Problemas de Muestreo. Soluciones

Estimación de la media muestral

1. La vida útil de las bombillas producidas por un fabricante en particular tiene una media de 1200 horas y una desviación estándar de 400 horas. La distribución de la población es normal. Suponga que compra nueve bombillas, que pueden considerarse como una muestra aleatoria de la producción del fabricante.

   a. ¿Cuál es la media de la vida media muestral?
   b. ¿Cuál es la varianza de la media muestral?
   c. ¿Cuál es el error estándar de la media muestral?
   d. ¿Cuál es la probabilidad de que, en promedio, esas nueve bombillas tengan una vida útil de menos de 1050 horas?

Solución

Nos dicen que la vida útil de las bombillas sigue una distribución $N(\mu = 1200, \sigma = 400)$. Tenemos una muestra aleatoria simple de $n = 9$ bombillas. Entonces la distribución de la variable media muestral será $N(\mu_{\overline{X}} = 1200, \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{9}} = \frac{400}{\sqrt{9}} = 133.3333333)$.

a. La media de la vida media muestral, será, pues, $\mu_{\overline{X}} = 1200$.
b. La varianza de la media muestral será $\sigma^2_{\overline{X}} = \left(\frac{400}{\sqrt{9}}\right)^2 = 17777.7777778$.
c. El error estándard de la media muestral es su desviación típica: $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{9}} = \frac{400}{\sqrt{9}} = 133.3333333$.
d. Nos piden la probabilidad siguiente: $P(\overline{X} < 1050)$. Estandarizando, obtenemos:

$$P(\overline{X} < 1050) = P\left(Z < \frac{1050 - 1200}{133.3333333}\right) = P(Z < -1.125) = 0.1302945.$$

2

. El precio medio de venta de los condominios para personas mayores en Green Valley durante un año fue de 172000 euros. La desviación estándar de la población fue de 20000 euros. Se obtuvo una muestra aleatoria de 100 nuevas unidades vendidas. a. ¿Cuál es la probabilidad de que el precio medio de la muestra supere los 168000 euros? a. ¿Cuál es la probabilidad de que el precio medio de la muestra esté entre los 170400 euros y los 173600 euros? a. ¿Cuál es la probabilidad de que el precio medio de la muestra esté entre los 171200 euros y los 172800 euros? a. Sin hacer los cálculos, indique en cuál de los siguientes rangos el precio de venta medio de la muestra es más probable que mienta: * 170400 a 172000 euros; * 171200 a 172800 euros; * 172000 a 173600 euros; * 172800 a 174400 euros.

Solución

Nos dicen que el precio de los condominios durante un año se distribuye según una variable P de media 172000 euros y desviación estándar $\sigma = 20000$ euros. Se consideró una muestra de $n = 100$ unidades vendidas. Usando el Teorema Central del Límite, podemos decir que la distribución del precio medio de la muestra elegida, $\overline{P}$ se distribuye aproximadamente según una normal $N\left(\mu_{\overline{P}} = 172000, \sigma_{\overline{P}} = \frac{\sigma}{\sqrt{n}} = \frac{20000}{\sqrt{100}} = 2000\right)$.

a. Nos piden $P(\overline{P} > 168000)$. Estandarizando,

$$P(\overline{P} > 168000) = P\left(Z > \frac{168000 - 172000}{2000}\right) = P(Z > -2) = 0.9772499.$$

b. Nos piden $P(170400 < \overline{P} < 173600)$. Estandarizando,

$$\begin{aligned} P(170400 < \overline{P} < 173600) =& P\left(\frac{170400 - 172000}{2000} < Z < \frac{173600 - 172000}{2000}\right) = P(-0.8 < Z < 0.8) \\ =& P(Z < 0.8) - P(Z < -0.8) = 0.7881446 - 0.2118554 = 0.5762892. \end{aligned}$$

c. Nos piden $P(171200 < \overline{P} < 172800)$. Estandarizando,

$$\begin{aligned} P(171200 < \overline{P} < 172800) =& P\left(\frac{171200 - 172000}{2000} < Z < \frac{172800 - 172000}{2000}\right) = P(-0.4 < Z < 0.4) \\ =& P(Z < 0.4) - P(Z < -0.4) = 0.6554217 - 0.3445783 = 0.3108435. \end{aligned}$$

d. Todos los rangos tienen la misma amplitud, 1600. El rango más probable es aquel que esté centrado en la media, 172000 y es el segundo, 171200 a 172800 euros;

3. El tiempo dedicado al estudio de los estudiantes en la semana anterior a los exámenes finales sigue una distribución normal con una desviación estándar de 8 horas. Se tomó una muestra aleatoria de cuatro estudiantes con el fin de estimar el tiempo medio de estudio de la población de todos los estudiantes.

   a. ¿Cuál es la probabilidad de que la media de la muestra supere la media de la población en más de 2 horas?
   b. ¿Cuál es la probabilidad de que la media de la muestra esté más de 3 horas por debajo de la media de la población?
   c. ¿Cuál es la probabilidad de que la media muestral difiera de la media poblacional en más de 4 horas?
   d. Suponga que se tomó una segunda muestra aleatoria (independiente) de 10 estudiantes. Sin hacer los cálculos, indique si las probabilidades en los incisos (a), (b) y (c) serían mayores, menores o iguales para la segunda muestra.

Solución

Nos dicen que la variable tiempo T que los estudiantes dedican al estudio en la semana anterior a los exámenes finales sigue una distribución normal con una desviación estándar de 8 horas: $T = N(\mu, \sigma = 8)$. Se tomó una muestra de $n = 4$ estudiantes. La distribución de la variable media $\overline{T}$ será $N\left(\mu, \sigma_{\overline{T}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{4}} = 4\right)$.

a. Nos piden $P(\overline{T} > \mu + 2)$. Estandarizando:

$$P(\overline{T} > \mu + 2) = P(\overline{T} - \mu > 2) = P\left(\frac{\overline{T} - \mu}{4} > \frac{2}{4}\right) = P(Z > 0.5) = 0.3085375.$$

b. Nos piden $P(\overline{T} < \mu - 3)$. Estandarizando:

$$P(\overline{T} < \mu - 3) = P(\overline{T} - \mu < -3) = P\left(\frac{\overline{T} - \mu}{4} < \frac{-3}{4}\right) = P(Z < -0.75) = 0.2266274.$$

c. Nos piden $P(|\overline{T} - \mu| > 4)$. Estandarizando:

$$\begin{aligned} P(|\overline{T} - \mu| > 4) =& P\left(\left|\frac{\overline{T} - \mu}{4}\right| > \frac{4}{4}\right) = P(|Z| > 1) = P(Z > 1) + P(Z < -1) \\ =& 0.1586553 + 0.1586553 = 0.3173105. \end{aligned}$$

d. En este caso, la desviación típica de la muestra $\sigma_{\overline{T}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{10}} = 2.5298221$ será más pequeña. Las probabilidades calculadas en (a), (b) y (c) serán mayores ya que los valores estandarizados serán más pequeños en valor absoluto y las áreas correspondientes serán más grandes.

Estimación de la proporción

4. En 1992, los canadienses votaron en un referéndum sobre una nueva constitución. En la provincia de Quebec, el 42.4% de los que votaron estaban a favor de la nueva constitución. Se tomó una muestra aleatoria de 100 votantes de la provincia.

   a. ¿Cuál es la media de la distribución de la proporción muestral a favor de una nueva constitución?
   b. ¿Cuál es la varianza de la proporción muestral?
   c. ¿Cuál es el error estándar de la proporción muestral?
   d. ¿Cuál es la probabilidad de que la proporción muestral sea superior a 0.5?

Solución

Consideramos la variable aleatoria proporción muestral $\hat{P}_X$ de votantes a favor de la nueva constitución. Como el tamaño de la muestra es grande $n = 100$, podemos aproximar la variable proporción muestral $\hat{P}_X$ por una distribución normal $N\left(\mu = p = 0.424, \sigma = \sqrt{\frac{p\cdot(1-p)}{n}} = \sqrt{\frac{0.424\cdot(1-0.424)}{100}} = 0.049419\right)$ usando el Teorema Central del Límite.

a. La media de la variable proporción muestral será, pues, $\mu_{\hat{P}_X} = 0.424$.
b. La varianza de la variable proporción muestral será: $\sigma^2_{\hat{P}_X} = \sigma^2 = \frac{p\cdot(1-p)}{n} = \frac{0.424\cdot(1-0.424)}{100} = 0.0024422$.
c. El error estándar de la proporción muestral es su desviación estándar: $\sigma = \sqrt{\frac{p\cdot(1-p)}{n}} = \sqrt{\frac{0.424\cdot(1-0.424)}{100}} = 0.049419$.
d. Nos piden $P(\hat{P}_X > 0.5)$. Estandarizando,

$$P(\hat{P}_X > 0.5) = P\left(Z > \frac{0.5 - 0.424}{0.049419}\right) = P(Z > 1.5378693) = 0.0620403.$$

5

. Una corporación recibe 120 solicitudes para puestos de recién graduados universitarios en negocios. Suponiendo que estos solicitantes puedan verse como una muestra aleatoria de todos esos graduados, ¿cuál es la probabilidad de que entre el 35% y el 45% de ellos sean mujeres si el 40% de todos los recién graduados universitarios en negocios son mujeres?

Solución

Consideramos la variable aleatoria proporción muestral $\hat{P}_X$ de mujeres en las solicitudes para puestos de recién graduados universitarios en negocios. Como el tamaño de la muestra es grande $n = 120$, podemos aproximar la variable proporción muestral $\hat{P}_X$ por una distribución normal $N\left(\mu = p = 0.4, \sigma = \sqrt{\frac{p\cdot(1-p)}{n}} = \sqrt{\frac{0.4\cdot(1-0.4)}{120}} = 0.0447214\right)$ usando el Teorema Central del Límite.

Nos piden $P(0.35 < \hat{P}_X < 0.45)$. Estandarizando,

$$\begin{aligned} P(0.35 < \hat{P}_X < 0.45) =& P\left(\frac{0.35-0.4}{0.0447214} < Z < \frac{0.45-0.4}{0.0447214}\right) = P(-1.118034 < Z < 1.118034) \\ =& P(Z < 1.118034) - P(Z < -1.118034) = 0.8682238 - 0.1317762 = 0.7364475. \end{aligned}$$

. Supongamos que el 50% de los australianos adultos creen que Australia debería postularse para albergar la próxima Copa del Mundo de rugby. Calcule la probabilidad de que más del 56% de una muestra aleatoria de 150 australianos adultos lo crea.

Solución

Consideramos la variable aleatoria proporción muestral $\hat{P}_X$ de australianos que creen que Australia debería postularse para albergar la próxima Copa del Mundo de rugby. Como el tamaño de la muestra es grande $n = 150$, podemos aproximar la variable proporción muestral $\hat{P}_X$ por una distribución normal $N\left(\mu = p = 0.5, \sigma = \sqrt{\frac{p\cdot(1-p)}{n}} = \sqrt{\frac{0.5\cdot(1-0.5)}{150}} = 0.0408248\right)$ usando el Teorema Central del Límite.

Nos piden $P(\hat{P}_X > 0.56)$. Estandarizando,

$$P(\hat{P}_X > 0.56) = P\left(Z > \frac{0.56 - 0.5}{0.0408248}\right) = P(Z > 1.4696938) = 0.0708223.$$

Estimación de la varianza

7. Todos los estudiantes de primer año que ingresan a una gran universidad deben recibir una prueba de matemáticas de 100 preguntas de opción múltiple. Inicialmente, en un estudio piloto, la prueba se aplicó a una muestra aleatoria de 20 estudiantes de primer año. Suponga que, para la población de todos los estudiantes de primer año que ingresan, la distribución del número de respuestas correctas sería normal con una varianza de 250.

   a. ¿Cuál es la probabilidad de que la varianza de la muestra sea menor que 100?
   b. ¿Cuál es la probabilidad de que la varianza muestral sea superior a 500?

Solución

Sea S^2 la varianza muestral de las puntuaciones de la prueba de la muestra de 20 estudiantes. Sabemos que la distribución de la variable $\frac{(n-1)\cdot S^2}{\sigma^2} = \frac{19\cdot S^2}{250}$ es una χ^2_{19}.

a. Nos piden $P(S^2 > 100)$:

$$P(S^2 < 100) = P\left(\frac{19 \cdot S^2}{250} < \frac{19 \cdot 100}{250}\right) = P(\chi^2_{19} < 7.6) = 0.0097373.$$

b. Nos piden $P(S^2 > 500)$:

$$P(S^2 > 500) = P\left(\frac{19 \cdot S^2}{250} > \frac{19 \cdot 500}{250}\right) = P(\chi^2_{19} > 38) = 0.0059347.$$

8

. Un proceso de producción fabrica componentes electrónicos con señales de temporización cuya duración sigue una distribución normal. Se tomó una muestra aleatoria de seis componentes y se midieron las duraciones de sus señales de tiempo. Calcular el porcentaje de la varianza de la población tal que la probabilidad de que la varianza de la muestra sea mayor que dicho porcentaje sea 0.05.

Solución

Sea S^2 la varianza muestral de las duraciones de los componentes electrónicos de la muestra de 6 componentes. Sabemos que la distribución de la variable $\frac{(n-1)\cdot S^2}{\sigma^2}$ es una χ_5^2.

Nos piden el porcentaje p tal que $P\left(S^2 > 0.01 \cdot p \cdot \sigma^2\right) = 0.05$:

$$P\left(S^2 > 0.01 \cdot p \cdot \sigma^2\right) = 0.05, \Rightarrow P\left(\frac{S^2}{\sigma^2} > 0.01 \cdot p\right) = 0.05, \Rightarrow P\left(\frac{5 \cdot S^2}{\sigma^2} > 5 \cdot 0.01 \cdot p\right) = 0.05, \Rightarrow$$
$$P\left(\frac{5 \cdot S^2}{\sigma^2} > 0.05 \cdot p\right) = 0.05, \Rightarrow P\left(\chi_5^2 > 0.05 \cdot p\right) = 0.05.$$

El percentil 0.95 para la distribución χ_5^2 vale 11.0704977. Por tanto,

$$0.05 \cdot p = 11.0704977, \Rightarrow p = \frac{11.0704977}{0.05} = 221.4099539.$$

El porcentaje será pues un 221.41%.

9

. Una compañía farmacéutica produce pastillas que contienen un ingrediente activo. A la empresa le preocupa el peso medio de este ingrediente por pastilla, pero también exige que la variación (en miligramos cuadrados) no sea superior a 1.5. Se selecciona una muestra aleatoria de 20 píldoras y la varianza de la muestra es 2.05. ¿Qué probabilidad hay de que se encuentre una varianza muestral tan alta o superior si la varianza de la población es, de hecho, 1.5? Suponga que la distribución de la población es normal.

Solución

Sea S^2 la varianza muestral de los pesos del ingrediente activo de la muestra de 20 píldoras. Sabemos que la distribución de la variable $\frac{(n-1)\cdot S^2}{\sigma^2} = \frac{19\cdot S^2}{1.5} = 12.6666667 \cdot S^2$ es una χ^2_{19}.

Nos piden la probabilidad $P(S^2 > 2.05)$:

$$P(S^2 > 2.05) = P(12.6666667 \cdot S^2 > 12.6666667 \cdot 2.05) = P(\chi^2_{19} > 25.9666667) = 0.1311205.$$

La probabilidad sería de un 13.11%, no es demasiado alto, deberíamos comprobar la muestra seleccionada.