TEMA 4: Estimación puntual. Insesgadez y mínima varianza

4.1. PLANTEAMIENTO DEL PROBLEMA DE ESTIMACIÓN

$(X_1, \ldots, X_n)$ muestra aleatoria simple de $X \to \{P_\theta;\ \theta \in \Theta\}$

Se trata de aproximar el verdadero valor del parámetro (o de alguna función paramétrica) a partir de las observaciones muestrales.

Estimador puntual: *Un estimador de θ es un estadístico, $T(X_1, \ldots, X_n)$, que toma valores en Θ.*

$\downarrow$ *Función de pérdida :* $L : \Theta \times \Theta \to \mathbb{R}$

$L(\theta, t)$: pérdida que conlleva estimar el parámetro por el valor t si su verdadero valor es θ.

Función de riesgo de un estimador: *Si $T(X_1, \ldots, X_n)$ es un estimador de θ, su función de riesgo bajo la función de pérdida L es la que asigna a cada valor del parámetro la pérdida media asociada al estimador*:

$$\begin{aligned} R_T^L : \Theta &\longrightarrow \mathbb{R} \\ \theta &\longmapsto R_T^L(\theta) = E_\theta\left[L(\theta, T(X_1, \ldots, X_n))\right]. \end{aligned}$$

Estimador óptimo bajo una función de pérdida L: *Es el estimador que minimiza uniformemente la función de riesgo. Esto es, un estimador del parámetro, $T(X_1, \ldots, X_n)$, es óptimo bajo la función de pérdida L si, para cualquier otro estimador, $T'(X_1, \ldots, X_n)$, se tiene:*

$$R_T^L(\theta) \leq R_{T'}^L(\theta), \quad \forall \theta \in \Theta.$$

ESTIMACIÓN DE MENOR ERROR CUADRÁTICO MEDIO

Función de pérdida cuadrática : $L : \Theta \times \Theta \longrightarrow \mathbb{R}$

$$(\theta, t) \longmapsto L(\theta, t) = (t - \theta)^2$$

Función de riesgo de un estimador $T(X_1, \ldots, X_n)$:

$$R_T^L(\theta) = E_\theta\left[(T(X_1, \ldots, X_n) - \theta)^2\right] \ \to \ \textit{error cuadrático medio de } T(X_1, \ldots, X_n)$$

$\downarrow$

$$E_\theta\left[T(X_1, \ldots, X_n)\right] = \theta \Rightarrow E_\theta\left[(T(X_1, \ldots, X_n) - \theta)^2\right] = Var_\theta[T(X_1, \ldots, X_n]$$

4.2. ESTIMACIÓN INSESGADA DE MÍNIMA VARIANZA

$(X_1, \ldots, X_n)$ muestra aleatoria simple de $X \to \{P_\theta;\ \theta \in \Theta\}$

Estimador de $g(\theta)$: estadístico $T(X_1, \ldots, X_n)$ con valores en $g(\Theta)$

Estimador insesgado: *Un estimador de* $g(\theta)$, $T(X_1, \ldots, X_n)$, *es insesgado (centrado) si*

$$E_\theta\left[T(X_1, \ldots, X_n)\right] = g(\theta), \quad \forall \theta \in \Theta.$$

Estimador insesgado uniformemente de mínima varianza (UMVUE): *Un estimador de* $g(\theta)$, $T(X_1, \ldots, X_n)$, *insesgado y de segundo orden, es UMVUE para* $g(\theta)$ *si para cualquier otro estimador insesgado de* $g(\theta)$, $T'(X_1, \ldots, X_n)$, *se tiene*:

$$Var_\theta[T(X_1, \ldots, X_n)] \leq Var_\theta[T'(X_1, \ldots, X_n)], \quad \forall \theta \in \Theta.$$

- *Unicidad*: El UMVUE de cualquier función paramétrica, si existe, es único.
- *Linealidad*: Si $T_i(X_1, \ldots, X_n)$ es el UMVUE para $g_i(\theta)$, $i = 1, 2$, y $a_1, a_2 \in \mathbb{R}$, entonces $a_1 T_1(X_1, \ldots, X_n) + a_2 T_2(X_1, \ldots, X_n)$ es el UMVUE para $a_1 g_1(\theta) + a_2 g_2(\theta)$.

Teorema de Rao-Blackwell: *Si* $T(X_1, \ldots, X_n)$ *es suficiente para* $\{P_\theta;\ \theta \in \Theta\}$ *y* $S(X_1, \ldots, X_n)$ *es un estimador insesgado de* $g(\theta)$ *de segundo orden:*

- $E\left[S(X_1, \ldots, X_n) \big/ T(X_1, \ldots, X_n)\right]$ *es estimador insesgado de* $g(\theta)$ *y de segundo orden.*
- $Var_\theta\left[E\left[S(X_1, \ldots, X_n) \big/ T(X_1, \ldots, X_n)\right]\right] \leq Var_\theta[S(X_1, \ldots, X_n)], \ \forall \theta \in \Theta.$

Teorema de Lehmann-Scheffé: *Sea* $T(X_1, \ldots, X_n)$ *un estadístico suficiente y completo para* $\{P_\theta;\ \theta \in \Theta\}$. *Si* $g(\theta)$ *admite un estimador insesgado de segundo orden,* $S(X_1, \ldots, X_n)$, *entonces existe el UMVUE de* $g(\theta)$ *y está dado por*

$$E\left[S(X_1, \ldots, X_n) \big/ T(X_1, \ldots, X_n)\right].$$

4.3. ESTIMADORES EFICIENTES

$(X_1, \ldots, X_n)$ muestra aleatoria simple de $X \to \{P_\theta;\ \theta \in \Theta\}$

$f_\theta \to$ función de densidad o función masa de probabilidad de X bajo P_θ

$f_\theta^n \to$ función de densidad o función masa de probabilidad de $(X_1, \ldots, X_n)$ bajo P_θ

Condiciones de regularidad de Fréchet-Cramér-Rao: *La familia de distribuciones $\{P_\theta;\ \theta \in \Theta\}$ es regular en el sentido de Fréchet-Cramér-Rao si satisface:*

i) *Θ es un intervalo abierto de $\mathbb{R}$.*

ii) *El conjunto de valores de la variable es independiente de θ:*

$$\forall \theta \in \Theta,\ \{x \ /\ f_\theta(x) > 0\} = \chi.$$

iii) *$\forall x \in \chi,\ f_\theta(x)$ es derivable respecto de θ y*

- $\displaystyle\sum_{x \in \chi} \frac{\partial f_\theta(x)}{\partial \theta} = \frac{\partial}{\partial \theta} \sum_{x \in \chi} f_\theta(x) = 0,\ \forall \theta \in \Theta$ *(si las distribuciones son discretas)*
- $\displaystyle\int_\chi \frac{\partial f_\theta(x)}{\partial \theta} dx = \frac{\partial}{\partial \theta} \int_\chi f_\theta(x) dx = 0,\ \forall \theta \in \Theta$ *(si las distribuciones son continuas).*

Función de información de Fisher: *Si $\{P_\theta;\ \theta \in \Theta\}$ es regular, se definen las funciones de información asociadas a X y a la muestra, respectivamente, como*:

$$I_X(\theta) = E_\theta\left[\left(\frac{\partial \ln f_\theta(X)}{\partial \theta}\right)^2\right], \qquad I_{X_1,\ldots,X_n}(\theta) = E_\theta\left[\left(\frac{\partial \ln f_\theta^n(X_1, \ldots, X_n)}{\partial \theta}\right)^2\right].$$

- $I_X \geq 0, \quad I_{X_1,\ldots,X_n} \geq 0.$
- $E_\theta\left[\dfrac{\partial \ln f_\theta(X)}{\partial \theta}\right] = 0,\ \ \forall \theta \in \Theta, \qquad I_X(\theta) = Var_\theta\left[\dfrac{\partial \ln f_\theta(X)}{\partial \theta}\right],\ \ \forall \theta \in \Theta.$
- $E_\theta\left[\dfrac{\partial \ln f_\theta^n(X_1, \ldots, X_n)}{\partial \theta}\right] = 0,\ \ \forall \theta \in \Theta,\ \ I_{X_1,\ldots,X_n}(\theta) = Var_\theta\left[\dfrac{\partial \ln f_\theta^n(X_1, \ldots, X_n)}{\partial \theta}\right],\ \ \forall \theta \in \Theta.$
- *Aditividad*: $I_{X_1,\ldots,X_n}(\theta) = n I_X(\theta), \quad \forall \theta \in \Theta.$

Estadístico regular: *Un estadístico $T(X_1,\ldots,X_n)$ es regular en el sentido de Fréchet-Cramér-Rao si* (en función de que las distribuciones sean discretas o continuas):

- $$\underbrace{\frac{\partial}{\partial\theta}\sum_{(x_1,\ldots,x_n)\in\chi^n} T(x_1,\ldots,x_n) f_\theta^n(x_1,\ldots,x_n)}_{E_\theta[T(X_1,\ldots,X_n)]} = \underbrace{\sum_{(x_1,\ldots,x_n)\in\chi^n} T(x_1,\ldots,x_n)\frac{\partial f_\theta^n(x_1,\ldots,x_n)}{\partial\theta}}_{E_\theta\left[T(X_1,\ldots,X_n)\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta}\right]}$$

- $$\underbrace{\frac{\partial}{\partial\theta}\int_{\chi^n} T(x_1,\ldots,x_n) f_\theta^n(x_1,\ldots,x_n)dx_1\cdots dx_n}_{E_\theta[T(X_1,\ldots,X_n)]} = \underbrace{\int_{\chi^n} T(x_1,\ldots,x_n)\frac{\partial f_\theta^n(x_1,\ldots,x_n)}{\partial\theta}dx_1\cdots dx_n}_{E_\theta\left[T(X_1,\ldots,X_n)\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta}\right]}.$$

Cota de Fréchet-Cramér-Rao: *Si $\{P_\theta;\ \theta\in\Theta\}$ es regular, $0<I_X(\theta)<+\infty$, $\forall\theta\in\Theta$ y $T(X_1,\ldots,X_n)$ es un estadístico regular, de segundo orden, insesgado en una función paramétrica derivable $g(\theta)$, se tiene:*

a) $Var_\theta\left[T(X_1,\ldots,X_n)\right] \geq \dfrac{(g'(\theta))^2}{I_{X_1,\ldots,X_n}(\theta)}, \quad \forall\theta\in\Theta.$

b) $\forall\theta\in\Theta\ /\ g'(\theta)\neq 0:$

$$Var_\theta[T] = \frac{(g'(\theta))^2}{I_{X_1,\ldots,X_n}(\theta)} \ \Leftrightarrow\ \exists a(\theta)\neq 0\ /\ P_\theta\left(\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta} = a(\theta)\left[T(X_1,\ldots,X_n)-g(\theta)\right]\right)=1.$$

Estimador eficiente: *Sea $\{P_\theta;\ \theta\in\Theta\}$ regular, con $0<I_X(\theta)<+\infty$, $\forall\theta\in\Theta$, y $g(\theta)$ una función paramétrica derivable. Un estimador de $g(\theta)$, $T(X_1,\ldots,X_n)$, se dice que es eficiente si es insesgado, regular y su varianza alcanza la cota para cualquier valor del parámetro*:

$$Var_\theta\left[T(X_1,\ldots,X_n)\right] = \frac{(g'(\theta))^2}{I_{X_1,\ldots,X_n}(\theta)}, \quad \forall\theta\in\Theta.$$

Caracterización de estimadores eficientes: *Sea $\{P_\theta;\ \theta\in\Theta\}$ regular, con $0<I_X(\theta)<+\infty$, $\forall\theta\in\Theta$, $g(\theta)$ una función paramétrica derivable, no constante, y $T(X_1,\ldots,X_n)$ un estimador de $g(\theta)$. Una condición necesaria y suficiente para que $T(X_1,\ldots,X_n)$ sea eficiente es*

$$\forall\theta\in\Theta\ \ \exists a(\theta)\neq 0 \text{ tal que } \begin{cases} i)\ P_\theta\left(\dfrac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta} = a(\theta)\left[T(X_1,\ldots,X_n)-g(\theta)\right]\right)=1 \\ \\ ii)\ I_{X_1,\ldots,X_n}(\theta) = a(\theta)g'(\theta). \end{cases}$$

Propiedades de los estimadores eficientes: $(T \equiv T(X_1, \ldots, X_n))$

- *Si T es eficiente para $g(\theta)$, $aT + b$ es eficiente para $ag(\theta) + b$ $(a \neq 0, b \in \mathbb{R})$ y solo este tipo de funciones paramétricas admite estimador eficiente.*
- *Si existe estimador eficiente para una función paramétrica, es único.*
- *Solo existen estimadores eficientes en familias de tipo exponencial:*

$$\text{T \textit{eficiente para alguna función paramétrica}}$$
$$\Downarrow$$
$$f_\theta^n(x_1, \ldots, x_n) = \exp\{Q(\theta)T(x_1, \ldots, x_n) + D(\theta) + S((x_1, \ldots, x_n))\}, \quad (x_1, \ldots, x_n) \in \chi^n.$$

- *Si T es un estimador eficiente para $g(\theta)$, T es un estadístico suficiente. Si, además, el conjunto imagen de $Q(\theta)$ contiene a un abierto de $\mathbb{R}$, también es completo. En tal caso, T es el UMVUE para $g(\theta)$.*