Suficiencia

Sea X una variable aleatoria k-dimensional con función de densidad $f(x,\theta)$ (discreta o absolutamente continua) donde, para fijar ideas, $x \in \mathbb{R}^k$, $\theta \in \Theta \subset \mathbb{R}^m$.

Sean $X_1, \dots, X_n$ i.i.d X (variables aleatorias muestrales, correspondientes a una muestra aleatoria simple de tamaño n)

Definición 1

Sea $T = T(X_1, \dots, X_n)$ un estadístico (posiblemente multidimensional). Diremos que T es suficiente para el modelo (o para el parámetro θ) si y sólo si la distribución conjunta de $(X_1, \dots, X_n)$ condicionada a T no depende de θ.

Definicion 2 (equivalente)

Sea $S = S(X_1, \dots, X_n)$ un estadístico cualquiera. Diremos que T es suficiente para el modelo si y solo si $E_\theta(S \mid T)$ (la esperanza de S condicionada a T) no depende del parámetro θ, sea cual sea el estadístico S.

Interpretación:

Si la distribución conjunta de $(X_1, \dots, X_n)$ condicionada a T no depende de θ, significa que conociendo T, el conocimiento de $X_1, \dots, X_n$ no aporta información suplementaria sobre θ.

Ejemplo 1:

Sea $X \sim \text{Poisson}(\lambda)$ $\lambda > 0$. Comprobar que $T = \sum_{i=1}^{n} X_i$ es suficiente para dicho modelo.

Hemos de hallar la ley conjunta de $(X_1, \dots, X_n)$ condicionada a T. Puesto que en este caso la distribución conjunta es discreta, las densidades de probabilidad son simplemente probabilidades condicionadas:

$$\underbrace{\tilde{f}(x_1, \dots, x_n, \lambda)}_{\substack{P_\lambda(X_1 = x_1, \dots, X_n = x_n) \\ \| \\ P_\lambda([X_1 = x_1] \cap [X_2 = x_2] \cap \dots \cap [X_n = x_n])}} = \prod_{i=1}^{n} \left(e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \right) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{x_1! \cdot \dots \cdot x_n!}$$

Sucesos independientes

siempre que $(x_1, \dots x_n) \in \mathbb{N}^n$ (cero en caso contrario)

Para calcular la conjunta $X_1, \dots, X_m$ condicionada a T, tendremos que calcular:

$$\tilde{f}_{(X_1,\dots,X_m)|T}(x_1,\dots,x_m,\lambda) = P_\lambda\left([X_1=x_1]\cap\dots\cap[X_m=x_m] \,\middle|\, [T=t]\right)$$

$$= \frac{P_\lambda\left([X_1=x_1]\cap\dots\cap[X_m=x_m]\cap[T=t]\right)}{P_\lambda(T=t)} \qquad (1)$$

para $(x_1,\dots,x_m)\in \mathbb{N}^m$ y $t\in\mathbb{N}$ (podemos elegir $\tilde{f}_{(X_1,\dots,X_m)|T}$ igual a cero en caso contrario).

Observar que $T=\sum_{i=1}^m X_i$, al ser suma de distribuciones de Poisson independientes, seguirá una distribución de Poisson de parametro suma de parámetros: $m\lambda$

Por tanto: $P_\lambda(T=t) = e^{-m\lambda}\dfrac{(m\lambda)^t}{t!}$

Por otra parte, si $\sum_{i=1}^m x_i = t$, entonces $[X_1=x_1]\cap\dots\cap[X_m=x_m]\subset[T=t]$

y por tanto al ser $[X_1=x_1]\cap\dots\cap[X_m=x_m]\cap[T=t]=[X_1=x_1]\cap\dots\cap[X_m=x_m]$ tendremos:

$$\boxed{\tilde{f}_{(X_1,\dots,X_m)|T}}(x_1,\dots,x_m,\lambda) = \frac{P_\lambda\left([X_1=x_1]\cap\dots\cap[X_m=x_m]\right)}{P_\lambda(T=t)}$$

$$= \frac{P_\lambda(X_1=x_1)\,P_\lambda(X_2=x_2)\cdot\ldots\cdot P(X_m=x_m)}{P_\lambda(T=t)}$$

$$= \frac{e^{-\lambda}\frac{\lambda^{x_1}}{x_1!}\, e^{-\lambda}\frac{\lambda^{x_2}}{x_2!}\cdot\ldots\cdot e^{-\lambda}\frac{\lambda^{x_m}}{x_m!}}{e^{-m\lambda}\frac{(m\lambda)^t}{t!}}$$

$$= \frac{e^{-m\lambda}\,\lambda^{\sum_{i=1}^m x_i}\, t!}{x_1!\cdot\ldots\cdot x_m!\; e^{-m\lambda}(m\lambda)^t} = \boxed{\frac{t!}{x_1!\cdot\ldots\cdot x_m!}\left(\frac{1}{m}\right)^t}$$

(ya que $\sum_{i=1}^m x_i = t$)

y en el caso de que $\sum_{i=1}^{m} x_i \neq t$ entonces:

$$[X_1 = x_1] \cap \dots \cap [X_m = x_m] \cap [T = t] = \phi$$

y por tanto $\tilde{f}_{(X_1,\dots,X_m)|T}(x_1,\dots,x_m \lambda) = 0$

En resumen, la densidad condicionada de $(X_1 \dots X_m)$ dado T es:

$$\boxed{\tilde{f}_{(X_1,\dots,X_m)|T=t}(x_1,\dots x_m) = \frac{t!}{x_1! \cdot \ldots \cdot x_m!} \left(\frac{1}{m}\right)^t \quad (2)}$$

siempre que $(x_1,\dots,x_m) \in \mathbb{N}^m$ y $t = \sum_{i=1}^{m} x_i$

y cero en caso contrario. Observar que dicha densidad no depende de λ: por tanto T es un estadístico suficiente para el modelo Poisson.

Para determinar si un estadístico es suficiente tenemos el criterio de factorización de Neyman-Fisher:

Si la función de densidad conjunta de la muestra $(X_1,\dots,X_m)$ puede factorizarse como:

función de $T(x_1 \dots x_m)$ y θ — función sólo de $x_1,\dots x_m$

$$\boxed{\tilde{f}(x_1,\dots,x_m,\theta) = g(T(x_1,\dots,x_m),\theta)\, h(x_1,\dots,x_m) \quad (3)}$$

entonces el estadístico: $T = T(X_1,\dots,X_m)$ es un estadístico suficiente para el modelo considerado (o para θ).

Ejemplo 2:

$X \sim$ Poisson de parámetro λ. Entonces, teniendo en cuenta (1) tenemos:

$$\tilde{f}(x_1,\dots,x_m,\lambda) = e^{-m\lambda} \lambda^{\sum_{i=1}^{m} x_i} \frac{1}{x_1! \cdot \ldots \cdot x_m} \quad (4)$$

Si introducimos:

$$T(x_1,\dots,x_m) = \sum_{i=1}^{m} x_i \qquad h(x_1,\dots,x_m) = \frac{1}{x_1! \cdot \ldots \cdot x_m!}$$

y $\quad g(T(x_1,\dots,x_n),\lambda) = e^{-n\lambda}\lambda^{T(x_1\dots x_n)}$

resulta que (4) puede escribirse como:

$$\boxed{\tilde{f}(x_1,\dots,x_n,\lambda) = g\left(\sum_{i=1}^{n} x_i,\lambda\right) h(x_1,\dots,x_n)}$$

y por el criterio de factorización de Neyman-Fisher podemos asegurar que $T(X_1,\dots,X_n) = \sum_{i=1}^{n} X_i$ es un estadístico suficiente para el modelo (ya lo habíamos comprobado a partir de la definición)

Ejemplo 3:

$X \sim$ Uniforme (α,β) $\quad \alpha<\beta$

$f(x,\alpha,\beta) = \frac{1}{\beta-\alpha}\mathbb{1}_{[\alpha,\beta]}(x)$

recordar $\mathbb{1}_A(x) = \begin{cases} 1 & \Leftrightarrow x\in A \\ 0 & \Leftrightarrow x\notin A\end{cases}$

$$\tilde{f}(x_1,\dots,x_n,\alpha,\beta) = \prod_{i=1}^{n}\left(\frac{1}{\beta-\alpha}\mathbb{1}_{[\alpha,\beta]}(x_i)\right) = \frac{1}{(\beta-\alpha)^n}\mathbb{1}_{(-\infty,\beta]}(x_{(n)})\cdot\mathbb{1}_{[\alpha,\infty)}(x_{(1)})$$

si introducimos $h(x_1\dots x_n) = 1$

$$g((x_{(1)},x_{(n)}),(\alpha,\beta)) = \frac{1}{(\beta-\alpha)^n}\mathbb{1}_{(-\infty,\beta]}(x_{(n)})\,\mathbb{1}_{[\alpha,\infty)}(x_{(1)})$$

resulta que

$$\boxed{\tilde{f}(x_1,\dots,x_n,\alpha,\beta) = g(\underbrace{(x_{(1)},x_{(2)})}_{T(x_1\dots x_n)},(\alpha,\beta))\, h(x_1\dots x_n)}$$

Por tanto el estadístico:

$$\boxed{T = T(X_1\dots X_n) = (X_{(1)},X_{(n)}) = (\min\{X_1\dots X_n\},\max\{X_1\dots X_n\})}$$

es un estadístico suficiente para el modelo Uniforme entre α y β.

Teorema de Rao-Blackwell

$X \sim f(x,\theta) \quad x \in \mathbb{R}^k \quad \theta \in \Theta \subset \mathbb{R}^m$

$X_1, \dots, X_n$ iid X (las variables aleatorias muestrales de una muestra aleatoria simple de tamaño n)

Sea $T = T(X_1 \dots X_n)$ un estadístico suficiente para el modelo

Sea $U = U(X_1, \dots, X_n)$ un estimador insesgado de $\psi(\theta)$ una función paramétrica de θ (posiblemente multidimensional), entonces

$\varphi(T) \equiv E_\theta(U \mid T)$ es un estadístico insesgado de $\psi(\theta)$ (observar que $\varphi(T)$ no depende de θ ya que T es un estadístico suficiente)

y además:

$$E_\theta\left(\|\varphi(T) - \psi(\theta)\|^2\right) \leq E_\theta\left(\|U - \psi(\theta)\|^2\right) \qquad (4)$$

(en particular si $\psi(\theta)$ es unidimensional podemos escribir:

$$\mathrm{var}_\theta(\varphi(T)) \leq \mathrm{var}_\theta(U)$$

Interpretación:

El Teorema de Rao-Blackwell permite obtener, a partir de un estimador insesgado de $\psi(\theta)$, otro estimador insesgado de $\psi(\theta)$ con riesgo (error cuadrático medio) menor. (Mejora de Rao-Blackwell)

Observar que $\varphi(T)$ no puede ser mejorado aplicando de nuevo Rao-Blackwell pues $E_\theta(\varphi(T) \mid T) = \varphi(T)$

Si se exige una nueva propiedad al estadístico T entonces podremos asegurar algo más: que la mejora obtenida por Rao-Blackwell es óptima. Previamente:
(y, esencialmente, única)

Definición 3

Sea $X \sim f(x,\theta) \quad x \in \mathbb{R}^k, \theta \in \Theta \subset \mathbb{R}^m \quad X_1 \dots X_n$ iid X

$S = S(X_1 \dots X_n)$ un estadístico diremos que es completo para el modelo anterior si y sólo si se cumple la propiedad siguiente:

$$\boxed{E_\theta(\varphi(S)) = 0 \quad \forall \theta \in \Theta} \implies \varphi(S) = 0 \text{ con probabilidad } 1 \quad (5)$$

(o bien casi seguramente)

Ejemplo 4:

$X \sim$ Poisson de parámetro λ $\quad X_1 \dots X_n$ i.i.d X

Si $T = T(X_1 \cdots X_n) = \sum_{i=1}^{n} X_i$ comprobar que es un estadístico completo para el modelo Poisson. En efecto como $T \sim$ Poisson $(n\lambda)$ resulta:

$$E_\lambda(\varphi(T)) = \sum_{k=0}^{\infty} \varphi(k)\, P(T=k) = \sum_{k=0}^{\infty} \varphi(k)\, e^{-n\lambda} \frac{(n\lambda)^k}{k!}$$

por tanto:

$$E_\lambda(\varphi(T)) = 0 \iff \sum_{k=0}^{\infty} \varphi(k)\, e^{-n\lambda} \frac{(n\lambda)^k}{k!} = 0$$

igual a:

$$e^{-n\lambda} \sum_{k=0}^{\infty} \varphi(k) \frac{(n\lambda)^k}{k!} = 0$$

equivalente a:

$$\sum_{k=0}^{\infty} \frac{\varphi(k)}{k!} n^k \lambda^k = 0$$

si esto ocurre para todo $\lambda > 0$ y si llamamos $a_k = \frac{\varphi(k)}{k!} n^k$ resulta:

$$\sum_{k=0}^{\infty} a_k \lambda^k = 0 \qquad \forall \lambda > 0$$

es decir tenemos una "serie de potencias" en λ igual a cero para todo valor de la variable λ : por propiedades de las series de potencias ello sólo es posible si $a_k = 0$ para todo $k \in \mathbb{N}$

(pensar en un polinomio de cualquier grado igual a cero para todo valor de su variable independiente; ¿cuales serían sus coeficientes?)

por tanto:

$$a_k = \frac{\varphi(k)}{k!} n^k = 0 \qquad \forall k \in \mathbb{N}$$

ello implica que $\varphi(k) = 0 \quad \forall k \in \mathbb{N}$

por tanto $\varphi(T) = 0$ para T natural, pero T toma valores naturales con probabilidad 1 ya que $T \sim$ Poisson $(n\lambda)$

$\boxed{T \text{ es completo para el modelo}}$

Ejemplos:

$X \sim \exp(\alpha)$ $\quad X_1, \dots, X_n$ iid X
$\alpha > 0$

Comprobar que $T = \sum_{i=1}^{n} X_i$ es completo para el modelo exponencial

$T \sim \text{Gamma}(\alpha, n)$

$$E_\alpha(\varphi(T)) = \int_0^\infty \varphi(t) \frac{1}{\Gamma(n)} t^{n-1} \alpha^n e^{-\alpha t} dt$$

Si $\quad E_\alpha(\varphi(T)) = 0 \quad \forall \alpha > 0$ entonces:

$$\int_0^\infty \varphi(t) \frac{1}{\Gamma(n)} t^{n-1} \alpha^n e^{-\alpha t} dt = 0 \qquad \forall \alpha > 0$$

igual a :

$$\int_0^\infty \varphi(t) t^{n-1} e^{-\alpha t} dt = 0 \qquad \forall \alpha > 0$$

ello es igual a decir que la transformada de Laplace de la función $\varphi(t) t^{n-1}$ es identicamente nula : por propiedades de la transformada de Laplace se deduce que la función es nula en el intervalo $[0, \infty)$ por tanto

$$\varphi(t) t^{n-1} = 0 \quad \forall t > 0$$

$\varphi(T) = 0$ para $T > 0$, lo que ocurre con probabilidad 1 para el modelo exponencial.

T es completo para el modelo

Nota: La transformada de Laplace de una función $g(t)$ es

$$\mathcal{L}(g)(s) = \int_0^\infty g(t) e^{-st} dt \quad (6)$$

(véase textos de análisis matemático standard o buscar en internet Wikipedia etc.. para propiedades)

puede entenderse como una "media ponderada" según la función e^{-st} (donde s es variable).

Teorema de Lehmann-Scheffé

$X \sim f(x,\theta) \quad x \in \mathbb{R}^k \quad \theta \in \Theta \subset \mathbb{R}^m$

$X_1, \dots, X_n$ iid X

$T = T(X_1, \dots, X_n)$ un estadístico suficiente y completo para el modelo

Sea $U = U(T)$ un estadístico función de T tal que

$E_\theta(U) = \psi(\theta)$ para una cierta función paramétrica (posiblemente multidimensional),

Entonces U es el estimador insesgado de $\psi(\theta)$ con UMR (unbiased minimum risk) mínimo riesgo cuadrático (y único en el sentido de que cualquier otro estimador insesgado de $\psi(\theta)$ con el mismo riesgo diferirá de U sólo en un conjunto de probabilidad cero).

Observación:

Como consecuencia de Lehmann-Scheffé, si partimos de un estimador insesgado de $\psi(\theta)$ y lo mejoramos (en el sentido de reducir su riesgo) a partir del Teorema de Rao-Blackwell, basándonos en un estadístico no sólo suficiente, sino completo, el estadístico "mejorado" así obtenido es el de mínimo riesgo cuadrático. (En el caso de que $\psi(\theta)$ sea un escalar, habremos obtenido el estimador UMV del parámetro $\psi(\theta)$). unbiased minimum variance

Nota:

en algunos textos UMV se indica por UMVU (uniformly minimum variance unbiased)

y UMR se indica por UMRU (uniformly minimum risk unbiased)

el adjetivo "uniformly" indica que la propiedad de mínimo riesgo (dentro de la clase de estimadores insesgados) es una propiedad que ocurre con independencia del valor de θ

Ejemplo 6

Hallar el estimador UMVU de $P(X=0)$ donde $X \sim \text{Poisson}(\lambda)$, basándonos en una muestra aleatoria simple de tamaño m, $X_1, \dots, X_m$ iid X.

Sabemos que $T = \sum_{i=1}^{m} X_i$ es un estadístico suficiente y completo.

Para hallar el estimador UMVU (uniformly minimum variance unbiased) de $P(X=0)$ bastara encontrar un estimador insesgado de dicha probabilidad y luego aplicar Rao-Blackwell y Lehmann-Scheffé.

Observar que

$$P(X=0) = e^{-\lambda}$$

Para hallar un estimador insesgado de $\psi(\lambda) = e^{-\lambda}$ consideraremos el indicador del suceso $X_1 = 0$, $\mathbb{1}_{\{X_1=0\}}$. Como todo indicador,

$$E(\mathbb{1}_A) = P(A), \quad \text{por tanto} \quad E_\lambda(\mathbb{1}_{\{X_1=0\}}) = P_\lambda(X_1=0) = e^{-\lambda}$$

por tanto $\mathbb{1}_{\{X_1=0\}}$ es un estimador insesgado de $\psi(\lambda)$

Apliquemos Rao-Blackwell:

$$\varphi(t) = E_\lambda\left(\mathbb{1}_{\{X_1=0\}} \mid T=t\right) = P_\lambda\left(X_1=0 \mid T=t\right) = \frac{P_\lambda([X_1=0] \cap [T=t])}{P(T=t)}$$

con $t \in \mathbb{N}$

El suceso $[X_1=0] \cap [T=t]$ es igual al suceso
$[X_1=0] \cap \left[\sum_{j=2}^{m} X_j = t\right]$

Observar además que $X_1 \sim P(\lambda)$, $T \sim P(m\lambda)$ y
$\sum_{j=2}^{m} X_j \sim P((m-1)\lambda)$, por tanto, al ser $[X_1=0]$ y $\left[\sum_{j=2}^{m} X_j = t\right]$ independientes, tendremos:

$$\varphi(t) = \frac{P_\lambda(X_1=0)\, P_\lambda\left(\sum_{j=2}^{m} X_j = t\right)}{P(T=t)} =$$

$$= \frac{e^{-\lambda} \quad e^{-(m-1)\lambda} \dfrac{((m-1)\lambda)^t}{t!}}{e^{-m\lambda} \quad \dfrac{(m\lambda)^t}{t!}} = \left(1 - \frac{1}{m}\right)^t$$

Por tanto el estadístico:

$$\left[\ \varphi(T) = \left(1 - \frac{1}{n}\right)^{T}\ \right] \qquad (7)$$

es un estimador insesgado de $e^{-\lambda}$ y como es función de un estadístico suficiente y completo del modelo, por Lehmann-Scheffé se tratará del estimador UMVU de $e^{-\lambda}$.

Observación: puesto que $T = \sum_{i=1}^{n} X_i = \frac{1}{n} \bar{X}_n$, vemos que (7) puede escribirse como:

$$\varphi(T) = \left(1 - \frac{1}{n}\right)^{\bar{X}_n n}$$

y como $\bar{X}_n \xrightarrow{C.P} \lambda$ (ley de los grandes números).

y $\lim_{n \to \infty} \left(1 - \frac{1}{n}\right)^{an} = e^{-a}$

se puede demostrar que $\left(1 - \frac{1}{n}\right)^{\bar{X}_n n} \xrightarrow{C.P.} e^{-\lambda}$

es decir se trataría de un estimador también consistente de $e^{-\lambda}$.

Ejemplo 7

Hallar el estimador UMVU de $P(X > a)$, donde $X \sim \text{Exponencial}(\alpha)$ y a es una constante positiva conocida, a partir de una muestra aleatoria simple de tamaño n.

Sabemos que $T = \sum_{i=1}^{n} X_i$ es un estadístico suficiente y completo para el modelo exponencial. Por tanto, por los teoremas de Rao-Blackwell y Lehmann-Scheffé bastará encontrar un estimador insesgado de $P(X>a)$. Observemos que

$$P(X>a) = \int_a^{\infty} \alpha e^{-\alpha x}\, dx = \left[-e^{-\alpha x}\right]_a^{\infty} = \lim_{x \to \infty} \underbrace{-e^{-\alpha x}}_{0} - (-e^{-\alpha a}) = e^{-\alpha a}$$

Un estimador insesgado de $P(X>a) = e^{-\alpha a}$ será, por ejemplo,

$\mathbb{1}_{\{X_1>a\}}$ ya que $E(\mathbb{1}_{\{X_1>a\}}) = P(X_1>a) = P(X>a) = e^{-\alpha a}$

(indicador de $\{X_1>a\}$)

Por tanto, el estimador UMVU buscado es:

$$\varphi(t) = E_\alpha(\mathbb{1}_{\{X_1>a\}} \mid T=t)$$

Ahora bien:

$$E_\alpha(\mathbb{1}_{\{X_1>a\}} \mid T=t) = P(X_1>a \mid T=t)$$

Habrá que hallar la densidad condicionada $X_1|T$.

Puesto que $X_1, \dots X_n$ son independientes, también serán variables independientes X_1 y $S=\sum_{i=2}^n X_i$. Escribamos su densidad conjunta, sabiendo que $X_1 \sim \text{Exponencial}(\alpha)$, $S \sim \text{Gamma}(\alpha, n-1)$ por tanto:

$$f_{(X_1,S)}(x_1, s) = \alpha e^{-\alpha x_1} \frac{\alpha^{n-1}}{\Gamma(n-1)} s^{n-2} e^{-\alpha s}$$

para $x_1>0$ $s>0$ (cero en caso contrario)

$$= \frac{\alpha^n}{\Gamma(n-1)} e^{-\alpha(x_1+s)} s^{n-2}$$

Efectuemos ahora el cambio de variable $(X_1, S) \to (X_1, T)$ definido por:

$$X_1 = X_1$$
$$T = X_1 + S \qquad (S = T - X_1)$$

El determinante del jacobiano del cambio es:

$$\det\begin{pmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_1}{\partial t} \\ \frac{\partial s}{\partial x_1} & \frac{\partial s}{\partial t} \end{pmatrix} = \det\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1$$

por tanto la densidad conjunta resultante será:

$$f_{(X_1,T)}(x_1,t) = f_{(X_1,S)}(x, t-x_1)\cdot 1$$ para $x_1 > 0$ y $t \geqslant x_1$ (cero en caso contrario)

por tanto:

$$f_{(X_1,T)}(x_1,t) = \frac{\alpha^m}{\Gamma(m-1)}(t-x_1)^{m-2}e^{-\alpha t}$$ $t \geqslant x_1 > 0$ (cero en caso contrario)

para calcular $f_{X_1|T}(x_1)$ necesitamos saber cual es la marginal de T.
Aunque dicha marginal podríamos hallarla calculando $\int_0^t f_{(X_1,T)}(x_1,t)\,dx_1$ el resultado ya lo sabemos: como $T=\sum_{i=1}^m X_i \sim \text{Gamma}(\alpha,m)$ resultará:

$$f_T(t) = \frac{\alpha^m}{\Gamma(m)}t^{m-1}e^{-\alpha t}$$ para $t>0$ (cero en caso contrario)

Por tanto,

$$f_{X_1|T}(x_1) = \frac{f_{(X_1,T)}(x_1,t)}{f_T(t)} = \frac{\alpha^m(t-x_1)^{m-2}e^{-\alpha t}/\Gamma(m-1)}{\alpha^m t^{m-1}e^{-\alpha t}/\Gamma(m)}$$

$$= \frac{\Gamma(m)}{\Gamma(m-1)}\frac{(t-x_1)^{m-2}}{t^{m-1}}$$ para $x_1 \in (0,t)$ (cero en caso contrario) $t>0$

$$= (m-1)\frac{1}{t}\left(1-\frac{x_1}{t}\right)^{m-2}$$ $x_1 \in (0,t)$ $t>0$ (cero en caso contrario)

Hay que hallar ahora:

$$P(X_1>a \mid T=t) = \int_0^\infty f_{X_1|T}(x_1)\,dx_1 =$$

$$= \int_a^t (m-1)\frac{1}{t}\left(1-\frac{x_1}{t}\right)^{m-2}dx_1$$ caso $a<t$

$$= 0$$ caso $a > t$

pero:

$$\int_a^t (n-1)\,\frac{1}{t}\left(1-\frac{x_1}{t}\right)^{n-2} dx_1 = \int_{a/t}^{1} (n-1)\,(1-y)^{n-2}\,dy = \left[-(1-y)^{n-1}\right]_{a/t}^{1}$$

haciendo $\frac{x_1}{t} = y$

$x_1 = ty$

$$= 0 + \left(1-\frac{a}{t}\right)^{n-1}$$

por tanto el estimador buscado es:

$$\varphi(T) = \begin{cases} 0 & T < a \\ \left(1-\frac{a}{T}\right)^{n-1} & T \geq a \end{cases}$$

es el estimador UMVU,
uniformemente insesgado y de mínima varianza.

Para completar el tema de suficiencia, agregaremos alguna definición y comentaremos alguna propiedad.

Definición 4

Sea $X \sim f(x,\theta)$ $x \in \mathbb{R}^k$, $\theta \in \Theta \subset \mathbb{R}^m$ $X_1,\dots,X_n$ iid X

$S = S(X_1,\dots,X_n)$ es un estadístico suficiente minimal si y solo si es en primer lugar un estadístico suficiente, y en segundo lugar, dado otro estadístico suficiente $T(X_1,\dots,X_n)$ para el mismo modelo, resulta que existe una aplicación medible g tal que:

$S = g(T)$ con probabilidad 1 (casi seguramente)

es decir si siempre será posible expresar S como función de cualquier otro estadístico suficiente.

Observación: podría demostrarse que $S = \sum_{i=1}^{n} X_i$ es un estadístico suficiente minimal del modelo exponencial, pero $T = (X_1, \sum_{i=2}^{n} X_i)$ es también un estadístico suficiente pero no es minimal (S puede expresarse como función de T, pero T no puede expresarse como función de S).

Definición 5

Sea $X \sim f(x,\theta)$ $x \in \mathbb{R}^k$ $\theta \in \Theta \subset \mathbb{R}^m$ $X_1,\dots,X_n$ iid X

$S = S(X_1,\dots,X_n)$ es un estadístico libre si y sólo si la distribución de S no depende de θ (o a libre distribución)

Ejemplo 8

Sea $X \sim N(\mu,1)$ $X_1,\dots,X_n$ iid X

$S = X_{(n)} - X_{(1)}$ Vamos a comprobar que es un estadístico libre.

Introducimos $Z_i = X_i - \mu$ $Z_i \sim N(0,1)$ y son todas ellas independientes. Entonces, la distribución de $Z_{(n)} - Z_{(1)}$ no dependerá de μ puesto que se calculará a partir de la distribución conjunta de $Z_1 \dots Z_n$, distribución conjunta que no depende de μ.

Por tanto, si observamos que $Z_{(n)} = X_{(n)} - \mu$ y $Z_{(1)} = X_{(1)} - \mu$, resulta que $X_{(n)} - X_{(1)} = Z_{(n)} - Z_{(1)}$ y deducimos pues que $X_{(n)} - X_{(1)}$, su distribución, no depende de μ: es pues un estadístico libre

---

Observación:

Un estadístico libre no nos aporta información útil para estimar el parámetro θ.

Algunos resultados

* Dado un modelo estadístico si S es un estadístico suficiente, y si g es una aplicación medible y biyectiva, cuyo dominio sea el espacio muestral de los valores de S, entonces $g(S)$ es también un estadístico suficiente.

* Si tenemos un modelo exponencial:

$$f(x;\theta) = A(x)K(\theta)\exp\left(\sum_{j=1}^{r} \theta_j T_j(x)\right) \qquad \theta \in \Theta \subset \mathbb{R}^r \quad x \in \mathbb{R}^k$$

y el espacio vectorial generado por Θ es todo $\mathbb{R}^r$

entonces $T(x) = (T_1(x),\dots,T_r(x))$ es un estadístico suficiente y minimal.

Si $\operatorname{int}(\Theta) \neq \emptyset$ podemos asegurar que T es completo.

Observar que si $\operatorname{int}(\Theta) \neq \emptyset$ entonces seguro que el espacio vectorial generado por Θ es todo $\mathbb{R}^r$ y por tanto se concluye que T es suficiente minimal y completo.