

1. La v.a. X = “ingresos (en miles de euros) de un habitante elegido al azar en una cierta ciudad” sigue una distribución de Pareto dada por la siguiente densidad:

$$f(x; \theta) = 3\theta^3 x^{-4}, \quad 0 < \theta < x < \infty.$$

Sea X_1, \dots, X_n una muestra aleatoria de X . Un economista sugiere que $T_n = \min(X_1, \dots, X_n)$ es un posible estimador de θ . ¿Es T_n un estimador consistente de θ ? ¿Es insesgado? En caso negativo, calcula el sesgo. **[2 p.]**

2. Sea X una v.a. con distribución $\text{Beta}(\theta, 1)$ cuya función de densidad es $f(x; \theta) = \theta x^{\theta-1} \mathbb{I}_{(0,1)}(x)$, para $\theta > 0$.

- (a) Calcula la cantidad de información de Fisher $I(\theta)$. Explica brevemente por qué es importante esta cantidad.
- (b) Calcula el estimador de máxima verosimilitud de θ (basado en muestras de tamaño n), demuestra que es asintóticamente normal e identifica completamente su distribución asintótica.
- (c) Calcula el estimador de θ por el método de los momentos e identifica completamente su distribución asintótica. Demuestra que la correspondiente varianza asintótica es mayor que la obtenida en el apartado (b).
- (d) Supongamos ahora que se desea contrastar $H_0 : \theta = 1$ frente a $H_1 : \theta = 2$ a partir de una muestra de tamaño 2, X_1, X_2 . Para ello se usa el test de región crítica

$$R = \{(x_1, x_2) : 4x_1x_2 \geq 3\}.$$

Calcula el nivel de significación de este test y la probabilidad de error de tipo 2.

[4 p.]

3. En una encuesta realizada a una muestra aleatoria de 1500 personas, el 43 % de los encuestados se mostraba de acuerdo con endurecer la ley antitabaco.

- (a) Calcula el intervalo de confianza de nivel 0.95 para la proporción p de personas en la población que están de acuerdo con endurecer la ley.
- (b) Según los resultados obtenidos, ¿existe evidencia estadística suficiente para afirmar que la mayoría de los ciudadanos se opone a endurecer la ley? Para responder a la pregunta, calcula aproximadamente el p-valor del test e interpreta el resultado.

[3p.]

4. Supongamos que se tiene en el directorio de trabajo un fichero llamado `datos` que consiste en una matriz de 200 filas y 10 columnas. Cada fila es una muestra aleatoria de tamaño 10 de la distribución $N(2, 1)$. Redacta un código en `R` que calcule las medias y las medianas muestrales de esas 200 muestras, las almacene en dos vectores llamados `medias` y `medianas`, respectivamente, y aproxime los errores cuadráticos medios de ambos estimadores del valor del parámetro $\theta = 2$. [1 p.]

INFORMACIÓN DE POSIBLE INTERÉS SOBRE DISTRIBUCIONES:

- **Distribución normal**, $N(\mu, \sigma)$, $\mu \in \mathbb{R}$, $\sigma > 0$.

Función de densidad: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right)$, $x \in \mathbb{R}$.

Momentos: $\mathbb{E}(X) = \mu$, $V(X) = \sigma^2$.

- **Distribución gamma**, $\gamma(a, p)$, $a > 0$, $p > 0$. Cuando $p = 1$ se denomina “distribución exponencial de parámetro a ”.

Función de densidad: $f(x) = \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}$, para $x > 0$. Aquí $\Gamma(p)$ denota la llamada “función gamma”, $\Gamma(p) = \int_0^\infty e^{-x} x^{p-1} dx$ que verifica $\Gamma(p+1) = p\Gamma(p)$ para $p > 0$.

Momentos: $\mathbb{E}(X) = \frac{p}{a}$, $V(X) = \frac{p}{a^2}$.

- **Distribución beta**, $\text{Beta}(a, b)$, $a > 0$, $b > 0$.

Función de densidad: $f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$, para $x \in (0, 1)$.

Momentos: $\mathbb{E}(X) = \frac{a}{a+b}$, $V(X) = \frac{ab}{(a+b+1)(a+b)^2}$.