Intervalos de confianza

Análisis estadístico de datos

2021

1. Considerar una variable aleatoria x que sigue una distribución normal con parámetro μ desconocido y $\sigma = 5.9$. En un experimento se miden los siguiente valores de $x : (67.6, 57.4, 63.0, 68.0, 63.1)$. Estimar el valor del parámetro μ y su intervalo con un nivel de confianza del 90%.

2. Considerar una muestra de 10 variables aleatorias $(x_1, \ldots, x_{10})$ que siguen una distribución normal con paramétro μ desconocido y $\sigma = 1.8$. Graficar el cinturón de confianza 1σ. Calcular el intervalo de μ en términos de la media muestral $\bar{x}$ y la desviación estándar σ. En un experimento se miden los siguients valores de (16.2, 12.4, 19.4, 17.3, 16.8, 24.4, 10.7, 18.1, 14.2, 14.8). A partir de los datos, estimar μ y su intervalo de confianza 1σ.

3. Simular una muestra de dos variables normales estándar. Calcular el intervalo de confianza 1σ del parámetro μ asumiendo que la desviación estándar $\sigma = 1$ es conocida. Verificar si el parámetro μ está contenido dentro del intervalo. Repetir la simulación 1000 veces para estimar la probabilidad de cobertura del intervalo. Comparar la probabilidad de cobertura con el nivel de confianza.

4. En un experimento se miden los valores $(13.4, 8.52, 12.7, 9.9, 12.8)$. Asumiendo que los datos siguen una distribución normal con parámetros μ y σ desconocidos, calcular el intervalo de Student al 90% de nivel de confianza en términos de la media muestral $\bar{x}$ y la desviación estándar muestral s.

5. Simular dos variables normales x_1 y x_2 con parámetros $\mu_1 = 10.7$, $\mu_2 = 8.3$, $\sigma_1 = 1.7$, $\sigma_2 = 2.4$, y correlación $\rho = 0.78$. Considerar la elipse de 95% de confianza del parámetro $\boldsymbol{\mu} = (\mu_1, \mu_2)$. Simular varios valores de x_1 y x_2 y verificar si las elipses correspondientes contienen a $\boldsymbol{\mu}$. Repetir la simulación 10.00 veces para estimar la probabilidad de cobertura y comparar con el nivel de confianza.

6. (opcional) Considerar la función de costo,

$$J(S,\beta) = 2\sum_{i=1}^{n}(\mu_i(S,\beta) - k_i) - k_i \log(\mu_i(S,\beta)/k_i),$$

con $\mu_i(S,\beta) = S\, {r_i}^{-\beta}$, dónde S y β son dos parámetros desconocidos. Los datos del experimento son el número de partículas medidas por un detector k_i ubicado a una distancia r_i,

detector	r_i	k_i
1	0.764	33
2	1.052	19
3	1.236	11

Dibujar la función de costo como función de S y β. Con una minimización numérica encontrar los estimadores de S y β. Dibujar la región de confianza de 1σ correspondiente a $J = J_{\min} + 1$, con $J_{\min}$ el mínimo de la función de costo.

7. **(Para entregar)** Considerar n=32 lanzamientos de una moneda cargada con probabilidad p de salir cara. El número de caras X es una variable aleatoria con distribución binomial $X \sim B(n,p)$. Considerar un tanda de lanzamientos en el que salen k caras. El estimador del parámetro p es $\hat{p} = k/n$. El intervalo de confianza estándar de p con un nivel de confianza del 95% es $\hat{p} \pm 1.96\sqrt{\hat{p}\,(1-\hat{p})/n}$, dónde $q = 1.96$ es el cuantil normal estándar para un 95% CL. Simular 10.000 tandas de lanzamientos asumiendo $p = 0.2$. Verificar para cada tanda si el intervalo incluye o no a $p = 0.2$. Estimar la cobertura del intervalo. Decidir si el intervalo es exacto y comparar la cobertura con respecto al nivel de confianza. (Opcional: Barrer p en el intervalo [0,1], graficar y ver si hay valores de p para los cuáles la cobertura es particularmente mala.)