Introducción a la Estadística y Ciencia de Datos

Guía de Actividades - Intervalos de Confianza

1. Un emisor transmite una señal de valor μ. El receptor recibe mediciones $X_i = \mu + \epsilon_i$ donde ϵ_i son los errores de medición, que pueden considerarse variables aleatorias independientes con distribución $\mathcal{N}(0, 3)$.

    a) Hallar un intervalo de confianza de nivel 0,95 para la señal transmitida μ basado en una muestra aleatoria de 9 mediciones recibidas.

    b) Si en la muestra de tamaño 9 se obtuvo un promedio de 3,48, hallar el intervalo de confianza de nivel 0,95 basado en la muestra observada.

    c) Calcular el número mínimo de mediciones que debe recibir el receptor para poder construir un intervalo de confianza de nivel 0,95 para μ cuya longitud sea la mitad de la del intervalo hallado en el ítem anterior.

2. Implementar la función `IC_mu_var_conocida(datos,var_0,nivel)` que tenga por argumento un conjunto de datos, el valor σ_0^2 (la varianza POBLACIONAL) y el nivel de confianza $1-\alpha$ deseado y devuelva una estimación por intervalos para μ de nivel $1-\alpha$, suponiendo que los datos provienen de una distribución normal.

Definición de Nivel de Cubrimiento Empírico: Sea $X_1, ..., X_n$ una muestra aleatoria de una población con distribución cuyo parámetro es $\theta \in \mathbb{R}$. Mediante un procedimiento de intervalos de confianza de nivel $1-\alpha$, se calcula un intervalo de confianza para la muestra. El nivel de cubrimiento empírico del procedimiento se define como la proporción de veces que los intervalos, construídos a partir de datos simulados, contienen a θ, en cierta cantidad de $Nrep$ replicaciones.

3. **Simulando.** Sean $X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2)$ variables aleatorias independientes con $\mu = 4$, $\sigma^2 = 9$.

    a) Guardar en el vector `datos_normales`, $n = 5$ datos generados a partir de una variable con distribución $\mathcal{N}(\mu, \sigma^2)$ siendo $\mu = 4$ y $\sigma^2 = 9$.

    b) Calcular la estimación por intervalo de confianza de nivel 0,95 para μ con los valores guardados en `datos_normales`.

    c) ¿Pertenece $\mu = 4$ al intervalo obtenido en el ítem anterior?

    d) Repetir $Nrep = 1000$ veces los tres ítems anteriores y registrar en qué proporción de las 1000 replicaciones el intervalo obtenido contiene al valor $\mu = 4$. ¿Se condice este valor con lo se esperaba?

4. El punto de ebullición del agua (en grados Fahrenheit) es una variable aleatoria con distribución normal. En un laboratorio se realizaron 16 experimentos independientes y se registraron los valores del punto de ebullición del agua, obteniendo un promedio de 202,38 y una varianza estimada de 439,75. En base a la información muestral, construir un intervalo de confianza de nivel 0,95 para

   a) el punto medio de ebullición del agua.

   b) la varianza del punto de ebullición del agua.

5. Sea X una variable aleatoria cuya función de densidad es de la forma

$$f_\theta(x) = \frac{3x^2}{\theta^3}\mathbb{I}_{[0,\theta]}(x), \quad \theta > 0.$$

   Utilizando los resultados de clase, hallar un intervalo de confianza de nivel 0,95 para θ basado en el estimador de máxima verosimilitud.

6. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población con densidad

$$f_\theta(x) = e^{-(x-\theta)}\mathbf{I}\{x > \theta\}, \quad \theta > 0.$$

   a) Hallar un intervalo de confianza de nivel 0,99 para θ.

   b) Hallar el intervalo de confianza de nivel 0,99 para θ de mínima longitud.

   c) Si en una muestra de tamaño 100 se obtuvo que $\min\limits_{i\in\{1,\ldots,n\}}\{x_i\} = 47{,}99$, hallar la estimación por intervalos de nivel 0,99 para θ usando lo hallado en el ítem anterior.

7. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población con distribución $\mathcal{G}(p)$. Hallar un intervalo de confianza de nivel asintótico 0,9 para p.

8. Luego de los cambios en las leyes de tránsito de cierta región, se desea estudiar la proporción de motoqueros que usan casco. Se tomó una muestra de 200 motoqueros, encontrando que 148 usaban casco. Construir un intervalo de confianza de nivel asintótico 0,95 para la proporción de motoqueros que usan casco, y en base a esta información muestral, calcular la estimación por intervalos.