Tema 4.Estimación puntual. Insesgadez y mínima varianza

A. Hermoso Carazo
Universidad de Granada

Curso 2020/2021

Tema 4.Estimación puntual. Insesgadez y mínima varianza A. Hermoso Carazo

Uno de los problemas típicos de inferencia en modelos paramétricos es la estimación puntual del parámetro o parámetros que describen la familia de distribuciones de la variable considerada. El problema consiste en aproximar el verdadero valor del parámetro mediante una función de la muestra, que se denomina *estimador*.

La diversidad de estimadores para un mismo parámetro obliga a establecer criterios que permitan seleccionar el óptimo en algún sentido.

Concretamente, aquí abordaremos el problema de estimación bajo el criterio de *minimizar el error cuadrático medio* lo que, en el caso de *estimadores insesgados*, que consideraremos en este tema, equivale a minimizar la varianza.

4.1. Planteamiento del problema

El problema de estimación puntual consiste en aproximar el verdadero valor de un parámetro (o de una función paramétrica) a partir de los valores muestrales observados. Para ello, se elige un estadístico muestral cuyas realizaciones particulares sean posibles valores de dicho parámetro.

Estimador puntual

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

- *Un estimador de θ es un estadístico, $T(X_1, \ldots, X_n)$, que toma valores en Θ.*
- *Un estimador de una función paramétrica $g(\theta)$ es un estadístico con valores en $g(\Theta)$.*

Nota: En la práctica, si Θ $(g(\Theta))$ es un intervalo abierto se admiten como estimadores los estadísticos que toman valores en la frontera. Si Θ $(g(\Theta))$ es discreto, se admiten como estimadores los que toman valores entre el máximo y el mínimo de Θ $(g(\Theta))$.

En el resto de este apartado nos referiremos a la estimación de θ, siendo todo directamente extensible a la estimación de $g(\theta)$ arbitraria.

Cuando se considera un estimador de θ, $T(X_1, \ldots, X_n)$, una vez observada la muestra y obtenida una realización $(x_1, \ldots, x_n) \in \chi^n$, el estimador toma un valor concreto, $T(x_1, \ldots, x_n) \in \Theta$, que es el que se toma como aproximación, y se denomina una *estimación del parámetro*:

$$\begin{array}{l} T(X_1,\ldots,X_n) \text{ estimador} \\ \qquad\quad \downarrow\ (X_1,\ldots,X_n) = (x_1,\ldots,x_n) \\ T(x_1,\ldots,x_n) \text{ estimación.} \end{array}$$

La gran diversidad de estimadores que pueden encontrarse para un mismo parámetro obliga a seleccionar uno que sea óptimo en algún sentido.

La selección se basa en la bondad de las estimaciones proporcionadas, $T(x_1, \ldots, x_n)$, $(x_1, \ldots, x_n) \in \chi^n$. Para valorar éstas, el problema de estimación debe especificar lo que se denomina una *función de pérdida: $L : \Theta \times \Theta \to \mathbb{R}$, tal que $L(\theta, t)$ cuantifica la pérdida que se obtendría si el verdadero valor del parámetro fuese θ y se estimase por t.*

Se supondrá que la función de pérdida es medible en el segundo argumento; esto es, fijado $\theta \in \Theta$, la función $L(\theta, \cdot) : \Theta \to \mathbb{R}$ es medible.

Una vez especificada la función de pérdida del problema, cada estimador, $T(X_1,\dots,X_n)$, llevará asociado un conjunto de posibles pérdidas (una para cada posible valor del parámetro y cada posible estimación):

$$L(\theta, T(x_1,\dots,x_n)), \quad \forall\theta\in\Theta,\ \forall(x_1,\dots,x_n)\in\chi^n.$$

Aunque lo ideal sería elegir un estimador que las minimice todas, en general esto no es posible, y se razona como sigue.

Para cada $\theta_0\in\Theta$, las pérdidas $L(\theta_0, T(x_1,\dots,x_n))$, $(x_1,\dots,x_n)\in\chi^n$, son los valores de la variable aleatoria $L(\theta_0, T(X_1\dots,X_n))$. Esta variable describe la pérdida que conllevaría el uso del estimador $T(X_1\dots,X_n)$ si θ_0 fuese el verdadero valor del parámetro. Para dar una medida global de su magnitud, en inferencia clásica suele usarse su valor medio (obviamente, hay que suponer que existe), surgiendo así la siguiente definición.

Función de riesgo de un estimador: *Sea $X\to\{P_\theta;\ \theta\in\Theta\}$. Si, para estimar θ, se considera una función de pérdida L, la función de riesgo de un estimador $T(X_1,\dots,X_n)$ bajo esa función de pérdida es la que asigna a cada valor del parámetro la pérdida media asociada*:

$$R_T^L:\Theta\longrightarrow\mathbb{R}$$
$$\theta\longmapsto R_T^L(\theta)=E_\theta[L(\theta,T(X_1,\dots,X_n))].$$

Así, una vez especificada la función de pérdida considerada, *el problema de estimación consiste en buscar un estimador que minimice uniformemente la función de riesgo; esto es, que, sea cual sea el valor del parámetro, el riesgo asociado sea lo menor posible.*

Esto es lo que se denomina un *estimador óptimo bajo la función de pérdida considerada*:

$$T(X_1,\dots,X_n)\ \text{óptimo bajo } L \Leftrightarrow$$
$$\Leftrightarrow \forall T'(X_1,\dots,X_n)\ \text{estimador de } \theta,\ R_T^L(\theta)\le R_{T'}^L(\theta),\quad \forall\theta\in\Theta.$$

Ejemplo 4.1.1: *Sea $X \to \{B(1,\theta);\ \theta \in [0,1]\}$ y supongamos que se desea estimar θ a partir de una muestra de tamaño 2, considerando la función de pérdida $L(\theta,t) = |\theta - t|$. Calcular y comparar las funciones de riesgo de los estimadores $T(X_1,X_2) = \overline{X}$ y $T'(X_1,X_2) = 1/2$.*

La pérdida asociada a $T(X_1,X_2) = \overline{X}$ si θ_0 fuese el verdadero valor del parámetro, es $L(\theta_0,\overline{X}) = |\theta_0 - \overline{X}|$. Esta pérdida es una variable aleatoria cuyos valores dependen de la muestra:

(X_1,X_2)	Probabilidad	$\overline{X}$	$L(\theta_0,\overline{X})$
$(0,0)$	$(1-\theta_0)^2$	0	θ_0
$(0,1)$	$\theta_0(1-\theta_0)$	$1/2$	$\lvert\theta_0 - 1/2\rvert$
$(1,0)$	$\theta_0(1-\theta_0)$	$1/2$	$\lvert\theta_0 - 1/2\rvert$
$(1,1)$	θ_0^2	1	$1-\theta_0$

- $P_{\theta_0}\big(L(\theta_0,\overline{X}) = \theta_0\big) = (1-\theta_0)^2,$
- $P_{\theta_0}\big(L(\theta_0,\overline{X}) = |\theta_0 - 1/2|\big) = 2\theta_0(1-\theta_0),$
- $P_{\theta_0}\big(L(\theta_0,\overline{X}) = 1-\theta_0\big) = \theta_0^2.$

Para dar una medida global de $L(\theta_0,\overline{X})$, usamos $E_{\theta_0}\left[L(\theta_0,\overline{X})\right]$, y haciendo esto para cada valor del parámetro, obtenemos la *función de riesgo de $\overline{X}$:*

$$R_{\overline{X}}(\theta) = E_\theta\left[L(\theta,\overline{X})\right] = \theta(1-\theta)^2 + |\theta - 1/2|2\theta(1-\theta) + (1-\theta)\theta^2, \quad \theta \in [0,1].$$

4.2. Estimación insesgada de mínima varianza

Estimador insesgado

Sea $X \to \{P_\theta;\ \theta \in \Theta\}$ y $(X_1,\ldots,X_n)$ una muestra aleatoria simple de X. Un estimador de θ ($g(\theta)$), $T(X_1,\ldots,X_n)$, es insesgado (o centrado) si:

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

Nota: Aquí nos centraremos exclusivamente en el caso uniparamétrico. Si el parámetro es k-dimensional, $\theta = (\theta_1,\ldots,\theta_k)$, un estimador también debe serlo, $T(X_1,\ldots,X_n) = \big(T_1(X_1,\ldots,X_n),\ldots,T_k(X_1,\ldots,X_n)\big)$, y la condición de insesgadez es

$$E_\theta\left[T_i(X_1,\ldots,X_n)\right] = \theta_i, \quad i = 1,\ldots,k, \quad \forall \theta \in \Theta.$$

Para el estimador $T'(X_1, X_2) = 1/2$, es claro que la pérdida asociada es constante para cada valor del parámetro, $L(\theta, T') = L(\theta, 1/2) = |\theta - 1/2|$ y, por lo tanto,

$$R_{T'}(\theta) = E_\theta\big[L(\theta, 1/2)\big] = |\theta - 1/2|, \quad \theta \in [0, 1].$$

Si comparamos los riesgos en distintos valores del parámetro, observamos que en algunos casos supondría menos riesgo usar $\overline{X}$, y en otros T'. Por ejemplo:

- $\theta = 0 \to R_T(0) = 0 < R_{T'}(0) = 1/2.$
- $\theta = 1/8 \to R_T(1/8) = 49/256 < R_{T'}(1/8) = 3/8.$
- $\theta = 1/2 \to R_T(1/2) = 1/4 > R_{T'}(1/2) = 0.$

Ya que no conocemos el valor del parámetro, no podemos afirmar que uno de estos estimadores sea mejor o peor que el otro.

Este ejemplo muestra que, en general, no es posible encontrar un estimador que minimice uniformemente la función de riesgo y el problema debe ser reconsiderado. La solución, bajo la perspectiva clásica, es restringir la búsqueda del estimador óptimo a determinadas clases de estimadores, que se fijan imponiendo propiedades deseables.

Aquí tratamos el problema de estimación considerando una *función de pérdida cuadrática*:

$$\begin{aligned} L : \Theta \times \Theta &\longrightarrow \mathbb{R} \\ (\theta, t) &\longmapsto L(\theta, t) = (t - \theta)^2. \end{aligned}$$

En este caso, la función de riesgo asociada a un estimador $T(X_1, \ldots, X_n)$ es

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

y se denomina *error cuadrático medio de* $T(X_1, \ldots, X_n)$.

Por tanto, el objetivo es encontrar un estimador que minimice uniformemente el error cuadrático medio.

Ya que, como acabamos de ver, en general esto no es posible, trataremos el problema de estimación restringiendo la búsqueda del estimador óptimo a la clase de estimadores con la propiedad de *insegadez*.

Para estos estimadores, como ahora veremos, el error cuadrático medio es la varianza, y por lo tanto, nuestro estudio se refiere al problema de *estimación insesgada de mínima varianza*.

Ejemplo 4.2.1: *La media y la cuasivarianza muestral son estimadores insesgados de la media y de la varianza poblacional, respectivamente, siempre que éstas existan.*

En efecto, según se probó en el Tema 1:

$$E_\theta\left[\overline{X}\right] = E_\theta[X], \quad \forall\theta \in \Theta, \qquad E_\theta\left[S^2\right] = Var_\theta[X], \quad \forall\theta \in \Theta.$$

Así, si, por ejemplo, $X \to \{\mathcal{P}(\lambda);\ \lambda > 0\}$, $\overline{X}$ y S^2 son estimadores insesgados de λ. Además, por la linealidad de la esperanza:

$$E_\lambda\left[\alpha\overline{X} + (1-\alpha)S^2\right] = \alpha E_\lambda\left[\overline{X}\right] + (1-\alpha)E\left[S^2\right] = \lambda, \quad \forall\lambda > 0,$$

de forma que si $\alpha \in (0,1)$ (para asegurar que sea estimador) $\alpha\overline{X} + (1-\alpha)S^2$ es también un estimador insesgado.

Esto prueba que un estimador insesgado no tiene por qué ser único.

En general, debido a la linealidad de la esperanza, la propiedad de insesgadez se mantiene al aplicar una misma transformación lineal a un estimador insesgado y a la función que se estima:

$$E_\theta\left[T\right] = g(\theta), \quad \forall\theta \in \Theta \ \Rightarrow\ E_\theta\left[aT + b\right] = ag(\theta) + b, \quad \forall\theta \in \Theta.$$

Sin embargo, la insesgadez no se mantiene, en general, al aplicar otro tipo de transformación. Por ejemplo, si $X \to \{\mathcal{P}(\lambda);\ \lambda > 0\}$, $\overline{X}$ es insesgado en λ pero, sin embargo, $\overline{X}^2$ no es insesgado en λ^2:

$$E_\lambda\left[\overline{X}^2\right] = Var_\lambda\left[\overline{X}\right] + \left(E_\theta\left[\overline{X}\right]\right)^2 = \frac{\lambda}{n} + \lambda^2 \neq \lambda^2.$$

La no unicidad de estimadores insesgados obliga a plantearse el problema de buscar el óptimo. Trabajando, como hemos indicado, bajo el criterio de minimizar el error cuadrático medio, si $T(X_1,\ldots,X_n)$ es un estimador insesgado de $g(\theta)$, su error cuadrático medio es:

$$\begin{aligned} E_\theta\left[\left(T(X_1,\ldots,X_n) - g(\theta)\right)^2\right] &= E_\theta\left[\left(T(X_1,\ldots,X_n) - E_\theta\left[T(X_1,\ldots,X_n)\right]\right)^2\right] \\ &= Var_\theta\left[T(X_1,\ldots,X_n)\right], \quad \forall\theta. \end{aligned}$$

Esto es, el error cuadrático medio asociado a un estimador insesgado es su varianza. Por tanto, el estimador insesgado óptimo será el que tenga menor varianza para cualquier valor del parámetro.

Nota: En adelante, un *estadístico de segundo orden* será el que tenga momento de segundo orden finito bajo todas las distribuciones de la familia:

$T(X_1,\ldots,X_n)$ es de segundo orden $\Leftrightarrow E_\theta\left[\left(T(X_1,\ldots,X_n)\right)^2\right] < +\infty, \ \forall\theta \in \Theta.$

Estimador insesgado uniformemente de mínima varianza (UMVUE)

Sea $(X_1, \ldots, X_n)$ *una muestra aleatoria simple de* $X \to \{P_\theta;\ \theta \in \Theta\}$ *y* $T(X_1, \ldots, X_n)$ *un estimador de* $g(\theta)$ *insesgado y de segundo orden.*

Diremos que $T(X_1, \ldots, X_n)$ *es UMVUE para* $g(\theta)$ *si para cualquier otro estimador insesgado de* $g(\theta)$*,* $T'(X_1, \ldots, X_n)$*, se tiene:*

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

Propiedades:

- *Unicidad*: *El UMVUE de un parámetro (o función paramétrica), si existe, es único.*
- *Linealidad*: *Si* $T_i(X_1, \ldots, X_n)$ *es 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 UMVUE para* $a_1 g_1(\theta) + a_2 g_2(\theta)$.

Aunque, en general, la determinación de UMVUE's no es fácil, el problema se simplifica cuando existe un estadístico suficiente y completo, como se deduce de los siguientes teoremas.

Teorema de Rao-Blackwell

Sea $(X_1, \ldots, X_n)$ *una muestra aleatoria simple de* $X \to \{P_\theta;\ \theta \in \Theta\}$*, y* $T \equiv T(X_1, \ldots, X_n)$ *un estadístico suficiente. Si* $S \equiv S(X_1, \ldots, X_n)$ *es un estimador insesgado de* $g(\theta)$*, de segundo orden:*

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

Demostración: La clave de este resultado radica en que, por ser T suficiente, $E[S/T]$ es independiente del parámetro y, por lo tanto, es un estadístico.

$E[S/T]$ toma valores en el rango de valores de S, de manera que es también estimador de $g(\theta)$. Además, de las propiedades de la esperanza condicionada, se deduce que este estimador es también insesgado y de segundo orden:

- $E_\theta\left[E[S/T]\right] = E_\theta[S] = g(\theta), \quad \forall \theta \in \Theta.$
- $(E[S/T])^2 \leq E\left[S^2/T\right]$. Por tanto:

$$E_\theta\left[(E[S/T])^2\right] \leq E_\theta\left[E\left[S^2/T\right]\right] = E_\theta\left[S^2\right] < +\infty, \quad \forall \theta \in \Theta.$$

Veamos finalmente que su varianza es menor o igual que la de S:

$$Var_\theta\left[E[S/T]\right] = E_\theta\left[(E[S/T])^2\right] - \left(E_\theta\left[E[S/T]\right]\right)^2 \leq E_\theta\left[S^2\right] - (E_\theta[S])^2 = Var_\theta[S], \quad \forall \theta \in \Theta.$$

El *Teorema de Rao-Blackwell* indica que todo estimador insesgado de segundo orden puede mejorarse, en el sentido de disminuir uniformemente su varianza, considerando su esperanza condicionada a un estadístico suficiente, T. Por tanto, la búsqueda del UMVUE debe restringirse a la clase de estimadores de la forma $E[S/T]$.

Ya que $E[S/T]$ es una función de T, resulta que el UMVUE, si existe, es función del estadístico suficiente T.

A continuación vemos que si el estadístico suficiente es, además, completo, la clase de estimadores de la forma $E[S/T]$ contiene un único elemento y, por tanto, este elemento es el UMVUE.

Teorema de Lehmann-Scheffé

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de $X \to \{P_\theta;\ \theta \in \Theta\}$, y $T \equiv T(X_1, \ldots, X_n)$ un estadístico suficiente y completo.

Si la función $g(\theta)$ admite algún estimador insesgado de segundo orden, $S \equiv S(X_1, \ldots, X_n)$, entonces existe el UMVUE de $g(\theta)$ y está dado por $E\big[S/T\big]$.

Demostración: Según el *Teorema de Rao-Blackwell*, el UMVUE, de existir, es de la forma $E\big[S'/T\big]$, donde S' es un estimador insesgado de segundo orden.

Probamos ahora que si T es, además de suficiente, completo, esa clase contiene un único elemento y es, por tanto, el UMVUE.

En efecto, si $E\big[S/T\big]$ y $E\big[S'/T\big]$ son dos elementos de esa clase, son insesgados y, por tanto:

$$E_\theta\Big[E\big[S/T\big] - E\big[S'/T\big]\Big] = 0, \quad \forall \theta \in \Theta.$$

Así, tenemos una función de T, $E\left[S/T\right] - E\left[S'/T\right]$, que tiene media nula bajo todas las distribuciones de la familia. Entonces, como T es completo, se tiene:

$$P_\theta\Big(E\left[S/T\right] = E\left[S'/T\right]\Big) = 1, \quad \forall \theta \in \Theta. \qquad \square$$

Métodos para el cálculo de UMVUE: Si la familia de distribuciones admite un estadístico suficiente y completo, $T \equiv T(X_1, \ldots, X_n)$, los teoremas anteriores proporcionan el camino para calcular el UMVUE (si existe) de una función paramétrica arbitraria, $g(\theta)$:

- *Método directo*: Buscar un estimador insesgado de segundo orden cualquiera y calcular su esperanza condicionada a T.
- *Método alternativo*: Considerar una función arbitraria, $h(T)$, e imponer que sea insesgada (al ser T completo, si existe, es única). Ya que $E\left[h(T)/T\right] = h(T)$, si esta función es un estimador de $g(\theta)$ y es de segundo orden, será el UMVUE.

Normalmente, el segundo método es más cómodo ya que evita el cálculo de las distribuciones condicionadas.

Ejemplo 4.2.2: *Usando el método directo, determinar el UMVUE para p basado en una muestra, $(X_1, \ldots, X_n)$, de $X \to \{B(1,p);\ p \in (0,1)\}$.*

El primer paso es buscar un estadístico suficiente y completo, problema resuelto en los ejemplos 3.1.3 y 3.2.1. Este estadístico es $T = \sum_{i=1}^{n} X_i$ y, por tanto, la búsqueda del UMVUE se restringe a funciones de T, cuya familia de distribuciones es:

$$T \to \left\{B(n,p);\ p \in (0,1)\right\}.$$

Método directo: Buscamos un estimador insesgado de segundo orden y calculamos su esperanza condicionada a T.

Para simplificar, elegimos una función de una sóla variable de la muestra; por ejemplo X_1 (insesgado y de segundo orden), y calculamos $E\left[X_1/T\right]$. Dado un valor de T, $t = 0, 1, \ldots, n$:

$$E[X_1/T=t] = (0 \times P(X_1 = 0/T = t)) + (1 \times P(X_1 = 1/T = t)) = P(X_1 = 1/T = t)$$

$$= \frac{P_p\left(X_1 = 1, T = t\right)}{P_p\left(T = t\right)} = \frac{P_p(X_1 = 1,\ \sum_{i=2}^{n} X_i = t-1)}{P_p\left(T = t\right)}.$$

Para calcular la probabilidad del numerador, tenemos en cuenta que $\sum_{i=2}^{n} X_i$ es independiente de X_1, y $\sum_{i=2}^{n} X_i \to B(n-1,p)$. Por tanto:

- $t = 0 \to P_p(X_1 = 1, \sum_{i=2}^{n} X_i = t-1) = 0.$
- $t = 1, \ldots, n:$

$$P(X_1 = 1/T = t) = \frac{p \times \binom{n-1}{t-1} p^{t-1}(1-p)^{n-t}}{\binom{n}{t} p^t (1-p)^{n-t}} = \frac{\binom{n-1}{t-1}}{\binom{n}{t}} = \frac{t}{n}.$$

Por tanto:

$$E[X_1/T = t] = \begin{cases} 0, & t = 0 \\ \frac{t}{n}, & t = 1, \ldots, n, \end{cases} \Rightarrow E[X_1/T] = \frac{T}{n} = \overline{X}.$$

Así, $\overline{X}$ es el UMVUE de p (notemos que p es la media de la variable).

Nota: En este caso, el problema se resuelve de forma inmediata observando que $E_\theta[T] = np$, de manera que $E_\theta[T/n] = p$, $\forall p \in [0,1]$. Entonces, como $T/n = \overline{X} \in [0,1]$, es estimador de p. También es de segundo orden y, por lo tanto, es el UMVUE.

Tema 4.Estimación puntual. Insesgadez y mínima varianza A. Hermoso Carazo
4.2. Estimación insesgada de mínima varianza

Ejemplo 4.2.3: *Usando el método alternativo, determinar el UMVUE de p basado en una muestra, $(X_1, \ldots, X_n)$, de $X \to \{B(1,p);\ p \in (0,1)\}$.*

Método alternativo: Dado un estadístico suficiente y completo, T, el método consiste en considerar una función genérica, $h(T)$, imponer la condición de insesgadez y resolver la ecuación resultante en h. Recordamos que, como T es completo, la función insesgada, si existe, es única. Una vez obtenida $h(T)$, hay que comprobar que es estimador y de segundo orden. En nuestro caso, el estadístico suficiente y completo es

$$T = \sum_{i=1}^{n} X_i \to \{B(n,p);\ p \in (0,1)\}.$$

Buscamos $h(T)\ /\ E_p[h(T)] = p$, $\forall p \in (0,1)$:

$$E_p[h(T)] = \sum_{t=0}^{n} h(t) P_p(T = t) = \sum_{t=0}^{n} h(t) \binom{n}{t} p^t (1-p)^{n-t} = p, \quad \forall p \in (0,1),$$

o, equivalentemente, sacando factor común $(1-p)^n$:

$$\sum_{t=0}^{n} h(t) \binom{n}{t} \left(\frac{p}{1-p}\right)^t = \frac{p}{(1-p)^n}, \quad \forall p \in (0,1).$$

Ahora resolvemos esta ecuación, buscando los valores $h(0), \ldots, h(n)$ que la satisfacen. Para ello, hacemos $y = \frac{p}{1-p}$, con lo que $p = \frac{y}{1+y}$, $1 - p = \frac{1}{1+y}$, y se tiene:

$$\sum_{t=0}^{n} h(t) \binom{n}{t} y^t = y(1+y)^{n-1}, \quad \forall y \in \mathbb{R}^+.$$

Para resolver esta ecuación, desarrollamos en serie de potencias el segundo miembro e identificamos coeficientes:

- $n \geq 2 \to \sum_{t=0}^{n} h(t) \binom{n}{t} y^t = \sum_{s=0}^{n-1} \binom{n-1}{s} y^{s+1} \overset{t=s+1}{=} \sum_{t=1}^{n} \binom{n-1}{t-1} y^t, \ \forall y \in \mathbb{R}^+.$

  Por tanto, $h(0) = 0$ y $h(t)\binom{n}{t} = \binom{n-1}{t-1}$, $t = 1, \dots, n$; esto es:

$$h(t) = \begin{cases} 0, & t = 0 \\ \dfrac{\binom{n-1}{t-1}}{\binom{n}{t}} = \dfrac{t}{n}, & t = 1, \dots, n. \end{cases} \Rightarrow h(T) = \frac{T}{n}.$$

- $n = 1 \to h(0) + h(1)y = y, \ \forall y \in \mathbb{R}^+ \Rightarrow h(0) = 0, \ h(1) = 1.$

  Ya que, en este caso, $T = 0, 1$, se tiene de nuevo $h(T) = T/n$.

En resumen, $h(T) = T/n = \overline{X}$ es la única función de T insesgada en p. Como es estimador y es de segundo orden, es el UMVUE.

Ejemplo 4.2.4: *Encontrar el UMVUE para θ, basado en un muestra aleatoria simple, $(X_1, \dots, X_n)$, de una variable $X \to \{U(0,\theta);\ \theta \in \mathbb{R}^+\}$.*

En los ejemplos 3.1.4 y 3.2.3, se probó que $T = \max X_i$ es suficiente y completo para esta familia de distribuciones. Por tanto, la búsqueda del UMVUE se restringe a funciones de T, cuya función de densidad (Ejemplo 1.4.4) es:

$$\forall \theta \in \mathbb{R}^+, \quad f_\theta^T(t) = \frac{nt^{n-1}}{\theta^n}, \quad 0 < t < \theta.$$

Consideramos una función genérica, $h(T)$, e imponemos que sea insesgada:

$$E_\theta[h(T)] = \int_0^\theta h(t) f_\theta^T(t)dt = \int_0^\theta h(t)\frac{nt^{n-1}}{\theta^n}dt = \theta, \quad \forall \theta \in \mathbb{R}^+.$$

o, simplificando la integral y pasando al último miembro las constantes:

$$\int_0^\theta h(t)t^{n-1}dt = \frac{\theta^{n+1}}{n}, \quad \forall \theta \in \mathbb{R}^+.$$

Notemos que ambos miembros de esta ecuación tienden a cero cuando θ tiende a cero, y son funciones derivables (*Teorema Fundamental del Cálculo*). Por tanto, tiene solución, y la obtenemos derivando en ambos miembros:

$$h(\theta)\theta^{n-1} = \left(\frac{\theta^{n+1}}{n}\right)' = \frac{(n+1)\theta^n}{n}, \quad \forall \theta \in \mathbb{R}^+ \Rightarrow h(\theta) = \frac{(n+1)\theta}{n}, \ \forall \theta \in \mathbb{R}^+.$$

Tenemos ya la forma de la función h en $\mathbb{R}^+$ (conjunto de posibles valores de T). Ahora, comprobamos que $h(T)$ es estimador de θ y que es de segundo orden:

- $T \in \mathbb{R}^+ \Rightarrow h(T) = \dfrac{(n+1)T}{n} \in \mathbb{R}^+$ (conjunto de valores de θ).

  Por tanto, $h(T)$ es estimador de θ.

- $\forall\theta,\ 0 < T < \theta \Rightarrow 0 < h(T) < \dfrac{(n+1)\theta}{n}$.

  Por tanto, para cada distribución de la familia, $h(T)$ está acotada y, consecuentemente, tiene momentos de todos los órdenes (en particular, de orden dos).

De todo esto se deduce que $h(T) = \dfrac{(n+1)T}{n}$, con $T = \max X_i$, es el UMVUE de θ.

Ejemplo 4.2.5: *Encontrar el UMVUE para para $1/\theta$, basado en una muestra aleatoria simple, $(X_1,\ldots,X_n)$, de $X \to \{U(0,\theta);\ \theta \in \mathbb{R}^+\}$.*

Basándonos de nuevo en el estadístico suficiente y completo $T = \max X_i$, consideramos una función genérica, $h(T)$, e imponemos que sea insesgada en $1/\theta$:

$$E_\theta\big[h(T)\big] = \int_0^\theta h(t) f_\theta^T(t)dt = \int_0^\theta h(t)\frac{nt^{n-1}}{\theta^n}dt = 1/\theta, \quad \forall\theta \in \mathbb{R}^+.$$

Simplificando la integral, la ecuación a resolver es:

$$\int_0^\theta h(t)t^{n-1}dt = \frac{\theta^{n-1}}{n}, \quad \forall\theta \in \mathbb{R}^+.$$

Notemos que si $n=1$, el primer miembro de esta ecuación tiende a cero cuando θ tiende a cero, mientras que el segundo miembro es $1/n$. Esto significa que la ecuación no tiene solución; por tanto, *no existe estimador insesgado de $1/\theta$ basado en una sóla observación de la variable.*

Nos centramos entonces en $n \geq 2$. En este caso, ambos miembros de la ecuación tienden a cero cuando θ tiende a cero y son derivables (*Teorema Fundamental del Cálculo*). Derivando en ambos miembros obtenemos:

$$h(\theta)\theta^{n-1} = \left(\frac{\theta^{n-1}}{n}\right)' = \frac{(n-1)\theta^{n-2}}{n}, \quad \forall\theta \in \mathbb{R}^+ \ \Rightarrow\ h(\theta) = \frac{(n-1)}{n\theta}, \quad \forall\theta \in \mathbb{R}^+.$$

Así, tenemos la forma de la función h sobre $\mathbb{R}^+$ (conjunto de posibles valores de T), que debe ser aplicada al estadístico T y comprobar si $h(T)$ es estimador de $1/\theta$ (que toma valores en $\mathbb{R}^+$) y de segundo orden:

- $T \in \mathbb{R}^+ \Rightarrow h(T) = \dfrac{n-1}{nT} \in \mathbb{R}^+$ (conjunto de valores de $1/\theta$).
  Por tanto, $h(T)$ es estimador de $1/\theta$.
- $E_\theta\left[h^2(T)\right] = \left(\dfrac{n-1}{n}\right)^2 \displaystyle\int_0^\theta \frac{nt^{n-3}}{\theta^n} dt = \frac{(n-1)^2}{n\theta^n} \int_0^\theta t^{n-3} dt < +\infty \Leftrightarrow n \geq 3,\ \forall\theta.$

Esto es, si $n = 2$ *ningún estimador insesgado de $1/\theta$ tiene varianza finita, y $h(T)$ sólo es UMVUE para $n \geq 3$.*

En resumen:

- Si $n = 1$, no existe ningún estimador insesgado.
- Si $n = 2$, ningún estimador insesgado tiene varianza finita y no tiene sentido la búsqueda del UMVUE.
- Si $n \geq 3$, el UMVUE de $1/\theta$ es $h(T) = \dfrac{n-1}{nT}$, con $T = \text{máx}\, X_i$.

Sea $(X_1 \dots X_n)$ un m.a.s. de $X \to \{N(\mu,\sigma^2),\ \mu\in\mathbb{R},\ \sigma^2\in\mathbb{R}^+\}$. Probar que

$$T(X_1\dots X_n) = \begin{cases} 1 & \bar{X} \leq 0 \\ 0 & \bar{X} > 0 \end{cases}$$

es un estimador insesgado de la función paramétrica $\Phi\left(\frac{-\mu\sqrt{n}}{\sigma}\right)$ con Φ la función de distribución de la $N(0,1)$.

$\Theta = \mathbb{R}\times\mathbb{R}^+$ $\qquad g(\mu,\sigma^2) = \Phi\left(\dfrac{-\mu\sqrt{n}}{\sigma}\right) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\frac{-\mu\sqrt{n}}{\sigma}} e^{-\frac{u^2}{2}}\, du$

El estimador toma valores 0 y 1 que:

(i) Si $\sigma = 0 \in \overline{\mathbb{R}^+}$ y $\mu > 0 \Rightarrow g(\mu,\sigma^2) = 0$

(ii) Si $\sigma = 0 \in \overline{\mathbb{R}^+}$ y $\mu < 0 \Rightarrow g(\mu,\sigma^2) = 1$ por ser Φ de distribución.

Luego T es un estimador de $g(\mu,\sigma^2)$.

Veamos si es insesgado:

$$E_\theta[T(X_1\cdots X_n)] = E_{(\mu,\sigma^2)}[T(X_1\cdots X_n)] = 1\cdot P_{(\mu,\sigma^2)}(\bar{X}\le 0) + 0\cdot P_{(\mu,\sigma^2)}(\bar{X}>0)$$

$$= P_{(\mu,\sigma^2)}(\bar{X}-\mu \le -\mu) = P_{(\mu,\sigma^2)}\left(\frac{\bar{X}-\mu}{(\sigma/\sqrt{n})} \le -\frac{\mu}{(\sigma/\sqrt{n})}\right) = P_{(\mu,\sigma^2)}\left(N(0,1) \le -\frac{\mu}{\sigma/\sqrt{n}}\right)$$

$$= \Phi\left(-\frac{\mu\sqrt{n}}{\sigma}\right) \quad \forall(\mu,\sigma^2)\in \mathbb{R}\times\mathbb{R}^+$$

Problema 2

Sea $(X_1,\ldots,X_n)$ una muestra aleatoria simple de una variable aleatoria $X \to \{B(1,p);\ p\in(0,1)\}$ y $T=\sum_{i=1}^{n} X_i$.

a) Probar que si $k\in\mathbb{N}$ y $k\le n$, el estadístico $\dfrac{T(T-1)\cdots(T-k+1)}{n(n-1)\cdots(n-k+1)}$ es un estimador insesgado de p^k. ¿Es este estimador el UMVUE?

b) Probar que si $k>n$, no existe ningún estimador insesgado para p^k.

c) ¿Puede afirmarse que $\frac{T}{n}\left(1-\frac{T}{n}\right)^2$ es insesgado para $p(1-p)^2$?

$T \to \{B(n,p);\ p\in(0,1)\} \longrightarrow$ *suficiente y completo.*

a) Notemos en primer lugar que el estadístico en cuestión, que notaremos $U(T)$, sólo está bien definido si $k\le n$, ya que si $k>n$, uno de los factores del denominador se anula.

$*$ *$U(T)$ es estimador de $p^k \in (0,1)$:*

$$U(T) = \frac{T}{n} \times \frac{T-1}{n-1} \times \cdots \times \frac{T-(k-1)}{n-(k-1)} \begin{cases} = 0, & T = 0, \ldots, k-1 \\ \in (0,1], & T = k, \ldots, n. \end{cases}$$

$*$ *U(T) es insesgado en p^k*: Dado $p \in (0,1)$:

$$E_p[U(T)] = \sum_{t=0}^{n} U(t) P_p(T=t) = \sum_{t=k}^{n} \frac{t(t-1)\cdots(t-k+1)}{n(n-1)\cdots(n-k+1)} \binom{n}{t} p^t (1-p)^{n-t}$$

$$= \sum_{t=k}^{n} \frac{(n-k)!}{(t-k)!(n-t)!} p^t (1-p)^{n-t} = \sum_{t=k}^{n} \binom{n-k}{t-k} p^t (1-p)^{n-t} = [t-k=s]$$

$$= \sum_{s=0}^{n-k} \binom{n-k}{s} p^{s+k} (1-p)^{n-k-s} = p^k \sum_{s=0}^{n-k} \binom{n-k}{s} p^s (1-p)^{n-k-s} = p^k.$$

$*$ *U(T) es el UMVUE de p^k:*

- $U(T)$ es insesgado en p^k.
- $(U(T))^2 \leq 1 \Rightarrow E_p\left[(U(T))^2\right] < +\infty, \ \forall p \in (0,1)$.
- $U(T)$ es función de T, que es suficiente y completo.

Relación 4. Estimación puntual. Insesgadez y mínima varianza A. Hermoso Carazo
└ Problema 2

b) *Probar que si $k > n$, no existe ningún estimador insesgado para p^k.*

Si $S(X_1, \ldots, X_n)$ es insesgado en p^k, $E_p[S(X_1, \ldots, X_n)] = \quad \forall p \in (0,1)$:

$$E_p[S(X_1,\ldots,X_n)] = \sum_{x_1,\ldots,x_n=0}^{1} S(x_1,\ldots,x_n) P_p(X_1 = x_1, \ldots, X_n = x_n)$$

$$= \sum_{x_1,\ldots,x_n=0}^{1} S(x_1,\ldots,x_n) p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i}.$$

Esta suma es un polinomio en p de grado $n < k$. Por tanto, no puede igualarse a p^k. No existe estimador insesgado de p^k.

c) *¿Puede afirmarse que $\frac{T}{n}\left(1 - \frac{T}{n}\right)^2$ es insesgado para $p(1-p)^2$?*

La función $p(1-p)^2 = p^3 - 2p^2 + p$ es un polinomio en p de grado 3; por tanto, no admite estimador insesgado si $n < 3$.

Para $n \geq 3$, si dicho estadístico fuese insesgado, al ser función de T (suficiente y completo) de segundo orden (T está acotado), sería el UMVUE. Aplicando el apartado $a)$ y la linealidad, el UMVUE para $p^3 - 2p^2 + p$ es

$$\frac{T(T-1)(T-2)}{n(n-1)(n-2)} - 2\frac{T(T-1)}{n(n-1)} + \frac{T}{n} \neq \frac{T}{n}\left(1 - \frac{T}{n}\right)^2.$$

Por lo tanto, no es insesgado.

Problema 3

Sea $(X_1,\ldots,X_n)$ una muestra aleatoria simple de una variable aleatoria $X \to \{\mathcal{P}(\lambda);\ \lambda>0\}$. Encontrar, si existe, el UMVUE para λ^s, siendo $s\in\mathbb{N}$ arbitrario.

$$T=\sum_{i=1}^{n}X_i \to \{\mathcal{P}(n\lambda);\ \lambda>0\} \longrightarrow \textit{suficiente y completo (R3-P2)}\,.$$

Consideramos una función arbitraria, $h(T)$, e imponemos que sea insesgada:

$$E_\lambda\big[h(T)\big]=\sum_{t=0}^{+\infty}h(t)P_\lambda(T=t)=\sum_{t=0}^{+\infty}h(t)\frac{e^{-n\lambda}(n\lambda)^t}{t!}=\lambda^s,\quad \forall\lambda>0.$$

$$\Downarrow$$

$$\sum_{t=0}^{+\infty}\frac{h(t)n^t}{t!}\lambda^t=e^{n\lambda}\lambda^s,\quad \forall\lambda>0.$$

$$\Downarrow$$

$$\sum_{t=0}^{+\infty}\frac{h(t)n^t}{t!}\lambda^t=\left(\sum_{k=0}^{+\infty}\frac{n^k\lambda^k}{k!}\right)\lambda^s=\sum_{k=0}^{+\infty}\frac{n^k\lambda^{k+s}}{k!}=[t=k+s]=\sum_{t=s}^{+\infty}\frac{n^{t-s}}{(t-s)!}\lambda^t,\ \forall\lambda>0.$$

$$\sum_{t=0}^{+\infty}\frac{h(t)n^t}{t!}\lambda^t=\sum_{t=s}^{+\infty}\frac{n^{t-s}}{(t-s)!}\lambda^t,\quad \forall\lambda>0.$$

Identificando coeficientes obtenemos:

$$h(t)=\begin{cases}0, & t<s,\\[2mm] \dfrac{t!}{n^s(t-s)!}=\dfrac{t(t-1)\cdots(t-s+1)}{n^s}, & t\geq s.\end{cases}$$

$$=\frac{t(t-1)\cdots(t-s+1)}{n^s},\ t\in\mathbb{N}\cup\{0\} \overset{T\in\mathbb{N}\cup\{0\}}{\Longrightarrow} h(T)=\frac{T(T-1)\cdots(T-s+1)}{n^s}$$

* $h(T)\geq 0 \longrightarrow$ estimador de $\lambda^s>0$.
* $E_\lambda\left[(h(T))^2\right]<+\infty,\ \forall\lambda>0$, ya que es un polinomio en T de grado s y T tiene momentos de todos los órdenes.

$$\text{UMVUE para } \lambda^s \ \to\ h(T)=\frac{T(T-1)\cdots(T-s+1)}{n^s}.$$

Nota: *En particular, tomando $s=1$ se deduce que $\overline{X}=T/n$ es el UMVUE para λ.*

Problema 4

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable con distribución uniforme discreta en los puntos $\{1, \ldots, N\}$, siendo N un número natural arbitrario. Encontrar el UMVUE para N.

$$P_N(X = x) = \frac{1}{N}, \quad x = 1, \ldots, N, \quad N \in \mathbb{N}.$$

$*$ *Estadístico suficiente y completo* $\longrightarrow$ $T = \max X_i$ (R3-P5).

$*$ *Distribución de T* $\longrightarrow$ $P_N(T = t) = \dfrac{t^n - (t-1)^n}{N^n}, \quad t = 1, \ldots, N.$

Consideramos una función arbitraria, $h(T)$, e imponemos insesgadez:

$$E_N[h(T)] = \sum_{t=1}^{N} h(t) P_N(T = t) = \sum_{t=1}^{N} h(t) \frac{t^n - (t-1)^n}{N^n} = N, \quad \forall N \in \mathbb{N}.$$

$$\Downarrow$$

$$\sum_{t=1}^{N} h(t)\big(t^n - (t-1)^n\big) = N^{n+1}, \quad \forall N \in \mathbb{N}.$$

$$\sum_{t=1}^{N} h(t)\big(t^n - (t-1)^n\big) = N^{n+1}, \quad \forall N \in \mathbb{N}.$$

$*$ $N = 1 \to h(1) = 1.$

$*$ $N \geq 2 \to \underbrace{\sum_{t=1}^{N} h(t)\big(t^n - (t-1)^n\big) - \sum_{t=1}^{N-1} h(t)\big(t^n - (t-1)^n\big)}_{h(N)\left(N^n - (N-1)^n\right)} = N^{n+1} - (N-1)^{n+1}$

$$h(N)\big(N^n - (N-1)^n\big) \Rightarrow h(N) = \frac{N^{n+1} - (N-1)^{n+1}}{N^n - (N-1)^n}, \quad \forall N \in \mathbb{N}.$$

$$\Downarrow \ T \in \mathbb{N}$$

$$h(T) = \frac{T^{n+1} - (T-1)^{n+1}}{T^n - (T-1)^n}.$$

$*$ $h(T)$ estimador de $N \in \mathbb{N}$:

$$T^{n+1} - (T-1)^{n+1} = TT^n - (T-1)(T-1)^n = T\big(T^n - (T-1)^n\big) + (T-1)^n$$
$$\geq T^n - (T-1)^n \Rightarrow h(T) \geq 1.$$

Si $h(T) \notin \mathbb{N}$ se toma como estimación el número natural más próximo.

$*$ $E_N[h^2(T)] < +\infty, \ \forall N \in \mathbb{N}$. Debido a que T, y por tanto $h(T)$, toma un número finito de valores.

UMVUE para N $\rightarrow$ $h(T) = \dfrac{T^{n+1} - (T-1)^{n+1}}{T^n - (T-1)^n}.$

Problema 5

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable X con función de densidad:

$$f_\theta(x) = \frac{1}{2\sqrt{x\theta}}, \quad 0 < x < \theta.$$

Calcular, si existe, el UMVUE para θ.

Ya que no se especifica el espacio paramétrico, tomamos el más general posible, $\Theta = \mathbb{R}^+$.

$$\chi_\theta = (0,\theta) \Rightarrow \chi = \bigcup_{\theta\in\mathbb{R}^+} (0,\theta) = \mathbb{R}^+ \longrightarrow \chi^n = (\mathbb{R}^+)^n.$$

Estadístico suficiente:

$$f_\theta^n(x_1,\ldots,x_n) = \frac{1}{2^n\sqrt{\prod_{i=1}^n x_i}} \times \frac{1}{\theta^{n/2}} I_{(0,+\infty)}(\theta - \text{máx}\, x_i), \quad x_1,\ldots,x_n > 0.$$

$$\Downarrow$$

$T = \text{máx}\, X_i$ suficiente.

Distribución de $T = \text{máx}\, X_i \longrightarrow F_\theta^T(t) = \big(F_\theta(t)\big)^n$:

$$F_\theta(x) = \int_0^x \frac{1}{2\sqrt{t\theta}}dt = \sqrt{\frac{x}{\theta}}, \;\; 0 < x < \theta \;\Rightarrow\; F_\theta^T(t) = (F_\theta(t))^n = \frac{t^{n/2}}{\theta^{n/2}}, \;\; 0 < t < \theta.$$

$$f_\theta^T(t) = \frac{nt^{n/2-1}}{2\theta^{n/2}}, \quad 0 < t < \theta.$$

Completitud de $T = \text{máx}\, X_i$: Sea g medible tal que $E_\theta\big[g(T)\big] = 0, \;\forall\theta \in \mathbb{R}^+$:

$$E_\theta[g(T)] = \int_0^\theta g(t)\frac{nt^{n/2-1}}{2\theta^{n/2}}dt = 0, \quad \forall\theta > 0 \Rightarrow \int_0^\theta g(t)t^{n/2-1}dt = 0, \quad \forall\theta > 0$$

$$\overset{TFC}{\Rightarrow} g(t)t^{n/2-1} = 0, \quad \forall t > 0 \Rightarrow g(t) = 0, \quad \forall t > 0.$$

$$\Downarrow$$

$$P_\theta\big(g(T) = 0\big) \geq P_\theta(T > 0) = 1, \quad \forall\theta > 0.$$

Cálculo del UMVUE: Buscamos $h(T)$ insesgada en θ:

$$E_\theta\big[h(T)\big] = \int_0^\theta h(t)\frac{nt^{n/2-1}}{2\theta^{n/2}}dt = \theta,\ \forall\theta > 0 \Rightarrow \int_0^\theta \frac{h(t)nt^{n/2-1}}{2}dt = \theta^{n/2+1},\ \forall\theta > 0.$$

Puesto que $\lim\limits_{\theta\to 0}\theta^{n/2+1} = 0,\ \forall n \in \mathbb{N}$, tiene sentido trabajar con n arbitrario:

$$\frac{h(\theta)n\theta^{n/2-1}}{2} = \left(\theta^{n/2+1}\right)' = (n/2+1)\theta^{n/2} \Rightarrow h(\theta) = \frac{(n+2)\theta}{n},\quad \forall\theta > 0.$$

$$\Downarrow\ T > 0$$

$$h(T) = \frac{(n+2)}{n}T.$$

* $h(T)$ estimador de $\theta > 0$:

  $T > 0 \Rightarrow h(T) > 0.$

* $E_\theta\big[(h(T))^2\big] < +\infty,\ \forall\theta \in \mathbb{R}^+$. Debido a que $T \in (0,\theta)$, y por tanto $h(T)$ está acotada para cada θ.

Problema 6

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable X con función de densidad

$$f_\theta(x) = \frac{\theta}{x^2}, \quad x > \theta.$$

Calcular, si existen, los UMVUE para θ y para $1/\theta$.

Ya que no se especifica el espacio paramétrico, tomamos el más general posible, $\Theta = \mathbb{R}^+$.

$$\chi_\theta = (\theta, +\infty) \Rightarrow \chi = \bigcup_{\theta \in \mathbb{R}^+} (\theta, +\infty) = \mathbb{R}^+ \longrightarrow \chi^n = (\mathbb{R}^+)^n.$$

Estadístico suficiente:

$$f_\theta^n(x_1, \ldots, x_n) = \frac{1}{\prod_{i=1}^n x_i^2} \times \theta^n I_{(0,+\infty)}\big(\text{mín}\, x_i - \theta\big), \quad x_1, \ldots, x_n > 0.$$

$$\Downarrow$$

$$T = \text{mín}\, X_i \text{ suficiente.}$$

UMVUE para θ: Buscamos $h(T)$ insesgada en θ:

$$E_\theta[h(T)] = \int_\theta^{+\infty} h(t) \frac{n\theta^n}{t^{n+1}} dt = \theta, \quad \forall \theta > 0 \ \Rightarrow \ \int_\theta^{+\infty} \frac{nh(t)}{t^{n+1}} dt = \theta^{1-n}, \quad \forall \theta > 0.$$

Puesto que $\lim_{\theta \to +\infty} \theta^{1-n} = 0 \Leftrightarrow n \geq 2$, se deduce que *para $n = 1$ no existe estimador insesgado de θ.*

$$\boxed{n \geq 2} \quad -\frac{nh(\theta)}{\theta^{n+1}} = (\theta^{1-n})' = (1-n)\theta^{-n} \ \Rightarrow \ h(\theta) = \frac{(n-1)\theta}{n}, \quad \forall \theta > 0$$

$$\Downarrow \ T > 0$$

$$h(T) = \frac{(n-1)T}{n}.$$

$*$ $h(T)$ estimador de $\theta > 0$:

$T > 0 \Rightarrow h(T) > 0.$

$*$ $E_\theta[(h(T))^2] < +\infty \Leftrightarrow E_\theta[T^2] = \int_\theta^{+\infty} \frac{n\theta^n t^2}{t^{n+1}} dt < +\infty \Leftrightarrow n+1-2 > 1 \Leftrightarrow n > 2.$

- $n = 1 \rightarrow$ *no existe ningún estimador insesgado de θ.*
- $n = 2 \rightarrow$ *no existe ningún estimador insesgado de θ con varianza finita.*
- $n \geq 3 \rightarrow \frac{(n-1)T}{n}$ *es el UMVUE para θ.*

- *UMVUE para* $1/\theta$: Buscamos $h(T)$ insesgada en $1/\theta$:

$$E_\theta[h(T)] = \int_\theta^{+\infty} h(t)\frac{n\theta^n}{t^{n+1}}dt = \frac{1}{\theta}, \quad \forall\theta > 0 \Rightarrow \int_\theta^{+\infty} \frac{nh(t)}{t^{n+1}}dt = \theta^{-n-1}, \quad \forall\theta > 0.$$

Puesto que $\lim\limits_{\theta\to+\infty} \theta^{-n-1} = 0$, $\forall n \geq 1$, tiene sentido trabajar con n arbitrario:

$$-\frac{nh(\theta)}{\theta^{n+1}} = (\theta^{-n-1})' = -(n+1)\theta^{-n-2} \Rightarrow h(\theta) = \frac{(n+1)}{n\theta}, \quad \forall\theta > 0.$$

$$\Downarrow \; T > 0$$

$$h(T) = \frac{(n+1)}{nT}.$$

$*$ $h(T)$ estimador de $1/\theta > 0$:

$T > 0 \Rightarrow h(T) > 0.$

$*$ $E_\theta\left[(h(T))^2\right] = \left(\frac{n+1}{n}\right)^2 \int_\theta^{+\infty} \frac{n\theta^n}{t^{n+3}}dt < +\infty, \; \forall\theta > 0.$

UMVUE para $1/\theta \to \dfrac{(n+1)}{nT}$.

Problema 7

Sea $X \to \{P_\theta;\ \theta \in \mathbb{R}\}$ *siendo* P_θ *una distribución con función de densidad*

$$f_\theta(x) = e^{\theta - x}, \quad x \geq \theta.$$

Dada una muestra aleatoria simple de tamaño arbitrario, encontrar los UMVUE de θ *y de* e^θ.

$*$ *Estadístico suficiente y completo* $\longrightarrow T = \min X_i$ (R3-P6).

$*$ *Distribución de* T $\xrightarrow{[1]}$ $f_\theta^T(t) = ne^{n(\theta - t)}, \quad t > \theta$

- *UMVUE para θ*: Buscamos $h(T)$ insesgada en θ:

$$E_\theta[h(T)] = \int_\theta^{+\infty} h(t) n e^{n(\theta-t)} dt = \theta, \ \forall\theta \in \mathbb{R} \Rightarrow \int_\theta^{+\infty} n h(t) e^{-nt} dt = \theta e^{-n\theta}, \ \forall\theta \in \mathbb{R}.$$

Puesto que $\lim\limits_{\theta\to+\infty} \theta e^{-n\theta} = 0, \ \forall n \geq 1$, tiene sentido trabajar con n arbitrario:

$$-nh(\theta)e^{-n\theta} = \left(\theta e^{-n\theta}\right)' = e^{-n\theta} - n\theta e^{-n\theta} = (1-n\theta)e^{-n\theta} \Rightarrow h(\theta) = \frac{n\theta - 1}{n} = \theta - \frac{1}{n}, \ \forall\theta \in \mathbb{R}.$$

$$\Downarrow \ T \in \mathbb{R}$$

$$h(T) = T - \frac{1}{n}.$$

$*$ $h(T)$ estimador de $\theta \in \mathbb{R}$.

$*$ $E_\theta\left[(h(T))^2\right] < +\infty, \ \forall\theta \in \mathbb{R} \Leftrightarrow E_\theta\left[T^2\right] = ne^{\theta}\int_\theta^{+\infty} t^2 e^{-nt} dt < +\infty, \ \forall\theta \in \mathbb{R}.$

$\theta > 0 \longrightarrow \int_\theta^{+\infty} t^2 e^{-nt} dt < \int_0^{+\infty} t^2 e^{-nt} dt < +\infty$ (integral tipo Γ).

$\theta < 0 \longrightarrow \int_\theta^{+\infty} t^2 e^{-nt} dt = \int_\theta^{0} t^2 e^{-nt} dt + \int_0^{+\infty} t^2 e^{-nt} dt < +\infty.$

UMVUE para $\theta \to T - \frac{1}{n}$.

- *UMVUE para e^θ*: Buscamos $h(T)$ insesgada en e^θ:

$$E_\theta[h(T)] = \int_\theta^{+\infty} h(t) n e^{n(\theta-t)} dt = e^\theta, \ \forall\theta \in \mathbb{R} \Rightarrow \int_\theta^{+\infty} n h(t) e^{-nt} dt = e^{(1-n)\theta}, \ \forall\theta \in \mathbb{R}.$$

Puesto que $\lim\limits_{\theta\to+\infty} e^{(1-n)\theta} = 0 \Leftrightarrow n \geq 2$, se deduce que *para $n = 1$ no existe estimador insesgado de e^θ*.

$\boxed{n \geq 2}$ $\quad -nh(\theta)e^{-n\theta} = \left(e^{(1-n)\theta}\right)' = (1-n)e^{(1-n)\theta} \Rightarrow h(\theta) = \frac{(n-1)e^\theta}{n}, \ \forall\theta \in \mathbb{R}$

$$\Downarrow \ T \in \mathbb{R}$$

$$h(T) = \frac{(n-1)e^T}{n}.$$

$*$ $h(T) \in \mathbb{R}^+$ estimador de $e^\theta \in \mathbb{R}^+$.

$*$ $E_\theta\left[(h(T))^2\right] = \frac{(n-1)^2}{n} e^{n\theta} \int_\theta^{+\infty} e^{(2-n)t} dt < +\infty, \ \forall\theta \in \mathbb{R} \Leftrightarrow n > 2.$

- $n = 1 \to$ *no existe ningún estimador insesgado de e^θ.*
- $n = 2 \to$ *no existe ningún estimador insesgado de e^θ con varianza finita.*
- $n \geq 3 \to \frac{(n-1)e^T}{n}$ *es el UMVUE para e^θ.*

4.3. Estimadores eficientes

En muchas ocasiones, es posible determinar cotas inferiores para la varianza de estimadores insesgados de un parámetro (o función paramétrica), que pueden simplificar notablemente la determinación del UMVUE ya que, si la varianza de un estimador alcanza una determinada cota inferior, tal estimador será, obviamente, el de mínima varianza.

En los casos en que no existe el UMVUE, estas cotas pueden usarse para comparar con ellas la varianza de los estimadores insesgados y medir si la diferencia es significativa.

4.3.1. Familias regulares en el sentido de Fréchet-Cramér-Rao

La cota de Fréchet-Cramér-Rao sólo es aplicable en familias de distribuciones que satisfacen las condiciones que especificamos a continuación.

Condiciones de regularidad de Fréchet-Cramér-Rao

Sea $\{P_\theta;\ \theta \in \Theta\}$ una familia de distribuciones uniparamétrica con funciones de densidad (o funciones masa de probabilidad) $\{f_\theta;\ \theta \in \Theta\}$. Se dice que esta familia es regular en el sentido de Fréchet-Cramér-Rao si:

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

$ii)$ *El conjunto donde f_θ no se anula 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*

- $\displaystyle\sum_{x\in\chi} \frac{\partial f_\theta(x)}{\partial\theta} = \frac{\partial}{\partial\theta}\sum_{x\in\chi} f_\theta(x) = 0,\ \forall\theta \in \Theta$ *(caso discreto).*
- $\displaystyle\int_\chi \frac{\partial f_\theta(x)}{\partial\theta}dx = \frac{\partial}{\partial\theta}\int_\chi f_\theta(x)dx = 0,\ \forall\theta \in \Theta$ *(caso continuo).*

Nota: Las dos primeras condiciones son las mismas que las de familias exponenciales (Θ abierto para la derivabilidad); aquí, además, se exige que f_θ sea derivable respecto de θ, y que la derivada permute con la suma o integral en χ. Ya que dicha suma o integral vale uno, $iii)$ puede expresarse como:

$$iii) \sum_{x\in\chi} \frac{\partial f_\theta(x)}{\partial\theta} = 0, \ \forall\theta\in\Theta \ \text{ ó } \int_\chi \frac{\partial f_\theta(x)}{\partial\theta} dx = 0, \ \forall\theta\in\Theta.$$

La cota de Fréchet-Cramér-Rao está basada en la siguiente función, que especifica una medida de la cantidad de información que proporciona la variable sobre el parámetro.

Función de información de Fisher

Sea X es una variable aleatoria con distribución en una familia regular, $\{P_\theta;\ \theta\in\Theta\}$. Se definen la función de información asociada a X y la función de información asociada a una muestra aleatoria simple, como:

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

siempre que tales esperanzas sean finitas.

Propiedades de la función de información:

$a)$ $I_X \geq 0$ ($\forall\theta,\ I_X(\theta)$ es la esperanza de una variable no negativa).

$b)$ $E_\theta\left[\frac{\partial \ln f_\theta(X)}{\partial\theta}\right] = 0, \quad \forall\theta\in\Theta$ y $I_X(\theta) = Var_\theta\left[\frac{\partial \ln f_\theta(X)}{\partial\theta}\right], \quad \forall\theta\in\Theta.$

$$* \ E_\theta\left[\frac{\partial \ln f_\theta(X)}{\partial\theta}\right] = \int_\chi \frac{\partial f_\theta(x)}{\partial\theta}\frac{1}{f_\theta(x)} f_\theta(x)dx = \int_\chi \frac{\partial f_\theta(x)}{\partial\theta} dx \overset{C.R.\ iii)}{=} 0.$$

$$* \ Var_\theta\left[\frac{\partial \ln f_\theta(X)}{\partial\theta}\right] = E_\theta\left[\left(\frac{\partial \ln f_\theta(X)}{\partial\theta}\right)^2\right] = I_X(\theta). \qquad \square$$

$c)$ $E_\theta\left[\frac{\partial \ln f_\theta^n(X_1,\dots,X_n)}{\partial\theta}\right] = 0, \ I_{X_1,\dots,X_n}(\theta) = Var_\theta\left[\frac{\partial \ln f_\theta(X_1,\dots,X_n)}{\partial\theta}\right]$

$$* \ f_\theta^n(X_1,\dots,X_n) = \prod_{i=1}^n f_\theta(X_i) \Rightarrow \ln f_\theta^n(X_1,\dots,X_n) = \sum_{i=1}^n \ln f_\theta(X_i) \Rightarrow$$

$$\Rightarrow \frac{\partial \ln f_\theta^n(X_1,\dots,X_n)}{\partial\theta} = \sum_{i=1}^n \frac{\partial \ln f_\theta(X_i)}{\partial\theta} \Rightarrow$$

$$\Rightarrow E\left[\frac{\partial \ln f_\theta^n(X_1,\dots,X_n)}{\partial\theta}\right] = \sum_{i=1}^n E_\theta\left[\frac{\partial \ln f_\theta(X_i)}{\partial\theta}\right] \overset{b)}{=} 0.$$

Entonces, como en $b)$, el momento no centrado de orden dos, $I_{X_1,\dots,X_n}(\theta)$, coincide con la varianza.

d) *Aditividad*: $I_{X_1,\ldots,X_n}(\theta) = nI_X(\theta), \quad \forall \theta \in \Theta.$

$*$ Ya que $\dfrac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta} = \sum_{i=1}^n \dfrac{\partial \ln f_\theta(X_i)}{\partial\theta}$ es suma de variables independientes:

$$I_{X_1,\ldots,X_n}(\theta) \overset{c)}{=} Var_\theta\left[\frac{\partial \ln f_\theta(X_1,\ldots,X_n)}{\partial\theta}\right] \overset{Ind.}{=} \sum_{i=1}^n Var_\theta\left[\frac{\partial \ln f_\theta(X_i)}{\partial\theta}\right]$$
$$\overset{b)}{=} \sum_{i=1}^n I_{X_i}(\theta) = nI_X(\theta). \qquad \square$$

4.3.2. Cota de Fréchet-Cramér-Rao. Estimadores eficientes

Como hemos indicado, la cota es aplicable sólo en familias regulares pero, además, los estimadores a los que se aplica deben satisfacer también una condición de regularidad.

Estadístico regular

Sea $(X_1,\ldots,X_n)$ una muestra aleatoria simple de una variable X con distribución en una familia regular. Un estadístico $T(X_1,\ldots,X_n)$ es regular en el sentido de Fréchet-Cramér-Rao si:

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

o cambiando las integrales por sumas y la función de densidad por la función masa de probabilidad en el caso de distribuciones discretas.

Nota: La integral del primer miembro es $E_\theta[T(X_1,\ldots,X_n)]$, de forma que la regularidad de un estadístico lleva implícito que exista su esperanza, y que ésta sea derivable respecto de θ. Por lo tanto, nos centramos sólo en la estimación de funciones paramétricas derivables.

Por otra parte, la integral del segundo miembro es:

$$\int_{\mathcal{X}^n} T(x_1,\ldots,x_n)\frac{\partial f_\theta^n(x_1,\ldots,x_n)}{\partial\theta}\frac{1}{f_\theta^n(x_1,\ldots,x_n)}f_\theta^n(x_1,\ldots,x_n)dx_1\cdots dx_n$$
$$=\int_{\mathcal{X}^n} T(x_1,\ldots,x_n)\frac{\partial \ln f_\theta^n(x_1,\ldots,x_n)}{\partial\theta}f_\theta^n(x_1,\ldots,x_n)dx_1\cdots dx_n$$
$$=E_\theta\left[T(X_1,\ldots,X_n)\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta}\right].$$

Por tanto, la condición de regularidad se expresa como:

$$E_\theta\left[T(X_1,\ldots,X_n)\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta}\right]=\frac{\partial E_\theta\left[T(X_1,\ldots,X_n)\right]}{\partial\theta},\ \forall\theta\in\Theta.$$

Cota de Fréchet-Cramér-Rao

Sea $(X_1,\ldots,X_n)$ una muestra aleatoria simple de una variable X con distribución en una familia regular, $\{P_\theta;\ \theta\in\Theta\}$, con $0<I_X(\theta)<+\infty$, $\forall\theta\in\Theta$. Si $T\equiv T(X_1,\ldots,X_n)$ es un estadístico regular, de segundo orden, insesgado en una función paramétrica derivable $g(\theta)$, se tiene:

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

b) *Si* $g'(\theta)\neq 0$:

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

1

[1]Desigualdad de Cauchy-Schwartz:

$i)\ E\left[T^2\right]<+\infty,\ E\left[V^2\right]<+\infty\Rightarrow\left(Cov\left[T,V\right]\right)^2\leq Var\left[T\right]Var\left[V\right].$

$ii)$ *Si T y V son no degeneradas :*

$$\left(Cov\left[T,V\right]\right)^2=Var\left[T\right]Var\left[V\right]\Leftrightarrow\exists a\neq 0,\ b\in\mathbb{R}\ /\ P(V=aT+b)=1.$$

Demostración: Para cada distribución de la familia (para cada $\theta \in \Theta$), se aplica la desigualdad de Cauchy-Schwartz a las variables

$$T \equiv T(X_1, \ldots, X_n); \qquad V_\theta \equiv \frac{\partial \ln f_\theta(X_1, \ldots, X_n)}{\partial \theta}.$$

- $E_\theta[T] = g(\theta), \quad E_\theta[V_\theta] = 0, \quad Var_\theta[V_\theta] = I_{X_1,\ldots,X_n}(\theta).$
- $Cov_\theta[T, V_\theta] = E_\theta[TV_\theta] \overset{T \text{ regular}}{=} g'(\theta).$

$a)$ DCS $\rightarrow (g'(\theta))^2 \leq Var_\theta[T] I_{X_1,\ldots,X_n}(\theta)$, y se divide por $I_{X_1,\ldots,X_n}(\theta) > 0$.

$b)$ Si $g'(\theta) \neq 0$, las variables T y V_θ son no degeneradas bajo P_θ:

- $Var_\theta[T] > 0$ por el apartado $a)$, ya que $(g'(\theta))^2 > 0$.
- $Var_\theta[V_\theta] = I_{X_1,\ldots,X_n}(\theta) > 0$.

Entonces, se da la igualdad en $a)$ si y sólo si

$$\exists a(\theta) \neq 0 \text{ y } b(\theta) \in \mathbb{R} \; / \; P_\theta\big(V_\theta = a(\theta)T + b(\theta)\big) = 1.$$

Ahora, tomando esperanzas en esta expresión de V_θ:

$0 = E_\theta[V_\theta] = a(\theta)E_\theta[T] + b(\theta) = a(\theta)g(\theta) + b(\theta) \;\rightarrow\; b(\theta) = -a(\theta)g(\theta).$

Por tanto, se de la igualdad en $a)$ si y sólo si

$$\exists a(\theta) \neq 0 \; / \; P_\theta\big(V_\theta = a(\theta)[T - g(\theta)]\big) = 1.$$

Esta cota proporciona, para $g(\theta)$ derivable, la varianza mínima de los estimadores insesgados que sean regulares. Por tanto, aquel cuya varianza la alcance para todo $\theta \in \Theta$, será el óptimo en esta clase de estimadores.

Estimador eficiente

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable X con distribución en una familia regular, $\{P_\theta;\; \theta \in \Theta\}$, con $0 < I_X(\theta) < +\infty$, $\forall \theta \in \Theta$, y sea $g(\theta)$ una función derivable.

Un estimador de $g(\theta)$, $T(X_1, \ldots, X_n)$, es eficiente si es insesgado, regular, y su varianza alcanza la cota para cualquier valor del parámetro:

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

Nota: Si $g(\theta)$ admite estimador eficiente, T, y $\exists \theta_0 \; / g'(\theta_0) = 0$, se tiene que $Var_{\theta_0}[T] = 0$; esto es, $\exists c \in \mathbb{R} \; / P_{\theta_0}(T = c) = 1$. Ya que $\{P_\theta;\; \theta \in \Theta\}$ es regular, el conjunto de valores de la muestra y, por tanto, el de T, no depende de θ; consecuentemente, $P_\theta(T = c) = 1, \; \forall \theta \in \Theta$. Entonces, por la insesgadez, $g(\theta) = E_\theta[T] = c, \; \forall \theta \in \Theta$. Por tanto, *excluido el caso trivial de $g(\theta)$ constante, si $g(\theta)$ admite un estimador eficiente, $g'(\theta) \neq 0, \; \forall \theta \in \Theta$.*

Caracterización de estimadores eficientes

Sea $(X_1,\ldots,X_n)$ una muestra aleatoria simple de una variable X con distribución en una familia regular, $\{P_\theta;\ \theta\in\Theta\}$, con $0<I_X(\theta)<+\infty$, $\forall\theta\in\Theta$. Sea $g(\theta)$ una función paramétrica derivable, no constante, y $T\equiv T(X_1,\ldots,X_n)$ un estimador de $g(\theta)$.

Una condición necesaria y suficiente para que T sea eficiente es:

$\forall\theta\in\Theta\ \ \exists a(\theta)\neq 0$ *tal que* :

$$i)\ P_\theta\Big(\frac{\partial\ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta}=a(\theta)\big[T(X_1,\ldots,X_n)-g(\theta)\big]\Big)=1.$$

$$ii)\ I_{X_1,\ldots,X_n}(\theta)=a(\theta)g'(\theta).$$

Demostración: Notando $V_\theta=\dfrac{\partial\ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta}$, de la relación lineal $i)$, se deduce que $a(\theta)$ es el coeficiente de regresión lineal de V_θ sobre T. Esto es, $\forall\theta\in\Theta$, $a(\theta)=Cov_\theta[T,V_\theta]/Var_\theta[T]$. Entonces, $ii)$ se puede especificar como:

$$Var_\theta[V_\theta]=\frac{Cov_\theta[T,V_\theta]}{Var_\theta[T]}g'(\theta)\ \to\ ii')\ Cov_\theta[T,V_\theta]=\frac{Var_\theta[T]Var_\theta[V_\theta]}{g'(\theta)},\ \ \forall\theta\in\Theta.$$

De esta forma, $i),\ ii)\Leftrightarrow i),\ ii')$.

$\Rightarrow)$ Si T es eficiente para $g(\theta)$, su varianza alcanza la *cota de Fréchet-Cramér-Rao* y, como $g'(\theta)\neq 0$, por el apartado $b)$ se tiene la relación lineal $i)$. La igualdad $ii')$ se deduce como sigue:

- $Cov_\theta[T,V_\theta]\overset{E_\theta[V_\theta]=0}{=}E_\theta[TV_\theta]\overset{\text{T regular}}{=}g'(\theta)$.
- $\dfrac{Var_\theta[T]Var_\theta[V_\theta]}{g'(\theta)}\overset{\text{Cota}}{=}\dfrac{(g'(\theta))^2}{I_{X_1,\ldots,X_n}(\theta)}\dfrac{Var_\theta[V_\theta]}{g'(\theta)}\overset{Var_\theta[V_\theta]=I(\theta)}{=}g'(\theta)$.

$\Leftarrow)$ Partimos ahora de $i)$, $ii')$ y probamos las condiciones de eficiencia:

- *T insesgado*:
  $0=E_\theta[V_\theta]\overset{i)}{=}a(\theta)\big[E_\theta[T]-g(\theta)\big],\ \forall\theta\in\Theta\overset{a(\theta)\neq0}{\Rightarrow}E_\theta[T]=g(\theta),\ \forall\theta\in\Theta.$
- *T de segundo orden*:
  $E_\theta[V_\theta^2]=I_{X_1,\ldots,X_n}(\theta)<+\infty,\ \forall\theta\in\Theta\overset{i)}{\Rightarrow}E_\theta[T^2]<+\infty,\ \forall\theta\in\Theta.$
- *T regular*:
  $(Cov_\theta[T,V_\theta])^2\overset{i)}{=}Var_\theta[V_\theta]Var_\theta[T]\overset{ii')}{=}Cov_\theta[T,V_\theta]\,g'(\theta).$
  De aquí se deduce que $E_\theta[TV_\theta]=Cov_\theta[T,V_\theta]=g'(\theta)$, y T es regular.

Entonces, por el apartado $b)$ de la cota, la relación lineal $i)$ implica que esta se alcanza para todo valor del parámetro y, por tanto, que T es eficiente.

Nota: Como se ha visto en la demostración, la relación lineal $i)$ especificada en la caracterización, junto con $ii)$, llevan implícito que T es insesgado, de segundo orden y regular. Por lo tanto, esta caracterización proporciona un camino simple para obtener estimadores eficientes; sólo hay que conseguir esta expresión, y comprobar que el estadístico $T(X_1, \ldots, X_n)$ que aparece es estimador de $g(\theta)$.

Tema 4.Estimación puntual. Insesgadez y mínima varianza A. Hermoso Carazo
└4.3. Estimadores eficientes
└4.3.2. Cota de Fréchet-Cramér-Rao. Estimadores eficientes

Propiedades de los estimadores eficientes:

- *Si $T \equiv T(X_1, \ldots, X_n)$ es eficiente para $g(\theta)$, $aT + b$ es eficiente para $ag(\theta) + b$ $(a \neq 0, b \in \mathbb{R})$ y sólo este tipo de funciones paramétricas admite estimador eficiente.*

Demostración: Si T es estimador de $g(\theta)$, $aT + b$ lo es de $ag(\theta) + b$. Además, teniendo en cuenta $i)$ y $ii)$ de la caracterización para T y g, obtenemos la correspondiente para $aT + b$ y $ag + b$ como sigue:

$$i)\ V_\theta = \frac{\partial \ln f_\theta^n(X_1, \ldots, X_n)}{\partial \theta} = a(\theta)\big[T - g(\theta)\big] = \frac{a(\theta)}{a}\big[(aT + b) - (ag(\theta) + b)\big].$$

$$ii)\ \frac{a(\theta)}{a}\big(ag(\theta) + b\big)' = a(\theta)g'(\theta) = I_{X_1,\ldots,X_n}(\theta).$$

Por otra parte, sea $g_1(\theta)$ otra función con estimador eficiente, T_1. Bajo cualquier distribución de la familia, V_θ es función lineal de T, con probabilidad uno, y también, con probabilidad uno, función lineal de T_1. Por tanto, existe una relación lineal entre T y T_1:

$$\forall \theta \in \Theta, \quad \exists a(\theta) \neq 0, \ b(\theta) \ / \ P_\theta\big(T_1 = a(\theta)T + b(\theta)\big) = 1.$$

Pero como T_1 es independiente de θ, $a(\theta)$ y $b(\theta)$ deben serlo también, por lo que $P_\theta\big(T_1 = aT + b\big) = 1, \quad \forall \theta \in \Theta$. Tomando esperanzas, se deduce que g_1 es función lineal de g, $g_1(\theta) = ag(\theta) + b, \ \forall \theta \in \Theta$. □.

- *Si existe estimador eficiente para una función paramétrica, es único.*

  Demostración: Si T y T' son eficientes para $g(\theta)$, según la propiedad anterior, $P_\theta(T' = aT + b) = 1, \ \forall\theta \in \Theta$ y la insesgadez de ambos obliga a que $g(\theta) = ag(\theta) + b, \ \forall\theta \in \Theta$. Por lo tanto, $a = 1, \ b = 0$, lo que implica que $P_\theta(T' = T) = 1, \ \forall\theta \in \Theta$. □

- *Sólo existen estimadores eficientes en familias de tipo exponencial.*

  Demostración: Sea T eficiente para $g(\theta)$. Especificando la relación lineal $i)$ de la caracterización en cada realización muestral, ($\mathbf{x} = (x_1, \ldots, x_n) \in \chi^n$), e integrando respecto de θ se deduce que:
  $$\begin{aligned}\ln f_\theta^n(\mathbf{x}) &= \int a(\theta)T(\mathbf{x})d\theta - \int a(\theta)g(\theta)d\theta + S(\mathbf{x})\\ &= Q(\theta)T(\mathbf{x}) + D(\theta) + S(\mathbf{x}), \quad \mathbf{x} \in \chi^n,\end{aligned}$$
  donde $Q(\theta) = \int a(\theta)d\theta, \ \ D(\theta) = -\int a(\theta)g(\theta)d\theta$ y $S(\mathbf{x})$ es la constante de integración. Por tanto:
  $$f_\theta^n(\mathbf{x}) = \exp\left\{Q(\theta)T(\mathbf{x}) + D(\theta) + S(\mathbf{x})\right\}, \quad \mathbf{x} \in \chi^n,$$
  lo que, junto con las condiciones de regularidad $i)$ y $ii)$ implica que la familia es de tipo exponencial (uniparamétrica).

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

  Demostración: La suficiencia y completitud se deducen del *Teorema de suficiencia y completitud en familias exponenciales*, según la expresión de $f_\theta^n(\mathbf{x})$ obtenida en la propiedad anterior. Además, como T es insesgado en $g(\theta)$ y de segundo orden, si es suficiente y completo, por los *teoremas de Rao-Blackwell y Lehmann-Sheffé* es el UMVUE.

Por lo tanto, la búsqueda de estimadores eficientes, cuando tenga sentido, proporciona un camino alternativo para la determinación de UMVUE (si T es estimador eficiente y es completo, es el UMVUE).

Sin embargo, incluso bajo condiciones de regularidad que permitan hablar de eficiencia, un UMVUE no tiene por qué ser eficiente, ya que su varianza, a pesar de ser la mínima, puede ser mayor que la cota.

Ejemplo 4.3.1: *Probar que la familia $\{B(k_0, p);\ p \in (0,1)\}$ es regular. Calcular la función de información asociada a una muestra aleatoria simple de una variable con distribución en dicha familia. Encontrar la clase de funciones paramétricas con estimador eficiente. Calcular la varianza de estos estimadores y comprobar que coincide con la cota correspondiente.*

Se trata de una familia de tipo discreto con funciones masa de probabilidad:

$$f_p(x) = P_p(X = x) = \binom{k_0}{x} p^x (1-p)^{k_0 - x}, \quad x = 0, 1, \ldots, k_0.$$

Condiciones de regularidad:

$i)$ $p \in (0,1)$ (intervalo abierto de $\mathbb{R}$).

$ii)$ $\{x \ /\ f_p(x) > 0\} = \{0, 1, \ldots, k_0\}$ (independiente de p).

$iii)$ $\forall x \in \chi,\ f_p(x)$ es derivable respecto de p y la suma de las derivadas es nula:

$$f_p(x) = \exp\left\{ \ln\binom{k_0}{x} + x \ln p + (k_0 - x)\ln(1-p) \right\}, \quad x = 0, 1, \ldots, k_0.$$

- $\dfrac{\partial f_p(x)}{\partial p} = \left[\dfrac{x}{p} - \dfrac{k_0 - x}{1-p}\right] f_p(x) = \dfrac{x - k_0 p}{p(1-p)} f_p(x), \quad x = 0, 1, \ldots, k_0.$
- $\displaystyle\sum_{x=0}^{k_0} \frac{\partial f_p(x)}{\partial p} = \sum_{x=0}^{k_0} \frac{x - k_0 p}{p(1-p)} f_p(x) = \frac{E_p[X - k_0 p]}{p(1-p)} = 0, \quad \forall p \in (0,1).$

4.3. Estimadores eficientes
4.3.2. Cota de Fréchet-Cramér-Rao. Estimadores eficientes

Función de información asociada a $(X_1, \ldots, X_n)$

$$I_{X_1,\ldots,X_n}(p) = n I_X(p) = n Var_p\left[\frac{\partial \ln f_p(X)}{\partial p}\right].$$

Partimos de $\dfrac{\partial f_p(x)}{\partial p} = \dfrac{x - k_0 p}{p(1-p)} f_p(x)$.

- $\dfrac{\partial \ln f_p(x)}{\partial p} = \dfrac{\dfrac{\partial f_p(x)}{\partial p}}{f_p(x)} = \dfrac{x - k_0 p}{p(1-p)}.$
- $I_X(p) = \dfrac{Var_p[X - k_0 p]}{p^2(1-p)^2} = \dfrac{Var_p[X]}{p^2(1-p)^2} = \dfrac{k_0 p(1-p)}{p^2(1-p)^2} = \dfrac{k_0}{p(1-p)}.$

$$I_{X_1,\ldots,X_n}(p) = \frac{n k_0}{p(1-p)}, \quad \forall p \in (0,1).$$

Antes del siguiente apartado, referido a la búsqueda de estimadores eficientes, recordemos que esta familia es de tipo exponencial (R3-P7) y, por tanto, tiene sentido, en principio, hablar de eficiencia. Además el estadístico $\sum_{i=1}^{n} X_i$ es suficiente y completo.

Clase de funciones paramétricas con estimador eficiente

Aplicamos el *Teorema de Caracterización*.

$$\frac{\partial \ln f_p^n(x_1,\ldots,x_n)}{\partial p} = \sum_{i=1}^n \frac{\partial \ln f_p(x_i)}{\partial p} = \sum_{i=1}^n \frac{x_i - k_0 p}{p(1-p)} = \frac{1}{p(1-p)}\left[\sum_{i=1}^n x_i - nk_0 p\right]$$
$$= \frac{n}{p(1-p)}\left[\overline{x} - k_0 p\right].$$

Comparando esta expresión con $i)$ de la caracterización tenemos:

$$a(p) = \frac{n}{p(1-p)}, \qquad T(X_1,\ldots,X_n) = \overline{X}, \qquad g(p) = k_0 p.$$

- $I_{X_1,\ldots,X_n}(p) = \dfrac{nk_0}{p(1-p)} = a(p)g'(p)$ (se cumple $ii)$ del T.C.).
- $T = \overline{X} \in [0, k_0]$ es estimador de $g(p) = k_0 p \in (0, k_0)$.

$$\Downarrow$$

$\forall a \neq 0,\ b \in \mathbb{R},\ a\overline{X} + b$ *es el estimador eficiente de* $ak_0 p + b$,

y sólo estas funciones (lineales de p*) tienen estimador eficiente.*

Nota: Como estos estimadores son función de $\sum_{i=1}^n X_i$, suficiente y completo, cada estimador eficiente es el UMVUE para la correspondiente función de p.

Varianza de los estimadores eficientes y comprobación de la cota

- $Var_p\left[a\overline{X}+b\right] = a^2 Var_p\left[\overline{X}\right] = a^2\dfrac{Var_p[X]}{n} = \dfrac{a^2 k_0 p(1-p)}{n}, \quad p \in (0,1).$
- Cota para la varianza de estimadores de $ak_0p + b$:

$$\frac{\left((ak_0p+b)'\right)^2}{I_{X_1,\ldots,X_n}(p)} = \frac{(ak_0)^2}{nk_0/p(1-p)}, \quad p \in (0,1).$$

Por tanto, queda comprobado que $Var_p\left[a\overline{X}+b\right]$ coincide con la cota correspondiente.

Ejemplo 4.3.2: *Probar que* $\{\mathcal{N}(\mu, \sigma_0^2);\ \mu \in \mathbb{R}\}$ *es una familia regular. Calcular la función de información asociada a una muestra aleatoria simple de una variable con distribución en dicha familia. Encontrar la clase de funciones paramétricas con estimador eficiente y calcular la varianza de cada uno de estos estimadores.*

Tema 4.Estimación puntual. Insesgadez y mínima varianza A. Hermoso Carazo
└4.3. Estimadores eficientes
└4.3.2. Cota de Fréchet-Cramér-Rao. Estimadores eficientes

Clase de funciones paramétricas con estimador eficiente

$$\frac{\partial \ln f_\mu^n(x_1,\ldots,x_n)}{\partial\mu} = \sum_{i=1}^n \frac{\partial \ln f_\mu(x_i)}{\partial\mu} = \frac{1}{\sigma_0^2}\sum_{i=1}^n (x_i-\mu) = \frac{1}{\sigma_0^2}\left[\sum_{i=1}^n x_i - n\mu\right]$$
$$= \frac{n}{\sigma_0^2}\left[\overline{x}-\mu\right], \quad (x_1,\ldots,x_n)\in\mathbb{R}.$$

$$a(\mu) = \frac{n}{\sigma_0^2}, \quad T(X_1,\ldots,X_n) = \overline{X}, \quad g(\mu)=\mu. \quad \text{(se cumple } i) \text{ del T.C.)}$$

- $I_{X_1,\ldots,X_n}(\mu) = \dfrac{n}{\sigma_0^2} = a(\mu)g'(\mu)$ (se cumple ii) del T.C.)
- $T=\overline{X}\in\mathbb{R}$ es estimador de $g(\mu)=\mu\in\mathbb{R}$.

$$\Downarrow$$

$\forall a \neq 0,\ b\in\mathbb{R},\ a\overline{X}+b$ *es el estimador eficiente de* $a\mu+b$,

y sólo estas funciones (funciones lineales de μ*) tienen estimador eficiente.*

Nota: Todos los estimadores eficientes son función de $\sum_{i=1}^n X_i$, suficiente y completo. Por tanto, cada uno es el UMVUE para la función paramétrica correspondiente.

Ejemplo 4.3.3: *Probar que la familia $\{\mathcal{N}(\mu_0, \sigma^2);\ \sigma^2 \in \mathbb{R}^+\}$ es regular. Calcular la función de información asociada a una muestra aleatoria simple de una variable con distribución en dicha familia. Encontrar la clase de funciones paramétricas con estimador eficiente y calcular la varianza de cada uno de estos estimadores.*

$$f_{\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu_0)^2/2\sigma^2}, \quad x \in \mathbb{R}.$$

Condiciones de regularidad:

i) $\sigma^2 \in \mathbb{R}^+ = (0, +\infty)$ (intervalo abierto de $\mathbb{R}$).

ii) $\{x \ /\ f_{\sigma^2}(x) > 0\} = \mathbb{R}$ (independiente de σ^2).

iii) $\forall x \in \mathbb{R},\ f_{\sigma^2}(x)$ es derivable respecto de σ^2 y la integral de la derivada es nula:

$$f_{\sigma^2}(x) = \exp\left\{-\ln(\sqrt{2\pi}) - \ln(\sigma^2)/2 - \frac{(x-\mu_0)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

- $\dfrac{\partial f_{\sigma^2}(x)}{\partial \sigma^2} = \dfrac{1}{2\sigma^2}\left[\dfrac{(x-\mu_0)^2}{\sigma^2} - 1\right] f_{\sigma^2}(x), \quad x \in \mathbb{R}.$
- $\displaystyle\int_{\mathbb{R}} \frac{\partial f_{\sigma^2}(x)}{\partial \sigma^2} dx = \frac{1}{2\sigma^2}\int_{\mathbb{R}} \left(\frac{(x-\mu_0)^2}{\sigma^2} - 1\right) f_{\sigma^2}(x) dx = \frac{1}{2\sigma^2} E_{\sigma^2}\left[\frac{(X-\mu_0)^2}{\sigma^2} - 1\right] = 0.$

Función de información asociada a $(X_1, \ldots, X_n)$

$$I_{X_1,\ldots,X_n}(\sigma^2) = nI_X(\sigma^2) = nVar_{\sigma^2}\left[\frac{\partial \ln f_{\sigma^2}(X)}{\partial \sigma^2}\right].$$

- $\dfrac{\partial f_{\sigma^2}(x)}{\partial \sigma^2} = \dfrac{1}{2\sigma^2}\left[\dfrac{(x-\mu_0)^2}{\sigma^2} - 1\right] f_{\sigma^2}(x), \quad x \in \mathbb{R}.$
- $\dfrac{\partial \ln f_{\sigma^2}(x)}{\partial \sigma^2} = \dfrac{\dfrac{\partial f_{\sigma^2}(x)}{\partial \sigma^2}}{f_{\sigma^2}(x)}$

Ejemplo 4.3.3: *Probar que la familia $\{\mathcal{N}(\mu_0, \sigma^2);\ \sigma^2 \in \mathbb{R}^+\}$ es regular. Calcular la función de información asociada a una muestra aleatoria simple de una variable con distribución en dicha familia. Encontrar la clase de funciones paramétricas con estimador eficiente y calcular la varianza de cada uno de estos estimadores.*

$$f_{\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu_0)^2/2\sigma^2}, \quad x \in \mathbb{R}.$$

Condiciones de regularidad:

$i)$ $\sigma^2 \in \mathbb{R}^+ = (0, +\infty)$ (intervalo abierto de $\mathbb{R}$).

$ii)$ $\{x \ /\ f_{\sigma^2}(x) > 0\} = \mathbb{R}$ (independiente de σ^2).

$iii)$ $\forall x \in \mathbb{R}$, $f_{\sigma^2}(x)$ es derivable respecto de σ^2 y la integral de la derivada es nula:

$$f_{\sigma^2}(x) = \exp\left\{-\ln(\sqrt{2\pi}) - \ln(\sigma^2)/2 - \frac{(x-\mu_0)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

- $\dfrac{\partial f_{\sigma^2}(x)}{\partial \sigma^2} = \dfrac{1}{2\sigma^2}\left[\dfrac{(x-\mu_0)^2}{\sigma^2} - 1\right] f_{\sigma^2}(x), \quad x \in \mathbb{R}.$
- $\displaystyle\int_{\mathbb{R}} \frac{\partial f_{\sigma^2}(x)}{\partial \sigma^2} dx = \frac{1}{2\sigma^2}\int_{\mathbb{R}}\left(\frac{(x-\mu_0)^2}{\sigma^2} - 1\right) f_{\sigma^2}(x)dx = \frac{1}{2\sigma^2} E_{\sigma^2}\left[\frac{(X-\mu_0)^2}{\sigma^2} - 1\right] = 0.$

Función de información asociada a $(X_1, \ldots, X_n)$

$$I_{X_1,\ldots,X_n}(\sigma^2) = nI_X(\sigma^2) = nVar_{\sigma^2}\left[\frac{\partial \ln f_{\sigma^2}(X)}{\partial \sigma^2}\right].$$

- $\dfrac{\partial f_{\sigma^2}(x)}{\partial \sigma^2} = \dfrac{1}{2\sigma^2}\left[\dfrac{(x-\mu_0)^2}{\sigma^2} - 1\right] f_{\sigma^2}(x), \quad x \in \mathbb{R}.$
- $\dfrac{\partial \ln f_{\sigma^2}(x)}{\partial \sigma^2} = \dfrac{\frac{\partial f_{\sigma^2}(x)}{\partial \sigma^2}}{f_{\sigma^2}(x)} = \dfrac{1}{2\sigma^2}\left[\left(\dfrac{x-\mu_0}{\sigma}\right)^2 - 1\right], \quad x \in \mathbb{R}.$
- $I_X(\sigma^2) = \dfrac{1}{4\sigma^4} Var_{\sigma^2}\left[\left(\dfrac{X-\mu_0}{\sigma}\right)^2\right] \left[\begin{array}{l} Y = \left(\dfrac{X-\mu_0}{\sigma}\right)^2 \to \chi^2(1) \\ Var_{\sigma^2}[Y] = 2 \end{array}\right] = \dfrac{1}{2\sigma^4}.$

$$I_{X_1,\ldots,X_n}(\sigma^2) = \frac{n}{2\sigma^4}, \quad \forall \sigma^2 \in \mathbb{R}^+.$$

En el Ejemplo 3.3.2 vimos que esta familia es exponencial (tiene sentido estudiar eficiencia), y que $\sum_{i=1}^{n}(X_i - \mu_0)^2$ es suficiente y completo.

Clase de funciones paramétricas con estimador eficiente

$$\frac{\partial \ln f^n_{\sigma^2}(x_1,\dots,x_n)}{\partial \sigma^2} = \sum_{i=1}^{n} \frac{\partial \ln f_{\sigma^2}(x_i)}{\partial \sigma^2} = \frac{1}{2\sigma^2}\left[\frac{\sum\limits_{i=1}^{n}(x_i-\mu_0)^2}{\sigma^2} - n\right]$$

$$= \frac{n}{2\sigma^4}\left[\frac{\sum\limits_{i=1}^{n}(x_i-\mu_0)^2}{n} - \sigma^2\right], \quad (x_1,\dots,x_n)\in\mathbb{R}.$$

$$a(\sigma^2)=\frac{n}{2\sigma^4}, \quad T(X_1,\dots,X_n)=\frac{\sum\limits_{i=1}^{n}(X_i-\mu_0)^2}{n}, \quad g(\sigma^2)=\sigma^2. \quad (i) \text{ del T.C.})$$

- $I_{X_1,\dots,X_n}(\sigma^2) = \frac{n}{2\sigma^4} = a(\sigma^2)g'(\sigma^2)$ (se cumple ii) del T.C.)
- $T\in\mathbb{R}^+$ es estimador de $g(\sigma^2)=\sigma^2\in\mathbb{R}^+$.

$$\Downarrow$$

$\forall a\neq 0,\ b\in\mathbb{R},\ aT+b$ es el estimador eficiente de $a\sigma^2+b$,

y sólo estas funciones (funciones lineales de σ^2) tienen estimador eficiente.

Nota: Todos los estimadores eficientes son función del suficiente y completo, $\sum\limits_{i=1}^{n}(X_i-\mu_0)^2$, y por tanto, todos son UMVUE para la correspondiente función.

Varianza de los estimadores eficientes

Coinciden con las cotas correspondientes:

$$Var_{\sigma^2}[aT+b] = \frac{((a\sigma^2+b)')^2}{I_{X_1,\dots,X_n}(\sigma^2)} = \frac{2a^2\sigma^4}{n}, \quad \forall\sigma^2\in\mathbb{R}^+.$$

Comprobación:

$$Y=\frac{(X-\mu_0)^2}{\sigma^2}\to\chi^2(1)\Rightarrow Var_{\sigma^2}[Y]=\frac{1}{\sigma^4}Var_{\sigma^2}\left[(X-\mu_0)^2\right]=2.$$

$$\Rightarrow Var_{\sigma^2}\left[(X-\mu_0)^2\right]=2\sigma^4.$$

$$Var_{\sigma^2}[aT+b]=a^2Var_{\sigma^2}[T]=a^2Var_{\sigma^2}\left[\frac{\sum\limits_{i=1}^{n}(X_i-\mu_0)^2}{n}\right]$$

$$=\frac{a^2}{n^2}Var_{\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2\right]=\frac{a^2}{n}Var_{\sigma^2}\left[(X-\mu_0)^2\right]=\frac{2a^2\sigma^4}{n}.$$

Problema 8

Sea X la variable que describe el número de fracasos antes del primer éxito en una sucesión de pruebas de Bernoulli independientes con probabilidad de éxito $\theta \in (0,1)$, y sea $(X_1,\dots,X_n)$ una muestra aleatoria simple de X.

$a)$ *Probar que la familia de distribuciones de X es regular y calcular la función de información asociada a la muestra.*

$b)$ *Especificar la clase de funciones paramétricas que admiten estimadores eficientes y los correspondientes estimadores.*

$c)$ *Calcular la varianza de cada estimador eficiente y comprobar que coincide con la correspondiente cota de Fréchet-Cramér-Rao.*

$d)$ *Calcular, si existen, los UMVUE para $P_\theta(X=0)$ y para $E_\theta[X]$ y decir si son eficientes.*

$$X \to G(\theta)$$

$$f_\theta(x) = P_\theta(X=x) = (1-\theta)^x\theta, \quad x \in \mathbb{N}\cup\{0\} \left\{ \begin{array}{l} E_\theta[X] = \dfrac{1-\theta}{\theta} \\ Var_\theta[X] = \dfrac{1-\theta}{\theta^2} \end{array} \right.$$

a) *Probar que la familia de distribuciones de X es regular y calcular la función de información asociada a la muestra.*

$i)$ $\Theta = (0,1) \to$ intervalo abierto de $\mathbb{R}$.

$ii)$ $\{x \,/\, f_\theta(x) > 0\} = \mathbb{N}\cup\{0\} \to$ independiente de θ.

$iii)$ $\forall x \in \mathbb{N}\cup\{0\}$, $f_\theta(x)$ es derivable respecto de θ y la suma de las derivadas es nula:

- $f_\theta(x) = \exp\{x\ln(1-\theta) + \ln\theta\}$.
- $\dfrac{\partial f_\theta(x)}{\partial\theta} = \left[-\dfrac{x}{1-\theta} + \dfrac{1}{\theta}\right] f_\theta(x) = -\dfrac{1}{1-\theta}\left[x - \dfrac{1-\theta}{\theta}\right] f_\theta(x)$.
- $\displaystyle\sum_{x=0}^{+\infty} \frac{\partial f_\theta(x)}{\partial\theta} = -\frac{1}{1-\theta}\left[E_\theta[X] - \frac{1-\theta}{\theta}\right] = 0, \quad \forall\theta\in(0,1).$

Función de información:

$$I_{X_1,\dots,X_n}(\theta) = nI_X(\theta) = nVar_\theta\left[\frac{\partial \ln f_\theta(X)}{\partial\theta}\right], \quad \theta\in(0,1).$$

$$\frac{\partial \ln f_\theta(x)}{\partial\theta} = \frac{\dfrac{\partial f_\theta(x)}{\partial\theta}}{f_\theta(x)} = -\frac{1}{1-\theta}\left[x - \frac{1-\theta}{\theta}\right] \Rightarrow I_X(\theta) = \frac{Var_\theta[X]}{(1-\theta)^2} = \frac{1}{\theta^2(1-\theta)}.$$

$$I_{X_1,\dots,X_n}(\theta) = \frac{n}{\theta^2(1-\theta)}, \quad \forall\theta\in(0,1).$$

b) *Especificar la clase de funciones paramétricas que admiten estimadores eficientes y los correspondientes estimadores.*

$$\frac{\partial \ln f_\theta^n(x_1,\ldots,x_n)}{\partial\theta} = \sum_{i=1}^n \frac{\partial \ln f_\theta(x_i)}{\partial\theta} = -\frac{1}{1-\theta}\left[\sum_{i=1}^n x_i - \frac{n(1-\theta)}{\theta}\right]$$

$$= -\frac{n}{1-\theta}\left[\overline{x} - \frac{1-\theta}{\theta}\right].$$

$$T(X_1,\ldots,X_n) = \overline{X}, \qquad g(\theta) = \frac{1-\theta}{\theta} = E_\theta[X], \qquad a(\theta) = -\frac{n}{1-\theta}.$$

$* \ a(\theta)g'(\theta) = \left(-\frac{n}{1-\theta}\right)\left(-\frac{1}{\theta^2}\right) = \frac{n}{\theta^2(1-\theta)} = I_{X_1,\ldots,X_n}(\theta) \Rightarrow \overline{X}$ es regular.

$* \ T(X_1,\ldots,X_n) = \overline{X} \geq 0 \Rightarrow$ estimador de $g(\theta) = E_\theta[X] \in \mathbb{R}^+$.

$$\Downarrow$$

$\forall a \neq 0,\ b \in \mathbb{R},\ a\overline{X} + b$ *es el estimador eficiente de* $aE_\theta[X] + b = a\frac{1-\theta}{\theta} + b$,

y sólo estas funciones admiten estimador eficiente.

c) *Calcular la varianza de cada estimador eficiente y comprobar que coincide con la correspondiente cota de Fréchét-Cramér-Rao.*

$\forall a \neq 0,\ b \in \mathbb{R},\ a\overline{X} + b$ *es el estimador eficiente de* $ag(\theta) + b$, $\quad g(\theta) = \frac{1-\theta}{\theta}$

$$Var_\theta[a\overline{X} + b] = a^2 Var_\theta[\overline{X}] = a^2\frac{Var_\theta[X]}{n} = a^2\frac{1-\theta}{n\theta^2}, \quad \theta \in (0,1).$$

Cota para la varianza de estimadores regulares insesgados en $ag(\theta) + b$:

$$\frac{((ag(\theta)+b)')^2}{I_{X_1,\ldots,X_n(\theta)}} = \frac{(ag'(\theta))^2}{I_{X_1,\ldots,X_n}(\theta)} \underset{g'(\theta) = -\frac{1}{\theta^2}}{\overset{\uparrow}{=}} \left(\frac{-a}{\theta^2}\right)^2 \frac{\theta^2(1-\theta)}{n} = a^2\frac{1-\theta}{n\theta^2}, \quad \theta \in (0,1).$$

d) *Calcular, si existen, los UMVUE para $P_\theta(X=0)$ y para $E_\theta[X]$ y decir si son eficientes.*

$\sum_{i=1}^{n} X_i$ *es suficiente y completo* (R3-P7) $\Rightarrow$ *Todo estimador eficiente (función de* $\sum_{i=1}^{n} X_i$*) es el UMVUE para la correspondiente función paramétrica* $\Rightarrow$ $\overline{X}$ *es el UMVUE para* $E_\theta[X]$.

UMVUE para $P_\theta(X=0)=\theta \longrightarrow$ No admite estimador eficiente.

Notamos $T_n=\sum_{i=1}^{n} X_i$, y buscamos $h(T_n)$ insesgada en θ.

$X \to BN(1,p) \Rightarrow T_n \Rightarrow BN(n,\theta) \Rightarrow P_\theta(T_n=t)=\binom{n+t-1}{t}(1-\theta)^t\theta^n, \ t\in\mathbb{N}\cup\{0\}.$

$$E_\theta[h(T_n)]=\sum_{t=0}^{+\infty} h(t)\binom{n+t-1}{t}(1-\theta)^t\theta^n=\theta, \ \ \forall\theta\in(0,1).$$

$$\Downarrow$$

$$\sum_{t=0}^{+\infty} h(t)\binom{n+t-1}{t}(1-\theta)^t\theta^{n-1}=1, \ \ \forall\theta\in(0,1).$$

Observamos que, como función de θ, los sumandos de esta ecuación son proporcionales a $P_\theta(T_{n-1}=t)=\binom{n-1+t-1}{t}(1-\theta)^t\theta^{n-1}, \ \ n\geq 2$

$\boxed{n\geq 2}$ Identificamos los sumandos de nuestra ecuación con $P_\theta(T_{n-1}=t)$:

$$h(t)\binom{n+t-1}{t}(1-\theta)^t\theta^{n-1}=P_\theta(T_{n-1}=t)=\binom{n-2+t}{t}(1-\theta)^t\theta^{n-1}, \ t\in\mathbb{N}\cup\{0\}.$$

$$\Downarrow$$

$$h(t)=\frac{\binom{n-2+t}{t}}{\binom{n+t-1}{t}}=\frac{n-1}{n-1+t}, \ \ t\in\mathbb{N}\cup\{0\} \overset{T_n\in\mathbb{N}\cup\{0\}}{\Rightarrow} h(T_n)=\frac{n-1}{n-1+T_n}.$$

$*$ $T_n\geq 0 \Rightarrow h(T_n)\in(0,1] \Rightarrow$ estimador de $\theta\in(0,1)$.

$*$ $E_\theta\left[(h(T_n))^2\right]<+\infty, \ \forall\theta\in(0,1)$, ya que está acotado.

$\boxed{n=1}$ $T_1=X$

$$E_\theta[h(X)]=\sum_{x=0}^{+\infty} h(x)(1-\theta)^x\theta=\theta \to \sum_{x=0}^{+\infty} h(x)(1-\theta)^x=1, \ \forall\theta\in(0,1).$$

Identificando coeficientes en estas dos series de potencias en $1-\theta$, se tiene $h(0)=1$ y $h(x)=0, \ x>0$; evidentemente, $h(X)$ es estimador de θ y de segundo orden.

$UMVUE$ para $\theta \longrightarrow \dfrac{n-1}{n-1+T_n}$ $(n\geq 2)$ $\qquad h(X)=\begin{cases} 1, & X=0 \\ 0, & X>0 \end{cases}$ $(n=1)$.

b) *Clase de funciones paramétricas con estimador eficiente y varianza de los estimadores:*

$$\frac{\partial \ln f_\theta^n(x_1,\dots,x_n)}{\partial\theta} = \sum_{i=1}^n \frac{\partial \ln f_\theta(x_i)}{\partial\theta} = \sum_{i=1}^n \ln x_i + \frac{n}{\theta} = \sum_{i=1}^n \ln x_i - \Big(-\frac{n}{\theta}\Big).$$

$$T(X_1,\dots,X_n) = \sum_{i=1}^n \ln X_i, \qquad g(\theta) = -\frac{n}{\theta}, \qquad a(\theta) = 1.$$

∗ $a(\theta)g'(\theta) = \frac{n}{\theta^2} = I_{X_1,\dots,X_n}(\theta) \Rightarrow \sum_{i=1}^n \ln X_i$ es regular.

∗ $\sum_{i=1}^n \ln X_i \leq 0 \Rightarrow$ estimador de $g(\theta) = -\frac{n}{\theta} \in \mathbb{R}^-$.

$$\Downarrow$$

$\forall a \neq 0,\ b \in \mathbb{R},\ a\sum_{i=1}^n \ln X_i + b$ *es el estimador eficiente de* $-a\frac{n}{\theta} + b$,

y sólo estas funciones admiten estimador eficiente.