Unidad nº6: *INTRODUCCIÓN al MUESTREO y ESTIMACIÓN PUNTUAL e INTERVALAR*

Ejercicios resueltos

Suponga que, en una población de seres humanos, el diámetro craneal sigue una distribución aproximadamente normal, con media de 185,6 mm y desviación estándar de 12,7 mm ¿Cuál es la probabilidad de que una muestra aleatoria de tamaño 10, de esta población, tenga una media mayor que 190 mm?

Solución:

Cuando se dice que la población sigue una distribución aproximadamente normal, se supone que la distribución muestral de $\overline{\mathrm{X}}$ tiene una distribución normal. También se sabe que la media y el error estándar de la distribución muestral de medias son iguales a 185,6 y $\frac{12,7}{\sqrt{10}} = 4,0161$ mm, respectivamente ⇒

$$\Rightarrow \mathrm{P}(\overline{\mathrm{X}} > 190) = \mathrm{P}\left(\mathrm{z} > \frac{190 - 185,6}{4,0161}\right) = \mathrm{P}(\mathrm{z} > 1,10) = \Phi(1,10) = 0,1357^{(1)}$$

De las puntuaciones de un test de agresividad donde, el 60% individuos alcanzaron un máximo de 40 puntos y el 30% obtuvieron puntuaciones comprendidas entre 40 y 90 puntos, se sabe que la media es 38,75 puntos y la desviación estándar es de 5 puntos.

a) A partir de una muestra aleatoria de tamaño 25. ¿Cuál es la probabilidad de que tengan un valor promedio superior a 40?

b) ¿Cuál es la probabilidad que tengan un valor promedio de entre 39 y 40,5 puntos, dada una muestra de 49 sujetos?

Solución:

a) $\mathrm{P}(\overline{\mathrm{X}} > 40) = \mathrm{P}\left(\mathrm{z} > \frac{40 - 38,75}{5/\sqrt{25}}\right) = \mathrm{P}(\mathrm{z} > 1,25) = \Phi(1,25) = 0,1056^{(1)}$

b) $\mathrm{P}(39 < \overline{\mathrm{X}} < 40,5) = \mathrm{P}\left(\frac{39 - 38,75}{5/\sqrt{49}} < \mathrm{z} < \frac{40,5 - 38,75}{5/\sqrt{49}}\right) = \mathrm{P}(0,35 < \mathrm{z} < 2,45) =$

$$= \Phi(0,35) - \Phi(2,45) = 0,3632^{(7)} - 0,0071^{(1)} = 0,3561$$

Se supone que, en poblaciones de individuos sanos, el cloruro sérico tiene una media de 100 mEq/L y una desviación estándar de 3 mEq/L.

a) ¿Qué proporción de la población tiene concentraciones de cloruro sérico mayores de 103 mEq/L?

b) Si se seleccionó una muestra de 36 individuos, ¿cuál es la probabilidad que, en promedio, los sujetos tengan entre 99 y 101,5 mEq/L?

Solución:

Sabiendo que $\mu = 100$ y $\sigma = 3$ (es decir, $N(100; 3)$) ⇒

a) $P(x > 103) = P\left(z > \frac{103 - 100}{3}\right) = P(z > 1) = \Phi(1) = 0{,}1587$ [1]

b) Para n = 36 individuos $\Rightarrow P(99 < \bar{X} < 101{,}5) = P\left(\frac{99 - 100}{3/\sqrt{36}} < z < \frac{101{,}5 - 100}{3/\sqrt{36}}\right) =$

$$= P(-2 < z < 3) = 1 - [\Phi(2) + \Phi(3)] = 1 - [0{,}0228\,^{(2)} + 0{,}0013\,^{(2)}] =$$

$$= 1 - 0{,}0241 = 0{,}9759$$

Nota: todos los ejercicios resueltos, presentados anteriormente, pueden realizarse utilizando el software estadístico InfoStat.

Ejercicios propuestos

1. Se sabe que la distribución de la temperatura del cuerpo humano en la población tiene una media de 37 ºC y una desviación estándar de 0,6 ºC. Si se toma una muestra aleatoria de 15 personas, calcule la probabilidad que la temperatura promedio sea menor a 36,7 ºC.

2. Una fábrica de coches lanza al mercado el modelo "Stat" del que se sabe que su peso sigue una distribución normal, con media 3100 kg. y desviación estándar de 130 kg.
   a) ¿Cuál es la probabilidad de que, al comprar un coche Stat, pese más de 3130 kg?
   b) ¿Cuál es la probabilidad de que, al comprar un coche, pese entre 2900 y 3500 kg?
   c) ¿Qué distribución seguirán las muestras de tamaño 100 de coches Stat?

3. Los sueldos de ciertos/as trabajadores/as están distribuidos normalmente con media de \$8000 y desviación estándar de \$2000. Si toma una muestra al azar de 25 trabajadores/as:
   a) Determine la probabilidad que un sueldo, elegido aleatoriamente, exceda los \$7700.
   b) ¿Cuál es la probabilidad de que el sueldo promedio exceda los \$7700?

4. Considérese que una población, en la cual se estudia una característica determinada, sigue una distribución normal de media $\mu = 12$ y desviación estándar $\sigma = 4$.
   a) Determine la probabilidad de que un sujeto, elegido al azar, tenga la característica superior a 14.
   b) Considerando una muestra aleatoria de tamaño, $n = 9$, cuál es la probabilidad:
      b.1) Que dicha muestra tenga un valor promedio superior a 14.
      b.2) ¿Y que, en dicha muestra, se encuentre un valor promedio entre 11 y 13?

5. En un servicio de atención al cliente, el tiempo de espera hasta recibir atención es una variable normal con media 11' y desviación estándar 2,5'. Se toman muestras aleatorias del tiempo de espera de clientes que llegan un día determinado.
   a) ¿Cuál es la probabilidad de que el tiempo medio de espera, de una muestra de 25 clientes, no supere los 10,5 minutos?
   b) ¿Cuál es la distribución de la media muestral, si se toman muestras aleatorias de 64 clientes?

6. Si las concentraciones de ácido úrico en sangre, en adultos normales, siguen una distribución aproximadamente normal, con media y desviación estándar de 5,35 y 1,85 mg/dL,

---

[1] Valor obtenido en la Tabla de la Distribución Normal Estandarizada (Fórmulas y Tablas III, 2ª ed., pág. 24).

respectivamente, encuentre la probabilidad de que una muestra probabilística de tamaño 9 proporcione una media:

a) Mayor que 6 mg/dL.
b) Entre 4,6 y 6,1 mg/dL
c) Menor que 5,2 mg/dL.
d) No menor a 5,5 mg/dL.
e) Entre 6,2 y 6,5 mg/dL.
f) No mayor a 4,8 mg/dL.

7. Si la concentración promedio de hierro en suero en hombres sanos es de 120 µg/100 ml y la desviación estándar es de 15 µg/100 ml, ¿cuál es la probabilidad de que una muestra al azar de 47 hombres sanos, tenga una media entre 115 y 125 µg/100 ml?

8. Si la droga (en miligramos) en cierto medicamento sigue una distribución $N(7{,}5;\, 0{,}5)$. Calcule la probabilidad de que en una muestra de tamaño 5 se obtenga una media mayor que 7.

9. Un ascensor limita el peso de sus 4 ocupantes a 290 kg. Si el peso de un individuo sigue una distribución normal $N(71;\, 7)$, calcule la probabilidad de que el peso de 4 individuos supere los 290 kg.

10. La edad en que contraen matrimonio los hombres de las Islas Galápagos es una variable aleatoria que se puede aproximar por una distribución normal con $\mu = 35$ años y $\sigma = 5$ años. Se elige aleatoriamente una muestra de 100 hombres de dicha isla. Sea $\overline{\mathrm{X}}$ la media muestral de la edad de casamiento:

a) ¿Cuáles son los valores de la media y el error estándar de $\overline{\mathrm{X}}$?
b) ¿Cuál es la probabilidad de que la edad media de casamiento de la muestra esté comprendida entre 34 y 36 años?
c) ¿Cuál es la probabilidad que la edad media de casamiento supere los 35,5 años?
d) ¿Cuál es la probabilidad que la edad media de casamiento no sea inferior a 33,5 años?

11. La duración de las baterías de un determinado modelo de balanza electrónica tiene una distribución normal con $\mu = 34{,}8$ h. y desviación estándar de 6,9 h. Si se toma una muestra aleatoria de 36 balanzas, diga cuál es la probabilidad que la duración media de las baterías de la muestra:

a) ¿Esté comprendida entre 33,8 y 34 h.?
b) ¿Sea mayor de 35 h.?
c) ¿Sea menor a 33,4 h.?
d) ¿Entre qué valores se encuentra el 50% central de la duración media de las baterías?

12. El peso de recién nacidos en una determinada población sigue una distribución $N(3600;\, 280)$. Se toma una muestra al azar de 196 neonatos y se calcula el peso promedio. ¿Cuál es la probabilidad de que esta media esté ente 3580 y 3620 g?

13. Se supone que la talla (en cm.) de neonatos de una determinada población sigue una distribución normal $N(50;\, 5{,}5)$. Se toma una muestra al azar de 144 de esos recién nacidos y se calcula la talla media. ¿Cuál es la probabilidad de que la media obtenida esté entre 49 y 51 cm?

14. El CI de los/as estudiantes de un centro especial se distribuye normalmente con media 85 puntos y desviación estándar 8 puntos. Si se toma una muestra, al azar, de 20 estudiantes:

a) ¿Cuál es la probabilidad que la media de la muestra sea como máximo de 80 puntos?
b) ¿Cuál es la probabilidad que la media aritmética se encuentre entre 78 y 81 puntos?
c) ¿Y la probabilidad que la media aritmética sea como mínimo de 88 puntos?
d) ¿Qué valor debería tomar la media aritmética para que, la probabilidad de obtenerlo en esa muestra, sea como máximo 0,75?

15. Las estaturas de ciertos estudiantes están distribuidas aproximadamente de forma normal con media de 175 cm. y desviación estándar de 7 cm. Si se extraen 200 muestras aleatorias de tamaño 25 de esta población, determine:
a) La media y el error estándar de la distribución muestral del promedio muestral.
b) El número de medias muestrales que caen entre 173,6 y 176,4 cm.
c) El número de medias muestrales que superan 177 cm.
d) El número de medias muestrales que no superan 173 cm.

16. Algunas estudiantes manifiestan, en general, tener dificultad para memorizar conceptos estadísticos. Experiencias anteriores han consistido en exponer 5 conceptos ante estudiantes durante 30 segundos al comienzo de la clase y luego preguntar por ellos al final de la misma, obteniéndose la siguiente distribución de probabilidad:

Nº de conceptos que recuerdan	0	1	2	3	4	5
P(x)	0,05	0,15	0,20	0,25	0,30	0,05

En una muestra aleatoria de 49 estudiantes, ¿cuál es la probabilidad de que en promedio recuerden por lo menos 3 conceptos sabiendo que $\sigma = 1{,}3$ conceptos?

17. En una casa de retiro, la edad de las personas tiene una media de 76 años y una desviación estándar de 10 años.

a) ¿Cuántas personas, con una probabilidad de 9,85%, superarán -en promedio- los 84 años?
b) Para una muestra de 400 personas, determine por debajo de qué valor se encuentra (aproximadamente) el 75% de las medias muestrales de las edades.
c) Y, para la misma muestra dada en el punto b), determine (aproximadamente) entre qué valores encierra el 80% central de las medias muestrales de las edades.

18. Una empresa eléctrica fabrica focos que tienen una duración, distribuida aproximadamente en forma normal, con media de 800 h. y desviación estándar de 40 h. Encuentre la probabilidad de que una muestra aleatoria de 16 focos, tenga una vida promedio:

a) Menor a 785 h.
b) Superior a 815 h.
c) De entre 780 y 820 h.
d) Menor a 770 o mayor a 830 h.
e) Entre qué horas se encuentra -aproximadamente- el 80% central de las medias muestrales.

ALGUNAS DUDAS FRECUENTES

¿Cuál es el principal riesgo que se puede cometer, si queremos hacer una inferencia, al realizar un muestreo?

Hay que ser cuidadoso respecto al tipo de muestreo si queremos hacer una inferencia pues, si no realizamos un muestreo probabilístico, cometemos el riesgo de que la muestra sea no-representativa y, por lo tanto, sesgada. Esto nos llevaría a tomar conclusiones erróneas.

¿Cuál es la diferencia entre las curvas del modelo "t" de Student, con las que se calculan probabilidades para estadísticos muestrales, y las distribuciones normales que consideran la población completa?, ¿acaso no son muy parecidas?

Al trabajar con muestras y, sobre todo, si son muestras de tamaño pequeño, es más probable que el promedio de sus valores dé como resultado un número grande. Así, son más probables los valores extremos y, por eso, las curvas del modelo "t" de Student consideran puntuaciones más dispersas, siendo entonces menos altas, pero con colas más anchas.

¿Por qué de pronto hablamos de "error estándar", de qué error se trata, de algún tipo de equivocación?

Toda inferencia (la pretensión de conocer valores poblacionales estimados) con base en los datos muestrales implica un salto desde lo conocido hasta lo desconocido. Entre todas las muestras posibles de determinado tamaño, no todas presentarán el mismo valor de media que el del total de todas las muestras. Cuánto se aleje la media particular de una muestra respecto de la media de todas las muestras es lo que llamamos *error estándar*; no una equivocación que pueda cometerse en su cálculo, sino un error teórico en un muestreo correctamente realizado (teniendo como meta de este, la mencionada coincidencia).

¿Qué relación hay entre los intervalos de confianza y las muestras?, ¿cómo es que luego de precisar un intervalo de confianza puede preguntarse por un número de muestras?

El intervalo de confianza para un estadístico muestral, como la media, es un recorte del eje horizontal en el que están distribuidas todas las medias de todas las muestras de un determinado tamaño. De esa forma, algunas medias corresponderán a muestras que pertenecerán a dicha porción del mencionado eje, pero, del mismo modo, otras muestras quedarán por fuera. Si el intervalo de confianza presenta un nivel del 95%, 5 de cada 100 muestras del mismo tamaño para esa variable presentarán medias menores o mayores que los límites de ese intervalo.