS

Página 266

1. Una ganadería tiene 3 000 vacas. Se quiere extraer una muestra de 120. Explica cómo se obtiene la muestra:

a) Mediante muestreo aleatorio simple.

b) Mediante muestreo aleatorio sistemático.

a) — Se numeran las vacas del 1 al 3 000.

— Se sortean 120 números de entre los 3 000.

— La muestra estará formada por las 120 vacas a las que correspondan los números obtenidos.

b) Coeficiente de elevación: $h = \dfrac{3000}{120} = 25$

— Se sortea un número del 1 al 25. Supongamos que sale el 9.

— Las vacas seleccionadas para la muestra serían las que correspondieran a los números 9, 34, 59, 84, 109, …, 2 984.

Página 267

2. Una ganadería tiene 2 000 vacas. Son de distintas razas: 853 de A, 512 de B, 321 de C, 204 de D y 110 de E. Queremos extraer una muestra de 120:

a) ¿Cuántas hay que elegir de cada raza para que el muestreo sea estratificado con reparto proporcional?

b) ¿Cómo ha de ser la elección dentro de cada estrato?

a) Llamamos n_1 al número de vacas que debemos elegir de raza A, n_2 al de raza B, n_3 al de C, n_4 al de D y n_5 al de E.

Ha de cumplirse que:

$$\frac{120}{2000} = \frac{n_1}{853} = \frac{n_2}{512} = \frac{n_3}{321} = \frac{n_4}{204} = \frac{n_5}{110}$$

Así, obtenemos:

$$n_1 = 51{,}18 \quad n_2 = 30{,}72 \quad n_3 = 19{,}26 \quad n_4 = 12{,}24 \quad n_5 = 6{,}6$$

La parte entera de estos número suma:

51 + 30 + 19 + 12 + 6 = 118. Faltan 2 para llegar a 120.

Por tanto, debemos elegir:

51 vacas de raza A, 31 vacas de B, 19 de C, 12 de D y 7 de E.

Página 268

1. Obtén aleatoriamente cuatro números enteros al azar entre 1 y 95.

Por ejemplo:

RAN# 0.226 × 95 + 1 = 22.47 → 22

RAN# 0.048 × 95 + 1 = 5.56 → 5

RAN# 0.277 × 95 + 1 = 27.315 → 27

RAN# 0.842 × 95 + 1 = 80.99 → 80

Hemos obtenido los números 5, 22, 27 y 80.

2. Obtén cinco números enteros elegidos aleatoriamente entre 1 y 800.

Por ejemplo:

RAN# 0.104 × 800 + 1 = 84.2 → 84

RAN# 0.098 × 800 + 1 = 79.4 → 79

RAN# 0.835 × 800 + 1 = 669 → 669

RAN# 0.449 × 800 + 1 = 360.2 → 360

RAN# 0.622 × 800 + 1 = 498.6 → 498

Hemos obtenido los números 79, 84, 360, 498 y 669.

Página 269

3. De una población de N = 856 elementos, deseamos extraer una muestra de tamaño n = 10.

Mediante el uso de números aleatorios, designa cuáles son los 10 individuos que componen la muestra.

Para multiplicar por 856 los números que aparezcan en pantalla, introducimos:

856 × × (factor constante)

Ahora recurrimos a los números aleatorios. Por ejemplo, podemos obtener:

RAN# 0.835 = 714.76 → 715

RAN# 0.419 = 358.664 → 359

RAN# 0.554 = 474.224 → 475

RAN# 0.567 = 485.352 → 486

RAN# 0.530 = 453.68 → 454

RAN# 0.057 = 48.792 → 49

RAN# 0.993 = 850.008 → 851

RAN# 0.396 = 338.976 → 339

RAN# 0.013 = 11.128 → 12

RAN# 0.636 = 544.416 → 545

Los individuos elegidos para la muestra serían los correspondientes a los números 12, 49, 339, 359, 454, 475, 486, 545, 715 y 851.

4. De una población de 543 individuos, queremos extraer una muestra de tamaño 40 mediante números aleatorios.

Obtén los cinco primeros elementos de dicha muestra.

Para multiplicar por 543 los números que aparezcan en pantalla, introducimos:

543 × × (factor constante)

Ahora recurrimos a los números aleatorios. Por ejemplo, podemos obtener:

RAN# 0.237 = 128.691 → 129

RAN# 0.071 = 38.553 → 39

RAN# 0.614 = 333.402 → 334

RAN# 0.497 = 269.871 → 270

RAN# 0.475 = 257.925 → 258

Los cinco primeros elementos de la muestra serían los correspondientes a los números 129, 39, 334, 270 y 258.

Página 272

EJERCICIOS Y PROBLEMAS PROPUESTOS

PARA PRACTICAR

Muestras

1 En cada uno de los casos que se mencionan a continuación, el colectivo ¿es población o es muestra?

Explica por qué.

a) Un campesino tiene 87 gallinas. Para probar la eficacia de un nuevo tipo de alimentación, las pesa a todas antes y después de los 30 días que dura el tratamiento.

b) Un granjero prueba con 100 de sus gallinas la eficacia de un nuevo tipo de alimentación.

a) Es **población**, porque pesa a todas las gallinas.

b) Es **muestra**, porque no pesa a todas las gallinas, sino solo a una parte de ellas.

2 **Un fabricante de elásticos quiere estudiar su resistencia a la rotura. Para ello, los estira hasta que se rompen y anota el grado de estiramiento que alcanzan sin romperse.**

¿Puede realizar dicho estiramiento sobre la población o es imprescindible realizarlo sobre la muestra? ¿Por qué?

Es imprescindible hacerlo sobre una muestra, porque interesa romper la menor cantidad de elásticos posible.

3 **Solo uno de los siguientes procedimientos nos permite obtener una muestra representativa. Di cuál es y, en los otros, estudia el sentido del sesgo y su importancia:**

a) Para estudiar las frecuencias relativas de las letras, se toman al azar 20 libros de la biblioteca de un centro escolar y se cuenta las veces que aparece cada letra en la página 20 de los libros seleccionados.

b) Para conocer la opinión de sus clientes sobre el servicio ofrecido por unos grandes almacenes, se selecciona al azar, entre los que poseen tarjeta de compra, a 100 personas entre las que han gastado menos de 1 000 € el último año, otras 100 entre las que han gastado entre 1 000 € y 5 000 € y 100 más entre las que han gastado más de 5 000 €.

c) Para calcular el número medio de personas por cartilla en un Centro de Salud de la Seguridad Social, los médicos toman nota de las cartillas de las personas que acuden a las consultas durante un mes.

a) Es una muestra representativa.

b) No es representativa, porque hay mucha más gente en un intervalo (por ejemplo, entre 1 000 € y 5 000 €) que en otro (más de 5 000 €), y hemos tomado el mismo número de representantes. Además, hay otra mucha gente sin tarjeta que no se ha tomado en cuenta.

c) No es representativa, ya que lo que más se va a ver son las cartillas que corresponden a familias numerosas. Está claro que, cuanta más gente tenga esa cartilla, más fácil es que ese mes se tome nota de ella.

4 **De un colectivo de 500 personas elige una muestra de 20 mediante:**

a) Un muestreo aleatorio sistemático.

b) Un muestreo aleatorio simple.

Utiliza la tecla RAN# de la calculadora.

Para los dos casos, numeramos a las personas del 1 al 500.

a) $h = \dfrac{500}{20} = 25$

Origen: 25 × RAN# + 1 = 14.075 (por ejemplo)

Deberemos elegir las personas cuyos números sean:

14, 39, 64, 89, 114, 139, 164, 189, 214, 239, 264, 289, 314, 339, 364, 389, 414, 439, 464, 489.

b) Con la tecla RAN# de la calculadora, hacemos: 500 × × RAN# = hasta obtener 20 resultados distintos.

5 **En un conjunto de 1 000 conductores hay:**

— 50 taxistas.

— 75 camioneros.

— 25 conductores de autobús.

El resto son conductores de vehículos corrientes y se reparten así:

— 250 con más de 20 años de experiencia.

— 425 con una experiencia de entre 5 y 20 años.

— 175 con una experiencia de 0 a 5 años.

Para confeccionar una muestra de 40 individuos mediante muestreo aleatorio estratificado proporcional, ¿cuántos hay que seleccionar de cada uno de los seis estratos?

Llamamos n_1 al número de taxistas que tendríamos que seleccionar, n_2 al número de camioneros, n_3 al número de conductores de autobuses, n_4 al número de conductores con más de 20 años de experiencia, n_5 al de conductores con una experiencia entre 5 y 20 años y n_6 al de conductores con una experiencia de 0 a 5 años. Entonces:

$$\frac{n_1}{50} = \frac{n_2}{75} = \frac{n_3}{25} = \frac{n_4}{250} = \frac{n_5}{425} = \frac{n_6}{175} = \frac{40}{1\,000}$$

Así, deberemos elegir:

n_1 = 2 taxistas

n_2 = 3 camioneros

n_3 = 1 conductor de autobús

n_4 = 10 conductores con más de 20 años de experiencia

n_5 = 17 con experiencia entre 5 y 20 años

n_6 = 7 con experiencia entre 0 y 5 años

6 S **En determinada provincia hay cuatro comarcas, C1, C2, C3 y C4, con un total de 1 500 000 personas censadas. De ellas, 300 000 residen en C1, 450 000 en C2 y 550 000 en C3. Se quiere realizar un estudio sobre las costumbres alimenticias en esa provincia basado en una muestra de 3 000 personas.**

a) ¿Qué tipo de muestreo deberíamos realizar si queremos que en la muestra resultante haya representación de todas las comarcas?

b) ¿Qué número de personas habría que seleccionar en cada comarca, atendiendo a razones de proporcionalidad?

c) ¿Cómo seleccionarías las personas en cada comarca?

Justifica las respuestas.

a) Deberíamos realizar un muestreo aleatorio estratificado.

b) El número de personas que residen en C4 es:

$$1\,500\,000 - (300\,000 + 450\,000 + 550\,000) = 200\,000$$

Llamamos n_1, n_2, n_3 y n_4 al número de personas que tendríamos que seleccionar en cada comarca (C1, C2, C3 y C4, respectivamente). Entonces:

$$\frac{n_1}{300\,000} = \frac{n_2}{450\,000} = \frac{n_3}{550\,000} = \frac{n_4}{200\,000} = \frac{3\,000}{1\,500\,000}$$

Por tanto, debemos elegir:

$n_1 = 600$ personas de C1

$n_2 = 900$ personas de C2

$n_3 = 1\,100$ personas de C3

$n_4 = 400$ personas de C4

c) Dentro de cada comarca, podríamos seleccionarlos mediante un muestreo aleatorio simple, o mediante un muestreo sistemático.

7 **En un centro de enseñanza con 981 alumnos y alumnas, se va a hacer un sondeo sobre tendencias políticas. Se va a escoger una muestra de 84 estudiantes. En el centro hay 5 cursos (1º, 2º, 3º, 4º y 5º) con un número de alumnos y alumnas en cada uno de ellos de 345, 234, 190, 140 y 72. ¿Cuántos alumnos deberemos escoger de cada curso si deseamos que el muestreo sea estratificado con reparto proporcional?**

$$\frac{84}{981} = \frac{a}{345} = \frac{b}{234} = \frac{c}{190} = \frac{d}{140} = \frac{e}{72}$$

Así: $a = 30$ $b = 20$ $c = 16$ $d = 12$ $e = 6$

8 **Queremos seleccionar una muestra de 50 alumnos de 2º de Bachillerato. En cada uno de los siguientes casos debes decidir si el muestreo debe ser aleatorio simple o estratificado por sexos (chicos-chicas) para estudiar las variables indicadas:**

a) Estatura.

b) Tiempo que emplean los alumnos en ir de su casa al instituto.

c) Agudeza visual (porcentaje de alumnado con gafas).

d) Incidencia de caries dental.

e) Práctica de fútbol.

f) Lectura de algún periódico.

a) En la estatura de chicos y chicas de esa edad suele haber diferencias significativas. El muestreo debe ser estratificado en este caso.

b) No

c) No

d) No

e) Sí, hay una gran diferencia entre el porcentaje de chicos y chicas que juegan al fútbol.

f) No

PARA PROFUNDIZAR

9 **Si cuentas el número de personas y el número de perros que viven en tu portal y todos los compañeros y compañeras hacéis lo mismo, obtendréis una muestra con la que podréis estimar el número de perros que hay en vuestra población.**

a) ¿Cómo es de fiable esta estimación?

b) ¿Es aleatoria la muestra que has utilizado?

c) ¿Se te ocurre un procedimiento mejor para seleccionar la muestra?

a) Es poco fiable.

b) La muestra no es aleatoria porque no la hemos elegido al azar entre los habitantes de la ciudad que se quiere estudiar.

Si en ese portal hay muchas viviendas, pueden representar, en el mejor de los casos, a las familias de ese barrio (céntrico o periférico, con ciertas características socioeconómicas, culturales…), pero no a los demás barrios de la población.

c) Utilizar una muestra de viviendas elegidas al azar entre las de esa población.

10 **Para hacer un estudio sobre los hábitos ecológicos de las familias de una ciudad, se han seleccionado, por sorteo, las direcciones, calle y número que serán visitadas. Si en un portal vive más de una familia, se sorteará entre ellas la que será seleccionada.**

¿Obtendremos con este procedimiento una muestra aleatoria?

(Piensa si tiene la misma probabilidad de ser incluida en la muestra una familia que vive en una vivienda unifamiliar, que otra que vive en un bloque de 32 viviendas).

Las familias que viven en viviendas unifamiliares tienen mayor probabilidad de ser elegidas.

11 **La validez de la información que nos proporciona una encuesta depende, en gran medida, de la cuidadosa elaboración del cuestionario.**

Algunas de las características que deben tener las preguntas, son:

— Ser cortas y con un lenguaje sencillo.

— Sus respuestas deben presentar opciones no ambiguas y equilibradas.

— Que no requieran esfuerzo de memoria.

— Que no levanten prejuicios en los encuestados.

Estudia si las siguientes preguntas son adecuadas para formar parte de una encuesta y corrige los errores que observes:

a) ¿Cuánto tiempo sueles estudiar cada día?

Mucho ☐ **Poco** ☐ **Según el día** ☐

b) ¿Cuántas veces fuiste al cine a lo largo del año pasado?

c) ¿Qué opinión tienes sobre la gestión del alcade?

Muy buena ☐ **Buena** ☐ **Indiferente** ☐

d) ¿Pierden sus hijos el tiempo viendo la televisión?

Sí ☐ **No** ☐

e) ¿En qué grado cree usted que la instalación de la planta de reciclado afectaría al empleo y a las condiciones de salud de nuestra ciudad?

a) y c) adecuadas (para mejorarlo, podríamos añadir en c) la opción: Mala ☐).

b) Cambiar por:

¿Con qué frecuencia vas al cine?

☐ Mucho ☐ Poco ☐ Nada

d) Cambiar por:

¿Con qué frecuencia ven sus hijos la televisión?

☐ Mucho ☐ Poco ☐ Nada

e) Cambiar por:

¿Instalaría una planta de reciclado en su ciudad?

☐ Sí ☐ No

12 **Se quieren realizar los siguientes estudios:**

a) Tipo de transporte que utilizan los vecinos de un barrio para acudir a su trabajo.

b) Estudios que piensan realizar los alumnos y alumnas de un centro escolar al terminar el Bachillerato.

c) Edad de las personas que han visto una obra de teatro en una ciudad.

d) Número de horas diarias que ven la televisión los niños y niñas españoles con edades comprendidas entre 5 y 10 años.

Di en cada uno de estos casos cuál es la población.

¿En cuáles de ellos es necesario recurrir a una muestra? ¿Por qué?

a) Población: personas de un barrio que realizan un trabajo fuera de su hogar.

b) Población: alumnos de ESO de un centro escolar.

c) Población: personas que han visto esa obra de teatro.

d) Población: niños españoles de entre 5 y 10 años de edad.

Es necesario recurrir a una muestra en los siguientes casos:

a) Sí, porque, en general, la población es grande y no es fácil controlarla. Puede no ser necesario en un barrio muy pequeño o con muy pocas personas que trabajen fuera de casa (por ejemplo, en zonas rurales agrícolas).

b) No es necesario.

c) Sí, porque la población es difícil de controlar.

d) Sí, porque la población es muy grande.