

Definición

- $\mathbf{X} \sim P_{\theta}$ con $\theta \in \Theta$.
- Se dice que $\mathbf{T} = t(\mathbf{X})$ es suficiente para θ si:

la distribución de $\mathbf{X} \mid \mathbf{T} = \mathbf{t}$ es independiente de θ para todo $\mathbf{t} \in \mathcal{T}$ tal que

$$\mathbb{P}_{\theta}(\mathbf{T} \notin \mathcal{T}) = 0 \text{ para todo } \theta \in \Theta$$

Rao Blackwell

Sea \mathbf{X} un vector de una distribución perteneciente a la familia $\mathcal{P} = \{P_{\theta} : \theta \in \Theta\}$. Sea \mathbf{T} un estadístico suficiente para θ y $\delta(\mathbf{X})$ un estimador de $q(\theta)$.

Sea $R(\theta, \delta) = \mathbb{E}_{\theta} L(\theta, \delta(\mathbf{X}))$ con $L(\theta, d)$ convexa en d .

Supongamos $\mathbb{E}_{\theta} |\delta(\mathbf{X})| < \infty$. Definamos un nuevo estimador

$$\eta(\mathbf{T}) = \mathbb{E}(\delta(\mathbf{X})|\mathbf{T}).$$

- Luego se tiene

- (i) $R(\theta, \eta) \leq R(\theta, \delta), \forall \theta \in \Theta$
- (ii) Cuando $L(\theta, d)$ es estrictamente convexa en d , la igualdad en (i) se satisface si y sólo si

$$P_{\theta}(\eta(\mathbf{T}) = \delta(\mathbf{X})) = 1 \quad \forall \theta \in \Theta$$

- (iii) Si $\delta(\mathbf{X})$ es insesgado, entonces $\eta(\mathbf{T})$ también lo es.

Teorema de Factorización

Sea \mathbf{X} un vector aleatorio con función de densidad o función de probabilidad puntual $p(\mathbf{x}, \theta)$, $\theta \in \Theta$.

El estadístico $\mathbf{T} = t(\mathbf{X})$ es suficiente para θ si y sólo si existen dos funciones g y h tales que

$$p(\mathbf{x}, \theta) = g(t(\mathbf{x}), \theta)h(\mathbf{x})$$

Otros resultados

Corolario.

Sea \mathbf{X} un vector aleatorio con función de densidad o función de probabilidad puntual $p(\mathbf{x}, \theta)$, $\theta \in \Theta$. Supongamos que la familia $\mathcal{P} = \{p(\mathbf{x}, \theta)\}$ tiene soporte común, independiente de θ .

Entonces, una condición necesaria y suficiente para que \mathbf{T} sea suficiente para θ es que fijados θ_1 y θ_2 el cociente

$$\frac{p(\mathbf{x}, \theta_1)}{p(\mathbf{x}, \theta_2)}$$

sea función de \mathbf{T} .

Otros resultados

Corolario.

Sea \mathbf{X} un vector aleatorio con función de densidad o función de probabilidad puntual $p(\mathbf{x}, \theta)$, $\theta \in \Theta$. Supongamos que la familia $\mathcal{P} = \{p(\mathbf{x}, \theta)\}$ tiene soporte común, independiente de θ .

Entonces, una condición necesaria y suficiente para que \mathbf{T} sea suficiente para θ es que fijados θ_1 y θ_2 el cociente

$$\frac{p(\mathbf{x}, \theta_1)}{p(\mathbf{x}, \theta_2)}$$

sea función de \mathbf{T} .

Lema. Si \mathbf{X} es un vector aleatorio con una distribución $F(\mathbf{x}, \theta)$, con $\theta \in \Theta$ si $\mathbf{T} = t(\mathbf{X})$ es un estadístico suficiente para θ y si m es una función biunívoca de \mathbf{T} entonces el estadístico $\mathbf{V} = m(\mathbf{T})$ también es suficiente para θ .

Minimal suficiente

Definición: Sea \mathbf{X} un vector aleatorio con distribución P_θ con $\theta \in \Theta$.

Se dice que un *estadístico* $\mathbf{T} = t(\mathbf{X})$ es *minimal suficiente* para θ si:

- \mathbf{T} es suficiente para θ .
- Dado cualquier otro estadístico $\mathbf{U} = g(\mathbf{X})$ suficiente para θ existe una función H tal que $\mathbf{T} = H(\mathbf{U})$.

Resultados

Teorema. Supongamos que \mathbf{X} tiene una distribución perteneciente a una familia finita de distribuciones $\mathcal{P} = \{P_{\theta_i} \mid 1 \leq i \leq k\}$ con densidades o probabilidades puntuales p_i , $1 \leq i \leq k$ todas con el mismo soporte.

Entonces el estadístico

$$\mathbf{T}(\mathbf{X}) = \left(\frac{p_2(\mathbf{X})}{p_1(\mathbf{X})}, \dots, \frac{p_k(\mathbf{X})}{p_1(\mathbf{X})} \right)$$

es minimal suficiente.

Resultados

Teorema. Supongamos que \mathbf{X} tiene una distribución perteneciente a una familia de distribuciones $\mathcal{P} = \{P_{\theta} \mid \theta \in \Theta\}$ con soporte común.

Sea $\mathcal{P}_0 \subset \mathcal{P}$.

Supongamos que

- $\mathbf{T} = t(\mathbf{X})$ es un estadístico minimal suficiente para \mathcal{P}_0
- $\mathbf{T} = t(\mathbf{X})$ es suficiente para \mathcal{P} ,

entonces \mathbf{T} es minimal suficiente para \mathcal{P} .

Completitud

Definición. Sea \mathbf{X} un vector aleatorio cuya distribución pertenece a una familia $\mathcal{P} = \{P_{\theta} : \theta \in \Theta\}$.

Un estadístico $\mathbf{T} = t(\mathbf{X})$ se dice completo si

$$\mathbb{E}_{\theta}(g(\mathbf{T})) = 0 \quad \forall \theta \in \Theta$$

implica que

$$\mathbb{P}_{\theta}(g(\mathbf{T}) = 0) = 1 \quad \forall \theta \in \Theta$$

Resultados

Teorema. Supongamos que \mathbf{X} tiene una distribución perteneciente a una familia de distribuciones $\mathcal{P} = \{P_{\theta} \mid \theta \in \Theta\}$.

Supongamos que

- existe un estadístico minimal suficiente para \mathcal{P}
- $\mathbf{T} = t(\mathbf{X})$ es suficiente y completo para \mathcal{P} ,

entonces \mathbf{T} es minimal suficiente para \mathcal{P} .

Resultados

Teorema de Basu. Supongamos que \mathbf{X} tiene una distribución perteneciente a una familia de distribuciones $\mathcal{P} = \{P_{\theta} \mid \theta \in \Theta\}$.

Supongamos que

- $\mathbf{T} = t(\mathbf{X})$ es suficiente y completo para \mathcal{P} .

Entonces,

Dado cualquier estadístico $\mathbf{U} = g(\mathbf{X})$ ancilario, o sea, tal que su distribución no depende de θ , se tiene que

\mathbf{U} es independiente de \mathbf{T} .

Teorema de Lehmann-Scheffé

- Sea \mathbf{X} un vector aleatorio de cuya distribución pertenece a la familia de distribuciones $\mathcal{P} = \{P_{\theta} \quad \theta \in \Theta\}$.
- Sea \mathbf{T} un estadístico suficiente y completo para θ .

Luego, dada una función $q : \Theta \rightarrow \mathbb{R}$, se tiene que

- Existe a lo sumo un estimador insesgado de $q(\theta)$, basado en \mathbf{T} .
- Si $\delta(\mathbf{T})$ es un estimador insesgado de $q(\theta)$, entonces $\delta(\mathbf{T})$ es IMVU.
- Si $\delta(\mathbf{X})$ es un estimador insesgado para $q(\theta)$, luego $\eta(\mathbf{T}) = \mathbb{E}(\delta(\mathbf{X})|\mathbf{T})$ es un estimador IMVU para $q(\theta)$.

Teorema

- Sea \mathbf{X} un vector aleatorio de cuya distribución pertenece a la familia de distribuciones $\mathcal{P} = \{P_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta\}$.
 - Sea \mathbf{T} un estadístico suficiente y completo para $\boldsymbol{\theta}$.
- (a) Para cada función U –estimable $q(\boldsymbol{\theta})$ existe un estimador insesgado que minimiza uniformemente el riesgo $R(\boldsymbol{\theta}, \delta) = \mathbb{E}_{\boldsymbol{\theta}} L(\boldsymbol{\theta}, \delta)$ para cualquier pérdida $L(\boldsymbol{\theta}, d)$ convexa en d .
En particular, ese estimador es IMVU.
- (b) El estimador IMVU de (a) es el único estimador insesgado que es función de \mathbf{T} y es el único estimador insesgado con mínimo riesgo si el riesgo es finito y $L(\boldsymbol{\theta}, d)$ es estrictamente convexa en d .

Definición.

Se dice que una familia de distribuciones continuas o discretas en \mathbb{R}^q , P_{θ} , donde $\theta \in \Theta \subset \mathbb{R}^p$ es una *familia exponencial a k parámetros* si su densidad o probabilidad puntual se puede escribir como

$$p(\mathbf{x}, \theta) = A(\theta) e^{\sum_{j=1}^k \eta_j(\theta) t_j(\mathbf{x})} h(\mathbf{x}) \quad (1)$$

donde

- $\eta_j : \Theta \rightarrow \mathbb{R}$,
- $A : \Theta \rightarrow \mathbb{R}^+$
- $t_j : \mathbb{R}^q \rightarrow \mathbb{R}$
- $h : \mathbb{R}^q \rightarrow \mathbb{R}^+$.

Resultados

Teorema. Una familia exponencial a k parámetros cuya función de densidad viene dada por (1) tiene como estadístico suficiente para θ el vector $\mathbf{T} = \mathbf{t}(\mathbf{X}) = (t_1(\mathbf{X}), \dots, t_k(\mathbf{X}))$.

Teorema. Sea $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ una muestra aleatoria de una distribución que pertenece a una familia exponencial a k parámetros, cuya función de densidad viene dada por (1).

Luego,

- la distribución conjunta de $\mathbf{X}_1, \dots, \mathbf{X}_n$ también pertenece a una familia exponencial a k parámetros y
- el estadístico suficiente para θ es el vector

$$\mathbf{T}^* = (T_1^*, \dots, T_k^*), \text{ donde } T_j^* = \sum_{i=1}^n t_j(\mathbf{X}_i), \quad 1 \leq j \leq k$$

Resultados

Teorema. Sea \mathbf{X} un vector cuya distribución pertenece a una familia exponencial a k parámetros cuya función de densidad satisface (1).

Luego, la función de densidad de los estadísticos suficientes $\mathbf{T} = (t_1(\mathbf{X}), \dots, t_k(\mathbf{X}))$ es de la forma

$$p_{\mathbf{T}}(t_1, t_2, \dots, t_k, \boldsymbol{\theta}) = A(\boldsymbol{\theta}) e^{\sum_{j=1}^k \eta_j(\boldsymbol{\theta}) t_j} h^*(t_1, \dots, t_k)$$

Por lo tanto, la familia de distribuciones de \mathbf{T} también forma una familia exponencial a k parámetros.

Lema

Sea $\mathbf{X} = (X_1, \dots, X_q)$ un vector aleatorio cuya distribución pertenece a una familia exponencial a un parámetro continua

$$p(\mathbf{x}, \theta) = A(\theta)e^{\eta(\theta)t(\mathbf{x})} h(\mathbf{x})$$

- $\theta \in \Theta \subset \mathbb{R}$, Θ abierto
- $\eta(\theta)$ infinitamente derivable.

Luego, si $m(\mathbf{x})$ es un estadístico tal que

$$\int |m(\mathbf{x})| p(\mathbf{x}, \theta) d\mathbf{x} < \infty \quad \forall \theta \in \Theta$$

entonces, la expresión

$$\int \dots \int m(\mathbf{x}) e^{\eta(\theta)t(\mathbf{x})} h(\mathbf{x}) dx_1 \dots dx_q$$

es infinitamente derivable y se puede derivar dentro del signo integral.

En el caso discreto, se reemplazan las integrales por sumatorias.

Teorema

Sea $\mathbf{X} = (X_1, \dots, X_q)$ un vector aleatorio cuya distribución pertenece a una familia exponencial a un parámetro con densidad dada por

$$p(\mathbf{x}, \theta) = A(\theta)e^{\eta(\theta)t(\mathbf{x})} h(\mathbf{x}) \quad \theta \in \Theta \subset \mathbb{R}$$

con Θ abierto y $\eta(\theta)$ infinitamente derivable. Luego se tiene:

(i) $A(\theta)$ es infinitamente derivable.

(ii)

$$\mathbb{E}_\theta(t(\mathbf{X})) = - \frac{A'(\theta)}{A(\theta)\eta'(\theta)}$$

(iii)

$$\text{VAR}_\theta(t(\mathbf{X})) = \frac{\frac{\partial \mathbb{E}_\theta(t(\mathbf{X}))}{\partial \theta}}{\eta'(\theta)}$$

Teorema

Sea una familia exponencial a k parámetros, discreta o continua con función de densidad dada por

$$p(\mathbf{x}, \boldsymbol{\theta}) = A(\boldsymbol{\theta}) e^{\sum_{j=1}^k \eta_j(\boldsymbol{\theta}) t_j(\mathbf{x})} h(\mathbf{x}).$$

Sea $\Lambda = \{\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_k)^T \in \mathbb{R}^k : \lambda_j = \eta_j(\boldsymbol{\theta}); \boldsymbol{\theta} \in \Theta\}$.

- a) Si Λ contiene $k + 1$ puntos $\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(k+1)}$ tales que $\{\boldsymbol{\lambda}^{(j)} - \boldsymbol{\lambda}^{(1)}, 2 \leq j \leq k + 1\}$ son linealmente independientes, entonces el estadístico suficiente $\mathbf{T} = (t_1(\mathbf{X}), \dots, t_k(\mathbf{X}))$ es minimal suficiente.
- b) Si Λ contiene una esfera en \mathbb{R}^k , entonces el estadístico suficiente $\mathbf{T} = (t_1(\mathbf{X}), \dots, t_k(\mathbf{X}))$ es completo.

Caracterización

Lema. Sea δ_0 un estimador insesgado de $q(\boldsymbol{\theta})$.

Dado cualquier otro estimador δ insesgado de $q(\boldsymbol{\theta})$, se cumple que

$$\delta = \delta_0 - U$$

con

$$\mathbb{E}_{\boldsymbol{\theta}}(U) = 0 \quad \forall \boldsymbol{\theta} \in \Theta$$

Caracterización

Teorema. Supongamos que \mathbf{X} es un vector aleatorio de cuya distribución pertenece a la familia $\mathcal{P} = \{P_{\theta} \mid \theta \in \Theta\}$.

Sean

$$\Delta = \{\delta(\mathbf{X}) : \mathbb{E}_{\theta} \delta^2(\mathbf{X}) < \infty\}$$

$$\mathcal{U} = \{\{\delta(\mathbf{X}) \in \Delta : \mathbb{E}_{\theta} \delta(\mathbf{X})\} = 0 \forall \theta \in \Theta\}.$$

Una condición necesaria y suficiente para que $\delta \in \Delta$, insesgado, sea IMVU para $q(\theta)$ es que

$$\text{Cov}(\delta, U) = \mathbb{E}_{\theta}(\delta U) = 0 \quad \forall \theta \in \Theta \quad \forall U \in \mathcal{U}$$

Teorema

Supongamos que

- \mathbf{X} es un vector aleatorio de cuya distribución pertenece a la familia $\mathcal{P} = \{P_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta\}$.
- $q(\boldsymbol{\theta})$ es U -estimable, o sea, existe un estimador insesgado $\delta(\mathbf{X})$ de $q(\boldsymbol{\theta})$.
- $L(\boldsymbol{\theta}, d)$ es tal que
 - $L(\boldsymbol{\theta}, d) \leq M$, para todo $\boldsymbol{\theta}$ y d .
 - $L(\boldsymbol{\theta}, q(\boldsymbol{\theta})) = 0$, para todo $\boldsymbol{\theta}$.

Sea $\boldsymbol{\theta}_0$ un valor arbitrario. Luego, existe una sucesión δ_n de estimadores insesgados tales que

$$\lim_{n \rightarrow \infty} R(\boldsymbol{\theta}_0, \delta_n) = 0$$

Hipotesis

Supongamos que $\mathbf{X} \sim f(\mathbf{x}, \theta)$, con $\theta \in \Theta$; Θ un conjunto abierto de \mathbb{R} .

- (A) El conjunto $\mathcal{S} = \{\mathbf{x} : f(\mathbf{x}, \theta) > 0\}$ es independiente de θ .
- (B) Para todo \mathbf{x} , $f(\mathbf{x}, \theta)$ es derivable respecto de θ .
- (C) Si $h(\mathbf{X})$ es un estadístico tal que $\mathbb{E}_\theta[|h(\mathbf{X})|] < \infty$ para todo $\theta \in \Theta$ entonces se tiene

$$\frac{\partial}{\partial \theta} \int h(\mathbf{x})p(\mathbf{x}, \theta)d\mathbf{x} = \int h(\mathbf{x})\frac{\partial f(\mathbf{x}, \theta)}{\partial \theta}d\mathbf{x}$$

(D)

$$0 < I(\theta) = \mathbb{E}_\theta \left[\left(\frac{\partial \log f(\mathbf{X}, \theta)}{\partial \theta} \right)^2 \right] < \infty$$

$I(\theta)$ se denomina *número de información de Fisher*.

Lema

Supongamos que se cumplan las condiciones A, B, C y D. Sea

$$\psi(\mathbf{x}, \theta) = \frac{\partial \log(f(\mathbf{x}, \theta))}{\partial \theta}$$

Entonces,

- (i) $\mathbb{E}_\theta \psi(\mathbf{X}, \theta) = 0$ y $\text{VAR}_\theta \psi(\mathbf{X}, \theta) = I(\theta)$.
- (ii) Si además existe la derivada segunda de $f(\mathbf{x}, \theta)$ respecto de θ y si para todo estadístico $h(\mathbf{X})$ tal que, $\mathbb{E}_\theta[|h(\mathbf{X})|] < \infty$ para todo $\theta \in \Theta$, se cumple que

$$\frac{\partial^2}{\partial \theta^2} \int h(\mathbf{x}) f(\mathbf{x}, \theta) d\mathbf{x} = \int h(\mathbf{x}) \frac{\partial^2 f(\mathbf{x}, \theta)}{\partial \theta^2} d\mathbf{x} \quad (2)$$

entonces

$$I(\theta) = -\mathbb{E}_\theta \frac{\partial^2 \log f(\mathbf{X}, \theta)}{\partial \theta^2} = -\mathbb{E}_\theta \frac{\partial \psi(\mathbf{X}, \theta)}{\partial \theta}$$

Teorema

Sea $\mathbf{X} = (X_1, \dots, X_q)$ un vector aleatorio cuya distribución pertenece a una familia exponencial a un parámetro con densidad dada por

$$f(\mathbf{x}, \theta) = A(\theta)e^{\eta(\theta)t(\mathbf{x})}h(\mathbf{x}) \quad \theta \in \Theta$$

con Θ un abierto en \mathbb{R} .

Supongamos que

- $\theta = \mathbb{E}_\theta(t(\mathbf{X}))$
- $\eta(\theta)$ es derivable.

Luego, si $T = t(\mathbf{X})$

$$I(\theta) = \frac{1}{\text{Var}_\theta(T)} = \eta'(\theta)$$

Teorema

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra aleatoria de una distribución con densidad $f(\mathbf{x}, \theta)$ con $\theta \in \Theta \subset \mathbb{R}$.

Luego, si se denomina

- $I_n(\theta)$ al número de información de $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ y
- $I_1(\theta)$ al número de información de \mathbf{X}_1 ,

entonces se tiene

$$I_n(\theta) = nI_1(\theta)$$

Teorema de Rao–Cramer

Bajo las condiciones A, B, C y D, si $\delta(\mathbf{X})$ es un estimador insesgado de $q(\theta)$ tal que $\mathbb{E}_\theta \delta^2(\mathbf{X}) < \infty$ se tiene

(a)

$$\text{VAR}_\theta(\delta(\mathbf{X})) \geq \frac{|q'(\theta)|^2}{I(\theta)}$$

(b) La desigualdad en (a) es una igualdad si y sólo si $\delta(\mathbf{X})$ es un estadístico suficiente de una familia exponencial, es decir, si y sólo si

$$f(\mathbf{x}, \theta) = A(\theta)e^{\eta(\theta)\delta(\mathbf{x})}h(\mathbf{x})$$

Sea $\mathcal{P} = \{P_{\theta}, \theta \in \Theta\}$ una familia de distribuciones.

Supongamos que para cada n se tiene un estimador $\delta_n(X_1, \dots, X_n)$ de $q(\theta)$ basado en una muestra aleatoria de tamaño n .

Sea $\mathcal{P} = \{P_{\theta}, \theta \in \Theta\}$ una familia de distribuciones.

Supongamos que para cada n se tiene un estimador $\delta_n(X_1, \dots, X_n)$ de $q(\theta)$ basado en una muestra aleatoria de tamaño n .

Definición. $\delta_n(X_1, \dots, X_n)$ es una *sucesión fuertemente consistente de estimadores* de $q(\theta)$ si

$$\lim_{n \rightarrow \infty} \delta_n(X_1, \dots, X_n) = q(\theta) \quad c.t.p.$$

o sea si $\mathbb{P}_{\theta}(\delta_n(X_1, \dots, X_n) \rightarrow q(\theta)) = 1$ para todo $\theta \in \Theta$.

Sea $\mathcal{P} = \{P_{\theta}, \theta \in \Theta\}$ una familia de distribuciones.

Supongamos que para cada n se tiene un estimador $\delta_n(X_1, \dots, X_n)$ de $q(\theta)$ basado en una muestra aleatoria de tamaño n .

Definición. $\delta_n(X_1, \dots, X_n)$ es una *sucesión fuertemente consistente de estimadores* de $q(\theta)$ si

$$\lim_{n \rightarrow \infty} \delta_n(X_1, \dots, X_n) = q(\theta) \quad \text{c.t.p.}$$

o sea si $\mathbb{P}_{\theta}(\delta_n(X_1, \dots, X_n) \rightarrow q(\theta)) = 1$ para todo $\theta \in \Theta$.

Definición. $\delta_n(X_1, \dots, X_n)$ es una *sucesión débilmente consistente de estimadores* de $q(\theta)$ si

$$\delta_n(X_1, \dots, X_n) \xrightarrow{P} q(\theta)$$

o sea, si para todo $\varepsilon > 0$ y $\theta \in \Theta$

$$\lim_{n \rightarrow \infty} P_{\theta}(|\delta_n(X_1, \dots, X_n) - q(\theta)| > \varepsilon) = 0.$$

Propiedades

Teorema. Sea, para todo n , $\delta_n = \delta_n(X_1, \dots, X_n)$ un estimador de $q(\theta)$ basado en una muestra aleatoria de tamaño n . Si

- $\text{VAR}_{\theta}(\delta_n) \rightarrow 0$
- $\mathbb{E}_{\theta}(\delta_n) \rightarrow q(\theta)$,

entonces $\delta_n(X_1, \dots, X_n)$ es débilmente consistente para $q(\theta)$.

Propiedades

Teorema. Sea, para todo n , $\delta_n = \delta_n(X_1, \dots, X_n)$ un estimador de $q(\theta)$ basado en una muestra aleatoria de tamaño n . Si

- $\text{VAR}_{\theta}(\delta_n) \rightarrow 0$
- $\mathbb{E}_{\theta}(\delta_n) \rightarrow q(\theta)$,

entonces $\delta_n(X_1, \dots, X_n)$ es débilmente consistente para $q(\theta)$.

Teorema. Sea $\delta_n(X_1, \dots, X_n)$ una sucesión de estimadores IMVU para $q(\theta)$, donde X_1, \dots, X_n es una muestra aleatoria de una distribución perteneciente a la familia $\mathcal{P} = \{P_{\theta}, \theta \in \Theta\}$. Luego,

- $\text{VAR}_{\theta}(\delta_n(X_1, \dots, X_n)) \rightarrow 0$ para todo $\theta \in \Theta$

de donde $\delta_n(X_1, \dots, X_n)$ es débilmente consistente para $q(\theta)$.

Consistencia de los estimadores de los momentos

- Sea X_1, \dots, X_n una muestra aleatoria de una distribución perteneciente a la familia $\mathcal{P} = \{P_\theta, \theta \in \Theta \subset \mathbb{R}\}$,
- $h(x)$ una función continua con valores en \mathbb{R}
- Supongamos que $g(\theta) = \mathbb{E}_\theta(h(X_1))$ es, como función de θ , continua y estrictamente monótona.

Sea el estimador de momentos $\hat{\theta}_n$ definido como la solución de

$$\frac{1}{n} \sum_{i=1}^n h(X_i) = \mathbb{E}_\theta(h(X_1)). \quad (3)$$

Luego con probabilidad 1 existe n_0 tal que para todo $n \geq n_0$ la ecuación (3) tiene una solución, $\hat{\theta}_n$, y $\hat{\theta}_n$ es fuertemente consistente para θ .

Consistencia de los estimadores de los máxima verosimilitud

Sea X_1, \dots, X_n i.i.d. $X_i \sim f(x, \theta)$ con $\theta \in \Theta$, donde Θ es un intervalo abierto de \mathbb{R} .

Supongamos que:

- $f(x, \theta)$ es derivable respecto de θ
- el conjunto $\mathcal{S} = \{x : f(x, \theta) \neq 0\}$ es independiente de θ .
- $\theta_1 \neq \theta_2$ implica que $f(x, \theta_1) \neq f(x, \theta_2)$.

Sea $\hat{\theta}_n$ el estimador de máxima verosimilitud de θ , que satisface

$$\sum_{i=1}^n \frac{\partial \log f(x_i, \hat{\theta}_n)}{\partial \theta} = 0 \quad (4)$$

Supongamos además que la ecuación (4) tiene a lo sumo una solución. Entonces,

$$\hat{\theta}_n \xrightarrow{a.s.} \theta$$

o sea, $\hat{\theta}_n$ es una sucesión de estimadores fuertemente consistente.

Estimadores asintóticamente normales

Sea $\mathcal{P} = \{P_{\theta}, \theta \in \Theta\}$ una familia de distribuciones.

Supongamos que para cada n se tiene un estimador $\delta_n(X_1, \dots, X_n)$ de $q(\theta)$ basado en una muestra aleatoria de tamaño n .

Definición 1. Diremos que $\delta_n(X_1, \dots, X_n)$ es una sucesión de *estimadores asintóticamente normal* si

$$\sqrt{n}(\delta_n(X_1, \dots, X_n) - q(\theta)) \xrightarrow{D} N(0, \sigma^2(\theta)) \quad (5)$$

Estimadores asintóticamente normales y eficientes

Definición 2. Supongamos que

- $X_i \sim f(x, \theta)$, $\theta \in \Theta \subset \mathbb{R}$
- El conjunto $\mathcal{S} = \{x : f(x, \theta) > 0\}$ es independiente de θ .
- Para todo x , $f(x, \theta)$ es derivable respecto de θ .
-

$$0 < I_1(\theta) = \mathbb{E}_\theta \left(\frac{\partial \log f(X, \theta)}{\partial \theta} \right)^2 < \infty$$

Se dice que $\delta_n(X_1, \dots, X_n)$ es una sucesión de *estimadores asintóticamente normal y eficiente* (A.N.E.) o simplemente una sucesión de *estimadores asintóticamente eficientes* si

- $\delta_n(X_1, \dots, X_n)$ es una sucesión de *estimadores asintóticamente normal*, o sea, cumple (5) y
- su varianza asintótica es igual a

$$\sigma^2(\theta) = \frac{[q'(\theta)]^2}{I_1(\theta)}$$

Teorema

Supongamos que $\mathbf{X}_1, \dots, \mathbf{X}_n$ son i.i.d. $\mathbf{X}_i \sim f(\mathbf{x}, \theta)$, con

- $\theta \in \Theta \subset \mathbb{R}$,
- Θ abierto

y que se cumple

- (A) El conjunto $\mathcal{S} = \{\mathbf{x} : f(\mathbf{x}, \theta) > 0\}$ es independiente de θ .
 (B) Para todo \mathbf{x} , $f(\mathbf{x}, \theta)$ es tres veces derivable respecto de θ y

$$\frac{\partial^3 f(\mathbf{x}, \theta)}{\partial \theta^3}$$

es continua en θ .

- (C) Si $h(\mathbf{X})$ es un estadístico tal que $\mathbb{E}_\theta |h(\mathbf{X})| < \infty$ para todo $\theta \in \Theta$ entonces $g(\theta) = \int h(\mathbf{x})f(\mathbf{x}, \theta)d\mathbf{x}$ es dos veces diferenciable y para $s = 1, 2$

$$\frac{\partial^s}{\partial \theta^s} \int h(\mathbf{x})f(\mathbf{x}, \theta)d\mathbf{x} = \int h(\mathbf{x}) \frac{\partial^s f(\mathbf{x}, \theta)}{\partial \theta^s} d\mathbf{x}$$

Teorema

$$(D) \quad 0 < I_1(\theta) = \mathbb{E}_\theta \left(\frac{\partial \log f(\mathbf{X}, \theta)}{\partial \theta} \right)^2 < \infty$$

(E) Para cada θ_0 fijo, existen $c > 0$ y una función $M(\mathbf{x}) > 0$ tales que

$$\left| \frac{\partial^3 \log f(\mathbf{x}, \theta)}{\partial \theta^3} \right| = \left| \frac{\partial^2 \psi(\mathbf{x}, \theta)}{\partial \theta^2} \right| \leq M(\mathbf{x})$$

$$\mathbb{E}_{\theta_0} M(\mathbf{X}_1) < \infty$$

para todo $\mathbf{x} \in \mathcal{S}$ y para todo $\theta \in \{\theta \in \Theta : |\theta_0 - \theta| < c\}$

Teorema

Sean

- $\hat{\theta}_n$ un estimador de máxima verosimilitud de θ consistente y
- $q(\theta)$ derivable con $q'(\theta) \neq 0$ para todo θ .

Entonces, $q(\hat{\theta}_n)$ es A.N.E. para estimar $q(\theta)$, o sea,

$$\sqrt{n} \left(q(\hat{\theta}_n) - q(\theta) \right) \xrightarrow{D} N \left(0, \frac{[q'(\theta)]^2}{I_1(\theta)} \right)$$

Teorema (Le Cam, 1953)

Supongamos que $\mathbf{X}_1, \dots, \mathbf{X}_n$ son i.i.d. $\mathbf{X}_i \sim f(\mathbf{x}, \theta)$, con

- $\theta \in \Theta \subset \mathbb{R}$,
- Θ abierto

y que se cumple

- (A) El conjunto $\mathcal{S} = \{\mathbf{x} : f(\mathbf{x}, \theta) > 0\}$ es independiente de θ .
 (B) Para todo \mathbf{x} , $f(\mathbf{x}, \theta)$ es dos veces derivable respecto de θ y

$$\frac{\partial^2 f(\mathbf{x}, \theta)}{\partial \theta^2}$$

es continua en θ .

- (C) Si $h(\mathbf{X})$ es un estadístico tal que $\mathbb{E}_\theta |h(\mathbf{X})| < \infty$ para todo $\theta \in \Theta$ entonces $g(\theta) = \int h(\mathbf{x})f(\mathbf{x}, \theta)d\mathbf{x}$ es dos veces diferenciable y para $s = 1, 2$

$$\frac{\partial^s}{\partial \theta^s} \int h(\mathbf{x})f(\mathbf{x}, \theta)d\mathbf{x} = \int h(\mathbf{x}) \frac{\partial^s f(\mathbf{x}, \theta)}{\partial \theta^s} d\mathbf{x}$$

Teorema (Le Cam, 1953)

$$(D) \quad 0 < I_1(\theta) = \mathbb{E}_\theta \left(\frac{\partial \log f(\mathbf{X}, \theta)}{\partial \theta} \right)^2 < \infty$$

(E) Para cada θ_0 fijo, existen $c > 0$ y una función $M(\mathbf{x}) > 0$ tales que

$$\left| \frac{\partial^2 \log f(\mathbf{x}, \theta)}{\partial \theta^2} \right| = \left| \frac{\partial \psi(\mathbf{x}, \theta)}{\partial \theta} \right| \leq M(\mathbf{x})$$
$$\mathbb{E}_{\theta_0} M(\mathbf{X}_1) < \infty$$

para todo $\mathbf{x} \in \mathcal{S}$ y para todo $\theta \in \{\theta \in \Theta : |\theta_0 - \theta| < c\}$

Teorema (Le Cam)

Sea $\delta_n(X_1, \dots, X_n)$ una sucesión de estimadores asintóticamente normales de $q(\theta)$, con q derivable y $q'(\theta) \neq 0$ para todo θ . Es decir,

$$\sqrt{n}(\delta_n(X_1, \dots, X_n) - q(\theta)) \xrightarrow{D} N(0, \sigma^2(\theta))$$

entonces

$$\sigma^2(\theta) \geq \frac{[q'(\theta)]^2}{I_1(\theta)}$$

excepto en un conjunto $\Theta_0 \subset \Theta$ tal que Θ_0 tiene medida 0.