

ESTADÍSTICA E INFERENCIA I

Primer Cuatrimestre — 2024

Primer Recuperatorio

APELLIDO Y NOMBRE:

1. Considerar muestras aleatorias de la distribución con densidad, para $0 < \theta < 1$,

$$f(x; \theta) = \frac{1}{\theta} x^{(\frac{1}{\theta}-1)} I_{[0,1]}(x), \quad \text{donde } I_{[0,1]}(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

- (a) Encontrar el estimador de momentos.
- (b) Encontrar el estimador de máxima verosimilitud.
- (c) Se sabe que la longitud de los ejes que fabrica un establecimiento siderúrgico sigue una distribución como $f(x; \theta)$. Se miden 20 ejes al azar y se obtienen las siguientes longitudes: (0.5, 0.7, 0.8, 0.95, 0.9, 0.6, 0.2, 0.85, 0.3, 0.2, 0.76, 0.55, 0.48, 0.8, 0.76, 0.13, 0.15, 0.67, 0.9, 0.95). Calcular los estimadores de θ según ambos métodos

2. Considerar los datasets de la distancia anual entre Urano y Saturno (en unidades astronómicas) y la generación de energía nuclear por hora en Brasil (kWh), encontrados en este sitio: <https://tinyurl.com/2k3h8bz4> (los datos están a mitad de página)

- (a) Encontrar el estimador plug-in del coeficiente de correlación entre ambos conjuntos de datos.
- (b) Estimar su error estándar usando bootstrap.
- (c) Encontrar un intervalo de confianza de nivel 0.95 usando los métodos normal, percentile bootstrap, basic bootstrap y BCa.

(Ayuda: si se usan funciones de librería `scipy.stats`, prestar atención a `vectorized=False`, `paired=True`)

3. Generar $n = 100$ samples de $X_1 \sim \mathcal{U}(0, 1)$, de $Z \sim \mathcal{N}(0, 0.1)$ y de $\varepsilon \sim \mathcal{N}(0, 1)$; a partir de ellas generar n samples de $X_2 := 0.5X_1 + Z$ y de

$$Y := 2 + 2X_1 + 0.3X_2 + \varepsilon.$$

- (a) Calcular la correlación entre las muestras de X_1 y X_2 , y graficar su distribución conjunta.
- (b) Ajusta un modelo de regresión lineal por mínimos cuadrados para predecir Y utilizando tanto X_1 como X_2 . Describir los resultados obtenidos, incluyendo los coeficientes de regresión $\hat{\beta}_0$, $\hat{\beta}_1$ y $\hat{\beta}_2$, y analizar la relación entre estos coeficientes y los verdaderos β_0 , β_1 y β_2 .

4. Nambé Mills es una fábrica de vajillas. En <https://tinyurl.com/nambeware> figuran, entre otros datos, el diámetro y el precio de la vajilla producida durante un período de tiempo fijo.

- (a) Graficar el precio y vs el diámetro x . A partir del gráfico, proponer un MLG.
- (b) Ajustar este MLG, es decir, encontrar estimaciones puntuales para los parámetros de regresión β .
- (c) Determinar el error estándar de cada parámetro de regresión.