Interrogación 1

$$\boxed{Var(x) = E(X^2) - E^2(X)}$$

$$\boxed{M_x(t) = E(e^{tx})}$$

$$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{X})^2$$

$$S_n^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{X})^2$$

Notación CASELLA:

$$\min\{x_i\} = x_{(1)}$$
$$\max\{x_i\} = x_{(n)}$$

La **función indicatriz** de un subconjunto A de $\mathcal{X}$ es la función $I: \mathcal{X} \to \{0,1\}$ dada por:

$$\boxed{I_A(x) \equiv I(x \in A) = \begin{cases} 1 & si\ x \in A \\ 0 & si\ x \notin A \end{cases}}$$

Para variables aleatorias continuas:

$$\boxed{E(x) = \int x \cdot f_x(x_i) \cdot dx}$$

Para variables aleatorias discretas:

$$\boxed{E(x) = \sum_{i=0}^{\infty} x_i \cdot f_x(x_i)}$$

Si $X_1, \ldots, X_n \sim \boldsymbol{iid}$:

$$\boxed{Var(\bar{X}) = \frac{Var(X)}{n}} \quad \wedge \quad \boxed{E(\bar{X}) = \mu}$$

Función de Máxima Verosimilitud:

$$\boxed{L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta)}$$

Notar que, para estimar con esta función, se debe derivar respecto al parámetro de interés, e igualar a 0

Definición: Sea $\pi(\theta)$ la distribución a *priori*, y la función distribución de la muestra $f(x|\theta)$, entonces la distribución ***posteriori***:

$$\boxed{\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)}}$$

Recordar:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Donde $m(x) = \int f(x|\theta)\pi(\theta)d\theta$

Constante de normalización $\boldsymbol{C}$ es tal que $m(x)$ cumple:

$$\boxed{\boldsymbol{C} \cdot m(x) = 1}$$

Luego:

$$\pi(\theta|x) \rightarrow \boldsymbol{C} \cdot \pi(\theta|x)$$

Dado un parámetro θ con función de densidad posteriori $\pi(\theta|x)$, su estimador de Bayes corresponde a:

$$\boxed{\hat{\theta}_B = E(\theta|x) = \int \theta \cdot \pi(\theta|x) \cdot d\theta}$$

Recordar el teorema de cambio de variable univariado (**Teorema 2.1.8 Casella & Berger**): Sea $X \sim f_X(x)$ y sea $Y = g(X)$. Supongamos que existe una partición A_0, A_1, ..., A_k de $\mathcal{X}$ tal que $P(X \in A_0) = 0$ y $f_X(x)$ es continua en cada A_i. Además, supongamos que existen funciones $g_1(x), ..., g_k(x)$, definidos en $A_1, ..., A_k$, respectivamente, que satisfacen:

i) $g(x) = g_i(x)$ para $x \in A_i$

ii) $g_i(x)$ es monótona en cada A_i

iii) El conjunto $\mathcal{Y} = \{y\colon y = g_i(x)$ para algún $x \in A_i\}$ es el mismo para cada $i = 1, ..., k$

iv) $g_i^{-1}(y)$ tiene derivada continua en $\mathcal{Y}$ en cada $i = 1, ..., k$

Entonces:

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right| & y \in \mathcal{Y} \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Estadísticos de Orden y sus distribuciones.

$$\boxed{f_{X_{(1)}}(x) = n \cdot f(x) \cdot \left(1 - F(x)\right)^{n-1}}$$

$$\boxed{f_{X_{(n)}}(x) = n \cdot f(x) \cdot \left(F(x)\right)^{n-1}}$$

$$\boxed{f_{X_{(i)}}(x) = \binom{n}{i} \cdot i \cdot f(x) \cdot \left(F(x)\right)^{i-1} \left(1 - F(x)\right)^{n-1}}$$

Recordamos que:

$$F(x) = \int_{-\infty}^{x} f_x(t) \cdot dt$$

Distribución Gamma:

$$\boxed{\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \cdot e^{-x} \cdot dx}$$

$$\boxed{\Gamma(n) = (n-1)!}$$

Interrogación 2

Definición:
El error cuadrático medio (ECM) de un estimador $T(X)$ de θ corresponde a:

$$\boldsymbol{ECM_\theta}\big(T(X)\big) = E_\theta\{(T(X)-\theta)^2\}$$
$$= Var_\theta\big(T(X)\big) + \big(\mathrm{Sesgo}_\theta T(X)\big)^2$$

Donde:

$$\mathrm{Sesgo}_\theta T(X) = E\big(T(X)\big) - \theta$$

$$\boxed{Sesgo_\theta T(X) = 0} \quad \Longrightarrow \quad \boxed{T(X) \text{ es } \textbf{insesgado}}$$

Mejores estimadores insesgados

Un estimador $W(\boldsymbol{X})$ es un <u>mejor estimador insesgado</u> si y sólo si:

$$E_\theta\big(W(\boldsymbol{X})\big) = \tau(\theta),$$

y, para todo estimador $T(\boldsymbol{X})$ tal que $E_\theta\big(T(\boldsymbol{X})\big) = \tau(\theta)$ se cumple que

$$\boxed{Var_\theta\big(W(\boldsymbol{X})\big) \le Var_\theta\big(T(\boldsymbol{X})\big)}$$

Para todo $\theta \in \Theta$.

El estimador $W(\boldsymbol{X})$ se dice también insesgado y de varianza uniformemente mínima, o **UMVUE**, de $\tau(\theta)$.

Teorema de Cramér-Rao:

Sea $X_1, \dots, X_n$ una muestra con función de densidad conjunta $f_\theta(\boldsymbol{x})$, y $T(\boldsymbol{X})$ un estimador tal que $E_\theta\big(T(\boldsymbol{X})\big)$ es diferenciable como función de θ. Suponga que la densidad de $\boldsymbol{X}$ satisface:

$$\frac{d}{d\theta}\int\cdots\int h(\boldsymbol{x}) f_\theta(\boldsymbol{x})\, d\boldsymbol{x} = \int\cdots\int h(\boldsymbol{x}) \cdot \frac{d}{d\theta} f_\theta(\boldsymbol{x})\, d\boldsymbol{x}$$

para toda función h con $E_\theta|h(\boldsymbol{x})| < \infty$. Entonces:

$$\boxed{Var_\theta\big(T(\boldsymbol{X})\big) \geq \frac{\left(\frac{d}{d\theta}E_\theta\big(T(\boldsymbol{X})\big)\right)^2}{E_\theta\left(\left(\frac{d}{d\theta}\log f_\theta(\boldsymbol{X})\right)^2\right)}}$$

Corolario:

Sea $\boldsymbol{X_1, \dots, X_n}$ **una muestra aleatoria** proveniente de una distribución con función de densidad $f_\theta(x)$, y sea $T(\boldsymbol{X})$ un estimador donde $E_\theta\big(T(\boldsymbol{X})\big)$ es diferenciable como función de θ. Si la densidad conjunta $\prod_{i=1}^n f_\theta(x_i)$ cumple las condiciones del *Teorema de Cramér-Rao*, entonces:

$$\boxed{Var_\theta\big(T(\boldsymbol{X})\big) \geq \frac{\left(\frac{d}{d\theta}E_\theta\big(T(\boldsymbol{X})\big)\right)^2}{\boldsymbol{n} \cdot E_\theta\left(\left(\frac{d}{d\theta}\log f_\theta(\boldsymbol{X})\right)^2\right)}}$$

Teorema:
Si $f_\theta(x)$ satisface:

$$\frac{d}{d\theta}\int \left(\frac{\partial}{\partial\theta}\log f_\theta(x)\right) f_\theta(x)\, dx = \int \frac{\partial}{\partial\theta}\left[\left(\frac{\partial}{\partial\theta}\log f_\theta(x)\right) f_\theta(x)\right] dx$$

entonces:

$$\boxed{E_\theta\left(\left(\frac{\partial}{\partial\theta}\log f_\theta(\boldsymbol{X})\right)^2\right) = -E_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f_\theta(\boldsymbol{X})\right)}$$

Familia Exponencial

Consideremos una familia de distribuciones indexadas por el parámetro θ,

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\},$$

para una variable aleatoria X, que toma valores en $\mathcal{X}$ con función de verosimilitud asociada $L(x, \cdot)$. Se dice que la log-verosimilitud de X pertenece a una **familia exponencial** si y sólo si:

- Existe un conjunto S, que no depende del valor de θ, tal que:

$$P_\theta(X \in S) = 1, \qquad \forall\theta \in \Theta$$

y

$$f_\theta(x) > 0, \qquad \forall x \in S$$

En otras palabras, existe un conjunto de posibles valores que puede tomar x ($\mathbb{N}, \mathbb{R}$, etc.) donde este conjunto NO depende del valor de θ.

Donde $f_\theta(\cdot)$ corresponde a la función densidad asociada a P_θ.

- Existen funciones $\omega_1, \dots, \omega_n$, con dominio Θ y $t_1, \dots, t_k$, con dominio $\mathcal{X}$, todas ellas real valoradas, tales que:

$$\log L(x,\theta) = \log h(x) + \log c(\theta) + \sum_{j=1}^{k} \omega_j(\theta) t_j(x)$$

con h, c funciones positivas, real valoradas.

Equivalentemente, según *Casella & Berger,* una densidad es pertenece a una familia exponencial si puede escribirse como:

$$\boxed{f_\theta(x) = h(x)c(\boldsymbol{\theta}) \exp\left\{\sum_{i=1}^{k} w_i(\boldsymbol{\theta}) t_i(x)\right\}}$$

Importante: Si una distribución constituye una familia exponencial, entonces se cumplen las condiciones del teorema de *Cramér-Rao*.

Principio de suficiencia:

Si $T(\boldsymbol{X})$ es un estadístico suficiente para θ, inferencia sobre este parámetro debe ser hecha solo sobre la base e $T(\boldsymbol{X})$.

Estadístico suficiente:

Un estadístico $T(\boldsymbol{X})$ es un estadístico suficiente para θ, si y sólo si la distribución condicional de la muestra $\boldsymbol{X}$ dado el valor de $T(\boldsymbol{X})$ NO depende e θ.

Teorema:

Sea f_θ la función densidad conjunta de $X_1, \dots, X_n$. El estadístico $T(\boldsymbol{X})$, con distribución q_θ, es suficiente para θ si y sólo si la razón

$$\boxed{\frac{f_\theta(x)}{q_\theta\big(T(x)\big)} \qquad \text{NO depende de } \theta}$$

es constante como función de θ, para todo valor de $x \in \mathcal{X}$.

Teorema de factorización (Fisher-Neyman):

Sea $f_{\boldsymbol{\theta}}(x)$ la función de densidad (probabilidad) conjunta de la muestra $\boldsymbol{X}$. Un estadístico $T(\boldsymbol{X})$ es un estadístico suficiente para $\boldsymbol{\theta}$, si y sólo si existen funciones g_θ y h, tales que, para todo $x \in \mathcal{X}$ y $\theta \in \Theta$,

$$\boxed{f_\theta(x) = g_\theta\big(T(x)\big) \cdot h(x)}$$

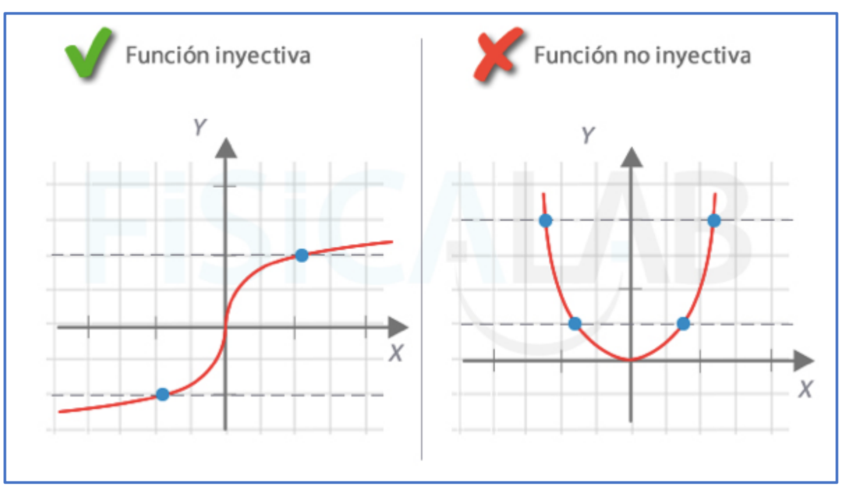
Resultado:

Suponga que $T(\boldsymbol{X})$ es un estadístico suficiente para θ, y defina el estadístico:

$$\boxed{T'(\boldsymbol{X}) = r \cdot \big(T(\boldsymbol{X})\big)},$$

con r una función inyectiva. Entonces, el estadístico $T'(\boldsymbol{X})$ también corresponde a un estadístico suficiente para θ.

Recordar:



Suficiencia en una familia exponencial:

Sea $X_1, \dots, X_n$ una muestra aleatoria proveniente de una distribución en una familia exponencial, con densidad (probabilidad) f_θ, de modo que:

$$f_\theta(x) = h(x) \cdot c(x) \cdot \exp\left\{\sum_{j=1}^{k} \omega_j(\theta) \cdot t_j(x)\right\}$$

Entonces,

$$\boxed{T(\boldsymbol{X}) = \left(\sum_{i=1}^{n} t_1(x_i), \cdots, \sum_{j=1}^{n} t_k(x_i)\right)}$$

es un estadístico suficiente para θ.

Estadístico suficiente minimal:

Un estadístico suficiente $T(\boldsymbol{X})$ se dice <u>suficiente minimal</u>, si y sólo si para cualquier otro estadístico suficiente $T'(\boldsymbol{X})$, existe una función v, tal que:

$$T(\boldsymbol{X}) = v\big(T'(\boldsymbol{X})\big)$$

Teorema sobre minimalidad:

Sea f_θ la función de densidad (probabilidad) de una muestra aleatoria $X_1, \dots, X_n$. Si se cumple que:

$$\boxed{\frac{f_\theta(x)}{f_\theta(y)} \text{ No depende de } \theta} \iff \boxed{T(\boldsymbol{x}) = T(\boldsymbol{y})}$$

Entonces:

$T(\boldsymbol{X})$ es suficiente minimal para θ.

Interrogación 3

Estadístico Ancilar:

Un estadístico $S(\boldsymbol{X})$ se denomina <u>ancilar para un parámetro θ</u> si y sólo si su distribución NO depende de este parámetro.

Ejercicio:

Sea $X_1, \dots, X_n$ una muestra aleatoria proveniente de una distribución Uniforme$(\theta, \theta+1)$, con $\theta \in \mathbb{R}$. Muestre que el rango:

$$R(\boldsymbol{X}) = X_{(n)} - X_{(1)}$$

es ancilar para θ.

<u>Solución:</u>

Es fácil ver que el rango de la uniforme será $(\theta+1)-(\theta) = 1$ y, por lo tanto, no dependerá de θ y el rango será ancilar. Sin embargo, lo demostraremos por definición:

Sea $X_1, \dots, X_n \sim iid$, con $x_{(i)},\ x_{(j)}$ estadísticos de orden, la función de distribución será:

$$f_{x_{(i)},x_{(j)}}(u,v) = \frac{n!}{(i-1)!\,(j-1-i)!\,(n-j)!} \cdot f_x(u) \cdot f_x(v) \cdot [F_x(u)]^{i-1}$$

$$\cdot [F_x(v) - F_x(u)]^{j-1-i} \cdot [1 - F_x(v)]^{n-j} \cdot I(-\infty < u < v < \infty)$$

Ahora, necesitamos que $i = 1$ y $j = n$, dado que eso estamos buscando.
Las funciones f_x y F_x serán, respectivamente:

$$f_x(x) = \frac{1}{\theta + 1 - \theta} \cdot I(\theta < x < \theta + 1) = I(\theta < x < \theta + 1)$$

$$F_x(x) = P(X \leq x) = \begin{cases} 0 & , \quad x \leq \theta \\ x - \theta & , \quad \theta < x < \theta + 1 \\ 1 & , \quad x \geq \theta + 1. \end{cases}$$

Revisar guía distribuciones

Reemplazamos al caso general y nos queda:

$$f_{x_{(1)},x_{(n)}}(u,v) = n(n-1)(v-u)^{n-2} \cdot I(\theta < u < v < \theta + 1)$$

Queremos la distribución de $R(\boldsymbol{X}) = X_{(n)} - X_{(1)}$.

Sean U y V dos variables aleatorias:

Función Biyectiva, es decir, conociendo R y M debo poder conocer U y V.

$$(U, V) \Rightarrow R(U, V)$$

$$(U, V) \Rightarrow M(U, V)$$

R = Rango

Función que pasa de dos dimensiones, a una dimensión, por lo tanto, por sí sola no me sirve, necesito una función Biyectiva.

M = Función inventada para cumplir biyectividad.

$$(u, v) \Rightarrow (r, m)$$

$$u = g_u^{-1}(r, m)$$

$$v = g_v^{-1}(r, m)$$

Funciones que me permiten conocer los valores de u y v conociendo r y m.

Teorema del Jacobiano:

"Derivadas parciales de las variables antiguas (u, v) con respecto a las nuevas (r, m)"

$$J = \begin{pmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial m} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial m} \end{pmatrix}$$

$$f_{R,M}(r, m) = |\det(J)| \cdot \underbrace{f_{U,V}(u, v)}_{de\ x_{(1)}, x_{(n)}}$$

Propuesta:

$$R = X_{(n)} - X_{(1)}$$

$$M = \frac{X_{(1)} + X_{(n)}}{2}$$

$$\boxed{\begin{aligned} U &= X_{(1)} \\ V &= X_{(n)} \end{aligned}}$$

$$\left(x_{(1)}, x_{(n)}\right) \Rightarrow (r, m)$$

$$\underbrace{x_{(1)} = \frac{2m - r}{2}}_{g_u^{-1}(r,m)} \wedge \underbrace{x_{(n)} = \frac{2m + r}{2}}_{g_v^{-1}(r,m)}$$

$$J = \begin{pmatrix} \frac{\partial x_{(1)}}{\partial r} & \frac{\partial x_{(1)}}{\partial m} \\ \frac{\partial x_{(n)}}{\partial r} & \frac{\partial x_{(n)}}{\partial m} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 1 \\ \frac{1}{2} & 1 \end{pmatrix} \Rightarrow \begin{aligned} \det(J) &= -1 \\ |\det(J)| &= 1 \end{aligned}$$

Luego,

$$f_{R,M}(r,m) = \underbrace{|\det(J)|}_{1} \cdot \underbrace{f_{U,V}(u,v)}_{de\ x_{(1)},x_{(n)}}$$

$$= n(n-1)(v-u)^{n-2} \cdot I(-\infty < u < v < \infty)$$

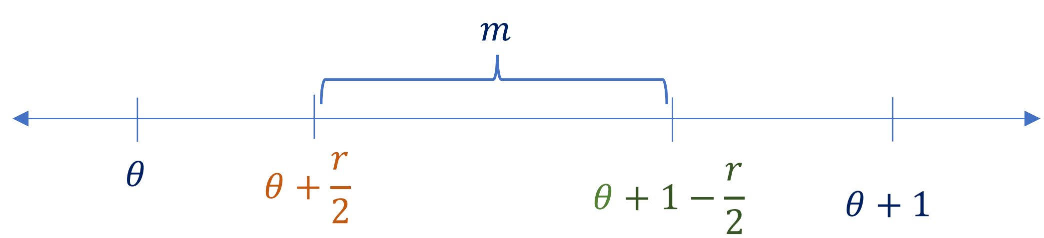
$$\left.\begin{aligned} u &= \frac{2m-r}{2} \\ v &= \frac{2m+r}{2} \end{aligned}\right\} \quad v - u = \frac{r}{2} + \frac{r}{2} = r$$

$$= n(n-1)r^{n-2} \cdot I\left(\theta < \frac{2m-r}{2} < \frac{2m+r}{2} < \theta + 1\right)$$

Como vemos, necesitamos que $r > 0$ para que se cumpla lo de la indicatriz.

Ahora, si despejamos m en ambos lados de la indicatriz, tenemos que:

$$\boxed{\begin{aligned} \theta + \frac{r}{2} &< m \\ \theta + 1 - \frac{r}{2} &> m \end{aligned}} \longrightarrow \boxed{1^{era} \text{ condición}}$$



¿Qué valores puede tomar r además de ser positiva?

$$\theta + \frac{r}{2} < \theta + 1 - \frac{r}{2}$$

$$\boxed{0 < r < 1} \longrightarrow \boxed{2^{da} \text{ condición}}$$

Entonces nos queda:

$$f_{R,M}(r,m) = n(n-1)r^{n-2} \cdot I(0 < r < 1) \cdot I\left(\theta + \frac{r}{2} < m < \theta + 1 - \frac{r}{2}\right)$$

$$f_R(r) = \int f_{R,M}(r,m) \cdot dm$$

2^{da} condición

Tomado $0 < r < 1$:

1^{era} condición

$$f_R(r) = \int_{\theta+\frac{r}{2}}^{\theta+1-\frac{r}{2}} n(n-1)r^{n-2} \cdot dm$$

$$= n(n-1)r^{n-1} \cdot \left(\left(\theta + 1 - \frac{r}{2}\right) - \left(\theta + \frac{r}{2}\right)\right)$$

$$= n(n-1)r^{n-1}(1-r)$$

Luego, $f_R(r) = n(n-1)r^{n-1}(1-r) \cdot I(0 < r < 1)$, NO depende de θ, luego:

$$R = X_{(n)} - X_{(1)}$$

Es ancilar para θ. Entonces, el rango no aporta información sobre θ.

Dependencia entre un estadístico suficiente minimal y uno ancilar:

Considere una muestra aleatoria $X_1, \dots, X_n$ proveniente de una distribución con función de probabilidad dada por:

$$P_\theta(X = \theta) = P_\theta(X = \theta + 1) = P_\theta(X = \theta + 2) = \frac{1}{3},$$

∴ ∈ a una Familia de localización, al no depender de θ.

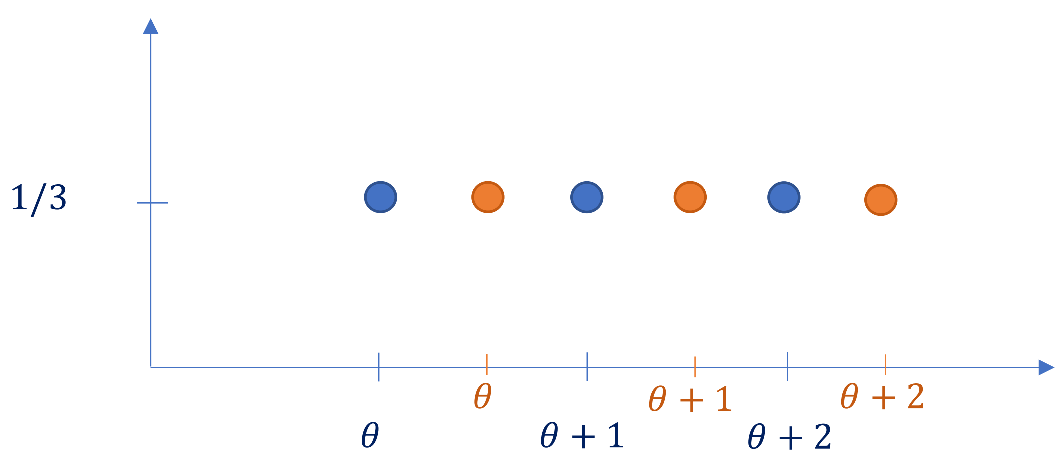
con θ valor desconocido. Considere el estadístico:

$$\boldsymbol{T}(\boldsymbol{X}) = \left(R(\boldsymbol{X}), M(\boldsymbol{X})\right) = \left(X_{(2)} - X_{(1)}, \frac{X_{(1)} + X_{(2)}}{2}\right).$$

Sabemos que $R(\boldsymbol{X})$ es ancilar para θ (1), y se puede demostrar que $T(\boldsymbol{X})$ es suficiente minimal (2). Claramente ellos no son independientes.

1) $R(\boldsymbol{X})$ es ancilar:

Recordatorio: Ancilaridad en página 8.



$$P(X = \theta) = P(X = \theta + 1) = P(X = \theta + 2) = \frac{1}{3}$$

$$P(X - \theta = 0) = P(X - \theta = 1) = P(X - \theta = 2) = \frac{1}{3}$$

$\Longrightarrow$ Familia de probabilidad es de localización. Entonces, por ejercicio propuesto (pág. 8):

$R(\boldsymbol{X}) = X_{(n)} - X_{(1)}$ es ancilar para θ.

2) $T(\boldsymbol{X})$ es suficiente:

$$X_{(2)} - X_{(1)} = r$$
$$X_{(1)} + X_{(2)} = 2m$$

Luego:

$$X_{(2)} = \frac{2m + r}{2} \quad \wedge \quad X_{(1)} = \frac{2m - r}{2} \quad //$$

Estadístico completo:

Sea $f_{\boldsymbol{\theta}}$ una familia de funciones de densidad (probabilidad) para un estadístico $T(\boldsymbol{X})$. La familia $f_{\boldsymbol{\theta}}$ se dice una <u>familia completa</u> si y sólo si, para toda función g,

$$E_{\theta}\left(g\big(T(\boldsymbol{X})\big)\right) = 0, \qquad \forall \boldsymbol{\theta} \in \Theta$$

si y sólo si:

$$P_{\boldsymbol{\theta}}\big(g\big(T(\boldsymbol{X})\big) = 0\big) = 1, \quad \forall \boldsymbol{\theta} \in \Theta$$

Es decir, $g\big(T(\boldsymbol{X})\big)$ tiene que ser 0.

En este caso, también se dice que el estadístico $T(\boldsymbol{X})$ es un <u>estadístico completo</u>.
Esta idea suele leerse, también, como: "El único estimados insesgado (pág. 3) del cero es el mismo cero".

Ejercicio:

Considere un estadístico T, con distribución Binomial(n, θ), $0 < \theta < 1$. Muestre que T corresponde a un estadístico completo.

Solución:

Primero, tenemos que llegar a un estadístico que distribuya binomial. Sabemos que dado $X \sim Bernoulli(\theta)$, entonces $\sum x_i \sim Binomial(n,\theta)$. Entonces probamos sacando un estadístico de Bernoulli:

$$x_1, \dots, x_n \sim Bernoulli(\theta)$$

$$\begin{aligned} f_\theta(x) &= \theta^{\sum x_i} \cdot (1-\theta)^{n-\sum x_i} \\ &= \left(\frac{\theta}{\theta-1}\right)^{\sum x_i} \cdot (1-\theta)^n \end{aligned}$$

Luego, por teorema de factorización (pág. 6):

$$T(X) = \sum x_i$$

y distribuye $Binomial(n,\theta)$.

Ahora, T será completo si y sólo si $\forall g$ se tiene que

$$E_\theta\big(g(T)\big) = 0$$

si y sólo si:

$$g(T) = 0 \qquad \forall\theta$$

Tomamos g arbitrario tal que:

$$E_\theta\big(g(T)\big) = 0$$

si y sólo si:

$$\sum_{t=0}^{n} g(t) \cdot P_\theta(T(\boldsymbol{X}) = t) = 0$$

Por definición de Esperanza Matemática

$$E[X] = \begin{cases} \sum_{i=1}^{n} X_i\, P(X_i) & \text{si X es una } \textbf{\textit{Variable Aleatoria Discreta}} \\ \int_{-\infty}^{+\infty} X_i\, f(X_i)\,dx & \text{si X es una } \textbf{\textit{Variable Aleatoria Continua}} \end{cases}$$

$$\sum_{t=0}^{n} g(t) \binom{n}{t} \theta^t (1-\theta)^{n-t} = 0 \qquad \boxed{T(X) \sim Binomial(n,\theta)}$$

$$(1-\theta)^n \sum_{t=0}^{n} g(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t = 0$$

Ahora, como θ no puede tomar el valor de 1 ($0 < \theta < 1$), se tiene que cumplir que la sumatoria sea igual a 0.

$$\sum_{t=0}^{n} g(t) \binom{n}{t} (r)^t = 0 \qquad \boxed{r = \frac{\theta}{1-\theta}}$$

$$g(0)\binom{n}{0} r^0 + g(1)\binom{n}{1} r^1 + \cdots + g(n)\binom{n}{n} r^n, \quad \forall\theta \to \forall r$$

Entonces todos mis coeficientes por separado tienen que ser igual a 0.
Como las combinatorias nunca serán 0, se tiene que cumplir que $\boldsymbol{g(t) = 0}$, $\forall \boldsymbol{t}$.
Dicho de otra manera:

$$P_\theta(g(T) = 0) = 1$$

Y, por lo tanto, se cumple la condición de completitud.

Teorema de Basú

Sea $T(\boldsymbol{X})$ un estadístico suficiente minimal (pág. 7) y completo (pág. 12). Entonces $T(\boldsymbol{X})$ es independiente de todo estadístico ancilar $S(\boldsymbol{X})$ (pág. 8).

Completitud en Familia exponencial

Sea $X_1, \dots, X_n$ una muestra aleatoria proveniente de una distribución en una familia exponencial, con función de densidad (probabilidad) dada por:

$$f_{\boldsymbol{\theta}}(x) = h(x) \cdot c(\boldsymbol{\theta}) \cdot \exp\left\{\sum_{j=1}^{k} \omega_j(\boldsymbol{\theta}) \cdot t_j(x)\right\},$$

donde $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$. Si el conjunto $\left\{\left(\omega_1(\boldsymbol{\theta}), \dots, \omega_k(\boldsymbol{\theta})\right) : \boldsymbol{\theta} \in \Theta\right\}$ contiene un conjunto abierto en $\mathbb{R}^k$, entonces, el estadístico:

$$T(\boldsymbol{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \dots, \sum_{i=1}^{n} t_k(X_i)\right)$$

es completo.

Ejercicio:

Sea $X_1, \ldots, X_n$ una muestra aleatoria proveniente de una distribución Normal(μ, σ^2), con σ conocido. Muestre que los estadísticos S^2 y $\bar{X}$ son independientes, donde:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \wedge \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Solución:

Utilizaremos el argumento de Basú.

Sabemos que Normal(μ, σ^2) con σ conocido es una familia exponencial (pág. 4). Debemos encontrar las $\omega_1, \ldots, \omega_k$ para revisar si sus valores posibles constituyen un conjunto abierto en $\mathbb{R}^k$.

$$f_{\boldsymbol{\mu}}(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}(x^2 - \underline{2x\mu} + \mu^2\right\}$$

Revisar pág. 5

$$\left\{\begin{array}{l} h(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}x^2\right\} \\ c(\mu) = \exp\left\{-\frac{1}{2\sigma^2}\mu^2\right\} \\ \underline{\omega_1(\mu) \cdot t_1(x) = \frac{\mu}{\sigma^2} \cdot x} \end{array}\right.$$

Y como $\mu \in \mathbb{R}$, entonces $\frac{\mu}{\sigma^2} \in \mathbb{R}$. Como $\mathbb{R}$ es abierto, se cumple la condición de Basú.

Ahora, por una parte, sabemos que:

1. $T(\boldsymbol{X}) = \bar{X}$ es suficiente minimal (pág. 7).
2. S^2 es ancilar (En toda familia de localización, al igual que el rango, S^2 es ancilar (pág. 8) para μ).

$$\Longrightarrow T(\boldsymbol{X}) \perp S^2$$
$$\bar{X} \perp S^2$$

$\perp$ = Independiente

"Un abierto es un conjunto, donde dado un punto, yo siempre me puedo abrir un poquito".

La condición es que los valores posibles de $\omega(\mu)$ formen un abierto.

$$\omega(\mu) = \frac{\mu}{\sigma^2}$$

Los valores posibles de μ son los posibles de $\mathbb{R} \rightarrow$ Pasa lo mismo con μ/σ^2.

Equivarianza frente a cambio de escala de medición:

El principio de Equivarianza bajo cambio de escala de medición indica que las inferencias basadas en observaciones que solo difieren en su escala de medida deben ser idénticas.

Equivarianza (invarianza) formal:

El principio de Equivarianza formal dice que, si asociamos a dos problemas un mismo modelo matemático, entonces debemos utilizar el mismo procedimiento de inferencia en ambos casos.

Principio de Equivarianza

Sea $Y = g(X)$ un cambio de escala en la medición de la variable aleatoria X, tal que los modelos formales para Y y X son los mismos. El principio de equivarianza exige, simultáneamente, el principio de equivarianza bajo cambio de escala de medición y el principio de equivarianza (invarianza) formal.

Ejercicio:

Considere una variable aleatoria X con distribución Binomial(n, θ), y la transformación:

$$Y = g(X) = n - X$$

Para cada uno de los siguientes estadísticos, estudie si cumple con el principio de equivarianza.

- $T_1(X) = X/n$
- $T_2(X) = \frac{9}{10}\frac{X}{n} + \frac{1}{20}$
- $T_3(X) = \frac{8}{10}\frac{X}{n} + \frac{2}{10}$

<u>Solución:</u>

Sea

$$X \sim Binomial(n, \theta)$$

Buscamos estimadores de θ que cumplan con el principio de equivarianza, bajo la transformación $Y = g(X) = n - X$.

$Y = n - X =$ número de fracasos, por eso, la Binomial de Y tendrá como segundo parámetro el $1 - \theta$, donde θ es la probabilidad de éxito.

$$X \sim Binomial(n, \theta) \Longrightarrow Y \sim Binomial(n, \delta), \quad con\ \delta = 1 - \theta$$

1) Equivarianza bajo cambio de escala

$$\hat{\theta} = T(X) \qquad\qquad \begin{aligned} \hat{\delta} &= T^*(Y) \\ \widehat{1-\theta} &= T^*(Y) \\ \hat{\theta} &= 1 - T^*(Y) \end{aligned}$$

El concepto dice que:

$$\begin{aligned} T(X) &= 1 - T^*(Y) \\ &= 1 - T^*(n - X) \qquad (1) \end{aligned}$$

2) Como ambos tienen distribuciones binomiales, entonces tienen el mismo modelo matemático (condición de equivarianza formal).

$$T(X) \Longrightarrow T^* = T$$

En (1):

$$\boxed{T(X) = 1 - T(n - x)}$$

Esta sería la condición para cualquier estadístico que cumple el principio de equivarianza.

<u>Propuesta 1:</u>

$$T(X) = \frac{X}{n}$$

$x \in \{0,1, \dots, n\}$.
Lado izquierdo de la condición:

$$T_1(X) = \frac{X}{n}$$

Lado derecho de la condición:

$$1 - T_1(n - X) = 1 - \frac{n - X}{n} = \frac{X}{n} = T_1(X)$$

$\therefore \mathit{Si\ cumple}.$

Propuesta 2:
Lado izquierdo:

$$T_2(X) = \frac{9}{10}\frac{X}{n} + \frac{1}{20}$$

Lado derecho:

$$= 1 - \left(\frac{9}{10}\frac{n - X}{n} + \frac{1}{20}\right)$$
$$= \frac{9}{10}\frac{X}{n} + \frac{1}{20} = T_2(X)$$

$\therefore \mathit{Si\ cumple}.$

Propuesta 3:
Lado izquierdo:

$$T_3(X) = \frac{8}{10}\frac{X}{n} + \frac{2}{20}$$

Lado derecho:

$$= 1 - \left(\frac{8}{10}\frac{n - X}{n} + \frac{2}{20}\right)$$
$$= \frac{8}{10}\frac{X}{n} \neq T_2(X)$$

$\therefore \mathit{NO\ cumple}.$

Grupos de transformaciones

Un conjunto de funciones, $\{g(x), g \in \mathcal{g}\}$, con dominio y recorrido el espacio muestral $\mathcal{X}$, se dice un grupo de transformaciones de $\mathcal{X}$ si y sólo si cumple:

i. Inverso: Para toda función $g \in \mathcal{g}$, existe una función $g' \in \mathcal{g}$ tal que:

$$g'\big(g(x)\big) = x, \quad \forall x \in \mathcal{X}$$

ii. Composición: Para cada par de funciones g y $g' \in \mathcal{g}$, existe una función $g'' \in \mathcal{g}$ tal que:

$$g'\big(g(x)\big) = g''(x), \quad \forall x \in \mathcal{X}$$

iii. Identidad: La función identidad, $g(x) = x, x \in \mathcal{X}$ está en $\mathcal{g}$.

Invarianza de una familia de modelos bajo un grupo de transformaciones

Sea:

$$\mathcal{F} = \{f_\theta(x); \theta \in \Theta\}$$

un conjunto de funciones de densidad (probabilidad) para una variable aleatoria X, y sea $\mathcal{g}$ un grupo de transformaciones del espacio muestral $\mathcal{X}$. Entonces el conjunto $\mathcal{F}$ es invariante bajo el grupo $\mathcal{g}$ si y sólo si para todo $\theta \in \Theta$ y $g \in \mathcal{g}$, existe un único valor $\theta' \in \Theta$ tal que:

$$Y = g(X)$$

tiene distribución $f_{\theta'}(y)$, cuando X tiene distribución $f_\theta(x)$.

Ejercicio:

Considere la familia de distribuciones $\mathcal{F}$ para una muestra aleatoria $\boldsymbol{X}$ de una población con distribución Normal(μ, σ^2), $\mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$. Muestre que $\mathcal{F}$ es invariante bajo el grupo de transformaciones $\mathcal{g} = \{g_a(x) = x + a,\ \ a \in \mathbb{R}\}$.

Solución:

$$\mathcal{F} = \left\{ f_\theta(x) = (2\pi\sigma^2)^{-\frac{n}{2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \sum (x_i - \mu)^2 \right\}, \mathfrak{u} \in \mathbb{R} \times \mathbb{R}^+ \right\}$$

Con $\mathfrak{u} =$ Espacio paramétrico.

$\boldsymbol{\theta} = (\mu, \sigma^2)$
$\mathcal{g} = \{g_a(x) \mid\ g_a(x) = x + a = (x_1 + a, \dots, x_n + a),\ \ a \in \mathbb{R}\}$

Previo

1) g_a es inverso de $g - a$, $\ \forall a \in \mathbb{R}$.
2) $(g_a \circ g_b)(x) = g_a(x + b) = x + (a + b) = g_{a+b}(x), \quad a + b \in \mathbb{R}$.
3) Identidad: $a = 0$.

Por lo tanto, $\mathcal{g}$ es un grupo de transformaciones.

Ahora,

$$X \sim Normal(\mu, \sigma^2) \Longrightarrow X + a \sim Normal(\mu + a, \sigma^2)$$

$$f_{\theta'}(y) = (2\pi\sigma^2)^{-n/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}(y_i - \mu')^2\right\}$$

Donde $\mu' = \mu + a$

$$\theta' = (\mu', \sigma' = \sigma) \in \mathbb{R} \times \mathbb{R}^+$$

Y la distribución de $X + a \in \mathcal{F}$.

Entonces, $\mathcal{F}$ es invariante bajo $\mathcal{g}$.

Ejercicio:

Suponga que, en el ejercicio anterior, se desea estimar μ.

1. Encuentre características que debe tener un estadístico para cumplir con el principio de equivarianza.
2. Estudie si el estadístico $T_1(\boldsymbol{X}) = \bar{X}$ cumple con dicho principio.
3. Repita para $T_2(\boldsymbol{X}) = \frac{1}{2}\left(X_{(1)} + X_{(n)}\right)$.

<u>Solución:</u>

1)

$$\mathcal{g} = \{g_a(x), a \in \mathbb{R}\}$$

$$\mathcal{F} = \{(Normal(\mu, \sigma^2) \times \cdots \times Normal(\mu, \sigma^2), \qquad (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+\}$$

I. Equivarianza bajo transformación $g_a \in \mathcal{g}$.

$$\boldsymbol{X} \sim Normal_n(\mu \perp, \sigma^2 \cdot I) \rightarrow \hat{\mu} = T(\boldsymbol{X})$$

$\perp$ = Vector de 1s

I = Identidad.

$$\boldsymbol{Y} = g_a(x), \quad a \in \mathbb{R}$$

$$\boldsymbol{Y} \sim Normal_n\big((\mu + a) \perp, \sigma^2 \cdot I\big) \rightarrow \widehat{\mu + a} = T^*(\boldsymbol{Y})$$

$$\hat{\mu} = Y^*(\boldsymbol{Y})$$

Buscamos igualar los $\hat{\mu}$:

$$\boxed{T(\boldsymbol{X}) = T^*(\boldsymbol{Y}) - a}$$

II. Equivarianza formal $T = T^*$.

$$T(\boldsymbol{X}) = T(\boldsymbol{Y}) - a$$
$$T(\boldsymbol{X}) = T(\boldsymbol{X} + a) - a$$

Entonces $\forall \boldsymbol{x} \in \mathbb{R}^n$:

$$\boxed{T(\boldsymbol{X}) = T(\boldsymbol{X} + a) - a}$$

2) $\bar{X}$

Lado izquierdo de la condición:

$$T(\boldsymbol{X}) = \bar{X}$$

Lado derecho de la condición:

$$\begin{aligned} T(\boldsymbol{X} - a) - a &= \overline{\boldsymbol{X} + a} - a \\ &= \bar{X} + a - a \\ &= \bar{X} \end{aligned}$$

$\therefore \bar{X}$ cumple con el principio de equivarianza bajo $\mathcal{g}$.

3)

Lado izquierdo de la condición:

$$T(\boldsymbol{X}) = \frac{1}{2}\left(x_{(1)} + x_{(n)}\right)$$

Lado derecho de la condición:

$$\begin{aligned} T(\boldsymbol{X} + a) - a &= \frac{1}{2}\left(y_{(1)} + y_{(n)}\right) - a \\ &= \frac{1}{2}\left(\left(x_{(1)} + a\right) + \left(x_{(n)} + a\right)\right) - a \\ &= \frac{1}{2}\left(x_{(1)} + x_{(n)}\right) \end{aligned}$$

$$\boxed{y = \boldsymbol{X} + a}$$

$\therefore \frac{1}{2}\left(x_{(1)} + x_{(2)}\right)$ cumple con el principio de equivarianza bajo $\mathcal{g}$.

Estimación Intervalar:

Una estimación intervalar para un parámetro real valorado θ corresponde a cualquier par de funciones $L(x_1, \dots, x_n)$ y $U(x_1, \dots, x_n)$ de una muestra x, que satisface:

$$L(x) \leq U(x)$$

para todo $x \in \mathcal{X}$. El intervalo aleatorio $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ se denomina un estimador intervalar para θ.

Probabilidad de cobertura:

Para un estimador intervalar $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ de un parámetro θ, se define su probabilidad de cobertura como la probabilidad de contener al verdadero valor de θ, es decir:

$$P_\theta(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})]$$

o, dicho de otra manera:

$$P_\theta\big((L(\boldsymbol{X}) \leq \theta) \cap (U(\boldsymbol{X}) \geq \theta)\big)$$

donde $L(\boldsymbol{X})$ y $U(\boldsymbol{X})$ son aleatorios.

Ejercicio:

Sea $X_1, \dots, X_n$ una muestra aleatoria proveniente de una distribución Uniforme$(0, \theta)$. Suponga que interesa saber un intervalo de confianza para θ, y se proponen los intervalos:

$$(1) \quad \big[aX_{(n)}, bX_{(n)}\big], \;\; 1 \leq a \leq b$$
$$(2) \quad \big[X_{(n)} + c, X_{(n)} + d\big], \;\; 0 \leq c \leq d$$

Obtenga la probabilidad de cobertura de cada uno de los intervalos propuestos.

Solución Intervalo (1):

$$X_1, \dots, X_n \sim Uniforme(0, \theta)$$

Tomamos a, b fijos, con $1 \leq a \leq b$.

Probabilidad de cobertura

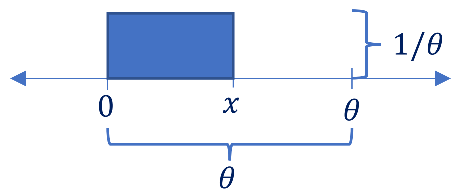
$$\begin{aligned}
&P_\theta\big(aX_{(n)} \leq \theta \leq bX_{(n)}\big) \\
&= P_\theta\left(a \leq \frac{\theta}{X_{(n)}} \leq b\right) \\
&= P_\theta\left(b \leq \frac{X_{(n)}}{\theta} \leq a\right) \\
&= P_\theta\left(\frac{\theta}{b} \leq X_{(n)} \leq \frac{\theta}{a}\right) \\
&= P_\theta\left(X_{(n)} \leq \frac{\theta}{a}\right) - P_\theta\left(X_{(n)} \leq \frac{\theta}{b}\right) = F_{X_{(n)}}\left(\frac{\theta}{a}\right) - F_{X_{(n)}}\left(\frac{\theta}{b}\right)
\end{aligned}$$

Sea

$$F_{X_{(n)}}(x) = P_\theta\big(X_{(n)} \leq x\big) = P_\theta(X_1 \leq x, \dots, X_n \leq x)$$

$$= \prod_{i=1}^{n} P_\theta(X_i \leq x) = \prod_{i=1}^{n} x \cdot \frac{1}{\theta} = \left(\frac{x}{\theta}\right)^n$$

Como es Uniforme$(0, \theta)$:



Entonces, $P(X_i \leq x)$ tiene que ser $x \cdot \frac{1}{\theta}$

Ahora, retomando:

$$= P_\theta\left(X_{(n)} \leq \frac{\theta}{a}\right) - P_\theta\left(X_{(n)} \leq \frac{\theta}{b}\right)$$

$$= \left(\frac{\theta/a}{\theta}\right)^n - \left(\frac{\theta/b}{\theta}\right)^n$$

$$P\ de\ cobertura = \boxed{\left(\frac{1}{a}\right)^n - \left(\frac{1}{b}\right)^n}$$

Nota: En este caso, la probabilidad de cobertura NO depende de θ.

<u>Solución Intervalo (2)</u>:

$$X_1, \dots, X_n \sim Uniforme(0, \theta)$$

Tomamos c, d fijos, con $0 \leq c \leq d$.

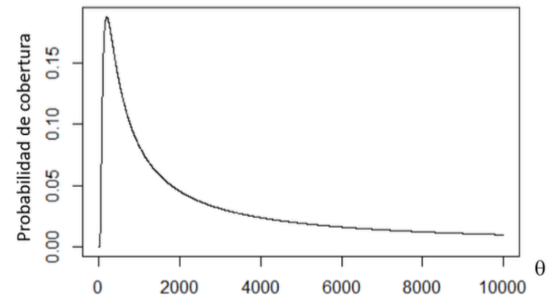
$$\begin{aligned}
&P_\theta\left(X_{(n)} + c \leq \theta \leq X_{(n)} + d\right) \\
&= P_\theta\left(c \leq \theta - X_{(n)} \leq d\right) \\
&= P_\theta\left(c - \theta \leq -X_{(n)} \leq d - \theta\right) \\
&= P_\theta\left(\theta - c \geq X_{(n)} \geq \theta - d\right) \\
&= P_\theta\left(\theta - d \leq X_{(n)} \leq \theta - c\right) \\
&= F_{X_{(n)}}(\theta - c) - F_{X_{(n)}}(\theta - d) \\
&= \left(\frac{\theta - c}{\theta}\right)^n - \left(\frac{\theta - d}{\theta}\right)^n \\
&= \left(1 - \frac{c}{\theta}\right)^n - \left(1 - \frac{d}{\theta}\right)^n
\end{aligned}$$

Probabilidad de cobertura

Nota: Ahora, mi intervalo si depende de θ, algo que nos molesta.

Como vemos, a medida que θ crece, nuestra probabilidad de cobertura podría ser muy baja, por lo que no nos interesa realmente tomar este intervalo.

Cobertura del intervalo $[X_{(n)} + c, X_{(n)} + d]$:



Coeficiente de confianza

Para un estimador intervalar $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ de un parámetro θ, se define su coeficiente de confianza como el ínfimo de las probabilidades de cobertura, es decir:

$$\inf_{\theta} P_\theta(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})])$$

inf = Ínfimo = Mayor cota inferior.

Ejercicio:

En el problema anterior, encuentre el coeficiente de confianza de cada estimador propuesto.

Solución:

(1) $P_\theta\left(aX_{(n)} \leq \theta \leq bX_{(n)}\right)$ no depende de θ, luego, es igual al coeficiente de confianza: $\left(\frac{1}{a}\right)^n - \left(\frac{1}{b}\right)^n$.

(2) $\inf_{\theta}\left\{\left(1-\frac{c}{\theta}\right)^n - \left(1-\frac{d}{\theta}\right)^n\right\} = \lim_{\theta\to\infty}\left\{\left(1-\frac{c}{\theta}\right)^n - \left(1-\frac{d}{\theta}\right)^n\right\} = 0$. Por lo tanto, no debiésemos utilizar este intervalo.

Conjunto de confianza

Para un parámetro cualquiera θ, se define un conjunto de confianza $1-\alpha$ como un conjunto $C(\boldsymbol{X})$ tal que:

$$\inf_{\theta} P_\theta\left(\theta \in C(\boldsymbol{X})\right) = 1 - \alpha$$

Pivote para un parámetro

Una variable aleatoria $Q(\boldsymbol{X}, \theta)$ corresponde a un pivote para θ si, y sólo si su distribución NO depende de ningún parámetro desconocido. Es decir, la distribución $Q(\boldsymbol{X}, \theta)$ es la misma para todo θ.

Ejercicio:

Encuentre pivotes en las siguientes situaciones:

1. Familia de distribuciones de localización, con parámetro de localización θ.
2. Familia de distribuciones de escala, con parámetro de escala σ.
3. Familia de distribuciones de localización y escala, con parámetros θ y σ.

Solución:

1. Familia de distribuciones de localización, con parámetro de localización θ:

Sea f una densidad arbitraria que no depende de parámetros desconocidos.
Sea $\theta \in \mathbb{R}$ y se define $f_\theta(x) = f(x - \theta)$.
Ejemplo:

$$f(x) = N(0,1) = \frac{1}{(2\pi)^{1/2}} \cdot \exp\left\{-\frac{1}{2}x^2\right\}$$

$\theta \in \mathbb{R}$

$$f_\theta(x) = f(x - \theta) = \frac{1}{(2\pi)^{1/2}} \cdot \exp\left\{-\frac{1}{2}(x - \theta)^2\right\}$$

Recordamos

Al evaluarla en $1 - \theta$, ahora la distribución sí depende de θ, formando una familia:

$$\{f_\theta(x) = Normal(\theta, 1), \qquad \theta \in \mathbb{R}\}$$

es familia de distribución de localización.

Lo que necesitamos:

$Q(\boldsymbol{X}, \theta)$ con distribución que no dependa de θ.

Entonces, ¿Cómo podemos, basándonos en $\boldsymbol{x}$ y $\boldsymbol{\theta}$ construir un $\boldsymbol{Q}$ que no dependa de θ?

ACLARACIÓN: f_θ = Función que depende de θ. f = Función que NO depende de θ.

Sabemos que $\underbrace{x-\theta}$ $\sim$ Normal$(0{,}1)$

$Q(\boldsymbol{X},\theta)$ es pivote para θ, pues su distribución no depende de θ.

Entonces es pivote para la Normal$(0{,}1)$.
Veamos ahora para el <u>caso general</u>:

$$f(\,)\ \text{No depende de } \theta$$
$$x \sim f_\theta(x) = f(x-\theta)$$

$z = x - \theta$
$x = z + \theta$
$\left|\frac{dx}{dz}\right| = |1|$

$$\begin{aligned} f_Z(z) &= \left|\frac{dx}{dz}\right| f_\theta(x) \\ &= 1 \cdot f_\theta(z+\theta) \\ &= f((z+\theta)-\theta) \\ &= f(z)\ \ \text{No depende de } \theta \end{aligned}$$

"Teorema del Jacobiano"

Entonces, si x está en familia de distribución de localización de parámetro $\theta \in \mathbb{R}$, $(x-\theta)$ será pivote para θ.

Hasta ahora, hemos trabajado con pivote para una sola observación. A continuación, veremos pivotes para conjunto de muestras:

$$x_1, \dots, x_n \sim f_\theta(x_i) = f(x_i - \theta)$$

Tenemos que la distribución de $X_i - \theta$ no depende de θ (Demostrado anteriormente).
Por lo tanto:

$$\sum(X_i - \theta)\ \text{No depende de } \theta$$

Entonces, tampoco dependen de θ:

$$\begin{aligned} &= \sum X_i - n\theta \\ &= \bar{X} - \theta. \end{aligned}$$

$$\therefore\ \bar{X} - \theta = Q(\boldsymbol{X},\theta) \text{ es pivote para } \theta.$$

Como Q no depende de θ, podemos decir que tiene una distribución F.

Siempre que tenemos un pivote, podemos escribir:

$$P\left(q_{\alpha/2} \leq Q(\boldsymbol{X}, \theta) \leq q_{1-\alpha/2}\right) = 1 - \alpha$$
$$P\left(q_{\alpha/2} \leq \bar{X} - \theta \leq q_{1-\alpha/2}\right) = 1 - \alpha$$
$$P\left(\underbrace{\bar{X} + q_{1-\alpha/2}}_{L(\boldsymbol{X})} \leq \theta \leq \underbrace{\bar{X} + q_{\alpha/2}}_{U(\boldsymbol{X})}\right) = 1 - \alpha$$

Entonces, nuestra probabilidad de cobertura es igual a $1 - \alpha$ y NO depende de θ, por lo que el coeficiente de confianza también es $1 - \alpha$.

Si:

$$L(\boldsymbol{X}) = \bar{X} + q_{1-\alpha/2}$$
$$U(\boldsymbol{X}) = \bar{X} + q_{\alpha/2}$$

Entonces:

$$P\left(L(\boldsymbol{X}) \leq \theta \leq U(\boldsymbol{X})\right) = 1 - \alpha$$

Y, $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ es intervalo con probabilidad de cobertura $(1 - \alpha)$ para θ. Dado que no depende de θ, también es intervalo de confianza, pues:

$$\inf_{\theta} P\left(L(\boldsymbol{X}) \leq \theta \leq U(\boldsymbol{X})\right) = 1 - \alpha$$

2. Familia de distribuciones de escala, con parámetro de escala σ:

Al igual que en una familia de localización, existe una función distribución f que no depende de ningún parámetro desconocido.

Tomamos $\sigma > 0$.

Entonces, ahora definimos una función de probabilidad que dependa de σ, pero, que va a estar basado en esta distribución f.

$$\left\{f_\sigma(x) = \frac{1}{\sigma} f\left(\frac{x}{\sigma}\right), \quad \sigma > 0\right\}$$

Familia de distribuciones de escala.

<u>Ejemplo</u>:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2} x^2\right\}$$

$$f_\sigma(x) = \frac{1}{\sigma} f\left(\frac{x}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2\sigma^2} x^2\right\}$$

Entonces, la familia

$$\{Normal(0, \sigma^2), \quad \sigma > 0\}$$

Es una familia de escala.

Ahora, proponemos:

$$\frac{x}{\sigma} = u \Longrightarrow x = u \cdot \sigma \Longrightarrow \left|\frac{dx}{du}\right| = \sigma$$

Teorema de cambio de variable:

$$\begin{aligned} f_U(u) &= \sigma \cdot f_\sigma(x) \\ &= \sigma \cdot \frac{1}{\sigma} \cdot f\left(\frac{x}{\sigma}\right) \\ &= f\left(\frac{u \cdot \sigma}{\sigma}\right) \\ &= f(u) \end{aligned}$$

$$\boxed{U = \frac{x}{\sigma}} = Q(X, \sigma) \sim f \longrightarrow$$

No depende de ningún parámetro desconocido.

Luego, x/σ es pivote para σ.

Si tenemos:

$$X_1, \dots, X_n \sim f_\sigma$$

$$\Longrightarrow \boxed{\frac{X_i}{\sigma} \ \text{ pivote para } i = 1, \dots, n}$$

Usualmente se toma:

$$\frac{\sum x_i^2}{\sigma^2}$$

$$P\left(q_{\alpha/2} < \frac{\sum x_i^2}{\sigma^2} < q_{1-\alpha/2}\right) = 1 - \alpha$$

$$P\left(\underbrace{\frac{\sum x_i^2}{q_{1-\alpha/2}}}_{L(\boldsymbol{X})} < \sigma^2 < \underbrace{\frac{\sum x_i^2}{q_{\alpha/2}}}_{U(\boldsymbol{X})}\right) = 1 - \alpha$$

Sea el intervalo aleatorio $[L(\boldsymbol{X}), U(\boldsymbol{X})]$. Su probabilidad de cobertura:

$$P\big(L(\boldsymbol{X}) < \sigma^2 < U(\boldsymbol{X})\big) = 1 - \alpha$$

Coeficiente de confianza

$$\inf_{\sigma>0} P\big(L(\boldsymbol{X}) < \sigma^2 < U(\boldsymbol{X})\big)$$
$$\inf_{\sigma>0}(1-\alpha) = 1-\alpha$$

Ejercicio:

Encuentre un pivote para θ en la distribución Exponencial de media $1/\theta$.

Solución:

$$X_1, \dots, X_n \sim Exp(\theta)$$

Primero mostramos que sea familia de escala:

$$f(x) = e^{-x}, \quad x > 0$$
$$\{f_\sigma(x) = \frac{1}{\sigma} e^{-x/\sigma}, \qquad \sigma > 0\}$$

$\therefore$ Exponencial$(1/\sigma)$ es familia de escala.

Tenemos que: $\frac{x_i}{\sigma}$ es pivote para σ.

Tomamos:

$$\frac{\sum x_i}{\sigma} \text{ pivote para } \sigma$$

Luego,

$$\underbrace{\sum x_i}_{U} \sim Gamma\left(n, \frac{1}{\sigma}\right)$$

$$V = \frac{U}{\sigma} \Longrightarrow u = \sigma \cdot v \Longrightarrow \left|\frac{du}{dv}\right| = \sigma$$

$$f_\sigma(u) = \left(\frac{1}{\sigma}\right)^n \frac{u^{n-1} \cdot e^{-u/\sigma}}{\Gamma(n)}, \quad u > 0$$

$$f_V(v) = \frac{\sigma \cdot \sigma^{-n} (\sigma_v)^{n-1} \cdot e^{-v}}{\Gamma(n)}, \quad v > 0$$

$$= \frac{1}{\Gamma(n)} \cdot v^{n-1} \cdot e^{-v}, \quad v > 0$$

$$V = \frac{\sum x_i}{\sigma} \sim Gamma(n, 1)$$

$$P\left(q_{\alpha/2} < \frac{\sum x_i}{\sigma} < q_{1-\alpha/2}\right) = 1-\alpha$$

$$P\left(\underbrace{\frac{\sum x_i}{q_{1-\alpha/2}}}_{L(X)} < \sigma < \underbrace{\frac{\sum x_i}{q_{\alpha/2}}}_{U(X)}\right) = 1-\alpha$$

$$q_{1-\alpha/2} = q\ Gamma(1-\alpha/2\,, n, 1)$$
$$q_{\alpha/2} = q\ Gamma(\alpha/2\,, n, 1)$$

Usando $\sigma = 1/\theta$ es posible obtener un intervalo de confianza $(1-\alpha)$ para θ.

Caso general:

Considere un estadístico $T(\boldsymbol{X})$ con función de densidad f_θ. En general, se busca factorizar su densidad en la forma:

$$f_\theta(t) = g\big(Q(t,\theta)\big)\left|\frac{\partial Q(t,\theta)}{\partial t}\right|,$$

con Q función monótona en t, para cada valor de θ.

Para encontrar un intervalo de$(1-\alpha)100\%$ de confianza, es posible plantear la igualdad.

$$P_\theta(a \le Q(t,\theta) \le b) \ge 1-\alpha$$

Teorema:

Sea $T(\boldsymbol{X})$ una variable aleatoria **continua** con función de distribución F_θ.
Sea $0 < \alpha < 1$ un valor constante. Para todo $t \in \mathrm{T}$, se definen $\theta_L(t)$ y $\theta_U(t)$ tales que:

1. Si F_θ es decreciente en θ para cada t.

$$F_{\theta_{U(t)}}(t) = \frac{\alpha}{2} \qquad\qquad F_{\theta_{L(t)}}(t) = 1-\frac{\alpha}{2}$$

2. Si F_θ es creciente en θ para cada t.

$$F_{\theta_{U(t)}}(t) = 1-\frac{\alpha}{2} \qquad\qquad F_{\theta_{L(t)}}(t) = \frac{\alpha}{2}$$

Entonces $[\theta_{L(t)}, \theta_{U(t)}]$ corresponde a un intervalo de confianza $(1-\alpha)100\%$ para θ.

Teorema:

Sea $T(\boldsymbol{X})$ una variable aleatoria **discreta** con función de distribución F_θ.
Sea $0 < \alpha < 1$ un valor constante. Para todo $t \in \mathrm{T}$, se definen $\theta_L(t)$ y $\theta_U(t)$ tales que:

1. Si F_θ es decreciente en θ para cada t.

$$P_{\theta_{U(t)}}(T \leq t) = \frac{\alpha}{2} \qquad\qquad P_{\theta_{L(t)}}(T \geq t) = 1 - \frac{\alpha}{2}$$

2. Si F_θ es creciente en θ para cada t.

$$P_{\theta_{U(t)}}(T \geq t) = 1 - \frac{\alpha}{2} \qquad\qquad P_{\theta_{L(t)}}(T \leq t) = \frac{\alpha}{2}$$

Entonces $[\theta_{L(t)}, \theta_{U(t)}]$ corresponde a un intervalo de confianza $(1-\alpha)100\%$ para θ.

Propuestos:

1. $X_1, \ldots, X_n \sim Normal(\theta, \sigma^2)$ con σ^2 desconocido. Encontrar un intervalo de confianza $(1-\alpha)$ para θ considerando que corresponde a una familia de localización.
2. $X_1, \ldots, X_n \sim Normal(\theta, \sigma^2)$ con σ^2 desconocido. Encontrar un intervalo de confianza $(1-\alpha)$ para θ.
3. En los casos anteriores, encuentre intervalos de confianza $(1-\alpha)$ de la forma $(-\infty, U(X)]$ y $[L(X), +\infty)$.

Modelo de Regresión Lineal simple:

El modelo de regresión lineal simple para explicar el comportamiento de una variable respuesta, Y, en términos de un predictor x, para n pares de observaciones de la forma $(x_i, Y_i), i = 1, \ldots, n$ corresponde a:

$$Y_i \sim Normal(\beta_0 + \beta_1 x_i, \sigma^2)$$

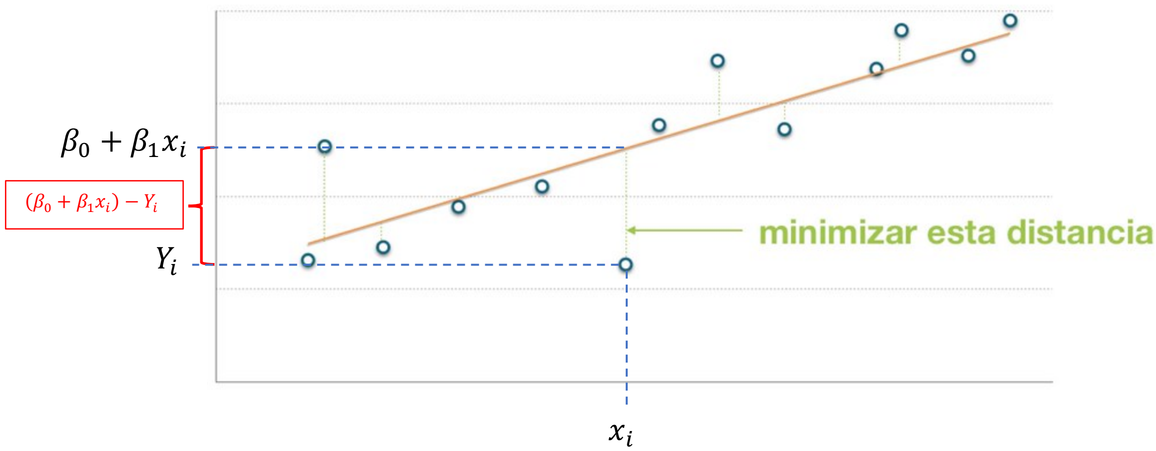
con $Y_1, \ldots, Y_n$ independientes.

Formulación alternativa:

El modelo puede ser reescrito como:

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\epsilon_i \sim Normal(0, \sigma^2), \quad i = 1, \dots, n$$



Estimadores de Mínimos Cuadrados Ordinarios de β_0 y β_1:

Como me interesan todas las distancias, buscaremos minimizar la sumatoria de éstas.

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \epsilon_i^2$$
$$= \sum_{i=1}^{n} \left(Y_i - (\beta_0 + \beta_1 x_i)\right)^2$$

Notar que elevamos al cuadrado para eliminar los signos negativos, solo nos interesa la distancia.

Para resolver esto, derivamos con respecto a β_0 y β_1 e igualamos a 0:

$$\left.\begin{aligned} \frac{\partial S}{\partial \beta_0} &= -2\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 x_i) = 0 \\ \frac{\partial S}{\partial \beta_1} &= -2\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 x_i)x_i = 0 \end{aligned}\right\} \text{Punto crítico}$$

Luego, se calcula la matriz Hessiana y mostramos que sea definida positiva (poque estamos minimizando) en el punto crítico.

Resoviendo este problema de minimización de manera analítica, se obtienen las expresiones para los valores buscados:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \qquad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{x}$$

Observaciones:

1) Alrededor de 1920, Fischer propuso la estimación máxima versímil. También, se da cuenta que los parámetros MV son:

$$\hat{\beta}_0^{MV} = \hat{\beta}_0$$
$$\hat{\beta}_1^{MV} = \hat{\beta}_1$$

$$\sigma_{MV}^2 = \frac{\sum_{i=1}^{n}\left(Y_i - (\beta_0 + \beta_1 x_i)\right)^2}{n}$$

2) Pero σ_{MV}^2 no es insesgado, entonces se toma uno corregido:

$$\hat{\sigma}^2 = \sigma_M^2 \cdot \frac{n}{n-2}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))}{n-2}$$

El $n-2$ viene de que estamos estimando 2 parámetros, β_0 y β_1.

Este último, viene siendo el estimador de σ^2 de mínimos cuadrados, es decir:

$$E(\hat{\sigma}^2) = \sigma^2$$

Ahora, si en mi gráfico todos los Y_i parecen estar dentro del mismo rango, es decir que para cualquier valor de x_i el valor de Y_i siempre estará en el mismo rango, podemos pensar que Y_i no depende de x_i. En este caso, el valor de β_1 debe ser igual a 0. Sin embargo, por ruido, podemos obtener valores de β_1 cercanos a 0, pero no el 0 mismo, lo que podría ser un erroe al interpretar el gráfico.

Es por esto que nos preguntamos,

$$¿\beta_1 \neq 0?$$

Para resolver esto, lo ideal es usar un test de hipótesis, pero con las herramientas que tenemos nosotros vamos a hacer un intervalo de confianza. Aquí, revisaremos a si el 0 está en el intervalo, o no. Si 0 está en el intervalo, no podríamos descartar que el valor de β_1 sea igual a 0. Por otro lado, si el intervalo no contiene al 0, entonces concluimos que β_1 es distinto de 0, y existiría una regresión.

La estrategia que tenemos, es la estrategia del pivote. Entonces necesitamos un pivote para β_1. ¿Qué era un pivote? Revisar página 25.

$$Q(\boldsymbol{Y}, \beta_1) \sim F \text{ no depende de parámetros desconocidos}$$

Como $\hat{\beta}_1 = S_{xy}/S_{xx}$:

Se puede demostrar (raramente) que:

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(Y_i - \bar{Y}) = \sum_{i=1}^{n} (x_i - \bar{x}) Y_i$$

Luego,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) Y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 = \sum_{i=1}^{n} \underbrace{\left(\frac{x_i - \bar{x}}{S_{xx}}\right)}_{cte} \cdot Y_i = \sum_{i=1}^{n} d_i \cdot Y_i \Longrightarrow \text{Combinación lineal de Normales}$$

$$\hat{\beta}_1 \sim Normal$$

$$E(\hat{\beta}_1) = E\left(\sum_{i=1}^{n} d_i \cdot Y_i\right) = \sum_{i=1}^{n} E(d_i \cdot Y_i) = \sum_{i=1}^{n} d_i \cdot E(Y_i)$$

$$= \sum_{i=1}^{n} d_i \cdot (\beta_0 + \beta_1 x_i) = \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{S_{xx}}\right) \cdot (\beta_0 + \beta_1 x_i)$$

$$= \beta_0 \frac{\overbrace{\sum_{i=1}^{n} (x_i - \bar{x})}^{0}}{S_{xx}} + \beta_1 \frac{\overbrace{\sum_{i=1}^{n} (x_i - \bar{x}) x_i}^{S_{xx}}}{S_{xx}}$$

$$= \boxed{\beta_1} \Longrightarrow \text{Es insesgado}$$

$$\boxed{S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i - \bar{x}) x_i}$$

$$Var(\hat{\beta}_1) = Var\left(\sum_{i=1}^{n} d_i \cdot Y_i\right) = \cdots = \frac{\sigma^2}{S_{xx}}$$

$$\hat{\beta}_1 \sim Normal\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$$

Tratamos de hacer el pivote estandarizando:

$$\frac{\hat{\beta}_1 - \beta_1}{\sigma/\sqrt{S_{xx}}} \sim Normal(0,1)$$

$$(\hat{\beta}_1 - \beta_1)\sqrt{S_{xx}} \sim Normal(0, \sigma^2)$$

¿Es esto igual al pivote $Q(\hat{\beta}_1, \beta_1)$ que buscamos?

NO, no es igual al pivote, ya que depende de σ, el cual es un parámetro desconocido.

¿Qué podemos hacer para solucionarlo?

$$\boxed{\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}/\sqrt{S_{xx}}} \sim t_{n-2}}$$

Este pivote, si cumple.

$$P\left(t_{n-2}^{\alpha/2} < Q(\boldsymbol{Y}, \beta_1) < t_{n-2}^{1-\alpha/2}\right) = 1 - \alpha$$

Por simetría se t_n:

$$P\left(-t_{n-2}^{1-\alpha/2} < \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}/\sqrt{S_{xx}}} < t_{n-2}^{1-\alpha/2}\right) = 1 - \alpha$$

$$\hat{\beta}_1 - \frac{t_{n-2}^{1-\alpha/2}\hat{\sigma}}{\sqrt{S_{xx}}} < \beta_1 < \hat{\beta}_1 + \frac{t_{n-2}^{1-\alpha/2}\hat{\sigma}}{\sqrt{S_{xx}}}$$

Con 95% de confianza, β_1 está entre $-0{,}55$ y $-0{,}26$. Dado que no contiene el valor de $\beta_1 = 0$, concluimos que el pH aporta para explicar el comportamiento de la concentración de Hg en el tejido de los peces.

- Se estima que la <u>media</u> del contenido de Hg en los peces de un lago cuyo pH es 1 unidad mayor que otro es 0.40 unidades menor (ppm, escala log).
- Se estima que, para un lago con un pH de 6, el contenido medio de Hg (log) corresponde a:

$$\hat{\beta}_0 + \hat{\beta}_1 \cdot 6 = -0,67,$$

es decir, a:

$$\exp\{-0,67\} = 0,51 \text{ ppm}$$

Pivotes

Funciones en donde su distribución no depende del parámetro.

Ejemplo común:

Sea $X_1, \dots, X_n \sim N(\mu, \sigma^2)$:

$$Q_1(\boldsymbol{X}, \mu) = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1) \qquad \sigma \text{ conocido}$$

$$Q_2(\boldsymbol{X}, \mu) = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{n-1} \qquad \sigma \text{ desconocido}$$

$$Q_3(\boldsymbol{X}, \sigma^2) = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$