XII EDICIÓN PREMIOS JORGE JUAN
ESTADÍSTICA

Sea $(X_1,...,X_n)$ una muestra aleatoria simple de la variable aleatoria *X*, con función de densidad:

$$f_\theta(x) = \frac{1}{\theta_2 - \theta_1}, \quad \theta_1 < x < \theta_2,$$

donde $\theta = (\theta_1, \theta_2) \in \Theta = \{(\theta_1, \theta_2) \in \mathbb{R}^2 \mid \theta_1 < \theta_2\}$. Demostrar:

a) $S(X_1,...,X_n) = \left(X_{(1)}, X_{(n)}\right)$, donde $X_{(1)} = \min\{X_1,...,X_n\}$ y $X_{(n)} = \max\{X_1,...,X_n\}$ es un estadístico suficiente para el parámetro θ.

b) La función de densidad de *S* es $g_\theta(y,t) = n(n-1)\dfrac{(t-y)^{n-2}}{(\theta_2-\theta_1)^n}, \quad \theta_1 < y < t < \theta_2$.

c) *S* es un estadístico completo para θ.

d) $T_1(X_1,...,X_n) = \dfrac{X_{(1)} + X_{(n)}}{2}$ es el estimador insesgado de mínima varianza de $g_1(\theta) = \dfrac{\theta_1 + \theta_2}{2}$.

e) $T_2(X_1,...,X_n) = \dfrac{(n+1)}{(n-1)}\left(X_{(n)} - X_{(1)}\right)$ es el estimador insesgado de mínima varianza de $g_2(\theta) = \theta_2 - \theta_1$.

SOLUCIONES

a) Para demostrar que $S\left(X_1,...,X_n\right)=\left(X_{(1)},X_{(n)}\right)$ es un estadístico suficiente para θ, utilizaremos el Teorema de Factorización. Para ello, calculemos la función de densidad de $(X_1,...,X_n)$:

$$f_\theta\left(x_1,...,x_n\right)=\prod_{i=1}^n f_\theta\left(x_i\right)=\frac{1}{\left(\theta_2-\theta_1\right)^n},\quad \theta_1<x_i<\theta_2,\ \forall i=1,...,n.$$

Para poder aplicar el Teorema de Factorización, debemos extender esta función a cualquier vector $\left(x_1,...,x_n\right)\in\mathbb{R}^n$. Esto lo conseguimos utilizando funciones indicadoras que valgan uno cuando $\theta_1<x_i<\theta_2,\ \forall i=1,...,n$, y cero en el resto. Dado que:

$$\theta_1<x_i<\theta_2,\ \forall i=1,...,n \Leftrightarrow x_{(1)}>\theta_1 \quad \text{y} \quad x_{(n)}<\theta_2,$$

tenemos:

$$f_\theta\left(x_1,...,x_n\right)=\frac{1}{\left(\theta_2-\theta_1\right)^n}\cdot I_{]-\infty,\theta_2[}\left(x_{(n)}\right)\cdot I_{]\theta_1,+\infty[}\left(x_{(1)}\right).$$

$S\left(X_1,...,X_n\right)=\left(X_{(1)},X_{(n)}\right)$ es un estadístico suficiente para el parámetro θ.

b) La función de distribución conjunta de *S*, $G_\theta\left(y,t\right)$, se obtiene a partir de la función de distribución de X, $F_\theta\left(x\right)$, de la forma que aparece en las notas:

$$G_\theta\left(y,t\right)=F_\theta\left(t\right)^n-\left(F_\theta\left(t\right)-F_\theta\left(y\right)\right)^n,\quad y<t.$$

Derivando respecto de *y* y de *t*, obtenemos la función de densidad conjunta:

$$g_\theta\left(y,t\right)=n\left(n-1\right)\left(F_\theta\left(t\right)-F_\theta\left(y\right)\right)^{n-2}f_\theta\left(y\right)f_\theta\left(t\right),\quad y<t.$$

Calculemos $F_\theta\left(x\right)$:

$$F_\theta\left(x\right)=\begin{cases}0, & x\le\theta_1\\ \dfrac{x-\theta_1}{\theta_2-\theta_1}, & \theta_1<x<\theta_2,\\ 1, & x\ge\theta_2.\end{cases}$$

Así, obtenemos:

$$g_\theta\left(y,t\right)=n\left(n-1\right)\left(\frac{t-\theta_1}{\theta_2-\theta_1}-\frac{y-\theta_1}{\theta_2-\theta_1}\right)^{n-2}\cdot\frac{1}{\left(\theta_2-\theta_1\right)^2},\quad \theta_1<y<t<\theta_2 \Leftrightarrow$$

$$g_\theta\left(y,t\right)=\frac{n\left(n-1\right)}{\left(\theta_2-\theta_1\right)^n}\cdot\left(t-y\right)^{n-2},\quad \theta_1<y<t<\theta_2.$$

c) Supongamos que $E_\theta\left(h(S)\right)=0, \forall\theta\in\Theta,$ siendo $h(S)$ una función del estadístico *S*. Esto significa que:

$$E_\theta\left(h(S)\right)=\int_{\theta_1}^{\theta_2}\int_y^{\theta_2} h(y,t)\cdot g_\theta(y,t)\,dtdy=0, \forall\theta_1<\theta_2 \Rightarrow$$
$$\int_{\theta_1}^{\theta_2}\int_y^{\theta_2} h(y,t)\cdot(t-y)^{n-2}\,dtdy=0, \forall\theta_1<\theta_2. \quad (1)$$

Utilicemos la Regla de Barrow, y sea $H(y,t)$ una primitiva de $h(y,t)\cdot(t-y)^{n-2}$, para cada $y\in\left]\theta_1,\theta_2\right[$. Entonces:

$$\int_y^{\theta_2} h(y,t)\cdot(t-y)^{n-2}\,dt=H(y,\theta_2)-H(y,y).$$

Sustituyendo en (1), obtenemos:

$$\int_{\theta_1}^{\theta_2} H(y,\theta_2)dy-\int_{\theta_1}^{\theta_2} H(y,y)dy=0, \forall\theta_1<\theta_2. \quad (2)$$

Sean ahora $G_1(y)$ primitiva de $H(y,\theta_2)$, $G_2(y)$ primitiva de $H(y,y)$, $y\in\left]\theta_1,\theta_2\right[$. Entonces, de (2), obtenemos:

$$G_1(\theta_2)-G_1(\theta_1)-G_2(\theta_2)+G_2(\theta_1)=0, \quad \forall\theta_1<\theta_2 \Rightarrow$$
$$G_1(\theta_2)-G_2(\theta_2)=G_1(\theta_1)-G_2(\theta_1), \quad \forall\theta_1<\theta_2.$$

Por lo tanto, la función $(G_1-G_2)(y)$ es constante $\forall y\in\mathbb{R}$, y su derivada será cero, es decir:

$$H(y,\theta_2)-H(y,y)=0, \quad \forall y\in\mathbb{R}, \quad \forall\theta_2>y.$$

Esto significa que $H(y,t)$ es constante, $\forall y\in\mathbb{R}, \forall t>y$, y su derivada será cero, es decir:

$$h(y,t)\cdot(t-y)^{n-2}=0, \forall y<t \Rightarrow h(y,t)=0, \forall y<t \Rightarrow h(S)\equiv 0.$$

Por lo tanto, S es un estadístico completo para θ.

d) Como $T_1(X_1,...,X_n)$ es función del estadístico suficiente y completo para θ, S, de acuerdo con el Teorema de Lehmann-Scheffé, para que sea el estimador insesgado de mínima varianza de $g_1(\theta)=\dfrac{\theta_1+\theta_2}{2}$ debe cumplir $E_\theta(T_1)=g_1(\theta), \forall\theta\in\Theta.$ Tenemos:

$$E_\theta(T_1)=E_\theta\left(\frac{X_{(1)}+X_{(n)}}{2}\right)=\frac{n(n-1)}{2(\theta_2-\theta_1)^n}\int_{\theta_1}^{\theta_2}\int_y^{\theta_2}(y+t)\cdot(t-y)^{n-2}\,dtdy. \quad (3)$$

$$\int_y^{\theta_2}(y+t)\cdot(t-y)^{n-2}\,dt \underset{z=t-y}{\underset{\downarrow}{=}} \int_0^{\theta_2-y}(z+2y)z^{n-2}dz=\left[\frac{z^n}{n}+2y\frac{z^{n-1}}{n-1}\right]_0^{\theta_2-y}=$$
$$\frac{(n-1)(\theta_2-y)^n+2yn(\theta_2-y)^{n-1}}{n(n-1)}.$$

Sustituyendo en (3), obtenemos:

$$E_\theta(T_1)=\frac{1}{2(\theta_2-\theta_1)^n}\int_{\theta_1}^{\theta_2}\left\{(n-1)(\theta_2-y)^n+2yn(\theta_2-y)^{n-1}\right\}dy=$$

$$\underset{z=\theta_2-y}{\underset{\downarrow}{=}}\quad\frac{1}{2(\theta_2-\theta_1)^n}\int_0^{\theta_2-\theta_1}\left\{(n-1)z^n+2n(\theta_2-z)z^{n-1}\right\}dz=$$

$$\frac{1}{2(\theta_2-\theta_1)^n}\left[(n-1)\frac{z^{n+1}}{n+1}+2n\left\{\theta_2\frac{z^n}{n}-\frac{z^{n+1}}{n+1}\right\}\right]_0^{\theta_2-\theta_1}=$$

$$\frac{1}{2(\theta_2-\theta_1)^n}\left[(n-1)\frac{(\theta_2-\theta_1)^{n+1}}{n+1}+2n\left\{\theta_2\frac{(\theta_2-\theta_1)^n}{n}-\frac{(\theta_2-\theta_1)^{n+1}}{n+1}\right\}\right]=....=\frac{\theta_1+\theta_2}{2}.$$

e) Razonando exactamente igual que en el apartado anterior, debemos comprobar que $E_\theta(T_2)=g_2(\theta),\forall\theta\in\Theta.$ Tenemos:

$$E_\theta(T_2)=E_\theta\left(\frac{(n+1)}{(n-1)}\left(X_{(n)}-X_{(1)}\right)\right)=\frac{(n+1)}{(n-1)}\cdot\frac{n(n-1)}{(\theta_2-\theta_1)^n}\int_{\theta_1}^{\theta_2}\int_y^{\theta_2}(t-y)\cdot(t-y)^{n-2}\,dtdy=$$

$$\frac{n(n+1)}{(\theta_2-\theta_1)^n}\int_{\theta_1}^{\theta_2}\int_y^{\theta_2}(t-y)^{n-1}\,dtdy=\frac{n(n+1)}{(\theta_2-\theta_1)^n}\int_{\theta_1}^{\theta_2}\left[\frac{(t-y)^n}{n}\right]_y^{\theta_2}dy=\frac{(n+1)}{(\theta_2-\theta_1)^n}\int_{\theta_1}^{\theta_2}(\theta_2-y)^n dy=$$

$$\frac{(n+1)}{(\theta_2-\theta_1)^n}\left[-\frac{(\theta_2-y)^{n+1}}{n+1}\right]_{\theta_1}^{\theta_2}=\theta_2-\theta_1.$$