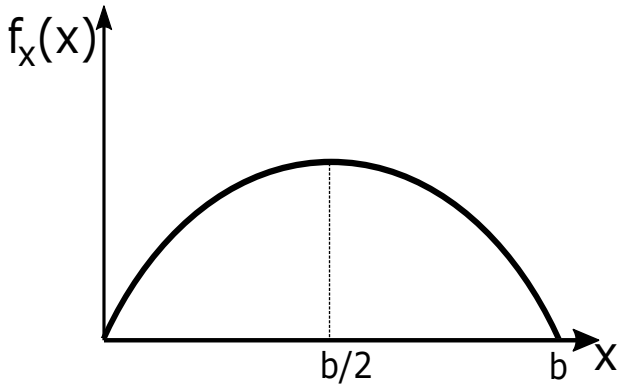
CPE (SEGUNDO CURSO)

PRÁCTICA 12 **SOLUCIONES** (**Curso 2024–2025**)

1.– Se sabe que la variable aleatoria X tiene la distribución de la figura (parábola).



Se pide:

a) Determinar la función de densidad y la función de distribución acumulada de X.

b) Dada una muestra de tamaño n determinar un estimador de b por el método de los momentos.

c) Estudiar, en lo posible, el sesgo y la consistencia del estimador.

——————————SOLUCIÓN——————————

Primero determinaremos el rango, función de densidad y función de distribución acumulada de X. El rango de X es $R_X = [0, b]$, la función de densidad es una parábola de la forma: $f_X(x) = a_0 + a_1 x + a_2 x^2$

La función de densidad debe verificar que pasa por los puntos $(0, 0)$ y $(b, 0)$. Además de verificarse que $\int_0^b f_X(x)dx = 1$

Al imponer las restricciones obtenemos $f_X(0) = a_0 = 0$, $f_X(b) = a_2 b^2 + a_1 b = 0$, $\int_0^b f_X(x)dx = \frac{a_2}{3} b^3 + \frac{a_1}{2} b^2 = 1$.

Por tanto:

$$f_X(x) = \frac{-6}{b^3} x^2 + \frac{6}{b^2} x$$

$$F_X(x) = \int_0^x f_X(x)dx = \frac{-6}{3b^3} x^3 + \frac{6}{2b^2} x^2 = \frac{-2}{b^3} x^3 + \frac{3}{b^2} x^2$$

.

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ \dfrac{-2}{b^3} x^3 + \dfrac{3}{b^2} x^2 & 0 \leq x \leq b \\ 1 & x \geq b \end{cases}$$

b) Dada la simetría en la función de densidad es directo determinar que

$$E[X] = b/2$$

Por tanto, un estimador de b obtenido por el método de los momentos será $\hat{b} = 2\bar{x}$

c)Primero analizaremos el sesgo

$$Sesgo = E[\hat{b}] - b$$

$$E[\hat{b}] = E[2\bar{x}] = 2E\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \frac{2}{n}E\left[\sum_{i=1}^{n} X_i\right] = \frac{2}{n}\sum_{i=1}^{n} E\left[X_i\right] =$$
$$= \frac{2}{n}\frac{n}{2}\frac{b}{} = b$$

$$Sesgo = E[\hat{b}] - b = 0$$

Por tanto el estimador es insesgado.

Para la analizar la consistencia calculamos primero la varianza

$$Var[\hat{b}] = Var[2\bar{x}] = 4\ Var\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \frac{4}{n^2}Var\left[\sum_{i=1}^{n} X_i\right] = \frac{4}{n^2}\sum_{i=1}^{n} Var\left[X_i\right] =$$
$$= \frac{4}{n^2}n\ \sigma_x^2 = \frac{4}{n}\sigma_x^2$$

$\lim_{n\to\infty} Var[\hat{b}] = 0$ y el estimador es consistente.

2.– La función

$$f_X(x) = \frac{(a+1)x^a}{\theta^{a+1}}, \quad 0 \le x \le \theta, \quad \theta > 0, \quad a > 0$$

es una función de densidad para cualesquiera valores de θ y a con las restricciones especificadas. Dada una muestra aleatoria de tamaño n de X y suponiendo $\theta =$ cte. deducir estimadores puntuales de a mediante todos los métodos que se conozcan.

Repítase el ejercicio suponiendo $a =$ cte. y encontrando estimadores puntuales de θ

——————————SOLUCIÓN——————————

1.- Estimadores de a

Calcularemos primero el estimador mediante el método de los momentos. La esperanza matemática de la variable es

$$E[X] = \int_0^\theta x\frac{(a+1)x^a}{\theta^{a+1}}dx = \frac{(a+1)}{\theta^{a+1}}\int_0^\theta x^{a+1}dx = \frac{a+1}{\theta^{a+1}}\frac{x^{a+2}}{a+2}\bigg|_0^\theta = \frac{\theta(a+1)}{a+2}$$

El estimador puntual que proporciona este método es el que proporciona la ecuación $\bar{x} = E[X]$. Por tanto

$$\bar{x} = \frac{\theta(\hat{a}+1)}{\hat{a}+2} \quad \Longrightarrow \quad \hat{a} = \frac{\theta - 2\bar{x}}{\bar{x} - \theta}$$

Si utilizamos ahora el método de máxima verosimilitud, debemos crear primero la función de verosimilitud

$$L = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{(a+1)x_i^a}{\theta^{a+1}} = \frac{(a+1)^n}{\theta^{n(a+1)}} \prod_{i=1}^{n} x_i^a$$

Tomando logaritmos

$$\ln L = n \ln(a+1) - n(a+1) \ln \theta + \sum_{i=1}^{n} a \ln x_i$$

Derivando con respecto a a

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a+1} - n \ln \theta + \sum_{i=1}^{n} \ln x_i$$

E igualando a cero

$$\frac{n}{\hat{a}+1} - n \ln \theta + \sum_{i=1}^{n} \ln x_i = 0 \quad \Longrightarrow \quad \hat{a} = \frac{n}{n \ln \theta - \sum_{i=1}^{n} \ln x_i} - 1$$

2.- Estimadores de θ

Utilizando el método de los momentos,

$$\bar{x} = \frac{\hat{\theta}(a+1)}{a+2} \quad \Longrightarrow \quad \hat{\theta} = \frac{\bar{x}(a+2)}{a+1}$$

Si utilizamos ahora el método de máxima verosimilitud, debemos crear primero la función de verosimilitud. Al igual que antes

$$L = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{(a+1)x_i^a}{\theta^{a+1}} = \frac{(a+1)^n}{\theta^{n(a+1)}} \prod_{i=1}^{n} x_i^a$$

Tomando logaritmos

$$\ln L = n \ln(a+1) - n(a+1) \ln \theta + \sum_{i=1}^{n} a \ln x_i$$

Derivando con respecto a θ

$$\frac{\partial \ln L}{\partial \theta} = \frac{-n(a+1)}{\theta}$$

y el problema no puede resolverse igualando a cero la derivada.

Si observamos la función de verosimilitud (que ha de maximizarse), es claro que esa función tendrá un máximo para el valor más pequeño de $\theta^{n(a+1)}$. Como $n > 1$ y $a > 0$ eso es lo mismo que decir para el valor más pequeño de θ. Pero θ es el extremo superior del rango de X. Luego si algún valor de la muestra estuviera fuera del rango, la correspondiente función de densidad sería nula y por consiguiente también lo sería la función de verosimilitud. Luego el valor de θ que hace máxima L es

$$\hat{\theta} = max\{x_i\}$$

.

3.– Se modela un sistema ecológico de manera que la especie animal dominante consume un tanto por uno, X, $0 \leq X \leq 1$, de los recursos totales del sistema que denotamos R. Se ha supuesto que X tiene la función de densidad de probabilidad:

$$f_X(x) = 2(1-x), \quad 0 \leq x \leq 1$$

a) Hallar la función de densidad de probabilidad de $W = XR$, el total de recursos consumido por la especie animal.

b) Suponiendo que se dispone de una muestra de W, $w_1, w_2, ...w_n$, determinar cómo se calcularía el estimador de máxima verosimilitud de R. Particularizar el resultado para el caso de una muestra de tamaño 1.

——————————SOLUCIÓN——————————

a) El total de los recursos consumidos por la especie animal es $W = XR$. Como esta relación es creciente en X,

$$f_W(w) = \left|\frac{dx}{dw}\right| f_X\left(\frac{w}{R}\right) = \frac{1}{R}2(1-\frac{w}{R}) = \frac{2(R-w)}{R^2}, \quad 0 \leq w \leq R$$

y es obvio que

$$\int_0^R \frac{2(R-w)}{R^2}\ dw = \frac{2}{R^2}\left[Rw - \frac{w^2}{2}\right]_0^R = 1$$

b) La verosimilitud de una muestra de W, como función del parámetro R, es

$$L(R|w_1, w_2, ...w_n) = \prod_{i=1}^{n} \frac{2(R-w_i)}{R^2}$$

Tomando logaritmos

$$\ln L(R|w_1, w_2, ...w_n) = \sum_{i=1}^{n} \ln \frac{2(R-w_i)}{R^2} = \sum_{i=1}^{n} \left[\ln(2(R-w_i)) - 2\ln R\right]$$

Derivando con respecto a R

$$\frac{d\ln L}{dR} = \sum_{i=1}^{n} \left(\frac{1}{R-w_i}\right) - \frac{2n}{R}$$

y la ecuación a resolver para calcular el estimador de máxima verosimilitud de R sería

$$\sum_{i=1}^{n} \left(\frac{1}{\hat{R}-w_i}\right) - \frac{2n}{\hat{R}} = 0$$

En general, esta ecuación se resolvería para un conjunto de valores $w_1, w_2, ...w_n$ mediante cualquier método numérico. Para el caso $n = 1$

$$\frac{1}{\hat{R}-w_1} - \frac{2}{\hat{R}} = 0 \implies \hat{R} = 2w_1$$

4.– Un conductor tarda un tiempo X en pagar un peaje. Se supone que X responde a una distribución exponencial de parámetro λ trasladada al rango $[3,\infty)$ (en segundos). Para estimar el parámetro λ se espera que paguen 100 conductores y se anota el tiempo que ha tardado el conductor más rápido. Esta operación se repite n veces para obtener los tiempos mínimos $y_1, y_2, ..., y_n$. Se pide:

a) Hallar la distribución de Y, tiempo mínimo de pago entre 100 conductores.

b) Establecer la ecuación que debe satisfacer el estimador de máxima verosimilitud de λ, basado en los datos $y_1, y_2, ..., y_n$.

——————————SOLUCIÓN——————————

Sea $x_1,\ x_2,\ ...,\ x_{100}$ el tiempo de pago de los 100 conductores. Sabemos que

$$f_X(x) = \lambda e^{-\lambda(x-3)}, \quad F_X(x) = 1 - e^{-\lambda(x-3)}, \quad x \geq 3$$

a) Sea Y el tiempo mínimo de pago entre 100 conductores. Calculemos la distribución de Y suponiendo a los 100 conductores independientes

$$P[Y \geq y] = \big[P[X \geq y]\big]^{100} = [1 - F_X(y)]^{100} = [1 - (1 - e^{-\lambda(y-3)})]^{100} = e^{-100\lambda(y-3)} = 1 - F_Y(y)$$

por tanto

$$\boxed{F_Y(y) = 1 - e^{-100\lambda(y-3)}, \quad y \geq 3}$$

b) Si tenemos una muestra de tamaño n la función de verosimilitud es

$$L(y_1, y_2, ..., y_n|\lambda) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \prod_{i=1}^{n} 100\lambda e^{-100\lambda(y_i-3)} = 100^n \lambda^n e^{-100\lambda \sum(y_i-3)}$$

Tomando logaritmos

$$\ln L(y_1, y_2, ..., y_n|\lambda) = n \ln 100 + n \ln \lambda - 100\lambda \sum(y_i - 3)$$

Derivando e igualando a cero,

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n}{\hat{\lambda}} - 100 \sum(y_i - 3) = 0 \quad \Rightarrow \quad \boxed{\hat{\lambda} = \frac{n}{100 \sum(y_i - 3)}}$$

5.– Determinar los estimadores de los parámetros de una distribución logarítmico normal por el método de máxima verosimilitud.

——————————SOLUCIÓN——————————

La función de densidad de una distribución lognormal puede escribirse como

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \ e^{-\frac{1}{2\sigma^2}\left(\ln(\frac{x}{m})\right)^2}$$

donde por σ representamos la desviación logarítmica y por m la mediana de la distribución. Con esta notación, la función de verosimilitud es

$$L = \prod_{i=1}^{n} \frac{1}{x_i\sigma\sqrt{2\pi}} \ e^{-\frac{1}{2\sigma^2}\left(\ln(\frac{x_i}{m})\right)^2} = \frac{1}{\sigma^n(\sqrt{2\pi})^n} \prod_{i=1}^{n} \frac{1}{x_i} \ e^{-\frac{1}{2\sigma^2}\left(\ln(\frac{x_i}{m})\right)^2}.$$

Tomando logaritmos

$$\ln L = -n\ln\sigma - n\ln\sqrt{2\pi} + \sum_{i=1}^{n} \ln\Big(\frac{1}{x_i}\Big) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} \big(\ln(\frac{x_i}{m})\big)^2 =$$

$$= -n\ln\sigma - n\ln\sqrt{2\pi} - \sum_{i=1}^{n} \ln x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} \big(\ln(\frac{x_i}{m})\big)^2$$

Derivando con respecto a σ y a m para hallar los valores que hacen máxima la función de verosimilitud obtenemos

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n} \big(\ln(\frac{x_i}{m})\big)^2$$

$$\frac{\partial \ln L}{\partial m} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} \Big[2\ln(\frac{x_i}{m})(-\frac{1}{m})\Big] = \frac{1}{m\sigma^2}\sum_{i=1}^{n} \Big[\ln x_i - \ln m\Big]$$

Si igualamos a cero esta última ecuación resulta

$$\sum_{i=1}^{n} \Big[\ln x_i - \ln \hat{m}\Big] = 0, \quad \Rightarrow \quad \hat{m} = e^{\frac{\sum \ln x_i}{n}}$$

De la primera ecuación, igualando igualmente a cero obtenemos

$$n\hat{\sigma}^2 = \sum_{i=1}^{n} \big(\ln(\frac{x_i}{\hat{m}})\big)^2, \quad \Rightarrow \quad \hat{\sigma}^2 = \frac{1}{n}\ \sum_{i=1}^{n} \big(\ln x_i - \ln \hat{m}\big)^2$$