Nombre:____ ***Curso***:_____

Estadística – Muestreo, Inferencia – 11/05/2023

EJERCICIO 1: [2,5]

a) [1] En un colegio hay 2 000 alumnos distribuidos en 5 cursos así: 400 en 1er curso, 380 en 2^{o}, 520 en 3^{o}, 360 en 4^{o} y 340 en 5^{o}. Se quiere seleccionar una muestra de 100 alumnos, utilizando la técnica de muestreo aleatorio estratificado con afijación proporcional y considerando cada curso como un estrato. ¿Cuál sería la composición de la muestra y cómo se seleccionaría?

b) [1,5] Las peores calificaciones de unos alumnos en un examen de muestreo e inferencia han sido 0, 2 y 4. Determine la varianza de las medias muestrales, con reemplazamiento, de tamaño dos.

EJERCICIO 2: [2,5]

Las naranjas de una cosecha tienen un peso medio de 145 gr. y una desviación típica de 30 gr. Las naranjas se ponen a la venta en cajas de 100 unidades.

a) [1] ¿Qué distribución siguen los pesos medios de las cajas?

b) [1,5] Halle la probabilidad de que el peso de una caja sea inferior a 14 kg.

EJERCICIO 3: [2,5]

En una población, una variable aleatoria sigue una ley normal con desviación típica 12. Con un nivel de confianza del 90% se ha construido un intervalo para la media poblacional, que ha resultado ser

$$I = (36.71, 43.29)$$

a) [0,5] ¿Cuál es la media muestral?

b) [2] ¿Cuál es el tamaño muestral?

EJERCICIO 4: [2,5]

Una tienda de ropa quiere estudiar la aceptación de un nuevo sistema de pago a través del teléfono móvil.

Para ello realiza una encuesta entre 200 de sus clientes elegidos al azar, resultando que 150 de ellos sí estarían dispuestos a usar el nuevo sistema de pago.

a) [1,5] Determine un intervalo de confianza al 97 % para estimar la proporción de clientes de esa tienda que estarían dispuestos a usar el nuevo sistema de pago.

b) [0,25] Con ese intervalo, ¿qué error máximo se cometería en dicha estimación?

c) [0,25] ¿Puede decirse que la proporción poblacional está en el intervalo antes determinado con una probabilidad $p = 0.97$?

d) [0,5] ¿Cómo mantener la amplitud del intervalo aumentando a la vez el nivel de confianza?

EJERCICIO 1:

a) Mostremos esquemáticamente las composiciones de la población y de la muestra:

Población				
1º	2º	3º	4º	5º
400	380	520	360	340
$N = 2000$				

Muestra				
1º	2º	3º	4º	5º
20	19	26	18	17
$n = 100$				

1º: $\frac{100}{2000} \cdot 400 = 20$, 2º: $\frac{100}{2000} \cdot 380 = 19$, 3º: $\frac{100}{2000} \cdot 520 = 26$, 4º: $\frac{100}{2000} \cdot 340 = 18$, 5º: $100 - 83 = 17$

Ahora elegimos en cada estrato, mediante muestreo aleatorio simple, el número de individuos señalados anteriormente.

b) La variable aleatoria en la población es $X = \{ 0 , 2 , 4 \}$

El conjunto de las muestras de tamaño $n = 2$, con reemplazamiento, y las medias muestrales son:

$$\mathfrak{M} = \left\{ \begin{matrix} (0,0) & (0,2) & (0,4) \\ (2,0) & (2,2) & (2,4) \\ (4,0) & (4,2) & (4,4) \end{matrix} \right\} \xrightarrow[medias]{} \overline{\mathbf{X}} = \left\{ \begin{matrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{matrix} \right\}$$

Así, la media y la varianza de las medias muestrales son:

$$\overline{\mu} = \frac{18}{9} = 2$$

$$\overline{\sigma}^2 = \frac{48}{9} - 2^2 = \frac{4}{3}$$

[**Nota**: Otro procedimiento que podemos usar, ya que no se pide siquiera calcular las muestras es calcular directamente con los tres valores de la población \mu y \sigma para usar a continuación las conocidas fórmulas $\overline{\mu} = \mu$ y $\overline{\sigma}^2 = \sigma/n$ donde aquí $n = 2$.]

EJERCICIO 2:

La variable **X** = "peso de las naranjas" tiene $\left\{ \begin{matrix} \mu & = & 145 \\ \sigma & = & 30 \end{matrix} \right.$.

Tamaño muestral: $n = 100$

a) La distribución de las medias muestrales $\overline{\mathbf{X}}$ es normal [*] con $\left\{ \begin{array}{l} \overline{\mu} = \mu = 145 \\ \overline{\sigma} = \frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{100}} = 3 \end{array} \right.$

[Por el TCL al ser $n = 100 > 30$ suficientemente grande]

b) La probabilidad pedida es:

$$\mathrm{p}\,(\overline{x} < 14) \overset{(*)}{=} \mathrm{p}\,(z < -1.67) = \mathrm{p}\,(z > 1.67) = 1 - 0.9525 = 0.0475$$

$(*)\ z = \frac{\overline{x} - \overline{\mu}}{\overline{\sigma}} = \frac{140 - 145}{3} \approx -1.67$

EJERCICIO 3:

La variable **X** es normal con $\begin{cases} \mu = ¿? \\ \sigma = 12 \end{cases}$

Al 90% se ha construido el intervalo confianza la media poblacional $I = (36.71, 43.29)$.

a) La media muestral está en el centro del intervalo al ser $I = (\overline{x} - E\,, \overline{x} - E)$. Así resulta ser:

$$\overline{x} = \frac{36.71 + 43.29}{2} = 40$$

b) Procedamos, antes de nada, a obtener el valor crítico bilateral asociado a la confianza:

$$1 - \alpha = 0.90 \rightarrow \frac{\alpha}{2} = 0.05 \rightarrow p\left(z < z_{\alpha/2}\right) = 0.9500 \xrightarrow{tabla} z_{\alpha/2} = 1.645$$

Observemos que el error máximo muestral es la mitad de la amplitud (L) del intervalo:

$$E = \frac{L}{2} = \frac{43.29 - 36.71}{2} = 3.29$$

Así que de la fórmula del error máximo $E = z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$ despejamos el tamaño muestral n:

$$3.29 = 1.645 \cdot \frac{12}{\sqrt{n}} \rightarrow \sqrt{n} = 1.645 \cdot \frac{12}{3.29} = 6 \rightarrow n = 36$$

EJERCICIO 4:

Estudiamos C = “número de clientes que acepta pago por el móvil”

Tamaño muestral: $n = 200$

Proporción muestral: $\tilde{p} = \dfrac{150}{200} = 0.75$

a) Procedamos, antes de nada, a obtener el valor crítico bilateral asociado a la confianza:

$$1 - \alpha = 0.97 \rightarrow \frac{\alpha}{2} = 0.015 \rightarrow p\left(z < z_{\alpha/2}\right) = 0.9850 \xrightarrow{tabla} z_{\alpha/2} = 2.17$$

El intervalo de confianza, para la proporción en la población, es:

$$I = (\tilde{p} - E\,, \tilde{p} + E) = (0.75 - 0.0664\,, 0.75 + 0.0664) = (0.6836\,, 0.8164)$$

donde

$$E = z_{\alpha/2}\sqrt{\frac{\tilde{p} \cdot \tilde{q}}{n}} = 2.17 \cdot \sqrt{\frac{0.75 \cdot 0.25}{200}} \approx 0.0664$$

b) El error máximo cometido es, como ya se ha obtenido arriba:

$$E_{max} = 0.0664$$

c) Podría decirse que es así en el sentido de que, con la confianza elegida, por cada 100 muestras que se tomen se construirán 100 intervalos de confianza de los que 97 contendrán a la proporción muestral y 3 no. El intervalo que hemos obtenido puede ser de esos 97% adecuados o no serlo.

Observemos que cada muestra proporciona un intervalo de confianza propio construido a partir de su proporción muestral.

d) Si aumenta el nivel de confianza, el valor crítico es mayor y, por ello, la amplitud del intervalo ($L = 2\,E$) es mayor también (al crecer el error máximo admisible). Para que no crezca la amplitud aumentamos convenientemente el tamaño muestral, pues al estar en el denominador el incremento de n conlleva una disminución del error máximo admisible.