ESTADÍSTICA I **Curso 2021/2022**
Grado en Matemáticas

Relación 3 de problemas

En los ejercicios que siguen (a menos que se indique lo contrario) los estimadores están basados en una muestra $X_1, \ldots, X_n$ de v.a.i.i.d. con las distribuciones indicadas. La abreviatura e.m.v. indicará "estimador de máxima verosimilitud".

1. Sea X una variable aleatoria con distribución de Pareto dada por la función de densidad:

$$f(x;\theta) = \theta x_0^{\theta} x^{-\theta-1}, \quad x \geq x_0, \quad \theta > 1,$$

donde x_0 se supone que es un valor conocido. Sea $X_1, \ldots, X_n$ una muestra de X.

(a) Calcula la cantidad de información de Fisher.
(b) Calcula el estimador de θ por el método de los momentos.
(c) Determina el estimador de máxima verosimilitud de θ, estudia su consistencia y determina completamente su distribución asintótica.

2. Supongamos que el error X cometido en la medición de una magnitud es una v.a. con distribución normal de media 0 y varianza θ. Por tanto, la función de densidad de X es

$$f(x;\theta) = \frac{1}{\sqrt{\theta}\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\theta}\right), \ x \in \mathbb{R}.$$

Se desea estimar θ a partir de una muestra $X_1, \ldots, X_n$ de X.

(a) Calcula el estimador de máxima verosimilitud, T_n.
(b) Prueba que T_n es insesgado y eficiente.
(c) Estudia la distribución asintótica de T_n.

Indicación: $\mathrm{E}_\theta(X^4) = 3\theta^2$.

3. Se dispone de un gran lote de piezas producidas en una cadena de montaje. Denotemos por p la proporción de piezas defectuosas en ese lote. Supongamos que se seleccionan al azar sucesivamente (con reemplazamiento) piezas del lote hasta que se encuentra una defectuosa. Sea X la variable aleatoria que indica el número de la extracción en la que aparece la primera pieza defectuosa.

(a) Calcula $\mathrm{P}(X = k)$ para $k = 1, 2, \ldots$. Obtén el estimador de p por el método de los momentos, a partir de una muestra $X_1, \ldots, X_n$ de la v.a. X.
(b) Calcula el estimador de p por el método de máxima verosimilitud. Calcula su distribución asintótica.

4. Estudia si es eficiente el e.m.v. del parámetro λ en una distribución de Poisson.

5. Un modelo probabilístico utilizado en ocasiones para la v.a. "velocidad del viento" es la distribución de Rayleigh cuya función de densidad es

$$f(x;\theta) = \frac{x}{\theta^2} e^{\frac{-x^2}{2\theta^2}} \mathbb{I}_{[0,\infty)}(x), \ \theta > 0.$$

La media y la varianza de esta distribución son, respectivamente, $\mu = \theta\sqrt{\frac{\pi}{2}}$, $\sigma^2 = \frac{4-\pi}{2}\theta^2$.

(a) Calcula el estimador de θ por el método de los momentos. Estudia si es asintóticamente normal y, en caso afirmativo, calcula su varianza asintótica.
(b) Calcula el estimador de θ por el método de máxima verosimilitud. ¿Es consistente?
(c) Estudia si el estimador de máxima verosimilitud es asintóticamente normal y, en caso afirmativo, calcula su varianza asintótica.

6. Sea el modelo definido por la densidad

$$f(x;\theta) = \frac{1}{2\theta} \exp\left(-\frac{|x|}{\theta}\right), \quad x \in \mathbb{R}, \ \theta > 0.$$

(a) Calcula el estimador de máxima verosimilitud, $\hat{\theta}_n$, de θ, basado en una muestra de tamaño n.
(b) ¿Es insesgado $\hat{\theta}_n$? Calcular su error cuadrático medio.
(c) ¿Es consistente $\hat{\theta}_n$? ¿Es eficiente?
(d) Obtener la distribución límite de $\sqrt{n}(\hat{\theta}_n - \theta)$.

7. El tiempo (en minutos) que una persona espera el autobús cada mañana tiene distribución uniforme sobre el intervalo $(0, \theta)$, $\theta > 0$. La función de densidad a priori sobre $\Theta = (0, \infty)$ es $\pi(\theta) = 192/\theta^4$, si $\theta \geq 4$ y 0 si $\theta < 4$. En tres mañanas, los tiempos de espera observados han sido de 5, 3 y 8 minutos. Calcula el estimador Bayes de θ.

8. Sea X una v.a. con distribución normal de media μ y varianza θ. Estamos interesados en la estimación de θ basados en muestras $X_1, \ldots, X_n$ de tamaño n. Calcula la cota de Fréchet-Cramer-Rao para estimadores insesgados de θ.

9. Sea $X_1, \ldots, X_n$ una muestra de una v.a. con función de densidad

$$f(x;\theta) = \theta x^{\theta-1}, \ 0 < x < 1, \ \theta > 0.$$

Sea $T_n(X_1, \ldots, X_n) = -(1/n)\sum_{i=1}^n \log X_i$.

(a) Prueba que $\mathrm{E}_\theta(T_n) = 1/\theta$, $\mathrm{Var}_\theta(T_n) = 1/(n\theta^2)$.
(b) ¿Es eficiente T_n como estimador de $1/\theta$?

10. El número de fallos que se producen anualmente en cierto mecanismo es una v.a. con distribución de Poisson de parámetro θ. El valor de θ no es conocido exactamente, pero se tiene cierta información a priori que permite considerarlo como una v.a. con distribución $\gamma(\alpha, \beta)$ (α y β son conocidos). Si $x_1, \ldots, x_n$ son observaciones independientes de la variable aleatoria "número de fallos", calcular la distribución a posteriori y obtener, a partir de ella, un estimador puntual para θ.

11. Sea X una variable aleatoria absolutamente continua con distribución uniforme en el intervalo $[0,\theta]$, donde $\theta > 0$. Sea $X_1,\ldots,X_n$ una muestra aleatoria de X.

(a) Escribe y dibuja las funciones de densidad f_θ y distribución F_θ de X. Determina y dibuja (aproximadamente) la función de logverosimilitud, $\log L_n$, como función de θ. Prueba entonces que el estimador de máxima verosimilitud de θ es $X_{(n)} = \max\limits_{i=1,\ldots,n} X_i$. Indicación: Para calcular el e.m.v. no hay que derivar $\log L_n$, sólo prestar atención al soporte de f_θ y al de $\log L_n$.

(b) Calcula $\tilde{\theta}_n$, el estimador de θ por el método de los momentos. ¿Es asintóticamente normal? ¿Cuál es su error cuadrático medio?

(c) Demuestra que $T_n = \dfrac{n+1}{n} X_{(n)}$ es un estimador insesgado de θ.

(d) Calcula $\mathbb{E}\left(\left(\dfrac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right)$. Teniendo en cuenta el valor de $\mathbb{V}(\tilde{\theta}_n)$, ¿es válida la cota de Fréchet-Cramer-Rao en este caso? ¿Por qué crees que es así? Indicación: Prueba a calcular $\dfrac{\partial}{\partial\theta}\displaystyle\int_{\mathbb{R}} f_\theta(x)\,dx$ metiendo la derivada dentro del signo integral.

(e) Se tiene cierta información que permite suponer que θ es una v.a. con distribución de Pareto de densidad

$$\pi(\theta) = \begin{cases} \dfrac{\alpha\kappa^\alpha}{\theta^{\alpha+1}} & \text{si } \theta \geq \kappa, \\ 0 & \text{si } \theta < \kappa, \end{cases}$$

con $\kappa = 2$. Demuestra que la distribución a posteriori de θ es una Pareto de parámetros $\alpha + n$ y $\max(x_{(n)}, 2)$.

(f) Escribe un código en R que extraiga una muestra aleatoria de tamaño $n = 50$ de la distribución uniforme en $[0,3]$, calcule los estimadores $\tilde{\theta}_n$ y T_n y dibuje un estimador kernel de la densidad. ¿Qué aspecto tendría este estimador kernel para un parámetro de suavizado h muy pequeño (infrasuavizado) y muy grande (sobresuavizado) comparado con la densidad f_3 original?

12. Al tirar una moneda trucada, la probabilidad desconocida de obtener cara es $\theta \in (0,1)$. A priori, pensamos que $\mathrm{E}(\theta) = 1/2$ y $\mathrm{Var}(\theta) = 1/12$. Queremos estimar θ tirando una vez la moneda. ¿Cuál es el estimador Bayes en función de si ha resultado cara o cruz?