ESTIMACIÓN DE DISTRIBUCIÓN DE PROBABILIDAD

ESTIMACIÓN DE PARÁMETROS

Notación:

- Individuo de interés: parámetros de distribución

$$\rightarrow f(x;\theta) = \begin{cases} \text{función de densidad cuando el parámetro es } \theta \\ \text{función de masa cuando el parámetro es } \theta \end{cases}$$

$\rightarrow$ $\Theta =$ conjunto de parámetros posibles
espacio de parámetros

$\rightarrow$ si $\theta \in \Theta$, el $sop_\theta = \{x : f(x;\theta) > 0\}$

Ejemplo

(1) X es Bernoulli(p)

$$f(x;p) = \begin{cases} p, & \text{si } x=1 \\ 1-p, & \text{si } x=0 \end{cases}$$

si $x \neq 0,1 \longrightarrow f(x;p) = 0$

$\Theta = [0,1]$

$$sop_p = \begin{cases} \{0,1\} & \text{si } p\neq 0 \text{ y } p \neq 1 \\ 0 & \text{si } p=0 \\ 1 & \text{si } p=1 \end{cases}$$

(2) X es Exp(λ)

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x>0 \\ 0 & x \leq 0 \end{cases}$$

$\Theta = (0,\infty)$

$\lambda \in \Theta$

$sop_\lambda = (0,\infty)$

(3) X es unif$(0,a)$

$$f(x;a) = \begin{cases} 0 & \text{si } x\leq 0 \\ 1/a & \text{si } x\in(0,a) \\ 0 & \text{si } x \geq a \end{cases}$$

$\Theta = (0,\infty)$

$a \in \Theta$

$sop_a = (0,a)$

(4) $\Theta = \{0,1\}$ ...

$$\longrightarrow f(x;0) = \begin{cases} 1 & \text{si } x\in(0,1) \\ 0 & \text{en el resto} \end{cases}$$

$$\searrow f(x;1) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{si } x\in(0,1) \\ 0 & \text{en el resto} \end{cases}$$

$$sop_\theta = (0,1)$$

Más notación

$\mathbb{E}_\theta(X)$ = esperanza de X suponiendo parámetro θ.

$\|$ si es continua

$$\int_{sop_\theta} x f(x;\theta)\,dx \Longrightarrow$$ función de θ, $\theta\in\Theta$

$$\searrow \sum_{x\in sop_\theta} x f(x;\theta)$$

$V_\theta(X)$ = varianza de X suponiendo parámetro θ

funciones de $\theta\in\Theta$

$\mathbb{P}_\theta(X\in A)$ = probabilidad con parámetro θ

Ejemplo:

(1) X es $Ber(p)$

$$p\in(0,1) \longrightarrow \mathbb{E}_p(X) = p$$

(2) X es $Exp(\lambda)$

$$\lambda>0 \longrightarrow \mathbb{E}_\lambda(X) = \frac{1}{\lambda}$$

3) X es $unif(0,a)$

$$a\in(0,\infty) \longrightarrow \mathbb{E}_a(\lambda) = \frac{a}{2}$$

(4) $\Theta = \{0,1\}$

$$\theta \in \{0,1\} \longrightarrow \mathbb{E}_\theta(X) = \begin{cases} 1/2 & \text{si } \theta = 0 \\ \int_0^1 x \frac{1}{2\sqrt{x}}\,dx = \mathbb{E}_1(X) = 1/3 & \text{si } \theta = 1. \end{cases}$$

Más notación: estimaciones y estimadores

$X \quad f(x;\theta) \quad \Theta$

muestra $x_1, \dots, x_n$

$\big\downarrow$ transformación con $h(x_1,\dots,x_n)$ $\quad h: \mathbb{R}^n \longrightarrow \mathbb{R}$

$\hat{\theta}$ = estimación de θ

→ clones indep.

$$\underbrace{h(x_1,\dots,x_n)}_{\text{variable aleatoria}} = \text{todas las posibles estimaciones de } \theta$$

⟶ ESTADÍSTICO ESTIMADOR ó ESTIMADOR

MÉTODOS DE CONSTRUCCIÓN DE ESTIMADORES

- → Momentos
- → Máxima verosimilitud

MÉTODO DE MOMENTOS

$X \quad f(x;\theta) \quad \Theta$

$x_1, \dots, x_n$ una muestra

→ función de θ

$$\underbrace{\bar{x} = \frac{x_1 + \cdots + x_n}{n}}_{\text{dato}} \quad \text{debe parecerse a} \quad \mathbb{E}_\theta(X) \longleftarrow \theta \in \Theta$$

↑ desconocido

$\big\Downarrow$ método

$$\boxed{\bar{x} \in \mathbb{E}_\theta(X)}$$

ECUACIÓN CON INCÓGNITA θ

Ejemplos

(1) X es uniforme(a)

$x_1, \dots, x_n$

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n}$$

método de momentos

muestra estimación

$$\bar{x} = \mathbb{E}_a(X) = \frac{a}{2} \Longrightarrow \hat{a} = 2\bar{x}$$

(*) ESTIMADOR POR MOMENTOS DE a: $\boxed{M_a = 2\bar{X}}$

(2) $\exp(\lambda)$

$x_1, \dots, x_n$ muestra

método de momentos

nuestra estimación

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} \qquad \bar{x} = E_\lambda(X) = \frac{1}{\lambda} \Longrightarrow \hat{\lambda} = \frac{1}{\bar{x}}$$

⊛ ESTIMADOR POR MOMENTOS DE λ : $\boxed{M_\lambda = \frac{1}{\bar{X}}}$

(3) X es $\mathrm{Ber}(p)$

$x_1, \dots, x_n$ muestra

método de momentos

nuestra muestra

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} \qquad \bar{x} = E_p(X) = p \Longrightarrow \hat{p} = \bar{x}$$

⊛ ESTIMADOR POR MOMENTOS DE p : $\boxed{M_p = \bar{X}}$

(4) (Mal ejemplo)

$$X \qquad f(x;\alpha) = \begin{cases} (\alpha+1)x^\alpha, & x \in (0,1) \\ 0 & \text{en el resto} \end{cases} \qquad \alpha \in (0,\infty) = \Theta$$

$$E_\alpha(X) = \int_0^1 x(\alpha+1)x^\alpha \, dx = \frac{\alpha+1}{\alpha+2}$$

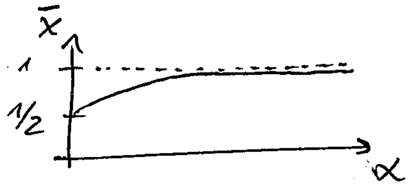
$$\bar{x} = \frac{\alpha+1}{\alpha+2} \Longrightarrow (\alpha+2)\bar{x} = \alpha+1 \Longrightarrow \alpha\bar{x} + 2\bar{x} = \alpha+1 \Longrightarrow \alpha = \frac{1-2\bar{x}}{\bar{x}-1}$$

$\Longrightarrow$ estimación de α = $\hat{\alpha} = \dfrac{1-2\bar{x}}{\bar{x}-1}$

si $\bar{x} < 1/2 \Longrightarrow \hat{\alpha} < 0$ y recordemos que $\alpha \in (0,\infty)$

el método de momentos no funcionaría

solo puedo despejar si $1/2 < \bar{x} < 1$

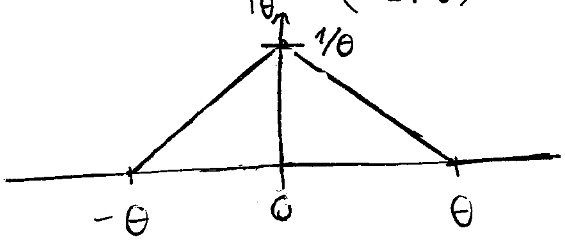


Ejemplo / ejercicio

$$X \qquad f(x;\theta) = \begin{cases} \frac{1}{\theta}\left(1-\frac{|x|}{\theta}\right) & \text{si } |x|<\theta \\ 0 & \text{si } |x|\geq\theta \end{cases}$$

$\theta \in \Theta = (0, \infty)$

$\text{sop}_\theta = (-\theta, \theta)$



datos: $x_1, \dots, x_n$

$$\bar{x} = \frac{x_1+\dots+x_n}{n}$$

$\bar{x} = \mathbb{E}_\theta(X) = 0$ para cualquier θ $\Longrightarrow$ no nos sirve para estimar el momento no centrado de grado 1.

Probamos con el momento de grado 2: $\overline{x^2} = \mathbb{E}_\theta(X^2)$

$$\overline{x^2} = \frac{x_1^2+\dots+x_n^2}{n}$$

$$\mathbb{E}_\theta(X^2) = \int_{-\theta}^{\theta} x^2 f(x;\theta)\,dx = \int_{-\theta}^{\theta} x^2 \cdot \frac{1}{\theta}\left(1-\frac{|x|}{\theta}\right)dx = 2\int_0^\theta x^2 \cdot \frac{1}{\theta}\left(1-\frac{x}{\theta}\right)d$$

$$= 2\int_0^1 y^2\theta^2 \cdot \frac{1}{\theta}(1-y)\,\theta\,dy = 2\theta^2\int_0^1 y^2 - y^3\,dy = \frac{\theta^2}{6}$$

cambio $y = \frac{x}{\theta} \rightarrow x = y\theta$

$dy = \frac{1}{\theta}dx \rightarrow dx = \theta dy$

$$\Longrightarrow \overline{x^2} = \frac{\theta^2}{6} \Longrightarrow \boxed{\hat{\theta} = \sqrt{6\,\overline{x^2}}}$$ estimación

En general:

$$M_\theta = \sqrt{6\,\frac{X_1^2+\dots+X_n^2}{n}} \longrightarrow \text{clones de } X$$

M_θ ⟶ estimador

Ejemplo: dos parámetros por momentos

X es $N(\mu,\sigma^2)$ $\qquad$ $\mathbb{E}_{(\mu,\sigma^2)}(X)=\mu$

$x_1,\dots,x_n$ $\qquad$ $\mathbb{E}_{(\mu,\sigma^2)}(X^2)=\sigma^2+\mu^2$

Sistema de dos ecuaciones:

$$\bar{x}=\mathbb{E}_{(\mu,\sigma^2)}(X)=\mu \Rightarrow \boxed{\hat{\mu}=\bar{x}}$$

$$\overline{x^2}=\mathbb{E}_{(\mu,\sigma^2)}(X^2)=\sigma^2+\mu^2 \Rightarrow \hat{\sigma}^2=\overline{x^2}-\hat{\mu}^2=\overline{x^2}-\bar{x}^2=V_x$$

$$\boxed{\hat{\sigma}^2=V_x}$$

MÉTODO DE MÁXIMA VEROSIMILITUD

X discreta finita $\qquad$ $f(x;\overset{\text{dado }\theta}{\theta})$ = probabilidad de obtener el valor x.

sop_θ = fijo finito

Dado x: $f(x;\theta)$ VEROSIMILITUD DE θ.

Dada la muestra $x_1,\dots,x_n$, llamamos VEROSIMILITUD de θ dada la muestra:

$$\text{VERO}(\theta;x_1,\dots,x_n)=\prod_{j=1}^{n} f(x_j;\theta)$$

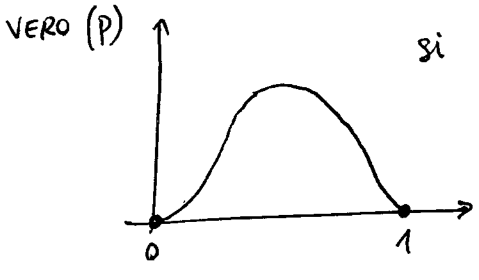
└→ función de verosimilitud está definida en Θ.

Ejemplo: X es $\text{Ber}(p)$ $\quad x_1,\dots,x_n$ (son 0's y 1's)

$\Theta=[0,1]$ $\qquad f(x;p)=\begin{cases} p, & x=1\\ (1-p), & x=0\end{cases}$ $\qquad n^{\circ}\{x_j=1\}=\sum_{j=1}^{n}x_j=n\bar{x}$

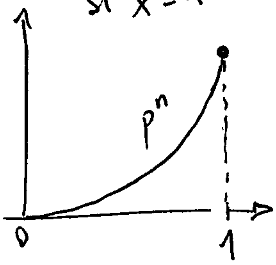
$$\text{VERO}(p;x_1\dots x_n)=\prod_{j=1}^{n}f(x;p)=p^{n^{\circ}x_j=1}\,(1-p)^{n^{\circ}x_j=0}=p^{n\bar{x}}(1-p)^{n(1-\bar{x})}$$

gráfica de $\text{VERO}(p;x_1,\dots,x_n)$

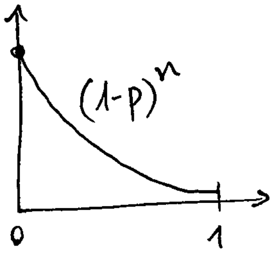


si $0<\bar{x}<1$

si $\bar{x}=1$



si $\bar{x}=0$



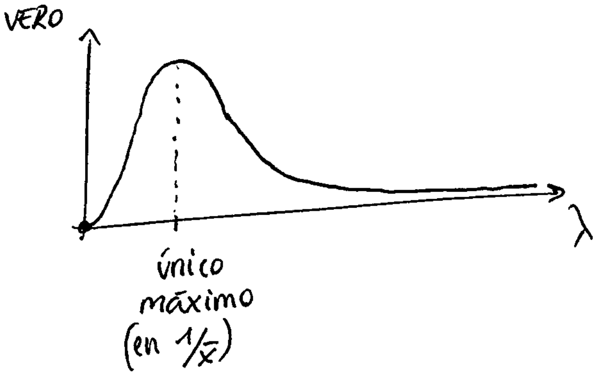
Ejemplo: X es $\mathrm{Exp}(\lambda)$ $\qquad \lambda \in \Theta = (0, \infty)$

$x_1, \dots, x_n$

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

$x_1, \dots, x_n > 0$

$$\mathrm{VERO}(\lambda; x_1, \dots, x_n) = \lambda^n e^{-\lambda(x_1 + \cdots + x_n)} = \lambda^n e^{-\lambda n \bar{x}}$$



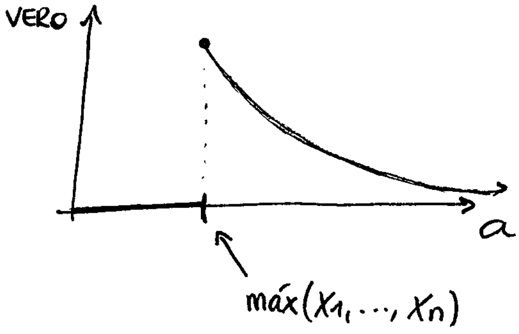
Ejemplo: X es uniforme$[0,a]$ $\qquad a \in (0, \infty)$

$$f(x;a) = \begin{cases} 1/a & , \text{ si } 0 \leq x \leq a \\ 0 & , \text{ en el resto} \end{cases}$$

muestra $x_1, \dots, x_n \rightarrow$ números positivos

$$\mathrm{VERO}(a; x_1, \dots, x_n) = \prod_{j=1}^{n} f(x_j; a) = \begin{cases} (1/a)^n & \text{si } x_j \leq a \text{ para } 1 \leq j \leq n \\ 0 & , \text{ en caso contrario} \end{cases} =$$

$$\underset{\uparrow \text{ paso clave}}{=} \begin{cases} (1/a)^n & \text{si } a \geq \max(x_1, \dots, x_n) \\ 0 & \text{si } a < \underset{\uparrow \text{dato}}{\max(x_1, \dots, x_n)} \end{cases}$$



Estimador de MÁXIMA VEROSIMILITUD

$X \qquad f(x;\theta) \qquad x_1,\dots,x_n$

$\theta \in \Theta \longrightarrow \text{VERO}(\theta; x_1,\dots,x_n)$

Si esta función tiene un ÚNICO MÁXIMO GLOBAL (en todo Θ) en el punto $\hat{\theta}$, a $\hat{\theta}$ lo llamamos ESTIMACIÓN DE MÁXIMA VEROSIMILITUD de θ dada $x_1,\dots,x_n$.

Ejemplos

(1) Ber(p) ▶ $0 < \bar{x} < 1$

$$\text{VERO}(p) = p^{n\bar{x}}(1-p)^{n(1-\bar{x})}$$

$$\left.\begin{array}{l} \text{VERO}(0) = 0 \\ \text{VERO}(1) = 0 \\ \text{VERO} \geq 0 \\ \text{continua y derivable} \end{array}\right\}$$

hay máximo global. Vamos a comprobar que hay un único punto crítico.

Simplificación habitual:

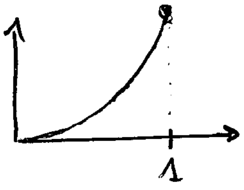
$\log\text{VERO}(p) = \ln \text{VERO}(p)$ puntos críticos de $\log\text{VERO}$

$\hookrightarrow n\bar{x}\ln p + n(1-\bar{x})\ln(1-p)$

$$0 = n\bar{x}\frac{1}{p} - \frac{n(1-\bar{x})}{1-p} \Rightarrow \boxed{\hat{p} = \bar{x}}$$ estimación por MÁXIMA VEROSIMILITUD

▶ $\bar{x} = 1$

$\text{VERO}(p) = p^n$

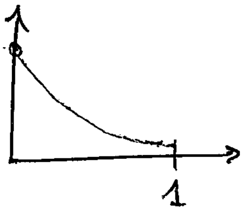


Alcanza máximo en $p = 1$

$\Rightarrow \boxed{\hat{p} = 1}$

▸ $\bar{x} = 0$

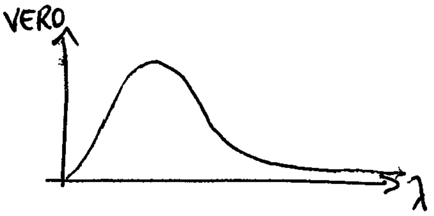
$\text{VERO} = (1-p)^n$



Alcanza máximo en $p = 0$

$\Rightarrow \boxed{\hat{p} = 0}$

(2) $X \sim \mathrm{Exp}(x)$ $\qquad x_1, \dots, x_n$

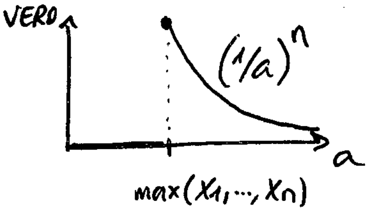


$$\mathrm{VERO}(\lambda) = \lambda^n e^{-\lambda n \bar{x}}$$

$$\log \mathrm{VERO}(\lambda) = n \log(\lambda) - \lambda n \bar{x}$$

$$\frac{d}{d\lambda} \log \mathrm{VERO}(\lambda) = \frac{n}{\lambda} - n\bar{x} \quad ; \quad \frac{n}{\lambda} - n\bar{x} = 0 \Rightarrow \boxed{\bar{\lambda} = 1/\bar{x}}$$

(3) $X \sim \mathrm{Unif}[0,a]$ $\qquad x_1, \dots, x_n > 0$



$$\Rightarrow \boxed{\hat{a} = \max(x_1, \dots, x_n)}$$

estimación por MÁXIMA VEROSIMILITUD de a dado $x_1, \dots, x_n$.

REGLA GENERAL: ESTIMADOR MV de a

$$\mathrm{EMV}_a = \max\underbrace{(X_1, \dots, X_n)}_{\text{clones de } X} \longleftarrow$$

aquí están todas las posibles estimaciones

$\downarrow$ variable aleatoria

en $X \sim \mathrm{Unif}[0,a$

NOTA:

$M_a = 2\bar{X}$

$\uparrow$ momentos

$$\mathrm{EMV}_p = \bar{X} = M_p \quad \text{en} \quad X \sim \mathrm{Ber}(p)$$

$$\mathrm{EMV}_\lambda = 1/\bar{X} = M_\lambda \quad \text{en} \quad X \sim \mathrm{Exp}(\lambda)$$

ROPIEDADES Y COMPARACIONES DE ESTIMADORES

$X \quad f(x;\theta) \qquad \theta \in \Theta$

Estimador $\longrightarrow T = h(X_1, \dots, X_n) \qquad h: \mathbb{R}^n \longrightarrow \mathbb{R}$

↑ recoge <u>todas</u> las posibles estimaciones

Decimos que T es ESTIMADOR INSESGADO de θ si se cumple

$\mathbb{E}_\theta(T) = \theta$ para todo $\theta \in \Theta$

Jn estimador insesgado en media no se equivoca.

<u>Ejemplo</u>

(1) $X \sim \text{Ber}(p) \qquad T = \bar{X}$

$$\underset{\mathbb{E}_p(T)}{\underset{\shortparallel}{\mathbb{E}_p(\bar{X})}} = \mathbb{E}_p(X) = p$$

T es insesgado

(2) Supongamos que el parámetro θ es la esperanza de X

$\mathbb{E}_\theta(X) = \theta$

estimador de θ: $T = \bar{X}$

$$\mathbb{E}_\theta(T) = \mathbb{E}_\theta(\bar{X}) \underset{\substack{\uparrow \\ \text{en general}}}{=} \mathbb{E}_\theta(X) = \theta \Rightarrow$$

La media muestral es estimador insesgado de $\theta = \mathbb{E}(X)$.

Si T no es insesgado decimos que es SESGADO.

$\text{sesgo}(T) = \mathbb{E}_\theta(T) - \theta$

(3) $X \sim \text{Unif}[0,a]$ $\qquad T_1 = 2X$

$\qquad\qquad\qquad\qquad T_2 = \max(x_1, \dots, x_n)$

$$\mathbb{E}_a(T_1) = \mathbb{E}_a(2\bar{X}) = 2\mathbb{E}_a(\bar{X}) = 2\,\mathbb{E}_a(X) = 2 \cdot \frac{a}{2} = a$$

$$\mathbb{E}_a(T_2) = \mathbb{E}_a(\max(x_1, \dots, x_n)) \Longrightarrow$$

$$F_{T_2}(x) = \left(\frac{x}{a}\right)^n \quad 0 \le x \le a$$

$$f_{T_2}(x) = \frac{n x^{n-1}}{a^n} \quad 0 \le x \le a$$

$$\Longrightarrow \mathbb{E}_a(\max(x_1, \dots, x_n)) = \int_0^a x \frac{n x^{n-1}}{a^n} \, dx = \frac{n}{n+1} a$$

$$\text{sesgo}(T_2) = \frac{n}{n+1} a - a = -\frac{a}{n+1}$$

T_2 es sesgado, pero asintóticamente insesgado

<u>Observación</u>: $T_3 = \frac{n+1}{n} T_2$ es insesgado. (de hecho, es el mejor estimador)

(4) $X \sim \text{Exp}(\lambda)$

$$T = 1/\bar{X} \begin{array}{l} \nearrow \text{momentos} \\ \searrow \text{max VERO} \end{array}$$

$$\mathbb{E}_\lambda(T) = \mathbb{E}_\lambda(1/\bar{X}) > \frac{1}{\mathbb{E}(X)} = \lambda$$

NOTA: Z variable aleatoria
$Z > 0$ no constante

$$1 < \mathbb{E}(Z) \cdot \mathbb{E}(1/Z)$$

T es siempre sesgado (sobreestimación)

demo de nota:

$$1 = \mathbb{E}(1) = \mathbb{E}\left(Z \cdot \frac{1}{Z}\right) = \mathbb{E}\left(\sqrt{Z} \cdot \frac{1}{\sqrt{Z}}\right) \le \mathbb{E}(Z)^{1/2} \, \mathbb{E}(1/Z)^{1/2}$$

Igualdad solo si Z constante.

VARIANZA DE UN ESTIMADOR

Eficiencia

- Varianza de estimador $V_\theta(T)$
- T es insesgado y apelo a Chebyshev

$$\mathbb{P}\left(|T-\underset{\underset{\mathbb{E}_\theta(T)}{\parallel}}{\theta}| \geq \varepsilon\right) \leq \frac{V_\theta(T)}{\varepsilon^2}$$

Si $V_\theta(T)$ pequeño, T en general se equivoca poco.

- T_1, T_2 son estimadores insesgados decimos que T_1 es más EFICIENTE que T_2 si:

$$V_\theta(T_1) < V_\theta(T_2) \quad \text{para todo } \theta \in \Theta$$

Ejemplo: X Unif$[0,a]$ $\qquad f(x;a) = \begin{cases} 1/a \ , & \text{si } x \in [0,a] \\ 0 \ , & \text{si no} \end{cases}$

$$T_1 = 2\bar{X} \qquad \mathbb{E}_a(T_1) = a$$

$$T_2 = \frac{n+1}{n}\overbrace{\max(x_1,\dots,x_n)}^{\hat{T}_2} \qquad \mathbb{E}_a(T_2) = a$$

Para comparar T_1, T_2 ¿$V_a(T_1)$, $V_a(T_2)$?

cuenta integrando la función de densidad

$$V_a(T_1) = 4V_a(\bar{X}) = \frac{4V_a(X)}{n} = \frac{4}{n}\cdot\frac{a^2}{12}$$

$$V_a(\hat{T}_2); \qquad F_{\hat{T}_2}(x) = \left(\frac{x}{a}\right)^n \longrightarrow f_{\hat{T}_2}(x) = n\,\frac{x^{n-1}}{a^n}$$

$$V_a(\hat{T}_2) = \underbrace{\int_0^a x^2\,\frac{nx^{n-1}}{a^n}\,dx}_{\mathbb{E}(\hat{T}_2^2)} - \mathbb{E}(\hat{T}_2)^2 = \text{res}$$

$$V_a(T_2) = \left(\frac{n+1}{n}\right)^2 V_a(\hat{T}_2) = \frac{a^2}{n(n+1)}$$

$\Rightarrow T_2$ es más eficiente que T_1.

Ejemplo: $X \sim \mathrm{Exp}(\lambda)$

RECUERDO:

$$\mathbb{E}_\lambda(X) = 1/\lambda$$

$$V_\lambda(X) = 1/\lambda^2$$

$$\min(X_1, \dots, X_n) \sim \mathrm{Exp}(n\lambda)$$

Estimadores de $1/\lambda$:

$$T_1 = \bar{X} \qquad \mathbb{E}(T_1) = 1/\lambda$$

$$T_2 = n \min(X_1, \dots, X_n) \qquad \mathbb{E}(T_2) = 1/\lambda$$

$$V(T_1) = V(\bar{X}) = \frac{V(X)}{n} = \frac{1}{n\lambda^2}$$

$$V(T_2) = n^2 V(\min(X_1, \dots, X_n)) = \frac{n^2}{n^2} \cdot \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

T_1 es más eficiente que T_2.

ESTIMADORES

$$X \qquad f(x;\overset{\in \Theta}{\theta}) \qquad X_1, \dots, X_n$$

- Límites de "calidad" de estimadores
- Lema de CRAMER-RAO

$$\left.\begin{array}{l}\text{VARIABLE DE INFORMACIÓN}\\ \text{CANTIDAD DE INFORMACIÓN}\end{array}\right\}\ (*)$$

Θ es un intervalo (a,b)

$f(x;\theta)$

$$\frac{\partial}{\partial\theta} f(x;\theta) \overset{\text{notación}}{=} \partial_\theta f(x;\theta)$$

Formamos una transformación de X: $Y :=$ variable de información de X

$$Y = \partial_\theta \ln f(X;\theta) = \frac{\partial_\theta f(X;\theta)}{f(X;\theta)}$$

Cantidad de información de $X := V_\theta(Y)$, y se denota como: $I_X(\theta)$

Ejemplos:

(1) X es $N(\overset{\text{fijo}}{\mu_0}, \overset{\text{parámetro } \theta=\sigma^2}{\sigma^2})$ $\qquad \Theta = (0,\infty)$

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}\sqrt{\theta}}\cdot \exp\left(-\frac{1}{2}\cdot\frac{(x-\mu_0)^2}{\theta}\right)$$

$$\partial_\theta \ln f(x;\theta) = \frac{-1}{2}\cdot\frac{1}{\theta} + \frac{1}{2}\cdot\frac{(x-\mu_0)^2}{\theta^2} = \frac{1}{2}\left(\frac{(x-\mu_0)^2-\theta}{\theta^2}\right)$$

$$\Rightarrow Y = \frac{1}{2}\cdot\frac{(X-\mu_0)^2-\theta}{\theta^2}$$

$$\mathbb{E}_\theta(Y) = \mathbb{E}_\theta\left(\frac{1}{2}\cdot\frac{(X-\mu_0)^2-\theta}{\theta^2}\right) = \frac{1}{2}\mathbb{E}_\theta\left(\frac{(X-\mu_0)^2-\theta}{\theta^2}\right) = \frac{1}{2\theta^2}\underbrace{\mathbb{E}_\theta\left((x-\mu_0)^2-\theta\right)}_{\substack{V_\theta(X)-\theta\\ \shortparallel\\ \theta}} \Rightarrow$$

$$\Rightarrow \mathbb{E}_\theta(Y) = 0$$

$$I_X(\theta) = V_\theta(Y) = \mathbb{E}_\theta(Y^2) - \overset{0}{\cancel{\mathbb{E}_\theta(Y)}}^2 = \mathbb{E}_\theta(Y^2) = \frac{1}{4}\cdot\frac{1}{\theta^4}\,\mathbb{E}_\theta\left(\left((X-\mu_0)^2-\theta\right)^2\right) =$$

$$= \frac{1}{4}\cdot\frac{1}{\theta^4}\cdot\mathbb{E}_\theta\left((X-\mu_0)^4 \underbrace{-2\underbrace{(X-\mu_0)^2}_{\theta}\theta}_{-2\theta^2} \underbrace{+\theta^2}_{+\ \theta^2}\right) = \frac{1}{4\theta^4}\cdot\left[\mathbb{E}_\theta\left((X-\mu_0)^4\right) - 2\theta^2 + \theta^2\right] =$$

NOTA: RECORDATORIO

$Z \sim N(0,1)$

$$\mathbb{E}(Z) = 0 \qquad \mathbb{E}(Z^2) = 1 \qquad \mathbb{E}(Z^3) = 0 \qquad \mathbb{E}(Z^4) = 3$$

Con la nota: $\mathbb{E}_\theta\left(\left(\frac{X-\mu_0}{\sqrt{\theta}}\right)^4\right) = 3 \Rightarrow \mathbb{E}_\theta\left((X-\mu_0)^4\right) = 3\theta^2$

$$\Rightarrow I_X(\theta) = V_\theta(Y) = \frac{1}{4\theta^4}\cdot\left(3\theta^2 - 2\theta^2 + \theta^2\right) = \frac{1}{2\theta^2}$$

alternativa: $V_\theta(Y) = \frac{1}{4\theta^4}\cdot V_\theta\left((X-\mu_0)^2\right) = \frac{1}{4\theta^2}\cdot V_\theta(\chi_1^2) = \frac{2}{4\theta^2} = \frac{1}{2\theta^2}$

(2) X es Exp( ) $\qquad \theta = 1/\lambda$

$$f(x;\theta) = \frac{1}{\theta}\cdot\exp\left(-\frac{1}{\theta}x\right)$$

$$\partial_\theta \ln f(x;\theta) = \partial_\theta\left(-\ln\theta - \frac{1}{\theta}x\right) = -\frac{1}{\theta} + \frac{x}{\theta^2} \Rightarrow Y = \frac{-1}{\theta} + \frac{X}{\theta^2} \Rightarrow$$

$$\Rightarrow Y = \frac{X-\theta}{\theta^2}$$

NOTA:
$\mathbb{E}(\exp) = \frac{1}{\lambda} \qquad V(\exp) = \frac{1}{\lambda^2}$

$\mathbb{E}_\theta(X) = \theta \quad ; \quad \mathbb{E}_\theta(Y) = \mathbb{E}\left(\frac{X-\theta}{\theta^2}\right) = 0$

$V_\theta(Y) = \frac{1}{\theta^4}\cdot\theta^2 = \frac{1}{\theta^2}$ = cantidad de información = $I_X(\theta)$

(3) X es Ber(p) $\qquad p \in (0,1) = \Theta$

$$f(x;p) = p^x\cdot(1-p)^{(1-x)} = \begin{cases} p, & x=1 \\ 1-p, & x=0 \end{cases}$$

$$\ln f(x;p) = x\ln p + (1-x)\ln(1-p)$$

$$\partial_\theta \ln f(x;p) = \frac{x}{p} - \frac{1-x}{1-p} = \frac{x-p}{p(1-p)} \Rightarrow Y = \frac{X-p}{p(1-p)}$$

NOTA: valores $\qquad$ probabilid

$$Y = \begin{cases} 1/p & \longrightarrow p \\ \frac{-1}{1-p} & \longrightarrow 1-p \end{cases}$$

$\Rightarrow \mathbb{E}(Y) = 1/p\cdot p + (1-p)\frac{-1}{1-p} =$

$\mathbb{E}_p(Y) = 0$

$$V_p(Y) = \frac{1}{(p(1-p))^2}\overbrace{V_p(X-p)}^{= V_p(X) = p(1-p)} = \frac{1}{p(1-p)} = I_X(p)$$

-EMA (Condiciones especiales) (A)

hipótesis

$f(x;\theta)$ $\quad\theta \in \Theta = (a,b)$ $\quad\text{sop}_\theta = \{a_1, \dots, a_M\}$

$\hookrightarrow$ variable finita $\longrightarrow$

$\theta \longmapsto f(x;\theta)$ derivable

Entonces:

$$\boxed{\mathbb{E}_\theta(Y) = 0 \quad \text{para todo } \theta}$$

(por tanto, obtenemos: $I_X(\theta) = \mathbb{E}_\theta(Y^2)$)

demostración

$$\sum_{j=1}^{M} f(a_j,\theta) = 1 \quad \text{para todo } \theta \in \Theta$$

derivamos respecto de θ

$$\sum_{j=1}^{M} \partial_\theta f(a_j,\theta) = 0 \quad \text{para todo } \theta \in \Theta$$

$$\sum_{j=1}^{M} \underbrace{\frac{\partial_\theta f(a_j,\theta)}{f(a_j,\theta)}}_{\text{valores de } Y} \cdot \underbrace{f(a_j,\theta)}_{\text{probab.}} = 0 = \mathbb{E}_\theta(Y) = 0 \quad \text{para todo } \theta \in \Theta$$

■

VARIABLE DE INFORMACIÓN de $(X_1, \dots, X_n)$

$$Z_n = \sum \partial_\theta \ln f(X_j;\theta)$$

(suma de variables de información de $X_1, \dots, X_n$)

LEMA (B) (Corolario del lema (A))

$\Theta = (a,b)$ $\qquad$ $sop_\theta = \{a_1, \dots, a_M\}$ para todo $\theta \in \Theta$

$\theta \longrightarrow f(x;\theta)$ derivable

Para la variable de información de la muestra

$$\mathbb{E}_\theta(Z_n) = 0 \quad \text{y} \quad V_\theta(Z_n) = n\, I_X(\theta) = n V_\theta(Y) \quad \forall \theta \in \Theta$$

LEMA (C) (Lema de CRAMER-RAO)

$\Theta = (a,b)$ $\qquad$ $sop_\theta = \{a_1, \dots, a_M\}$ $\forall \theta \in \Theta$

$\theta \longrightarrow f(a_j;\theta)$ es derivable con $1 \leq j \leq M$

Sea T un estadístico INSESGADO de X, entonces:

$$\underbrace{\boxed{V_\theta(T) \geqslant \frac{1}{n I_X(\theta)}}}_{\longrightarrow \text{ cota de CRAMER-RAO}} \quad \forall \theta \in \Theta$$

demostración

T insesgado $\quad \Big| \quad T = h(X_1, \dots, X_n) \qquad h: \mathbb{R}^n \longrightarrow \mathbb{R}$

Para todo $\theta \in \Theta$ $\qquad$ $\theta = \mathbb{E}_\theta(T) = \sum \overbrace{h(x_1, \dots, x_n)}^{\text{valores de } T} \cdot \overbrace{f(x_1;\theta) \cdot \dots \cdot f(x_n;\theta)}^{\text{probabil.}}$

$(x_1, \dots, x_n) \in sop_\theta^n \longrightarrow = \sum_{x_n \in sop_\theta} \sum_{x_{n-1} \in sop_\theta} \cdots \sum_{x_1 \in sop_\theta}$

$\longrightarrow$ un total de M^n sumandos

derivamos respecto de θ:

$$1 = \sum h(x_1, \dots, x_n) \cdot \begin{bmatrix} \partial_\theta f(x_1;\theta) \cdot f(x_2;\theta) \cdots f(x_n;\theta) + \\ + f(x_1;\theta) \cdot \partial_\theta f(x_2;\theta) \cdot \dots \cdot f(x_n;\theta) + \\ \dots + f(x_1;\theta) \cdot \dots \cdot \partial_\theta f(x_n;\theta) \end{bmatrix}$$

$\longrightarrow = \frac{\partial_\theta f_1(x_1;\theta)}{f_1(x_1;\theta)} \cdot f_1(x_1;\theta) \cdots f(x_n;\theta)$

el mismo con el resto

$$\Rightarrow 1 = \sum \underbrace{h(x_1, \dots, x_n)}_{\text{valores de } T} \underbrace{\left(\sum_{j=1}^n \partial_\theta \ln f(x_j;\theta) \right)}_{\text{valores de } Z_n} \underbrace{f_1(x_1;\theta) \cdots f(x_n;\theta)}_{\text{prob.}} \Rightarrow 1 = \mathbb{E}_\theta(T . Z_n) \ \forall \theta \in$$

$(Z_n) = 0$ $\Rightarrow$ $\quad 1 = \mathbb{E}_\theta(T Z_n) - \mathbb{E}(T) . \mathbb{E}(Z_n) = cov_\theta(T, Z_n)$

Finalmente, $1 = |cov_\theta(T, Z_n)| \leq \sqrt{V_\theta(T)} \sqrt{V_\theta(Z_n)} \Rightarrow V_\theta(T) \geqslant \frac{1}{n I_X(\theta)}$

$\longrightarrow = \sqrt{n I_X(\theta)}$

Ejemplos de Cramer-Rao

(1) Ber(p)

$$\frac{1}{n I_X(p)} = \frac{p(1-p)}{n} \Bigg]$$ Todo estimador $\nearrow$ insesgado de p tiene varianza $\geqslant \frac{p(1-p)}{n}$

Ahora $T = \bar{X}$ $\quad \mathbb{E}_p(\bar{X}) = \mathbb{E}_p(X) = p$ insesgado

$$V_p(\bar{X}) = \frac{1}{n} V_p(X) = \frac{1}{n}\left(\mathbb{E}_p(\underset{\overset{\shortparallel}{X}}{X^2}) - \mathbb{E}_p(X)^2\right) = \frac{1}{n}(p - p^2) =$$

$$= \frac{1}{n}\, p(1-p)$$

$\bar{X}$ es un estimador más eficiente. (puede no ser el único)

(2) $X \sim N(\mu_0, \sigma^2)$ $\longrightarrow \theta = \sigma^2$ $\qquad I_X(\theta) = \frac{1}{2\theta^2} \qquad \Theta = (0, \infty)$

μ_0: dato conocido

En general: variable U

$$\mathbb{E}(\bar{U}) = \mathbb{E}(U) \quad ; \quad V(\bar{U}) = \frac{V(U)}{n}$$

$$\mathbb{E}(S_U^2) = V(U)$$

insesgados

Tomamos $T = S^2$ $\qquad \mathbb{E}_\theta(S^2) = \theta = \sigma^2$

¿Es más eficiente? ¿$V_\theta(S^2)$?

Fisher-Cochran (variables normales)

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \quad ; \quad V(\chi^2_{n-1}) = 2(n-1)$$

$$V_\theta(S^2) = \frac{2(n-1)\theta^2}{(n-1)^2} = \frac{2\theta^2}{n-1} > \underbrace{\frac{2\theta^2}{n} = \frac{1}{n I_X(\theta)}}_{\text{cota de Cramer-Rao}}$$

Ejercicio X $\quad f(x;\alpha) = \begin{cases} \alpha x^{\alpha-1} & , \text{ si } \quad x \in (0,1) \\ 0 & , \text{ si } \quad \text{no} \end{cases} \qquad \alpha \in \Theta = (0,\infty)$

1) Preliminar, calcular $\mathbb{E}_\alpha((\log X)^K)$

2) Cantidad de información y cota de Cramer-Rao

3) Sea $T = \frac{1}{n}\sum_{i=1}^{n} \ln x_i$. Comprobar que es más eficiente.

$y = \ln\frac{1}{x} \to x = e^{-y}$

cambio variable

$$1) \quad \mathbb{E}_\alpha\left((\log X)^K\right) = \int_0^1 (\ln X)^K \, \alpha X^{\alpha-1} \, dx = \alpha \int_0^\infty (-1)^K y^K e^{\boxed{-y(\alpha-1)}=t} dy =$$

$$= \frac{\alpha(-1)^K}{(\alpha-1)^{K+1}} \underbrace{\int_0^\infty t^K . e^{-t} dt}_{\Gamma(K+1) = K!} = \frac{\alpha(-1)^K}{(\alpha-1)^{K+1}} K!$$

(2) Variable de información

$$Y = \frac{\partial}{\partial\theta} \ln f(X;\alpha) = \frac{\partial}{\partial\theta}\left[\ln(\alpha x^{\alpha-1})\right] = \frac{\partial}{\partial\theta}\left[\ln\alpha + (\alpha-1)\ln x\right]$$

$$= \frac{1}{\alpha} + \ln x = \frac{\alpha \ln x + 1}{\alpha}$$

Cantidad de información

$$I_X(\alpha) = V_\alpha(Y) = V_\alpha(\ln x) = \mathbb{E}_\alpha((\ln x)^2) - \mathbb{E}_\alpha(\ln x)^2$$

VARIABLE DE INFORMACIÓN PARA X CONTINUA

$\theta \in \Theta = (a,b)$

sop = $\text{sop}_\theta \longleftarrow$ fijo (no dep. de θ)

$\theta \longrightarrow f(x;\theta)$ derivab

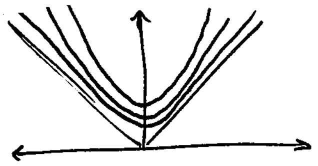
Queremos replicar la demostración de $\mathbb{E}_\theta(Y) = 0$

$$\int_{sop} f(x;\theta)\,dx = 1 \qquad \forall\theta\in\Theta$$

sop ← no depende de θ

$$\Theta \longrightarrow \int_{sop} f(x;\theta)\,dx$$

En general, $\frac{d}{d\theta}\int \neq \int \frac{\partial}{\partial\theta}$ como por ejemplo:



sucesión de curvas derivables que tienden a una no derivable

Idea

Si fuera:

$$\frac{d}{d\theta}\int_{sop} f(x;\theta)\,dx = \int_{sop}\frac{\partial}{\partial\theta} f(x;\theta)\,dx$$

entonces

$$0 = \int_{sop}\frac{\partial}{\partial\theta} f(x;\theta)\,dx = \int_{sop}\underbrace{\frac{\partial}{\partial\theta}\ln f(x;\theta)}_{\text{valores}}\,\underbrace{f(x;\theta)\,dx}_{\text{prob.}} = \mathbb{E}_\theta(Y)$$

y ya estaría

$\boxed{\text{LEMA A para continuas}}$

Derivadas bajo signo integral (Feynman)

Ejemplo: $\displaystyle\int_0^1 \frac{x^{11}-1}{\ln x}\,dx$

truco: $\displaystyle I(t) = \int_0^1 \frac{x^t-1}{\ln x}\,dx$ $\qquad x^t = e^{t\ln x}$

$$I(0) = 0 \; ; \quad I'(t) = \int_0^1 \frac{\cancel{\ln x}\,e^{t\ln x}}{\cancel{\ln x}}\,dx = \int_0^1 x^t\,dx = \frac{1}{t+1}$$

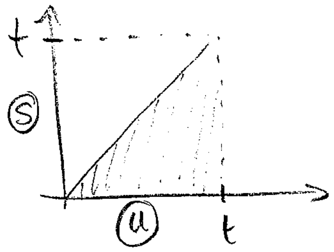
$$\Rightarrow I(t) = \int \frac{1}{t+1}\,dt = \ln(t+1) + C_1 \quad \text{y} \quad C_1 = 0 \quad \text{porque } I(0) = 0$$

$$\Rightarrow I(11) = \frac{1}{12}$$

Aproximar por 2ª derivada

$$g(t) - g(0) = \int_0^t g'(u)\,du$$

$$g'(u) - g'(0) = \int_0^u g''(s)\,ds$$



$$g(t) = g(0) + g'(0)t + \underbrace{\int_0^t\left(\int_0^u g''(s)\,ds\right)du}_{= \int_0^t g''(s)(t-s)\,ds}$$

$$\left|g(t) - g(0) - g'(0)t\right| \le \max_{0\le s\le t}|g''(s)| \cdot \frac{t^2}{2}$$

Más general:

$$\left|g(\theta+\delta) - g(\theta) - g'(\theta)\overset{\to\,\theta+\delta-\theta}{\delta}\right| \le \left(\max_{\theta\le\varphi\le\theta+\delta}|g''(\varphi)|\right)\cdot\frac{\delta^2}{2}$$

Aplicamos a $f(x;\theta)$:

$$\left|f(x;\theta+\delta) - f(x;\theta) - \frac{\partial f}{\partial\theta}(x;\theta)\,\delta\right| \le \overbrace{\left(\max_{\theta\le\varphi\le\theta+\delta}\partial_\theta f(x;\varphi)\right)}^{(*)}\cdot\frac{\delta^2}{2}$$

$x\in$ sop $\qquad [\theta,\theta+\delta]\subset\Theta$

$$\left|\underbrace{\int_{sop} f(x;\theta+\delta)\,dx}_{=0} - \underbrace{\int_{sop} f(x;\theta)\,dx}_{=0} - \delta\int_{sop}\frac{\partial}{\partial\theta}f(x;\theta)\,dx\right| \le \frac{\delta^2}{2}\cdot\int_{sop}(*)\,dx$$

$$\Rightarrow \left|\delta\int_{sop}\frac{\partial}{\partial\theta}f(x;\theta)\,dx\right| \le \frac{\delta^2}{2}\int_{sop}(*)\,dx \;\Rightarrow\; \left|\int_{sop}\frac{\partial}{\partial\theta}f(x;\theta)\,dx\right| \le \frac{\delta}{2}\int_{sop}(*)\,dx$$

haciendo $\delta\to 0 \Rightarrow \displaystyle\int_{sop}\frac{\partial}{\partial\theta}f(x;\theta)\,dx = 0$

$$\| $$

$$\int_{sop}\partial_\theta \ln f(x;\theta)\, f(x;\theta)\,dx = \mathbb{E}_\theta(Y)$$

Hipótesis del TEOREMA:

X ; $f(x;\theta)$ continua $\qquad \theta\in\Theta=(a,b)$

$\qquad\qquad \theta\longrightarrow f(x;\theta)$ es C^2

I compacto

$I\subset\Theta$; $\displaystyle\int_{sop}\max_{\varphi\in I}\left|\partial^2_{\theta\theta}f(x;\varphi)\right|dx < \infty$

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

COMPORTAMIENTO ASINTÓTICO DE ESTIMADORES

$X \quad f(x;\theta) \qquad x_1, \dots, x_n$

Tenemos una sucesión T_n de estimadores de θ tal que:

$$T_n = h_n(x_1, \dots, x_n)$$

DEFINICIÓN: Decimos que T_n es SUCESIÓN CONSISTENTE de estimadores de θ si: $\lim_{n\to\infty} \mathbb{P}(|T_n - \theta| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0$

- Desigualdad de Chebyshev y consistencia

Si T_n es insesgado, la desigualdad de Chebyshev da

$$\mathbb{P}(|T_n - \theta| > \varepsilon) \leq \frac{V_\theta(T_n)}{\varepsilon^2}$$

Por tanto si T_n son insesgados y $\lim_{n\to\infty} V_\theta(T_n) = 0$, entonces T_n es consistente.

Ejemplo: Supongamos $\theta = \mathbb{E}(X)$.

En este caso

$$\mathbb{E}_\theta(\bar{X}_n) = \mathbb{E}_\theta(X) = \theta \quad \to \text{insesgado}$$

$$T_n = \underset{\uparrow \atop \frac{X_1 + \cdots + X_n}{n}}{\bar{X}_n}$$

$$V_\theta(\bar{X}_n) = \frac{V_\theta(X)}{n}$$

T_n insesgado y $V_\theta(T_n) \xrightarrow[n\to\infty]{} 0 \implies T_n$ consistente

- Distribución EXACTA de $\mathbb{P}(|T_n-\theta|>\varepsilon)$

(interesa la velocidad con que esta probabilidad tinde a 0)

Ejemplo: X es $N(\mu, \overset{\text{conocido}}{\sigma_0^2}) \Rightarrow \bar{X}_n$ es $N\left(\mu, \frac{\sigma_0^2}{n}\right)$

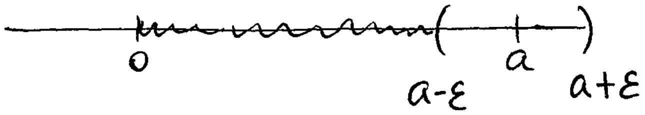
$T_n = \bar{X}_n$ Entonces: $\mathbb{P}(|T_n-\mu|>\varepsilon) = \mathbb{P}(|\bar{X}_n-\mu|>\varepsilon) =$

$$= \mathbb{P}\left(\underbrace{\frac{|\bar{X}_n-\mu|}{\sigma_0/\sqrt{n}}}_{N(0,1)} > \frac{\varepsilon\sqrt{n}}{\sigma_0}\right) = 2\left(1-\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma_0}\right)\right)$$

nos da el tamaño de muestra que necesitamos para tener un error menor que ε.

Ejemplo: $X \sim \text{Unif}[0,a]$

$M_n = \max(X_1,\dots,X_n)$ ¿ $\mathbb{P}_a(|M_n-a|>\varepsilon)$?



$$\mathbb{P}_a(|M_n-a|>\varepsilon) = \mathbb{P}_a(M_n<a-\varepsilon) =$$

$$= \left[\frac{1}{a}(a-\varepsilon)\right]^n = \underbrace{\left(1-\frac{\varepsilon}{a}\right)}_{<1}^n \Rightarrow \lim_{n\to\infty}(|M_n-a|>\varepsilon) = 0$$

M_n es sucesión consistente

↓ Comentario/"truquillo"

$$\mathbb{P}\Big(|M_n-a| > \underbrace{\frac{a\varepsilon}{n}}_{\overset{\shortparallel}{\varepsilon'}}\Big) = \Big(1-\underbrace{\left(\frac{a\varepsilon}{n}\right)}_{\overset{\shortparallel}{\varepsilon'}}\cdot\frac{1}{a}\Big)^n \xrightarrow[n\to\infty]{} e^{-\varepsilon}$$

$$\mathbb{P}\left(\frac{n}{a}|M_n-a|>\varepsilon\right) \xrightarrow[n\to\infty]{} \underbrace{e^{-\varepsilon}}_{\hookrightarrow \exp(1)}$$

Z es $\text{Exp}(\lambda)$

$\mathbb{P}(Z>\varepsilon) = e^{-\lambda\varepsilon}$

$$\Rightarrow \frac{n}{a}|M_n-a| \xrightarrow[\text{distr.}]{} \exp(1)$$

COMPORTAMIENTO ASINTOTICO DE ESTIMADORES POR MOMENTOS

X $\qquad\qquad$ $X_1, \dots, X_n$

$\mu = \mathbb{E}(X)$ $\qquad$ $\sigma^2 = \mathbb{V}(X)$ $\qquad$ $X_1, \dots, X_n$ clones de X

TCL dice:
$$\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} \xrightarrow{\text{distr.}} N(0,1)$$

y esto lo escribimos así:
$$\sqrt{n}\left(\overline{X_n} - \mu\right) \xrightarrow{\text{distr}} N(0, \sigma^2)$$

Estimador por momentos: $\quad X \quad f(x;\theta)$

$$\overline{X_n} = \mathbb{E}_\theta(X) \longleftarrow \text{función de } \theta$$

despejamos θ: $\boxed{M_\theta = g(\overline{X_n})}$

Ejemplos: $M_\theta = \overline{X_n}$; $\quad M_\theta = \dfrac{1}{\overline{X_n}}$; $\quad M_\theta = \sqrt{\overline{X_n^2}}$

TEOREMA (método delta): Z_n variables aleatorias con valores en un intervalo (a,b)

Las Z_n cumplen esta condición: $\sqrt{n}\left(Z_n - \alpha\right) \xrightarrow[n\to\infty]{\text{distr.}} N(0, \beta^2)$

$g : (a,b) \longrightarrow \mathbb{R}$, $\quad g$ es $C^2((a,b))$, $\quad g'(\alpha) \neq 0$

Entonces: $\sqrt{n}\left(g(Z_n) - g(\alpha)\right) \xrightarrow[n\to\infty]{\text{distr.}} N\left(0, |g'(\alpha)|^2 \cdot \beta^2\right)$

Idea demostración

$g(x) \cong g(x) + g'(\alpha)(x-\alpha)$; $\quad Z_n \cong \alpha + \frac{\beta}{\sqrt{n}} \overset{\longrightarrow N(0,1)}{Y}$

para la idea: $g(x) \boxed{=} g(\alpha) + g'(\alpha)(x-\alpha)$ $\quad$ y $\quad Z_n \boxed{=} \alpha + \frac{\beta}{\sqrt{n}} Y$

$g(Z_n) = g(\alpha) + g'(\alpha)(Z_n - \alpha) = g(\alpha) + g'(\alpha)\frac{\beta}{\sqrt{n}} Y$

$g(Z_n)$ es $N\left(g(\alpha), (g'(\alpha))^2 \frac{\beta^2}{n}\right)$

Aplicación

X $\quad f(x;\theta)$ $\quad X_1, X_2, \dots$

$$\sqrt{n}\left(\overline{X}_n - \underset{\substack{\uparrow \\ \mathbb{E}(X)}}{\mu}\right) \longrightarrow N(0, \underset{\substack{\uparrow \\ \mathbb{V}(X)}}{\sigma^2}) \qquad g'(\mu) \neq 0$$

Método delta: $\sqrt{n}\left(g(\overline{X}_n) - g(\mu)\right) \longrightarrow N\left(0, |g'(\mu)|^2 \cdot \sigma^2\right)$

↓ Muchos estimadores por momentos se escriben $g(\overline{X}_n)$, y esto nos dice que son asintóticamente normales ↑

Ejemplo: X es $\mathrm{Exp}(\lambda)$ $\qquad \mathbb{E}(X) = 1/\lambda \qquad \mathbb{E}(X) = 1/\lambda^2$

Usamos como estimador de λ: $\overline{X} \underset{\substack{\uparrow \\ \text{momentos}}}{=} \mathbb{E}(X) = 1/\lambda$

$M_{\lambda,n} = \dfrac{1}{\overline{X}_{(n)}}$ Método delta con $g(x) = 1/x$

$$\text{TLC} \Rightarrow \sqrt{n}\left(\overline{X}_n - \underset{\substack{\| \\ \mathbb{E}(X)}}{\frac{1}{\lambda}}\right) \longrightarrow N(0, \underset{\substack{\| \\ \mathbb{V}(X)}}{1/\lambda^2})$$

Por el método delta: $\sqrt{n}\left(\dfrac{1}{\overline{X}_n} - \lambda\right) \longrightarrow N\left(0, \underset{\substack{\uparrow \\ |g'(1/\lambda)|^2 = \lambda^4}}{\lambda^4} \cdot \dfrac{1}{\lambda^2}\right)$

$$\Rightarrow \sqrt{n}\left(M_{\lambda,n} - \lambda\right) \xrightarrow[n\to\infty]{\text{distr.}} N(0, \lambda^2)$$

Esto significa que para $\varepsilon > 0$ fijo, el margen de error:

$$\mathbb{P}\left(|M_{\lambda,n} - \lambda| > \varepsilon\right) = \text{probabilidad de equivocarse al estimar en un error mayor que } \varepsilon.$$

$$\| $$

$$\mathbb{P}\left(\frac{|M_{\lambda,n} - \lambda|}{\lambda/\sqrt{n}} \geqslant \frac{\varepsilon\sqrt{n}}{\lambda}\right) \cong 2\left(1 - \Phi\left(\frac{\varepsilon\sqrt{n}}{\lambda}\right)\right)$$

$= \lambda^2 \rightarrow \text{std} = \lambda$

Ejemplo: $X \sim Ber(p)$

$$\text{odds} = \frac{p}{1-p} := \text{parámetro a estimar}$$

$$T_n = \frac{\overline{X_n}}{1-\overline{X_n}}$$

- TCL $\quad \sqrt{n}\left(\overline{X_n} - p\right) \xrightarrow[n\to\infty]{distr.} N\left(0, p(1-p)\right)$
- Método delta, $\quad g(x) = \dfrac{x}{1-x}$

$$\Rightarrow \sqrt{n}\left(\frac{\overline{X_n}}{1-\overline{X_n}} - \frac{p}{1-p}\right) \longrightarrow N\left(0, \frac{p}{(1-p)^3}\right)$$

$$\sqrt{n}\left(T_n - \text{odds}\right) \longrightarrow N\left(0, \frac{p}{(1-p)^3}\right)$$

MÉTODO DELTA si $g'(\alpha) = 0$

$$\sqrt{n}\left(Z_n - \alpha\right) \xrightarrow[n\to\infty]{} N\left(0, \beta^2\right)$$

Idea: $\quad g(x) - g(\alpha) \cong \frac{1}{2} g''(\alpha)(x-\alpha)^2 \qquad$ (ahora $g''(\alpha) \neq 0$)

$$Z_n \cong \alpha + \frac{\beta}{\sqrt{n}} Y$$

Y ⟶ normal(0,1)

idea ⇒ "="

$$\Rightarrow g(Z_n) - g(\alpha) = \frac{1}{2} g''(\alpha)\left(Z_n - \alpha\right)^2 = \frac{1}{2} g''(\alpha) \frac{\beta^2}{n} Y^2$$

Y^2 ⟶ χ^2_1

Entonces:

$$n\left(g(Z_n) - g(\alpha)\right) = \frac{1}{2} \underbrace{g''(\alpha)} \beta^2 \chi^2_1$$

⟶ si esto es positivo siempre estamos estimando para arriba.
Si es negativo, siempre estimamos para abajo.
hay sesgo, error sistemático

MÉTODO DELTA si $g'(\alpha) = 0$, $g''(\alpha) \neq 0$ y $g \in C^3$

$$\boxed{n\left(g(Z_n) - g(\alpha)\right) \cong \frac{1}{2} g''(\alpha) \beta^2 \chi_1^2}$$

INTERVALOS DE CONFIANZA

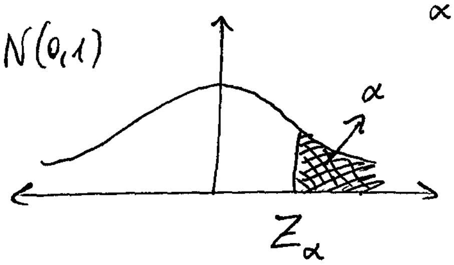
Ejemplo básico:

X es $N(\underset{\uparrow\ \text{a estimar}}{\mu}, \overset{\text{conocido}}{\sigma_0^2})$ $\qquad \underbrace{X_1, \dots, X_n}_{\text{clones}}$

$\bar{X}$ es $N\left(\mu, \frac{\sigma_0^2}{n}\right) \qquad \frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}} \sim N(0,1)$

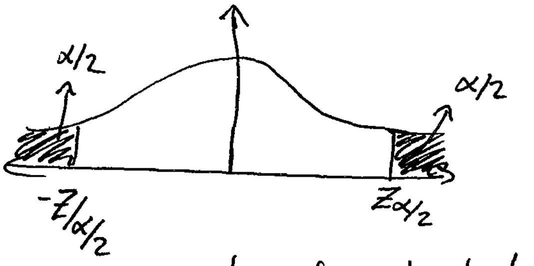
Notación percentiles:

α pequeño



$\phi(Z_\alpha) = 1-\alpha \implies Z_\alpha = \phi^{-1}(1-\alpha)$

$\implies \mathbb{P}\left(\left|\frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}}\right| > Z_{\alpha/2}\right) = \alpha$



por el valor absoluto

$$\mathbb{P}\left(\mu - Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}} \leq \bar{X} \leq \mu + Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right) = 1-\alpha$$

(PROBABILIDAD ANTES DE LA MUESTRA)

Prob: antes de la muestra

Estad: tenemos una muestra

Estadística: muestra $x_1, \dots, x_n \longrightarrow \bar{x}$

Con confianza $1-\alpha$ se tiene $\mu - Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}} \leq \bar{X} \leq \mu + Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}$

ponemos a μ en el papel central: $\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{}}$

Al intervalo $\left(\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}, \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$ lo llamamos INTERVALO DE CONFIANZA μ de $N(\mu, \sigma_0^2)$ con σ_0^2 conocido.

lamaño del intervalo: $2 Z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$

Dos grados de libertad $\begin{cases} \alpha \\ n \end{cases}$

Si aumentamos confianza $(1-\alpha)$ con n fijo
$\rightarrow$ crece $(1-\alpha)$
$\rightarrow$ decrece α
$\rightarrow$ crece Z_α y $Z_{\alpha/2}$
$\rightarrow$ crece el intervalo

Si aumentamos n con α fijo $\longrightarrow$ decrece el intervalo

Tamaño de la muestra (para un error determinado)

Fijamos $\overbrace{Z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}}^{\text{error}}$ (con α fijado)

Esto determina un n mínimo.

Ejemplo: error permitido es de 0'01 y $\alpha = 5\%$ $(Z_{\alpha/2} = 1'96)$
$\sigma_0 = 1$

$$\Longrightarrow n \geq \left(Z_{\alpha/2} \frac{\sigma_0}{0'01}\right)^2 \Longrightarrow n \geq 38416$$

Ejemplo: $X \sim N(\mu, \sigma^2)$ $\quad x_1, \dots, x_n$
$\rightarrow \bar{x} = \frac{x_1 + \dots + x_n}{n}$
$\rightarrow s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Intervalo para μ e intervalo para σ^2?

Probabilidad

$\bar{X} \sim N(\mu, \sigma^2/n)$ $\quad \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ $\quad \underbrace{\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\text{con confianza } (1-\alpha)}$

$\boxed{\text{Para } \mu:}$ $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ es t_{n-1} de Student

variables aleatorias

$$\Longrightarrow \mathbb{P}\left(-t_{n-1,\alpha/2} \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq t_{n-1,\alpha/2}\right) = 1-\alpha$$

⊛ $\bar{x}$ estimador insesgado de μ

Deutro del ámbito estadístico (ya tenemos una muestra)

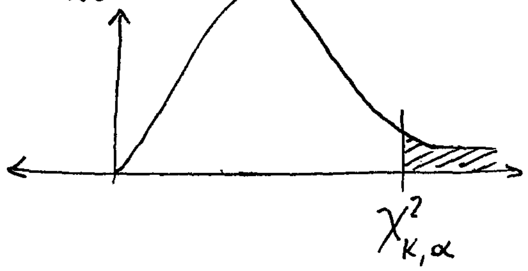
$$\boxed{\bar{x} - \frac{s}{\sqrt{n}} t_{n-1,\alpha/2} \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}} t_{n-1,\alpha/2}}$$

Intervalo para μ de $N(\mu,\sigma^2)$ con confianza (1-

▶ $\boxed{\text{Para } \sigma^2:}$ S^2 estimador de σ^2 (insesgado)

$$(n-1)\frac{S^2}{\sigma^2} \text{ es } \chi^2_{n-1}$$

Percentiles:



$$\Rightarrow \mathbb{P}\Big(\underbrace{\chi^2_{n-1,1-\alpha/2}}_{} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{n-1,\alpha/2}\Big) = 1-\alpha$$

↳ debido a de χ^2_K no es simétrica

Intervalo para σ^2 de $N(\mu,\sigma^2)$ con confianza $(1-\alpha)$

de la muestra

$$\boxed{\frac{(n-1)}{\chi^2_{n-1,\alpha/2}} \cdot S^2 \leq \sigma^2 \leq \frac{(n-1)}{\chi^2_{n-1,1-\alpha/2}} \cdot S^2}$$

Ejemplo: $N(\mu, \sigma^2)$

α	$Z_{\alpha/2}$
5%	1'96
1%	2'58
0'1%	3'29

$n = 30$ $\qquad \bar{x} = 1'3$ $\qquad S^2 = 0'36$

Intervalo al 95% de μ y de σ^2

$$\mu = 1'3 \pm \underbrace{t_{29,\,2'5\%}}_{\overset{\shortparallel}{2'045\text{ (excel)}}} \cdot \frac{0'06}{\sqrt{30}}$$

$$\frac{29 \cdot 0'36}{45'72} \le \sigma^2 \le \frac{29 \cdot 0'36}{16'04}$$

INTERVALOS ASINTÓTICOS DE CONFIANZA

Idea: Tenemos un estadístico H_n (que usa $x_1, \dots, x_n$)

$$\sqrt{n}\,(H_n - a) \xrightarrow{distr.} N(0, b^2)$$

$$\mathbb{P}\left(-Z_{\alpha/2} \le \frac{H_n - a}{b/\sqrt{n}} \le Z_{\alpha/2}\right) \approx 1 - \alpha$$

un igual a partir de ahora

$\Downarrow$

Intervalo: $\quad a = h_n \pm \frac{b}{\sqrt{n}} Z_{\alpha/2} \quad \left(h_n - \frac{b}{\sqrt{n}} Z_{\alpha/2} \le a \le h_n + \frac{b}{\sqrt{n}} Z_{\alpha/2}\right)$

Confianza $1-\alpha$

h_n → muestra de $H_n(x_1, \dots, x_n)$

Ejemplo: X es $Ber(p)$ $\qquad x_1, \dots, x_n$ muestra $\qquad \bar{x} =$ proporción de 1's.

$$\sqrt{n}\,(\bar{X}_n - p) \longrightarrow N(0, p(1-p))$$

Intervalo de confianza $1-\alpha$: $\underbrace{\bar{x} - \frac{\sqrt{p(1-p)}}{\sqrt{n}} Z_{\alpha/2}}_{} \le p \le \underbrace{\bar{x} + \frac{\sqrt{p(1-p)}}{\sqrt{n}} Z_{\alpha/2}}_{}$

como depende de p arreglamos el intervalo

alternativas: ① $p(1-p) \le \frac{1}{4} \Rightarrow \sqrt{p(1-p)} \le \frac{1}{2}$

Confianza $1-\alpha$:

$$\bar{x} - \frac{1/2}{\sqrt{n}} Z_{\alpha/2} \le p \le \bar{x} + \frac{1/2}{\sqrt{n}} Z_{\alpha/2}$$

② Utilizar $\bar{x}$ como estimador:

$\longrightarrow$ media de la muestra pequeña y barata

$$\bar{X} - \frac{\sqrt{\bar{x}(1-\bar{x})}}{\sqrt{n}} Z_{\alpha/2} \le p \le \bar{x} + \frac{\sqrt{\bar{x}(1-\bar{x})}}{\sqrt{n}} Z_{\alpha/2}$$

$\downarrow$
media de
la muestra
de verdad (grande)

TAMAÑO DE MUESTRA

BERNOULLI:

Margen de error (absoluto), dado en porcentaje

error = ± 1% = ±0'01

$\alpha = 5\%$

$Z_{\alpha/2} = 1'96$

tamaño de la muestra conservador: $\frac{1}{2\sqrt{n}} \cdot 1'96 \le 0'01$

Alternativa: muestra previa $\longrightarrow$ estimación $\hat{p} = 0'3$

$$\frac{\sqrt{0'3(1-0'3)}}{\sqrt{n}} \cdot 1'96 \le 0'01$$

Ejemplo: Intervalo $(1-\alpha)$

X es Poisson (λ)

$E_\lambda(X) = \lambda$ $\qquad$ $V_\lambda(X) = \lambda$

$$\sqrt{n}\,(\bar{X}-\lambda) \longrightarrow N(0,\lambda)$$

Intervalo: $$\bar{x} - \sqrt{\lambda}\,\frac{Z_{\alpha/2}}{\sqrt{n}} \le \lambda \le \bar{x} + \sqrt{\lambda}\,\frac{Z_{\alpha/2}}{\sqrt{n}}$$

$\sqrt{\lambda} \longrightarrow \sqrt{\bar{x}}$

$\longrightarrow \sqrt{\bar{x}}$ (solución ingenieril)

INTERVALOS DE DOS POBLACIONES

Normales / Bernouillis

Ejemplo básico: $X \quad N(\mu_1, \sigma_1^2) \quad X_1, \dots, X_{n_1} \quad n_1$ tamaño

$Y \quad N(\mu_2, \sigma_2^2) \quad Y_1, \dots, Y_{n_2} \quad n_2$ tamaño

- Intervalo para $\mu_1 - \mu_2 \;\big|\; \sigma_1^2, \sigma_2^2$ son conocidas

$\bar{X} - \bar{Y}$, y como son indep. $\bar{X}$ e $\bar{Y}$

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

$$\bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

Probabilidad:

$$\mathbb{P}\left(-z_{\alpha/2} \leq \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

Confianza $1-\alpha$:

$$-z_{\alpha/2} \leq \frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{\sqrt{\dots}} \leq z_{\alpha/2} \quad \Rightarrow$$

$$\Rightarrow \bar{x} - \bar{y} - z_{\alpha/2}\sqrt{\dots} \leq \mu_1 - \mu_2 \leq \bar{x} - \bar{y} + z_{\alpha/2}\sqrt{\dots}$$

- Intervalo para $\mu_1-\mu_2$ con σ_1^2, σ_2^2 desconocidas
  pero $\sigma_1=\sigma_2=\sigma$ desconocid

$$\begin{aligned} & (n_1-1)S_1^2 = \sigma_1^2\chi^2_{n_1-1} \\ + \ & (n_2-1)S_2^2 = \sigma_2^2\chi^2_{n_2-1} \end{aligned}$$

independientes

$$(n_1-1)S_1^2 + (n_2-1)S_2^2 = \sigma^2\chi^2_{n_1-1+n_2-1} \quad (1)$$

Introducimos un $S_p^2 \leftarrow$ promedio $\Longrightarrow \; S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1-1+n_2-1)} \quad (2)$

(1) con (2) : $(n_1-1+n_2-1)S_p^2 = \sigma^2\chi^2_{n_1-1+n_2-1}$

Con esto obtenemos :

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad \text{es} \quad t_{n_1-1+n_2-1}$$

Intervalo:

$$\bar{x}-\bar{y}-\hat{S}_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{n_1-1+n_2-1;\,\alpha/2} \le \mu_1-\mu_2 \le \bar{x}-\bar{y}+\hat{S}_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,t_{n-1+n_2-1;\,\alpha/2}$$

COMPARACION DE PROPORCIONES

X es $\mathrm{Ber}(p_1)$ $\quad n_1 \quad$ $\bar{x}$ - proporción de 1's de

Y es $\mathrm{Ber}(p_2)$ $\quad n_2 \quad$ $\bar{y}$ - proporción de 1's de

Intervalo: $\qquad \bar{x} \sim N(p_1, \frac{p_1(1-p}{n_1}$

$$p_1 - p_2 = (\bar{x} - \bar{y}) \pm Z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$$\boxed{\begin{array}{l} p_1 \approx \bar{x} \\ p_2 \approx \bar{y} \end{array}}$$

CONTRASTE DE HIPÓTESIS

Ejemplo inicial: Moneda que se supone equilibrada

$\bar{x}$ = proporción de caras

$n=10 \longrightarrow$ #caras $=6 \longrightarrow \bar{x}=60\%$

$n=100 \longrightarrow$ #caras $=60 \longrightarrow \bar{x}=60\%$

$n=1000 \longrightarrow$ #caras $=600 \longrightarrow \bar{x}=60\%$

- Ingredientes del contraste de hipótesis:

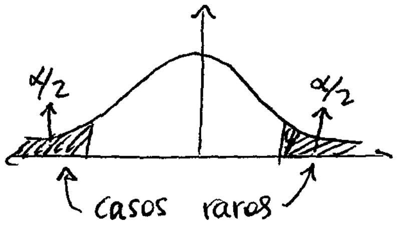
$$X \qquad N(\underset{\substack{\uparrow \\ \text{desconocido}}}{\mu}, \sigma^2) \quad \longleftarrow \text{conocido } (\sigma^2)$$

(1) HIPÓTESIS NULA: $H_0 : \mu = \mu_0$ (hipótesis de partida)

Nosotros vamos a hacer un experimento para ver si esta hipótesis se cumple o no.

(2) TEST/CONTRASTE: Damos $\alpha > 0$, próximo a cero

Si H_0 es cierto $\Longrightarrow \bar{x} \in \left(\mu_0 - \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma, \mu_0 + \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma\right)$



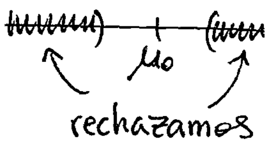
con probabilidad $1-\alpha$.

Hacemos experimento: si $\bar{x} \notin \left(\mu_0 - \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma, \mu_0 + \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma\right)$, entonces rechazamos la hipótesis H_0, en caso contrario Aceptamos H_0.

(3) HIPÓTESIS ALTERNATIVA: $H_1 : \mu \neq \mu_0$

(4) REGIÓN DE RECHAZO: R_α

$$R_\alpha = \left(-\infty, \mu_0 - \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma\right) \cup \left(\mu_0 + \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma, \infty\right)$$



(5) ERRORES :

5.1. ERROR DE TIPO 1 : rechazar H_0 si es cierta

5.2. ERROR DE TIPO 2 : aceptar H_0 cuando es falsa

En el ejemplo ⟶ $\mathbb{P}(\text{error de tipo } 1) = \alpha$ (pequeño)

Si lo rechazamos, tenemos CONFIANZA de que lo hemos rechazado bien.

Si aceptamos, es porque no nos queda más remedio.

Ejemplo: $N(\mu, 4)$ $n=100$ $\alpha = 1\%$ $\alpha = 5\%$

$H_0 \Rightarrow \mu_0 = 1$

$$R_\alpha : |\bar{x} - 1| > Z_{\alpha/2} \cdot \frac{\overset{\swarrow \sigma}{2}}{\underset{\leftarrow \sqrt{100}}{10}}$$

Hacemos un experimento y obtenemos $\bar{x} = 1'4$:

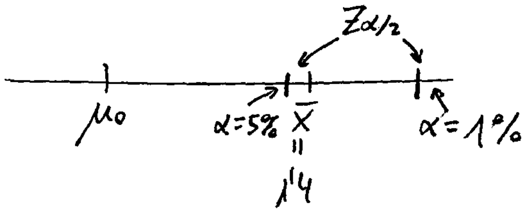
para $\alpha = 1\%$ ⟶ Rechazo $|\bar{x} - 1| > 2'58 . (0'2) = 0'52$

para $\alpha = 5\%$ ⟶ Rechazo $|\bar{x} - 1| > 1'96 . (0'2) = 0'39$

α = nivel de significación

Con significación de 1% ⟹ aceptamos

Con significación de 5% ⟹ rechazamos



El P-VALOR es el valor de α tal que $\bar{x}$ está en el borde de la región de rechazo.

$$|\bar{x} - \mu_0| = \frac{Z_{\alpha/2}}{\sqrt{n}} \sigma \Rightarrow |\bar{x} - \mu_0| \frac{\sqrt{n}}{\sigma} = Z_{\alpha/2} \Rightarrow$$

$$\Rightarrow \boxed{\alpha = 2\left(1 - \phi\left(|\bar{x} - \mu_0| \frac{\sqrt{n}}{\sigma}\right)\right)}$$

⟶ cuanto menor sea, más seguro estás de haber rechazado bien

⊛ Cuanto menor es el p-valor, mejor ha sido el experimento esta-

Ejemplo: $Ber(p)$ $\bar{x}$ = proporción en la muestra

$H_0 : p = p_0$

α: nivel de significación

n: tamaño de la muestra

Región de rechazo: $R_\alpha = \left\{ |\bar{x} - p_0| > \frac{Z_{\alpha/2}}{\sqrt{n}} \sqrt{p_0(1-p_0)} \right\}$

Retomamos el ejemplo inicial: ejemplo de "mosqueo"

$n = 10$ $\#caras = 6$

$n = 100$ $\#caras = 60$ $\bar{x} = 60\%$ en las tres muestras

$n = 1000$ $\#caras = 600$

$H_0 \Rightarrow p = p_0 = 0'5$

p-valor: $\underset{|\bar{x} - p_0|}{\overset{}{0'1 \,\|}} = \frac{Z_{\alpha/2}}{\sqrt{n}} \cdot \frac{1}{2} \longrightarrow \alpha = 2\left(1 - \phi\left(0'2\sqrt{n}\right)\right)$

p-valor para $n=10$ $\#caras = 6 \Rightarrow \sim 52$

p-valor para $n=100$ $\#caras = 60 \Rightarrow \sim 4\%$

p-valor para $n=1000$ $\#caras = 600 \Rightarrow \sim 2.1$

Hemos visto <u>TEST BILATERALES</u>.

$\bar{X} \sim N(\mu, \sigma^2)$ ← conocido

tenemos μ_0 de referencia

$H_0 \equiv \mu \leq \mu_0$

$(H_1 \equiv \mu > \mu_0)$

Región de rechazo: $\boxed{R_\alpha : \bar{x} > \mu_0 + \frac{\sigma}{\sqrt{n}} Z_\alpha}$

y esto de aquí son TEST UNILATERALES

Diseño de test unilateral

Referencia μ_0

$N(\mu, \sigma^2)$ con σ^2 conocido

$\bar{x}$ más grande que μ_0.

¿$H_0 : \mu \leq \mu_0$? ó ¿$H_0 : \mu \geq \mu_0$? → Planteamos $H_0 : \mu \leq \mu_0$ porque el test lo que hace bien es rechazar.

TESTS GENERALES

Ejemplo base: control de calidad

Producción: cajas con 10000 tornillos

Control de calidad: tomamos una caja y una muestra de 10 tornillos.

La caja es buena si la proporción de defectuosos es ≤5%

Test: Si hay ≥2 tornillos defectuosos de los 10 de la muestra declaramos la caja como no válida.

Nos interesa la probabilidad de rechazar cajas buenas.

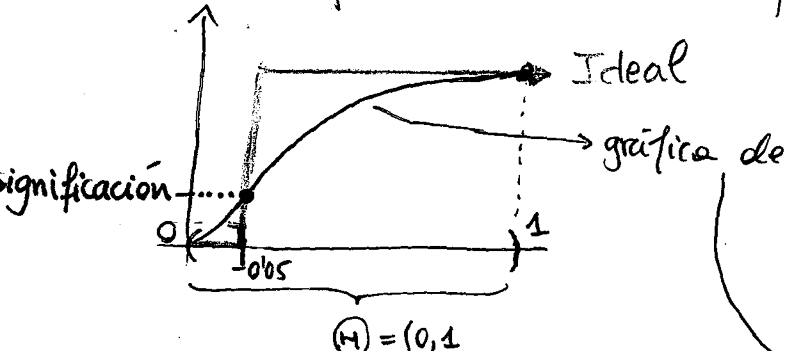
Notación: $p =$ prop. de defectuosas en la caja

Hipótesis nula: $H_0 : p \leq 5\%$

$n = 10$

Rechazamos si #defectuosas ≥ 2

Función de potencia del test: $p \in \underset{(0,1)}{\overset{\shortparallel}{\Theta}} \longrightarrow \beta(p) := \mathbb{P}_{\underset{\nwarrow \text{test}}{p}}(\text{rechazar})$



En este caso:

$$\beta(p) = \mathbb{P}_p(\#\text{defectuosos}_{\to \text{en muestra}} \geq 2) =$$

$$= \mathbb{P}(\text{Bin}(10, p) \geq 2) =$$

$$= 1 - \mathbb{P}(\text{Bin}(10,p) = 1) - \mathbb{P}(\text{Bin}(10,p) = 0)$$

$$= 1 - (1-p)^{10} - 10p(1-p)^9$$

Probabilidad de error de tipo 1:

$$\underset{\hookrightarrow \text{ cuando } H_0 \text{ es bueno}}{\mathbb{P}_{\square}(\text{rechazar})} = \mathbb{P}_p(\text{rechazar}) \quad \text{con} \quad p \in \underset{(0{,}5\%)}{\underset{\shortparallel}{\Theta_0}}$$

En este caso:

$$\text{Significación} = \mathbb{P}_{5\%}(\text{rechazar}) = 1 - (95\%)^{10} - 10p(1-p)^{9} \approx 8\%$$

<u>Ejemplo</u>: Propongo el siguiente test

$$X \sim N(\mu, \underset{\uparrow\!\_\_ \text{ conocido}}{\sigma^2}) \qquad \Theta = \mathbb{R}$$

$$\underset{\nearrow\, \text{hipótesis nula}}{\Theta_0} = \{\mu \leq \mu_0\} = (-\infty, \mu_0)$$

Queremos func. de potencia y significación.

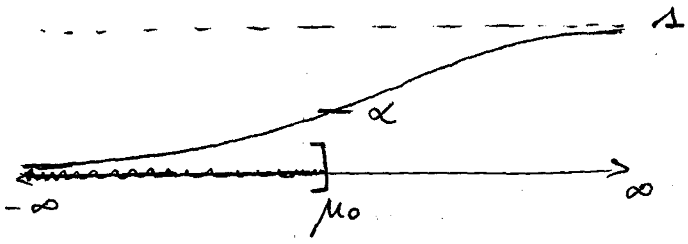
n = tamaño de la muestra

Rechazamos cuando: $\bar{x} \geq \mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}}$

Función de potencia:

$$\beta(\underset{\underset{\in\mathbb{R}}{\uparrow}}{\mu}) = \mathbb{P}_\mu\Big(\overbrace{\bar{X} \geq \mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}}}^{\text{rechazar}}\Big) = \mathbb{P}_\mu\Big(\underbrace{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}_{N(0,1)} \geq \frac{\mu_0-\mu}{\sigma/\sqrt{n}} + Z_\alpha\Big) =$$

$$= 1 - \Phi\Big(\frac{\mu_0-\mu}{\sigma/\sqrt{n}} + Z_\alpha\Big)$$



$$\beta(\mu_0) = 1 - \Phi(Z_\alpha) = 1 - (1-\alpha) = \alpha$$

$$\text{Significación} = \sup_{\mu\in\Theta_0} \beta(\mu)$$

$$\Rightarrow \text{Significación} = \beta(\mu_0) = \alpha$$

Notación:

Θ = espacio de parámetros

ipótesis nula $\nearrow$ $H_0 = \theta \in \Theta_0 \subset \Theta$

Rechazo $\begin{cases} n = \text{tamaño de la muestra} \\ T \text{ estadístico} \\ \text{Rechazo si } T \in \mathcal{R} \end{cases}$

$\uparrow$ región de rechazo

Función de potencia $\beta(\theta) = \mathbb{P}_\theta(T \in \mathcal{R})$

Significación del test: $\sup_{\theta \in \Theta_0} \beta(\theta)$

Ejemplo: Unif $[0, a]$

$\Theta = (0, \infty)$

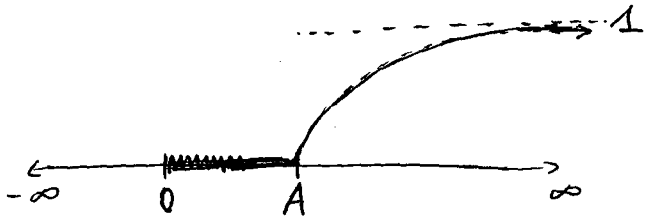
$\Theta_0 = (0, A)$

$H_0: a \leq A$

Rechazamos si $\max(x_1, \ldots, x_n) > A$

$$\beta(a) = \mathbb{P}_a\left(\max(x_1, \ldots, x_n) > A\right) = \begin{cases} 0 & \text{si } a < A \\ \circledast & \text{si } a > A \end{cases}$$

$$\circledast = \mathbb{P}_a\left(\max(x_1, \ldots, x_n) > A\right) = \mathbb{P}_a(A < \max(\cdots) < a) =$$
$$= 1 - \mathbb{P}_a\left(\max(\cdots) < A\right) = 1 - \left(\frac{A}{a}\right)^n \quad \text{si } a > A$$



Significación = 0

↳ jamás cometeríamos un error de tipo 1

Ejemplo: Unif$[0,a]$ $\quad a \in \Theta = (0,\infty)$ $\quad x_1,\dots,x_n$ muestra

$n=10$

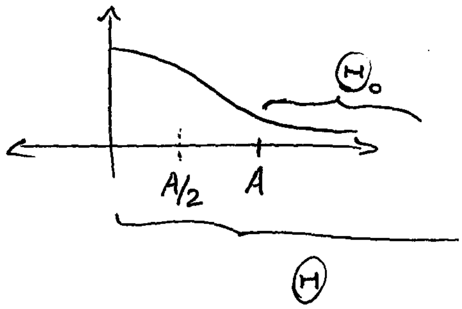
$$\underset{\text{hipótesis nula}}{H_0} = a \geq \underset{\text{dato}}{A}$$

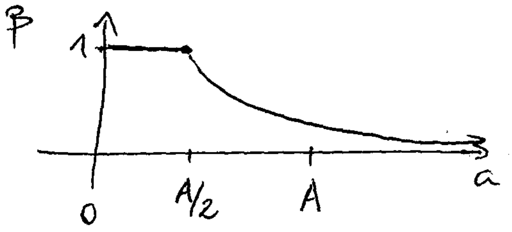
Rechazo si $\underbrace{\max(x_1,\dots,x_n) \leq \frac{A}{2}}_{\text{región de rechazo}}$

1) Función de potencia

$a \in \Theta$; $\quad \beta(a) = \mathbb{P}_a(\text{rechazar})$

$$\beta(a) = \begin{cases} 1 & \text{si } a \leq A/2 \quad \left(\begin{array}{l}\text{en este caso} \\ x_1,\dots,x_n \leq a \leq A/2\end{array}\right) \\ \left(\frac{A/2}{a}\right)^n & \text{si } a \geq A/2 \end{cases}$$





Significación :

$$\sup_{a\in\Theta_0} \mathbb{P}_a(\text{rechazar}) = $$
$$= \sup_{a\in\Theta_0} \beta(a) = \beta(A) = \left(\frac{1}{2}\right)^n \approx$$

$n=10$ $\searrow \approx 0'1\%$

Nueva región de rechazo porque la significación es muy pequeña:

$$\beta(a) = \begin{cases} 1 & a \leq ta \\ \left(\frac{tA}{a}\right)^n & a \geq tA \end{cases}$$

significación: $\underset{a=A}{\beta(A)} = t^n = t^{10} = 5\% \longrightarrow$ despejar t

$\hookrightarrow t = 74'1\%$

Otro contraste: Unif $[0,a]$ $\Theta_0 = [a \geq A] = (A, \infty)$

Nos mosqueamos si hay "muchos" "pequeños". $x_1, \dots, x_n$ muestra $n=10$

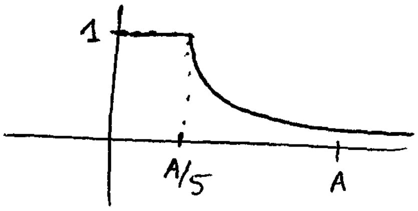
Rechazo si $\#\{x_j < \frac{A}{5}\} \geq \frac{n}{2}$

1) Función de potencia 2) Significación

$$\mathbb{P}_a(\text{rechazar}) = \mathbb{P}\left(\text{Bin}(\overset{n}{\overset{\|}{10}}, \underbrace{\tfrac{A/5}{a}}_{\downarrow}) \geq \overset{n/2}{\overset{\|}{5}}\right) \quad \text{si } a \geq \frac{A}{5}$$

$$\mathbb{P}_a(X < A/5) = \frac{A/5}{a}$$

$$\Rightarrow \beta(a) = \begin{cases} 1 & \text{si } a \leq A/5 \\ \mathbb{P}(\text{Bin}(10, \frac{A/5}{a}) \geq 5) & \text{si } a \geq A/5 \end{cases}$$



$$\text{Significación} = \sup_{a \in \Theta_0} \beta(a) = \beta(A) = \mathbb{P}(\text{Bin}(10, 1/5) \geq 5) = 3'28\%$$

MÉTODO DE RAZÓN DE VEROSIMILITUDES (TEST DE RAZÓN DE VEROS.)

(Método general para construir contrastes)

X $f(x;\theta)$ $\theta \in \Theta$

$H_0 \equiv \theta \in \Theta_0 \subset \Theta$ $x_1, \dots, x_n$ muestra de X.

$$\text{VERO}(\theta; x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i;\theta)$$

Si $\sup_{\theta \in \Theta_0} \text{VERO}(\dots)$ es pequeña, los $\theta \in \Theta_0$ son poco creíbles.
Son poco consistentes con la muestra.

Se toma $c \in (0,1)$ donde c se denomina CALIBRE.

Rechazo Θ_0 si $$\frac{\overset{\Theta_0}{\sup_{\theta \in \Theta_0}} \text{VERO}(\dots)}{\sup_{\theta \in \Theta} \text{VERO}(\dots)} \leq c$$

Ejemplos de test de razón de verosimilitudes con calibre c.

① $X \sim N(\mu, \overset{\text{conocido}}{\sigma^2})$ Para simplificar $\sigma = 1$

n = tamaño de la muestra

$x_1, \dots, x_n$

calibre $\longrightarrow c \in (0,1)$

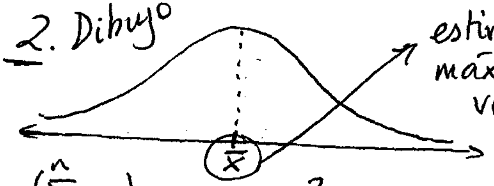
$$\boxed{H_0 \equiv \mu \leq \mu_0}$$

1. Función de verosimilitud

$$VERO(\mu; x_1, \dots x_n) = \prod_{i=1}^{n} f(x_i, \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\mu)^2} =$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

2. Dibujo



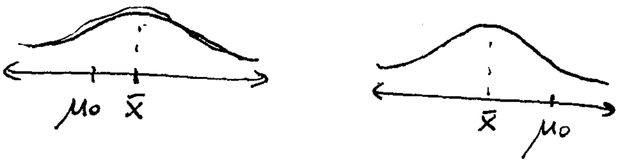
estimador de máxima verosimilitud

Observemos $\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n} x_i^2 - 2\left(\sum_{i=1}^{n} x_i\right)\mu + n\mu^2 \longleftarrow$ mínimo en $\mu = \bar{x}$

3. $\Theta = (-\infty, \infty)$ $\Theta_0 = (-\infty, \mu_0)$

$$\sup_{\mu \in \Theta} VERO(\mu; x_1, \dots, x_n) = VERO(\bar{x}; x_1, \dots, x_n)$$

¿ $\sup_{\mu \in \Theta_0} VERO(\mu; x_1, \dots, x_n)$? $\xrightarrow{\text{dos casos}}$



$$= \begin{cases} VERO(\mu_0, x_1, \dots, x_n) & \text{si } \mu_0 < \bar{x} \\ VERO(\bar{x}; x_1, \dots, x_n) & \text{si } \mu_0 > \bar{x} \end{cases}$$

4. Razón de verosimilitudes = RV

$$RV = \begin{cases} \dfrac{VERO(\bar{x})}{VERO(\mu_0)} = \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x})^2\right)\right) & \text{si } \mu_0 < \bar{x} \\ 1 & \text{si } \mu_0 \geq \bar{x} \end{cases}$$

5. Simplificación

$$\sum_{i=1}^{n}(x_i-\mu_0)^2 = \cancel{\sum_{i=1}^{n} x_i^2} - 2\left(\sum_{i=1}^{n} x_i\right)\mu_0 + n\mu_0^2 = -2n\bar{x}\mu_0 + n\mu_0^2$$

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = \cancel{\sum_{i=1}^{n} x_i^2} - 2\left(\sum_{i=1}^{n} x_i\right)\bar{x} + n\bar{x}^2 = -2n\bar{x}\bar{x} + n\bar{x}^2$$

$$\Rightarrow -2n\bar{x}\mu_0 + n\mu_0^2 + 2n\bar{x}\bar{x} - n\mu_0^2 - n\bar{x}^2 =$$

$$= n\left(-2\bar{x}\mu_0 + \mu_0^2 + 2\bar{x}^2 - \bar{x}^2\right) = n\left(-2\bar{x}\mu_0 + \mu_0^2 + \bar{x}^2\right) =$$

$$= n(\mu_0-\bar{x})(\mu_0+\bar{x}-2\bar{x}) = n(\mu_0-\bar{x})^2$$

$$\Rightarrow RV = \begin{cases} \exp\left(-\frac{1}{2}n(\mu_0-\bar{x})^2\right) & \text{si } \mu_0 < \bar{x} \\ 1 & \text{si } \mu_0 \geq \bar{x} \end{cases}$$

6. Región de rechazo: (Test)

Rechazamos cuando $RV \leq C \Rightarrow$

$$\Rightarrow \begin{cases} \exp\left(-\frac{1}{2}n(\mu_0-\bar{x})^2\right) \leq C \;\wedge\; \mu_0 \leq \bar{x} \quad ⊛ \\ \text{ó cuando} \\ \underline{1 \leq C} \;\wedge\; \mu_0 \geq \bar{x} \end{cases}$$

⟶ esto no pasa nunca

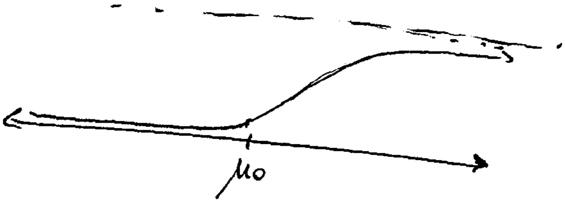
$$⊛ \quad \bar{x} \geq \mu_0 \;\wedge\; (\mu_0-\bar{x})^2 \geq \frac{2}{n}\ln\frac{1}{C} \iff \boxed{\bar{x} > \mu_0 + \sqrt{\frac{2}{n}\ln\frac{1}{c}}}$$

7. Función de potencia y significación

$$\beta(\mu) = \mathbb{P}_\mu(\text{rechazar}) \qquad \mu \in \Theta = (-\infty, \infty)$$

$$\beta(\mu) = \mathbb{P}_\mu\left(\bar{x} > \mu_0 + \frac{1}{\sqrt{n}}\sqrt{2\ln\frac{1}{c}}\right) = \mathbb{P}_\mu\left(\frac{\bar{x}-\mu}{\underset{\swarrow\ \sigma}{1/\sqrt{n}}} > \frac{\mu_0-\mu}{1/\sqrt{n}} + \sqrt{2\ln\frac{1}{c}}\right)$$

$$= 1 - \Phi\left(\frac{\mu_0-\mu}{1/\sqrt{n}} + \sqrt{2\ln\frac{1}{c}}\right)$$



Significación $= \sup\limits_{\mu\in\Theta_0} \beta(\mu) = \beta(\mu_0) = 1 - \Phi\left(\sqrt{2\ln\frac{1}{c}}\right)$