# Tercer Parcial

25 de mayo de 2021

**Indicaciones generales**

- Este es un examen **individual** con una duración de **120 minutos: de 1:00 a 3:00 p.m.**.
- Sólo se permite el uso de calculadoras como medio electrónico. Los celulares deben estar apagados durante todo el examen.
- Las cámaras deben estar activas durante todo el examen.
- Puede usar una única hoja con apuntes. El uso de libros u otro recurso "analógico" diferente no está permitido.
- Cualquier incumplimiento de lo anterior conlleva a la anulación del examen.
- Las respuestas deben estar totalmente justificadas.
- Al finalizar, suba a eaulas un **único** archivo .pdf con su solución. En caso de problemas con la plataforma envíe su archivo por el chat privado de Teams a martin.andrade@urosario.edu.co.
- ¡Suerte y ánimo!

1. (**15 pts**) Suponga que $X_1, X_2, \ldots, X_n$ son variables aleatorias i.i.d. con función de densidad:

$$f_X(x|\sigma) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}, \quad x \in \mathbb{R}, \quad \sigma > 0.$$

Encuentre dos estimadores de $\sigma$ utilizando el método de los momentos y el método de máxima verosimilitud. Observación: Note que la distribución de las variables es simétrica (su media es 0 para todo $\sigma > 0$) entonces hay que utilizar el segundo momento para la primera parte.

2. (**10 pts**) Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de una distribución con función de densidad:

$$f_Y(y|\theta) = \begin{cases} \theta y^{\theta-1}, & 0 < y < 1, \\ 0 & \text{d.l.c} \end{cases}$$

donde $\theta > 0$.

- Muestre que $\sum_{i=1}^{n} -ln(Y_i)$ es suficiente para $\theta$.
- Suponga que usted sabe que:

$$\mathbb{E}\left(\left(2\theta \sum_{i=1}^{n} -ln(Y_i)\right)^{-1}\right) = \frac{1}{2(n-1)}.$$

Halle el MVUE de $\theta$.

3. (**15 pts**) Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de tamaño $n$ de una población con función de densidad

$$\mathrm{f}_Y(y;\theta) = (1+\theta y)/2, \quad -1 < y < +1; \quad -1 < \theta < +1$$

Derive la *estructura general* de la región de rechazo de la prueba más potente en el nivel $\alpha$ para $\mathrm{H}_0 : \theta = 0$ contra $\mathrm{H}_\mathrm{a} : \theta = 0.50$. Observación: este punto no pide encontrar un valor específico de $k$, si no la forma de la región de rechazo, dado un valor específico de $\alpha$.

4. (**10 pts**)

Sea $\{X_n\}_{n\in\mathbb{N}^*}$ una sucesión variables aleatorias exponenciales con parámetro $\lambda = n$ ($X_n$ es una variable aleatoria con distribución exponencial con parámetro $\lambda = n$).

- Muestre que $\{X_n\}_{n\in\mathbb{N}^*}$ converge en probabilidad a un $c \in \mathbb{R}$ y determine el valor de $c$.
- Se dice que $\{X_n\}_{n\in\mathbb{N}^*}$ converge en distribución a $X$ (variable aleatoria), denotado $X_n \xrightarrow{d} X$ si:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

para todo $x$ donde $F_X(x)$ es continua. Teniendo en cuenta esta definición muestre que $X_n \xrightarrow{d} X$, donde $X$ es la variable aleatoria que toma como único valor el 0 con probabilidad 1.

Solución

Daniel Forero C
25. Mayo. 2021

1. (**15 pts**) Suponga que $X_1, X_2, \ldots, X_n$ son variables aleatorias i.i.d. con función de densidad:

$$f_X(x|\sigma) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}, \quad x \in \mathbb{R}, \quad \sigma > 0.$$

Encuentre dos estimadores de $\sigma$ utilizando el método de los momentos y el método de máxima verosimilitud. Observación: Note que la distribución de las variables es simétrica (su media es 0 para todo $\sigma > 0$) entonces hay que utilizar el segundo momento para la primera parte.

Sol. Metodo de los momentos.

Para el segundo momento calculamos:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_x(x|\delta)dx = \int_{-\infty}^{\infty} \frac{1}{2\delta} x^2 e^{\frac{-|x|}{\delta}} dx$$

$$= \int_{-\infty}^{0} \frac{1}{2\delta} x^2 e^{\frac{x}{\delta}} dx + \underbrace{\int_{0}^{\infty} \frac{1}{2\delta} x^2 e^{-\frac{x}{\delta}} dx}_{\substack{u=-x \\ du=-dx}}$$

$$= \int_{-\infty}^{0} \frac{1}{2\delta} x^2 e^{\frac{x}{\delta}} dx - \int_{0}^{-\infty} \frac{1}{2\delta} u^2 e^{\frac{u}{\delta}} du$$

$$= 2\int_{-\infty}^{0} \frac{1}{2\delta} x^2 e^{\frac{x}{\delta}} dx = \frac{1}{\delta} \int_{-\infty}^{0} x^2 e^{\frac{x}{\delta}} dx \rightsquigarrow z = \frac{x}{\delta} \rightsquigarrow x = \delta z$$

$$dz = \frac{dx}{\delta} \rightsquigarrow dx = \delta dz$$

$$= \frac{1}{\delta}\int_{-\infty}^{0} \delta^2 z^2 e^{z} \delta dz = \delta^2 \int_{-\infty}^{0} z^2 e^{z} dz = \delta^2 \left[ z^2 e^{z} - 2e^{z} z + 2e^{z} \right] \Big|_{-\infty}^{0}$$

$$= \delta^2 [2-0] = 2\delta^2$$

Por lo tanto, si $\mu_2 = E(X^2) = 2\delta^2$, entonces $\hat{\delta}^2 = \frac{\mu_2}{2} \Rightarrow \hat{\delta} = \sqrt{\frac{\mu_2}{2}}$

Metodo máx. verosimilitud.

$$L(x_1, \ldots, x_n | \delta) = \prod_{i=1}^{n} f_x(x_i|\delta) = \prod \left( \frac{1}{2\delta} e^{-|x_i|/\delta} \right) = \frac{1}{2^n \delta^n} e^{-\frac{1}{\delta}\sum |x_i|}$$

$$\Rightarrow \ell(x_1, \ldots, x_n|\delta) = \ln\left(\frac{1}{2^n \delta^n}\right) - \frac{1}{\delta}\sum |x_i| = -n\ln(2) - n\ln(\delta) - \frac{1}{\delta}\sum|x_i|$$

Se deriba respecto a $\delta$:

$$\frac{\partial \ell}{\partial \delta} = -\frac{n}{\delta} + \frac{1}{\delta^2}\sum_{i=1}^{n}|x_i| = 0 \rightsquigarrow \sum|x_i| = \delta n \Rightarrow \hat{\delta} = \frac{\sum|x_i|}{n}$$

Vemos que $\hat{\delta}$ es un máximo de $\ell$ pues su segunda deribada es:

$$\frac{\partial^2 \ell}{\partial^2 \delta} = \frac{n}{\delta^2} - \frac{2}{\delta^3} \cdot \sum_{i=1}^{n}|x_i| \Rightarrow \frac{\partial^2 \ell}{\partial^2 \hat{\delta}^2} = n \cdot \frac{n^2}{(\sum|x_i|)^2} - 2\sum|x_i| \cdot \frac{n^3}{(\sum|x_i|)^3}$$

$$= \frac{n^3}{(\sum|x_i|)^2} - \frac{2n^3}{(\sum|x_i|)^2} = \frac{n^3}{(\sum|x_i|)^2} < 0 \rightsquigarrow \text{Máx.}$$

Luego, $\hat{\delta} = \frac{\sum|x_i|}{n}$ es el MLE.

2. (10 pts) Sea $Y_1, Y_2, \dots, Y_n$ una muestra aleatoria de una distribución con función de densidad:

$$f_Y(y|\theta) = \begin{cases} \theta y^{\theta-1}, & 0 < y < 1, \\ 0 & \text{d.l.c} \end{cases}$$

donde $\theta > 0$.

- Muestre que $\sum_{i=1}^{n} -ln(Y_i)$ es suficiente para $\theta$.
- Suponga que usted sabe que:

$$\mathbb{E}\left(\left(2\theta \sum_{i=1}^{n} -ln(Y_i)\right)^{-1}\right) = \frac{1}{2(n-1)}.$$

  Halle el MVUE de $\theta$.

Sol. Calculamos la verosimilitud:

$$L(x_1, \dots, x_n | \theta) = \prod_{i=1}^{n} f_y(y_i|\theta) = \prod \theta y_i^{\theta-1} = \theta^n (\prod y_i)^{\theta-1}$$

Observamos que si $U = \sum -\ln(y_i) = \sum \ln\left(\frac{1}{y_i}\right) = \ln\left[(\prod y_i)^{-1}\right]$. Así: $e^U = (\prod y_i)^{-1}$

Luego $\prod y_i = e^{-U}$

Por tanto $L(y_1, \dots, y_n|\theta) = \theta^n e^{-U(\theta-1)}$

Ahora, si tomamos $g(U, \theta) = \theta^n e^{-U(\theta-1)}$, $h(y_1, \dots, y_n) = 1$ ; entonces:

$$L(y_1, \dots, y_n|\theta) = g(U, \theta) h(y_1, \dots, y_n)$$

Por teo. de factorización $U$ es suficiente para $\theta$.

Ahora:

$$\mathbb{E}\left[(2\theta)^{-1} U^{-1}\right] = \frac{1}{2(n-1)}$$

$$(2\theta)^{-1} \mathbb{E}\left(\frac{1}{U}\right) = \frac{1}{2(n-1)}$$

$$\mathbb{E}\left(\frac{1}{U}\right) = \frac{\theta}{n-1}$$

Por tanto, si $V = \frac{n-1}{U}$, entonces $\mathbb{E}(V) = \theta$

Entonces $V$ es estadístico insesgado de $\theta$.

Así, como $V$ es función de $U$, q' es suficiente, entonces $V$ es el MVUE de $\theta$.

3. (15 pts) Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de tamaño $n$ de una población con función de densidad

$$f_Y(y;\theta) = (1+\theta y)/2, \quad -1 < y < +1; \quad -1 < \theta < +1$$

Derive la *estructura general* de la región de rechazo de la prueba más potente en el nivel $\alpha$ para $H_0 : \theta = 0$ contra $H_a : \theta = 0.50$. Observación: este punto no pide encontrar un valor específico de $k$, si no la forma de la región de rechazo, dado un valor específico de $\alpha$.

Sol. Tenemos $H_a: \theta_o = 0$

$H_0: \theta_a = 0.5$

Calculamos la verosimilitud:

$$L(y_1, \ldots, y_n \mid \theta) = \prod_{i=1}^{n} f_y(y_i \mid \theta) = \prod \frac{1+\theta y_i}{2} = \frac{1}{n^2} \prod 1+\theta y_i$$

Por el lema Neyman-Pearson, la region de rechazo de la prueba mas potente en nivel $\alpha$, es:

$$\frac{L(\theta_o)}{L(\theta_a)} < k \rightsquigarrow \frac{\frac{1}{n^2}\prod 1+\theta_o y_i}{\frac{1}{n^2}\prod 1+\theta_a y_i} < k$$

$$\frac{\prod 1+\theta_o y_i}{\prod 1+\theta_a y_i} < k$$

Para $\theta_o = 0$, $\theta_a = 0.5$, tenemos:

$$\frac{\prod(1)}{\prod 1+y_i/2} < k$$

$$1 < k \cdot \prod(1+y_i/2)$$

$$\prod(1+\frac{y_i}{2}) > \frac{1}{k} = k'$$

Por tanto, si $U = \prod(1+y_i/2)$ (estadistico) entonces la región de rechazo tiene la forma:

$$RR = \{U > k'\}$$

Ahora, para un nivel $\alpha$:

$$\alpha = P(U \in RR \mid \theta = 0) = P(U > k' \mid \theta = 0)$$

$$\rightsquigarrow \mu = E(y) = \int_{-1}^{1} \frac{y(1+\theta y)}{2} dy = \frac{1}{2}\int_{-1}^{1}(y+\theta y^2)\,dy = \frac{1}{2}\left[\frac{y^2}{2} + \frac{\theta}{3}y^3\right]\Big|_{-1}^{1} = \ldots$$

$$\ldots = \frac{1}{2}\left[\frac{1}{2} + \frac{\theta}{3} - \frac{1}{2} + \frac{\theta}{3}\right] = \frac{\theta}{3}$$

$$\approx P\left(\frac{U - \theta/3}{\delta/\sqrt{n}} > \frac{k' - \theta/3}{\delta/\sqrt{n}}\right) = P\left(Z > \frac{\sqrt{n}}{\delta} k'\right)$$

$$\Rightarrow Z_\alpha = \frac{\sqrt{n}}{\delta} k'$$

$$k' = \frac{\delta}{\sqrt{n}} Z_\alpha$$

Así $RR = \left\{U > \frac{\delta}{\sqrt{n}} Z_\alpha\right\}$ para un nivel $\alpha$.

4. (10 pts)

Sea $\{X_n\}_{n\in\mathbb{N}^*}$ una sucesión variables aleatorias exponenciales con parámetro $\lambda = n$ ($X_n$ es una variable aleatoria con distribución exponencial con parámetro $\lambda = n$).

- Muestre que $\{X_n\}_{n\in\mathbb{N}^*}$ converge en probabilidad a un $c \in \mathbb{R}$ y determine el valor de $c$.
- Se dice que $\{X_n\}_{n\in\mathbb{N}^*}$ converge en distribución a $X$ (variable aleatoria), denotado $X_n \xrightarrow{d} X$ si:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

para todo $x$ donde $F_X(x)$ es continua. Teniendo en cuenta esta definición muestre que $X_n \xrightarrow{d} X$, donde $X$ es la variable aleatoria que toma como único valor el 0 con probabilidad 1.

Sol. La funcion de distribución acumulada de $X_n \sim e^{\lambda=n}$ es:

$$F_x(x) = 1 - e^{-nx} \quad , x \geq 0$$

Sea $\varepsilon > 0$:

$$P(|X_n - c| > \varepsilon) = 1 - P(|X_n - c| \leq \varepsilon)$$

$$= 1 - P(-\varepsilon \leq X_n - c \leq \varepsilon)$$

$$= 1 - P(X_n \leq \varepsilon) + P(X_n \leq c - \varepsilon)$$

$$= 1 - F_{X_n}(c+\varepsilon) + F_{X_n}(c-\varepsilon)$$

Para $c = 0$ tenemos:

$$c - \varepsilon < 0 \Rightarrow F_{X_n}(c - \varepsilon) = 0$$
$$c + \varepsilon > 0 \Rightarrow F_{X_n}(c+\varepsilon) = 1 - e^{-n(c+\varepsilon)}$$

Así, $P(|X_n - 0| > \varepsilon) = 1 - (1 - e^{-n(0+\varepsilon)}) = e^{-n\varepsilon} \xrightarrow[n\to\infty]{} 0$

Por tanto, $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} P(|X_n - 0| > \varepsilon) = 0$$

Luego $X_n$ converge en probabilidad a $c = 0$.

Como $F_{X_n}(x) = 1 - e^{-nx}$, $x \geq 0$, entonces:

$$\lim_{n\to\infty} F_{X_n}(x) = 1 - 0 = 1 \quad , \forall x \geq 0.$$

Por tanto, $F_X(x) = 1$, $\forall x \geq 0$.
luego, $\forall x < 0$, $F_{X_n}(x) = 0$, así:

$$\lim_{n\to\infty} F_{X_n}(x) = 0.$$

Por tanto, $F_X(x) = 0$, $\forall x < 0 \Rightarrow F_X(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$, Así, $X$ toma a 0 para valor con probabilidad 1.