PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA

INTERROGACIÓN 1
Inferencia Estadística - EYP2114 - EYP2127

Profesora	:	Ana María Araneda
Ayudante	:	Sebastián Guerra
Fecha	:	5 de septiembre de 2023

1. **[30 %]** Considere una muestra aleatoria $X_1, \ldots, X_n$ de una distribución con función de densidad dada por:

$$f_\theta(x) = \exp\{-(x-\theta)\}\, I_{(\theta,\infty)}(x).$$

a) Encuentre un estadístico suficiente minimal para θ.

Solución: La función de densidad conjunta corresponde a:

$$\begin{aligned} f_{\boldsymbol{X}}(x_1,\ldots,x_n) &= \prod_{i=1}^{n} \exp\{-(x_i-\theta)\} I_{(\theta,\infty)}(x_i) \\ &= \exp\left\{-\sum_{i=1}^{n} x_i + n\theta\right\} I_{(\theta,\infty)}(x_{(1)}), \end{aligned}$$

donde $x_{(1)}$ corresponde al mínimo de la muestra.
Sean $\boldsymbol{x} = (x_1,\ldots,x_n)$ e $\boldsymbol{y} = (y_1,\ldots,y_n)$ dos puntos arbitrarios del espacio muestral, y considere la razón:

$$\begin{aligned} \frac{f_{\boldsymbol{X}}(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{y})} &= \frac{\exp\{-\sum_{i=1}^{n} x_i + n\theta\} I_{(\theta,\infty)}(x_{(1)})}{\exp\{-\sum_{i=1}^{n} y_i + n\theta\} I_{(\theta,\infty)}(y_{(1)})} \\ &= C(\boldsymbol{x},\boldsymbol{y}) \frac{I_{(\theta,\infty)}(x_{(1)})}{I_{(\theta,\infty)}(y_{(1)})}. \end{aligned}$$

Esta cantidad no depende de θ, **para todo** θ**, ssi** $x_{(1)} = y_{(1)}$. Luego,

$$T(\boldsymbol{X}) = X_{(1)}$$

es un estadístico suficiente minimal para θ.

[3,0] Debe incluir el punto en negrita. Se pide justificar todos los pasos.

b) Considere el vector aleatorio $\boldsymbol{Y} = (Y_1, \ldots, Y_{n-1})$, donde $Y_i = X_{(n)} - X_{(i)}, i = 1, \ldots, n-1$. Determine si $\boldsymbol{Y}$ es un estadístico ancilar para θ.

<u>Solución:</u> Dado que:

$$f_\theta(x) = f_Z(x - \theta),$$

donde $f_Z(z) = e^{-z} I_{\mathbb{R}^+}(z)$, **donde** f_Z **no depende de** θ, $\{f_\theta, \theta \in \mathbb{R}\}$ corresponde a una familia de localización, y podemos considerar $\mathbf{X_i} = \mathbf{Z_i} + \theta$, **donde** $\mathbf{Z_1}, \ldots, \mathbf{Z_n}$ **son independientes y su función de densidad no depende de** θ.

Por otra parte, de la definición de $Z_1, \ldots, Z_n$ se tiene que $\mathbf{X_{(i)}} = \mathbf{Z_{(i)}} + \theta, \mathbf{i} = \mathbf{1}, \ldots, \mathbf{n}$. Por ello, la distribución de cada $Y_i, i = 1, \ldots, n$, puede obtenerse como:

$$\begin{aligned} F_{Y_i}(y) &= P(Y_i \leq y) \\ &= P(X_{(n)} - X_{(i)} \leq y) \\ &= P(Z_{(n)} + \theta - (Z_{(i)} + \theta) \leq y) \\ &= P(Z_{(n)} - Z_{(i)} \leq y). \end{aligned}$$

Dado que la distribución de $Z_1, \ldots, Z_n$ no depende de θ, la distribución de $Z_{(1)}, \ldots, Z_{(n)}$ tampoco depende de θ, por lo que F_{Y_i} no depende de $\theta, i = 1, \ldots, n$. Luego, $Y_i, \imath = 1, \ldots, n-1$ es ancilar para θ, y por ello $\boldsymbol{Y}$ también lo es.

[3,0] Debe incluir los puntos en negrita. Se pide justificar todos los pasos.

[1,0] punto base

2. **[35 %]** Considere la muestra $X_1, \ldots, X_n$, donde X_1 sigue una distribución Normal$(0, 1 - \theta^2)$, $|\theta| < 1$, y, para $i = 2, \ldots, n$ se cumple que:

$$X_i | X_1 = x_1, \ldots, X_{i-1} = x_{i-1} \sim \text{Normal}(\theta x_{i-1}, 1).$$

Encuentre un estadístico suficiente minimal para θ.

Puede utilizar que:

- Si $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ es un vector aleatorio, su función de densidad puede obtenerse como:

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = f(x_1) f(x_2|x_1) f(x_3|x_1, x_2) \ \ldots \ f(x_n|x_1, \ldots, x_{n-1}).$$

- Si $X \sim \text{Normal}(\mu, \sigma^2)$ $\mu \in \mathbb{R}, \sigma > 0$, su función de densidad corresponde a:

$$f_X(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\} I_{\mathbb{R}}(x).$$

Solución: La densidad conjunta corresponbde a:

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}) &= \left(\frac{1}{2\pi(1-\theta^2)}\right)^{1/2} \exp\left\{-\frac{1}{2(1-\theta^2)}x_1^2\right\} \prod_{i=2}^{n} \left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{1}{2}(x_i-\theta x_{i-1})^2\right\} \\ &= \left(\frac{1}{2\pi(1-\theta^2)}\right)^{1/2} \left(\frac{1}{2\pi}\right)^{(n-1)/2} \exp\left\{-\frac{1}{2(1-\theta^2)}x_1^2\right\} \\ &\quad \times \exp\left\{-\frac{1}{2}\left(\sum_{i=2}^{n} x_i^2 - 2\theta \sum_{i=2}^{n} x_i x_{i-1} + \theta^2 \sum_{i=2}^{n} x_{i-1}^2\right)\right\} \qquad \textbf{[1,0]} \end{aligned}$$

Para encontrar un estadístico suficiente minimal, tomamos dos puntos cualquiera en el espacio muestral, $\boldsymbol{x}$, $\boldsymbol{y}$, y revisamos las condiciones para que la razón $f_{\boldsymbol{X}}(\boldsymbol{x})/f_{\boldsymbol{X}}(\boldsymbol{y})$ no dependa de θ, para todo $\theta \in \mathbb{R}$:

$$\begin{aligned} \frac{f_{\boldsymbol{X}}(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{y})} &= \frac{\exp\left\{-\frac{1}{2(1-\theta^2)}x_1^2\right\} \exp\left\{-\frac{1}{2}\left(\sum_{i=2}^n x_i^2 - 2\theta \sum_{i=2}^n x_i x_{i-1} + \theta^2 \sum_{i=2}^n x_{i-1}^2\right)\right\}}{\exp\left\{-\frac{1}{2(1-\theta^2)}y_1^2\right\} \exp\left\{-\frac{1}{2}\left(\sum_{i=2}^n x_i^2 - 2\theta \sum_{i=2}^n y_i y_{i-1} + \theta^2 \sum_{i=2}^n y_{i-1}^2\right)\right\}} \\ &= h(\boldsymbol{x}, \boldsymbol{y}) \frac{\exp\left\{-\frac{1}{2(1-\theta^2)}x_1^2 + \theta \sum_{i=2}^n x_i x_{i-1} - 1/2 \cdot \theta^2 \sum_{i=2}^n x_{i-1}^2\right\}}{\exp\left\{-\frac{1}{2(1-\theta^2)}y_1^2 + \theta \sum_{i=2}^n y_i y_{i-1} - 1/2 \cdot \theta^2 \sum_{i=2}^n y_{i-1}^2\right\}} \qquad \textbf{[2,0]} \end{aligned}$$

Para que esta razón no dependa de θ se requiere:

$$\begin{aligned} x_1^2 &= y_1^2 \\ \sum_{i=2}^{n} x_i x_{i-1} &= \sum_{i=2}^{n} y_i y_{i-1} \\ \sum_{i=2}^{n} x_{i-1}^2 &= \sum_{i=2}^{n} y_{i-1}^2. \qquad \textbf{[2,0]} \end{aligned}$$

Luego, un estadístico suficiente minimal para θ corresponde a:

$$T(\boldsymbol{X}) = \left(X_1^2,\ \sum_{i=2}^{n} X_i X_{i-1},\ \sum_{i=2}^{n} X_{i-1}^2\right). \qquad \textbf{[1,0]}$$

[1,0] punto base

3. **[35 %]** Considere una muestra aleatoria bivariada, $(X_1, Y_1), \ldots, (X_n, Y_n)$, donde cada observación sigue una distribución Uniforme en el círculo de radio θ centrado en el origen. Esto es:

$$f_{X_i,Y_i}(x_i, y_i) = \frac{1}{\pi\theta^2} I_{(0,\theta)}((x_i^2 + y_i^2)^{1/2}).$$

Encuentre un estadístico suficiente minimal y determine si es completo.

Puede utilizar que:

- La densidad de la variable aleatoria $Z_i = (X_i^2 + Y_i^2)^{1/2}$ está dada por:

$$f_{Z_i}(z) = \frac{2z}{\theta^2} I_{(0,\theta)}(z),$$

$i = 1, \ldots, n$.

- Si $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución con función de densidad f y función de distribución F, la función de densidad del máximo $X_{(n)}$ está dada por:

$$f_{X_{(n)}}(x) = nf(x)\left(F(x)\right)^{n-1}.$$

Solución: La densidad conjunta corresponde a:

$$\begin{aligned} f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y}) &= \left(\frac{1}{\pi\theta^2}\right)^n \prod_{i=1}^n I_{(0,\theta)}((x_i^2 + y_i^2)^{1/2}) \\ &= \left(\frac{1}{\pi\theta^2}\right)^n I_{(0,\theta)}\left(\max_{i=1,\ldots,n}(x_i^2 + y_i^2)^{1/2}\right) \qquad \textbf{[0,5]} \end{aligned}$$

Para encontrar un estadístico suficiente minimal, tomamos dos puntos cualquiera en el espacio muestral, $(\boldsymbol{x},\boldsymbol{y})$ y $(\boldsymbol{u},\boldsymbol{v})$, y revisamos las condiciones para que la razón $f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})/f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{u},\boldsymbol{v})$ no dependa de θ, para todo $\theta \in \mathbb{R}^+$:

$$\frac{f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})}{f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{u},\boldsymbol{v})} = \frac{I_{(0,\theta)}\left(\max_{i=1,\ldots,n}(x_i^2 + y_i^2)^{1/2}\right)}{I_{(0,\theta)}\left(\max_{i=1,\ldots,n}(u_i^2 + v_i^2)^{1/2}\right)}. \qquad \textbf{[0,5]}$$

Esta razón no depende de θ ssi:

$$I_{(0,\theta)}\left(\max_{i=1,\ldots,n}(x_i^2 + y_i^2)^{1/2}\right) = I_{(0,\theta)}\left(\max_{i=1,\ldots,n}(u_i^2 + v_i^2)^{1/2}\right).$$

Luego, un estadístico suficiente minimal corresponde a:

$$T(\boldsymbol{X},\boldsymbol{Y}) = \max_{i=1,\ldots,n}(X_i^2 + Y_i^2)^{1/2}. \qquad \textbf{[1,0]}$$

Para verificar si es completo: consideramos las variables aleatorias $Z_i = X_i^2 + Y_i^2)^{1/2}, i = 1, \ldots, n$. Se indica que la función de densidad de cada una de ellas corresponde a:

$$f_Z(z) = \frac{2z}{\theta^2} I_{(0,\theta)}(z),$$

con lo que es posible obtener su función de distribución como:

$$F_Z(z) = \int_0^z \frac{2u}{\theta^2} I_{(0,\theta)}(u)\, du = \frac{z^2}{\theta^2},$$

para $0 < z < \theta$, $F_Z(z) = 0$ si $z \leq 0$ y $F_Z(z) = 1$ para $z \geq \theta$ **[1,0]**. Luego, la función de densidad del máximo está dada por:

$$\begin{aligned} f_{Z_{(n)}}(z) &= n\, \frac{2z}{\theta^2} \left(\frac{z^2}{\theta^2}\right)^{n-1} I_{(0,\theta)}(z) \\ &= 2n \left(\frac{z}{\theta}\right)^{2n-1} I_{(0,\theta)}(z). \qquad \textbf{[1,0]} \end{aligned}$$

Tomamos una función g tal que:

$$\begin{aligned} \mathrm{E}_\theta(g(Z_{(n)})) = 0 \qquad &\Longleftrightarrow \qquad \int_0^\theta g(t) 2n \left(\frac{t}{\theta}\right)^{2n-1} dt = 0 \qquad \textbf{[1,0]} \\ &\Longleftrightarrow \qquad \int_0^\theta g(t) t^{2n-1} dt = 0 \\ &\Longleftrightarrow \qquad g(\theta)\theta^{2n-1} = 0, \qquad \textbf{[0,5]} \end{aligned}$$

para todo valor de $\theta > 0$, lo que ocurre ssi $P(g(Z_{(n)}) = 0) = 1$. Luego, el estadístico suficiente minimal $Z_{(n)}$ es completo. **[0,5]**

[1,0] punto base