

Tema 4

ESTIMACIÓN PUNTUAL. INSESGADEZ Y MÍNIMA VARIANZA.

4.1. Planteamiento del problema

En el tema anterior se trataron aspectos relacionados con la reducción de los datos de una muestra en términos de un estadístico. En este tema se aborda el problema de estimación del parámetro de un modelo estadístico paramétrico (o de una función de dicho parámetro) mediante un estadístico conveniente.

En general se tiene una v.a. X con distribución en una familia de distribuciones paramétricas (es decir conocidas salvo por un parámetro), $X \rightsquigarrow F \in \{F_\theta, \theta \in \Theta\}$. El objetivo es determinar la distribución de la v.a. que se estudia, que en el caso paramétrico se reduce a conocer el parámetro, es decir, inferir el verdadero valor de θ , θ_0 .

Para ello lo que se hace es coger una muestra, X_1, \dots, X_n , m.a.s. de X , y en base a la información que proporciona la muestra se aproxima el valor de θ . Por lo tanto, el problema es escoger estadísticos, $T(X_1, \dots, X_n)$, que para valores concretos de la muestra proporcionen buenas aproximaciones del parámetro θ .

Con dicho fin se van a escoger unos estadísticos particulares, que se denominan *estimadores*, de forma que cuando se sustituya la muestra aleatoria, X_1, \dots, X_n , por sus observaciones, $T(x_1, \dots, x_n)$ proporcionen una buena aproximación del parámetro desconocido.

En ocasiones en vez de estimar el parámetro θ , interesará estimar una *función paramétrica*, es decir, una transformación del parámetro, $g : \Theta \rightarrow \mathbb{R}$. En ese caso se buscará $g(\theta_0)$, el verdadero valor de $g(\theta)$, en vez de θ_0 , el verdadero valor de θ .

Ejemplo: Sea $X \rightsquigarrow B(k_0, p)$ y $p \in (0, 1)$. A partir de una m.a.s. de X indicar algún estadístico que se pueda usar para inferir p y alguna función paramétrica que pueda ser

de interés.

A la aproximación que se obtiene del parámetro a través de un estimador se la denomina estimación puntual porque lo que se obtiene al aplicarla es un valor concreto para el parámetro desconocido.

El parámetro desconocido puede ser unidimensional o multidimensional. En general se van a estudiar casos unidimensionales y, si es fácil, se generalizará al caso multidimensional.

4.1.1. Estimador

Sea X una v.a. con función de distribución en una familia de distribuciones paramétricas, $F \in \{F_\theta : \theta \in \Theta\}$ y (X_1, \dots, X_n) una m.a.s. de X .

Definición: Un estimador de θ es un estadístico, $T(X_1, \dots, X_n)$, que toma valores en Θ .

$$T : \mathcal{X}^n \rightarrow \Theta.$$

Por tanto la diferencia entre un estimador y un estadístico es que el estimador es un estadístico que en lugar de estar definido sobre \mathbb{R}^k , se le exige que tome valores en el espacio paramétrico Θ , es decir, donde toma los valores el parámetro desconocido.

Si el estimador está definido con espacio de llegada $g(\Theta)$, es decir, $T : \mathcal{X}^n \rightarrow g(\Theta)$, T es un estimador de la función paramétrica $g(\theta)$.

Para valores concretos de la muestra, x_1, \dots, x_n , $T(x_1, \dots, x_n)$ es una estimación puntual de θ o de $g(\theta)$, según el caso.

Ejemplos:

1. Sea X_1, \dots, X_n una m.a.s. de $X \rightsquigarrow \{P(\lambda), \lambda > 0\}$. \bar{X} es un estimador de λ . Es más, cualquier función medible de la muestra que sea independiente del parámetro λ y tome valores positivos es un estimador del parámetro.
2. Sea X_1, \dots, X_n una m.a.s. de $X \rightsquigarrow \{B(1, p), p \in (0, 1)\}$. \bar{X} es un estimador de p . Es más, cualquier función medible de la muestra que sea independiente del parámetro p y tome valores en el intervalo $(0, 1)$ es un estimador del parámetro.

La no unicidad del estimador de un parámetro plantea el problema de encontrar el mejor estimador. Para ello hay que establecer criterios de selección entre los estimadores para encontrar el mejor en algún sentido. Una opción es seleccionar el estimador en base a una función denominada de pérdida.

4.1.2. Función de pérdida y función de riesgo

Definición: A cualquier función $L : \Theta \times \Theta \rightarrow \mathbb{R}$, que verifique las siguientes propiedades:

- (i) $L(\theta, T) \geq 0, \forall \theta \in \Theta, T \in \Theta$.
- (ii) $L(\theta, T) = 0$, si $T = \theta$.
- (iii) $L(\theta, T) \leq L(\theta, T')$ si la distancia de T a θ es menor que la distancia de T' a θ .

se la denomina *función de pérdida*.

$L(\theta, t)$ sería la pérdida que conlleva estimar el parámetro por el valor t si su verdadero valor es θ usando la función de pérdida L .

Ejemplos:

1. $L(\theta, T) = |\theta - T|$ (error absoluto de estimación).
2. $L(\theta, T) = (\theta - T)^2$ (error cuadrático de estimación).
3. $L(\theta, T) = \left| \frac{\theta - T}{\theta} \right|$ (error relativo de estimación).

Dado un estimador $T(X_1, \dots, X_n)$ de θ , la función $L(\theta, T(X_1, \dots, X_n))$, para cada $\theta \in \Theta$, es una variable aleatoria, siempre que L sea Borel-medible.

Definición: Se define la pérdida media o *función riesgo* de un estimador como la función (del parámetro θ) que asigna a cada valor del parámetro la pérdida media asociada al estimador bajo la función de pérdida L .

$$R_T^L(\theta) = E_\theta[L(\theta, T)].$$

En particular, la función riesgo de una función paramétrica $g(\theta)$ se definiría como:

$$R_{g,T}^L(\theta) = E_\theta[L(g(\theta), T)].$$

Definición: Se dice que un estimador $T(X_1, \dots, X_n)$ es *óptimo bajo una función de pérdida* $L(\theta, T)$ si dicho estimador minimiza uniformemente la función de riesgo $R_T(\theta)$; esto es, un estimador, T , tal que, para cualquier otro, T' se tiene

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

En general, el estimador óptimo no tiene porqué existir. Al no tener asegurada la existencia del estimador óptimo basado en la función de riesgo, el problema de estimación se puede reconsiderar mediante una de las dos siguientes vías:

1. Restringir la clase de estimadores imponiendo propiedades deseables de los mismos.
2. Introducir una nueva medida de la función de riesgo que permita ordenar totalmente la clase de todos los estimadores.

Estimación de menor error cuadrático medio (ECM)

Un criterio de comparación usual en múltiples ámbitos y aplicaciones de la estadística es el llamado criterio de menor *error cuadrático medio*. Se considera como función de pérdida:

$$L(\theta, T) = (\theta - T)^2$$

y como función de riesgo:

$$R_T(\theta) = E_\theta [(\theta - T)^2] = ECM_T(\theta).$$

Propiedades:

- El criterio del ECM tiene ventajas desde el punto de vista del manejo analítico, frente a otras funciones de riesgo.
- El ECM se interpreta como el grado de dispersión del estimador en torno al verdadero valor del parámetro, θ .
- El ECM puede descomponer en términos de la varianza y una función denominada sesgo:

$$ECM_T(\theta) = Var_\theta(T) + B_T^2(\theta)$$

donde $B_T(\theta)$ es la función denominada sesgo que se define como:

$$B_T(\theta) = E_\theta[T] - \theta.$$

- Si el estimador considerado verifica la propiedad de insesgadez, es decir, $E_\theta[T] = \theta$ o $B_T(\theta) = 0$, se verifica que el ECM coincide con la varianza del estimador.

$$ECM_T(\theta) = E_\theta [(\theta - T)^2] = Var_\theta(T).$$

4.2. Estimación insesgada de mínima varianza

Como se ha visto en la sección anterior, existe una relación sencilla e intuitiva que liga al ECM, la varianza y el sesgo de un estimador. En el problema de la búsqueda de un estimador óptimo, en algún sentido, se explota dicha relación dentro de la clase de estimadores que verifican ciertas propiedades, como la propiedad de insesgadez.

4.2.1. Estimador insesgado

Sea, como siempre a lo largo de este tema, X una v.a. con función de distribución en una familia de distribuciones paramétricas, $F \in \{F_\theta : \theta \in \Theta\}$ y (X_1, \dots, X_n) una m.a.s. de X .

Definición: Un estimador $T(X_1, \dots, X_n)$ de θ es *insesgado* o centrado en el parámetro θ si su sesgo asociado es idénticamente nulo o, equivalentemente, si:

$$E_\theta[T(X_1, \dots, X_n)] = \theta, \quad \forall \theta \in \Theta.$$

Si el estimador $T^*(X_1, \dots, X_n)$ es de una función paramétrica de θ , $g(\theta)$, se dice que es insesgado en $g(\theta)$ si:

$$E_\theta[T^*(X_1, \dots, X_n)] = g(\theta), \quad \forall \theta \in \Theta.$$

Ejemplo: Sea X_1, \dots, X_n una m.a.s. de alguna población. Probar que si existe la media de la población, $E_\theta X$, la media muestral es un estimador insesgado de la media poblacional, y si existe la varianza de la población, $Var_\theta X$, la cuasivarianza muestral es un estimador insesgado de ella.

Notas:

- Si $\theta = (\theta_1, \dots, \theta_k)$, un estimador insesgado de dicho vector de parámetros es un vector donde cada elemento del mismo es un estimador insesgado para cada parámetro θ_i componente de θ , es decir: Un estimador $T = (T_1, \dots, T_k)$ es insesgado en θ si se verifica

$$E_\theta[T_i] = \theta_i, \quad \forall i = 1, \dots, k.$$

- Para un estimador insesgado se verifica que

$$ECM_T(\theta) = Var_\theta(T(X_1, \dots, X_n)), \quad \forall \theta \in \Theta$$

- Si T es un estimador insesgado de $\theta \nRightarrow h(T)$ es estimador insesgado de $h(\theta)$, siendo h cualquier función.

Ejemplo: Sea X_1, \dots, X_n una m.a.s. de alguna población. Probar que, en general, la cuasidesviación típica no es un estimador insesgado para $\sigma_\theta = \sqrt{Var_\theta X}$.

Si h es una función lineal dicha implicación si se cumple, es decir, la insesgadez no se mantiene bajo transformaciones, en general, pero si se mantiene si la transformación es lineal.

- No tiene porque existir algún estimador insesgado de un parámetro.

Ejemplos:

1. Sea X_1, \dots, X_n , con $n \geq 2$, una m.a.s. de una distribución binomial $B(1, p)$. Probar que $\frac{T^2 - T}{n^2 - n}$ con $T = \sum_{i=1}^n X_i$ es un estimador insesgado para la función paramétrica $g(p) = p^2$. Probar, además, que para $n = 1$ no existe un estimador insesgado de p^2 .
 2. Sea X una v.a. con distribución $\mathcal{P}(\lambda)$, ¿existe algún estimador insesgado para la función paramétrica $1/\lambda$ basado en una muestra de tamaño 1?
- Un estimador insesgado no tiene porque ser único.

Ejemplo: Sea X una v.a. con distribución $\mathcal{P}(\lambda)$ y X_1, \dots, X_n una m.a.s. de X . Probar que \bar{X} y S^2 son estimadores insesgados de λ .

Es más, si T_1 y T_2 son estimadores insesgados de θ , entonces $\alpha T_1 + (1 - \alpha)T_2$ con $0 \leq \alpha \leq 1$ es un estimador insesgado de θ .

Se plantea, por tanto, el problema de seleccionar un estimador óptimo dentro de la clase de estimadores insesgados. Nuestro criterio de búsqueda de los mejores estimadores será seleccionar los que tengan mínima varianza, ya que son los que están menos dispersos respecto a la media.

4.2.2. Estimador insesgado uniformemente de mínima varianza (UMVUE)

Definición: Sea X una v.a. con función de distribución en una familia de distribuciones paramétricas, $F \in \{F_\theta : \theta \in \Theta\}$ y (X_1, \dots, X_n) una m.a.s. de X . Un estimador de $g(\theta)$, $T(X_1, \dots, X_n)$, insesgado y con momento de segundo orden finito, se dice que es un *estimador insesgado uniformemente de mínima varianza* (“UMVUE”) para $g(\theta)$ si para cualquier otro estimador de $g(\theta)$, $T'(X_1, \dots, X_n)$, se tiene:

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

Teorema (Unicidad del UMVUE): El UMVUE, si existe, es único, es decir, si hay dos UMVUEs son iguales con probabilidad 1.

Teorema (Linealidad del UMVUE): Si T_1 es UMVUE para una cierta función de θ , $g_1(\theta)$ y T_2 es UMVUE para otra cierta función de θ , $g_2(\theta)$ con $\theta \in \Theta$. Entonces:

1. λT_1 es UMVUE para $\lambda g_1(\theta)$ y λT_2 es UMVUE para $\lambda g_2(\theta)$,
2. $T_1 + T_2$ es UMVUE para $g_1(\theta) + g_2(\theta)$,

siendo λ cualquier valor real.

Una vez definido el UMVUE y estudiada algunas de sus propiedades, el siguiente paso es tener un método que permita su obtención de forma lo más sencilla posible. Para ello primero se estudian los Teoremas de Rao-Blackwell y Lehmann-Scheffé.

Teorema (Rao-Blackwell): Sea X una v.a. con función de distribución en una familia de distribuciones paramétricas, $F \in \{F_\theta : \theta \in \Theta\}$, y (X_1, \dots, X_n) una m.a.s. de X . Si $T(X_1, \dots, X_n)$ un estadístico suficiente para la familia, $\{F_\theta : \theta \in \Theta\}$ y $S(X_1, \dots, X_n)$ es un estimador insesgado de $g(\theta)$ con momento de segundo orden finito, entonces:

1. $E[S(X_1, \dots, X_n)/T(X_1, \dots, X_n)]$ es estimador insesgado de $g(\theta)$ y tiene momento de segundo orden finito.
2. $Var_\theta(E[S(X_1, \dots, X_n)/T(X_1, \dots, X_n)]) \leq Var_\theta(S(X_1, \dots, X_n)), \forall \theta \in \Theta$.

Teorema (Lehmann-Scheffé): Sea $T(X_1, \dots, X_n)$ es un estadístico suficiente y completo para la familia de distribuciones consideradas, $\{F_\theta : \theta \in \Theta\}$. Si $g(\theta)$ admite un estimador insesgado de segundo orden, $S(X_1, \dots, X_n)$, entonces existe el UMVUE de $g(\theta)$, que viene dado por

$$E[S(X_1, \dots, X_n)/T(X_1, \dots, X_n)].$$

Métodos para el cálculo del UMVUE: Sea $T(X_1, \dots, X_n)$ un estadístico suficiente y completo para $\{F_\theta : \theta \in \Theta\}$. El UMVUE para $g(\theta)$, si existe, se puede determinar mediante los dos siguientes procedimientos:

- Buscar cualquier estimador insesgado de $g(\theta)$ con momento de segundo orden finito, $S(X_1, \dots, X_n)$. Entonces $E[S(X_1, \dots, X_n)/T(X_1, \dots, X_n)]$ es el UMVUE.
- Buscar una función $h(T)$, siendo T un estadístico suficiente y completo, tal que $E_\theta[h(T)] = g(\theta) \forall \theta \in \Theta$, es decir, que sea insesgada en $g(\theta)$, que sea un estimador y que tenga momento de segundo orden finito. Entonces $E[h(T(X_1, \dots, X_n))/T(X_1, \dots, X_n)] = h(T(X_1, \dots, X_n))$ es el UMVUE.

Si no existiera un estadístico suficiente y completo, esto no implica que no exista UMVUE, sólo que habrá que calcularlo de otra forma.

Ejemplos

1. Sea X_1, \dots, X_n una m.a.s. con distribución $B(1, p)$, $p \in (0, 1)$. Encontrar el UMVUE para p .
2. Sea X_1, \dots, X_n una m.a.s. con distribución $\mathcal{U}(0, \theta)$, $\theta \in \mathbb{R}^+$. Encontrar el UMVUE para θ y $1/\theta$.
3. Sea X_1, \dots, X_n una m.a.s. con distribución $P(\lambda)$, $\lambda \in \mathbb{R}^+$. Encontrar, si existe, el UMVUE para la función paramétrica λ^s , $s \in \mathbb{N}$.

4.3. Estimadores eficientes

En esta sección del tema se va a estudiar el concepto de estimador eficiente con idea de, posteriormente, estudiar la relación que existe entre dichos estimadores y el UMVUE. En toda esta sección consideraremos un espacio paramétrico unidimensional para poder trabajar con el concepto de varianza en lugar del de matriz de varianzas-covarianzas.

Para dar la definición de estimador eficiente, previamente, se van a estudiar condiciones de regularidad.

4.3.1. Condiciones de regularidad de Fréchet-Crámer-Rao

Sea X una v.a. con distribución de probabilidad en la familia de distribuciones $\{F_\theta : \theta \in \Theta\}$. Sea $f_\theta(x)$ la fdp o fmp, según el caso, para cada $\theta \in \Theta \subseteq \mathbb{R}$. Se dice que esta familia de distribuciones cumple las *condiciones de regularidad* si:

- (i) Θ es un intervalo abierto de \mathbb{R} .
- (ii) El conjunto de valores de la variable, $\{x : f_\theta(x) > 0\} = \mathcal{X}$, es independiente de θ .
- (iii) $\forall x \in \mathcal{X}$, $f_\theta(x)$ es derivable respecto de θ y se verifica que

$$\int_{\mathcal{X}} \frac{\partial}{\partial \theta} f_\theta(x) dx = \frac{\partial}{\partial \theta} \int_{\mathcal{X}} f_\theta(x) dx = 0 \quad \left(\sum_{\mathcal{X}} \frac{\partial}{\partial \theta} f_\theta(x) = \frac{\partial}{\partial \theta} \sum_{\mathcal{X}} f_\theta(x) = 0 \right). \quad \forall \theta \in \Theta.$$

Esta tercera condición es equivalente a comprobar que

$$E_\theta \left[\frac{\partial \ln f_\theta(x)}{\partial \theta} \right] = 0.$$

Teorema: Si $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta\}$ cumple las condiciones de regularidad, entonces la familia de distribuciones asociada a la m.a.s. (X_1, \dots, X_n) de X también las cumple.

4.3.2. Función de información de Fisher

Definición: Sea $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta\}$, cuya familia de distribuciones es regular. Se definen las funciones

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

que se denominan *función de información de Fischer* asociada a X y a la muestra, respectivamente.

Propiedades:

- (i) $I_X \geq 0$.
- (ii) $E_\theta \left[\frac{\partial \ln f_\theta(X)}{\partial \theta} \right] = 0$ y $Var_\theta \left[\frac{\partial \ln f_\theta(X)}{\partial \theta} \right] = I_X(\theta)$, $\forall \theta \in \Theta$.
- (iii) $E_\theta \left[\frac{\partial \ln f_\theta^n(X_1, \dots, X_n)}{\partial \theta} \right] = 0$ y $Var_\theta \left[\frac{\partial \ln f_\theta^n(X_1, \dots, X_n)}{\partial \theta} \right] = I_{(X_1, \dots, X_n)}(\theta)$, $\forall \theta \in \Theta$.
- (iv) $I_{X_1, \dots, X_n}(\theta) = nI_X(\theta)$, $\forall \theta \in \Theta$. (Aditividad)

4.3.3. Desigualdad de Fréchet-Cramér-Rao

Definición: Sea (X_1, \dots, X_n) una m.a.s. de $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta\}$, cuya familia de distribuciones es regular. Un estadístico $T(X_1, \dots, X_n)$ se dice que es *regular* en el sentido de Fréchet-Cramér-Rao si verifica:

- Caso discreto

$$\frac{\partial}{\partial \theta} \sum_{(x_1, \dots, x_n) \in \mathcal{X}^n} T(x_1, \dots, x_n) f_\theta^n(x_1, \dots, x_n) = \sum_{(x_1, \dots, x_n) \in \mathcal{X}^n} T(x_1, \dots, x_n) \frac{\partial f_\theta^n(x_1, \dots, x_n)}{\partial \theta}$$

- Caso continuo

$$\frac{\partial}{\partial \theta} \int_{\mathcal{X}^n} T(x_1, \dots, x_n) f_\theta^n(x_1, \dots, x_n) dx_1 \cdots dx_n = \int_{\mathcal{X}^n} T(x_1, \dots, x_n) \frac{\partial f_\theta^n(x_1, \dots, x_n)}{\partial \theta} dx_1 \cdots dx_n$$

Es decir,

$$\frac{\partial}{\partial \theta} E_\theta [T(X_1, \dots, X_n)] = E_\theta \left[T(X_1, \dots, X_n) \frac{\partial \ln f_\theta^n(X_1, \dots, X_n)}{\partial \theta} \right]$$

Teorema (cota de Fréchet-Cramér-Rao): Sea (X_1, \dots, X_n) una m.a.s. de $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta\}$, cuya familia de distribuciones es regular con $0 < I_X(\theta) < +\infty \forall \theta \in \Theta$. Si $T(X_1, \dots, X_n)$ es un estadístico regular, de segundo orden e insesgado en una función paramétrica derivable $g(\theta)$, entonces se tiene

$$(i) \quad Var_\theta T(X_1, \dots, X_n) \geq \frac{(g'(\theta))^2}{I_{X_1, \dots, X_n}(\theta)} \quad \forall \theta \in \Theta.$$

- (ii) Para los puntos θ_0 tales que $g'(\theta_0) \neq 0$, se dará la igualdad si y sólo si $\exists a(\theta_0) \neq 0$ tal que

$$P_{\theta_0} \left[\left. \frac{\partial \ln f_\theta(X_1, \dots, X_n)}{\partial \theta} \right|_{\theta=\theta_0} = a(\theta_0) [T(X_1, \dots, X_n) - g(\theta_0)] \right] = 1$$

4.3.4. Estimador eficiente

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

$$\text{Var}_\theta(T(X_1, \dots, X_n)) = \frac{[g'(\theta)]^2}{I_{X_1, \dots, X_n}(\theta)}, \quad \forall \theta \in \Theta.$$

El estimador eficiente no tiene porque existir.

Lema: Sea $\{F_\theta : \theta \in \Theta\}$ regular, $0 < I_X(\theta) < +\infty \forall \theta \in \Theta$ y $g(\theta)$ una función paramétrica derivable. Entonces $g(\theta)$ admite un estimador eficiente $T(X_1, \dots, X_n)$ si:

- $g(\theta)$ es constante y en tal caso $T(X_1, \dots, X_n)$ es degenerado.
- $g(\theta)$ es estrictamente monótona: $g'(\theta) > 0 \forall \theta \in \Theta$ o $g'(\theta) < 0 \forall \theta \in \Theta$.

Teorema (caracterización de estimadores eficientes): Sea (X_1, \dots, X_n) una m.a.s. de $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta\}$ regular, con $0 < I_X(\theta) < \infty \forall \theta \in \Theta$, $g(\theta)$ una función paramétrica derivable, con $g'(\theta) \neq 0, \forall \theta \in \Theta$ y $T(X_1, \dots, X_n)$ es un estimador de $g(\theta)$. Una condición necesaria y suficiente para que T sea eficiente es:

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

Corolario 1: Si $T(X_1, \dots, X_n)$ es un estimador eficiente para $g(\theta)$, con $g'(\theta) \neq 0$, las únicas funciones paramétricas que admiten estimadores eficientes son las de la forma $ag(\theta) + b$ y los correspondientes estimadores eficientes son $aT + b$, con probabilidad 1, bajo todas las distribuciones de la familia.

Corolario 2: Si una función paramétrica admite dos estimadores eficientes, estos son iguales con probabilidad 1, bajo todas las distribuciones de la familia.

Corolario 3: Sólo existen estimadores eficientes en familias de tipo exponencial.

Corolario 4: Si $T(X_1, \dots, X_n)$ es eficiente para $g(\theta)$, entonces $T(X_1, \dots, X_n)$ es suficiente. Si además la imagen de $Q(\theta) = \int a(\theta) d\theta$ contiene a un abierto de \mathbb{R} , entonces $T(X_1, \dots, X_n)$ es completo.

Corolario 5: Si $T(X_1, \dots, X_n)$ es eficiente para $g(\theta)$, entonces $T(X_1, \dots, X_n)$ es el UMVUE de $g(\theta)$.

El recíproco no es cierto, es decir, si $T(X_1, \dots, X_n)$ es el UMVUE para $g(\theta)$ eso no implica que $T(X_1, \dots, X_n)$ sea eficiente para dicha función paramétrica.

Ejemplo: Buscar la clase de funciones paramétricas que admiten estimadores eficientes para las siguientes familias de distribuciones y calcular dichos estimadores:

1. $\{B(k_0, p), p \in (0, 1)\},$
2. $\{\mathcal{N}(\mu, \sigma_0^2), \mu \in \mathbb{R}\},$
3. $\{\mathcal{N}(\mu_0, \sigma^2), \sigma^2 > 0\}.$