Estadística I

Grado en Matemáticas, UAM, 2017-2018

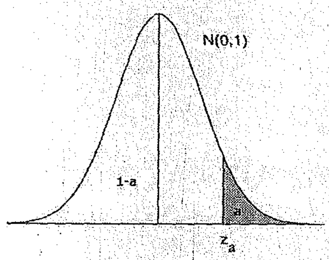
Listado de los intervalos de confianza más habituales

Notación para percentiles

A. Percentiles de la normal estándar

Sea Z una variable normal estándar. Para $\alpha \in (0, 1/2]$ se denota por z_α el valor real tal que

$$\mathbf{P}(Z > z_\alpha) = \alpha\,.$$



Nótese que

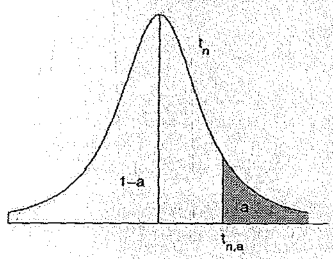
$$\mathbf{P}(|Z| < z_{\alpha/2}) = 1 - \alpha\,.$$

Para calcular $z_\alpha = \Phi^{-1}(1-\alpha)$ en excel: `=inv.norm.estand(1-`α`)`.

B. Percentiles de la t de Student con n grados de libertad

Sea Z una $\textsc{stu}(n)$. Para $\alpha \in (0, 1/2]$, denotamos por $t_{\{n\,;\,\alpha\}}$ al valor tal que

$$\mathbf{P}(Z > t_{\{n\,;\,\alpha\}}) = \alpha\,.$$

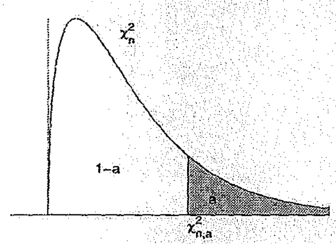


Nótese además que $\mathbf{P}(|Z| < t_{\{n\,;\,\alpha/2\}}) = 1 - \alpha$.
Para calcular $t_{\{n;\alpha\}}$ en excel: `=inv.t(1-`α`;`n`)`.

C. Percentiles de la χ^2_n

Sea Z una chi cuadrado con n grados de libertad. Para $\alpha \in (0,1)$, denotamos $\chi^2_{\{n\,;\,\alpha\}}$ al valor tal que

$$\mathbf{P}(Z > \chi^2_{\{n\,;\,\alpha\}}) = \alpha\,.$$



Obsérvese que

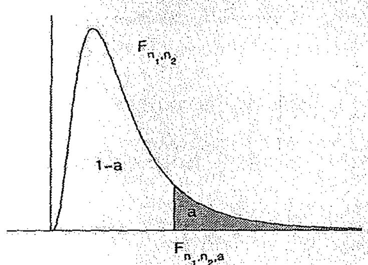
$$\mathbf{P}(\chi^2_{\{n\,;\,1-\alpha/2\}} < Z < \chi^2_{\{n\,;\,\alpha/2\}}) = 1 - \alpha\,.$$

Para calcular $\chi^2_{\{n;\alpha\}}$ en excel: `=inv.chicuad(1-`α`;`n`)`.

D. Percentiles de la F de Fisher-Snedecor con n y m grados de libertad

Sea Z una $F_{n,m}$. Para $\alpha \in (0,1)$, denotamos por $F_{\{n,m;\alpha\}}$ al valor tal que

$$\mathbf{P}(Z > F_{\{n,m\,;\,\alpha\}}) = \alpha\,.$$



Obsérvese que

$$\mathbf{P}(F_{\{n,m\,;\,1-\alpha/2\}} < Z < F_{\{n,m\,;\,\alpha/2\}}) = 1 - \alpha\,.$$

Para calcular $F_{\{n,m;\alpha\}}$ en excel: `=inv.f(1-`α`;`n`;`m`)`.

Dada una muestra $(x_1, \ldots, x_n)$ de tamaño n de la variable X, llamamos $\overline{x}$ a la media muestral y s^2 a la cuasivarianza muestral.

Los intervalos que siguen son con confianza $1-\alpha$.

Normal $\mathcal{N}(\mu, \sigma^2)$

A Intervalo para la media μ

Caso 1. Suponiendo que σ^2 es conocida:

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \stackrel{\mathrm{d}}{=} \mathcal{N}(0,1) \quad \longrightarrow \quad \text{Intervalo: } I = \left(\overline{x} \pm z_{\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right).$$

Caso 2. Con σ^2 desconocida:

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \stackrel{\mathrm{d}}{=} \mathrm{STU}(n-1) \quad \longrightarrow \quad \text{Intervalo: } I = \left(\overline{x} \pm t_{\{n-1\,;\,\alpha/2\}}\,\frac{s}{\sqrt{n}}\right)$$

B. Intervalo para la varianza σ^2

$$\frac{(n-1)\,S^2}{\sigma^2} \stackrel{\mathrm{d}}{=} \chi^2_{n-1} \quad \longrightarrow \quad \text{Intervalo: } I = \left(\frac{(n-1)\,s^2}{\chi^2_{\{n-1;\,\alpha/2\}}},\ \frac{(n-1)\,s^2}{\chi^2_{\{n-1;\,1-\alpha/2\}}}\right)$$

Proporción p

Para n grande,

$$\frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \stackrel{\mathrm{d}}{\approx} \mathcal{N}(0,1) \quad \longrightarrow \quad \text{Intervalo: } I = \left(\overline{x} \pm z_{\alpha/2}\sqrt{\frac{\overline{x}\,(1-\overline{x})}{n}}\right)$$

Poisson (λ)

Para n grande,

$$\frac{\overline{X}-\lambda}{\sqrt{\lambda/n}} \stackrel{\mathrm{d}}{\approx} \mathcal{N}(0,1) \quad \longrightarrow \quad \text{Intervalo: } I = \left(\overline{x} \pm z_{\alpha/2}\sqrt{\frac{\overline{x}}{n}}\right)$$

Dos normales, $X_1 = \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 = \mathcal{N}(\mu_2, \sigma_2^2)$

DATOS:

- muestra de tamaño n_1 de la variable X_1, con media muestral $\overline{x}_1$ y cuasivarianza muestral s_1^2.
- muestra de tamaño n_2 de la variable X_2, con media muestral $\overline{x}_2$ y cuasivarianza muestral s_2^2.

Para la diferencia de medias $\mu_1 - \mu_2$:

- *si σ_1, σ_2 conocidas:*

$$I = \left((\overline{x}_1 - \overline{x}_2) \pm z_{\alpha/2} \sqrt{\tfrac{\sigma_1^2}{n_1} + \tfrac{\sigma_2^2}{n_2}} \right)$$

- *si σ_1, σ_2 desconocidas, pero $\sigma_1 = \sigma_2$:*

$$I = \left((\overline{x}_1 - \overline{x}_2) \pm t_{\{n_1+n_2-2;\, \alpha/2\}} \cdot s_p \sqrt{\tfrac{1}{n_1} + \tfrac{1}{n_2}} \right)$$

donde

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)}.$$

- *si σ_1, σ_2 desconocidas, pero $\sigma_1 \neq \sigma_2$:*

$$I = \left((\overline{x}_1 - \overline{x}_2) \pm t_{\{f;\, \alpha/2\}} \sqrt{\tfrac{s_1^2}{n_1} + \tfrac{s_2^2}{n_2}} \right)$$

donde f es el entero más próximo a $\dfrac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$.

Para el cociente de varianzas σ_1^2/σ_2^2:

$$I = \left(\frac{s_1^2/s_2^2}{F_{\{n_1-1,\, n_2-1\,;\, \alpha/2\}}}, \frac{s_1^2/s_2^2}{F_{\{n_1-1,\, n_2-1\,;\, 1-\alpha/2\}}} \right)$$

Comparación de proporciones p_1, p_2

DATOS:

- muestra de tamaño n_1 de la variable $X_1 \sim \text{BER}(p_1)$, media muestral $\overline{x}_1$.
- muestra de tamaño n_2 de la variable $X_2 \sim \text{BER}(p_2)$, media muestral $\overline{x}_2$.

Tanto n_1 como n_2 son grandes.

Para $p_1 - p_2$:

$$I = \left((\overline{x}_1 - \overline{x}_2) \pm z_{\alpha/2} \sqrt{\tfrac{\overline{x}_1(1-\overline{x}_1)}{n_1} + \tfrac{\overline{x}_2(1-\overline{x}_2)}{n_2}} \right)$$

Estadística I
Grado en Matemáticas, UAM, 2018-2019

Contrastes de hipótesis (paramétricas)

Contrastes para UNA distribución

Normal $N(\mu;\sigma^2)$

Hipótesis nula H_0		Región de rechazo R
$\mu = \mu_0$	σ conocida	$\lvert\overline{x} - \mu_0\rvert > z_{\alpha/2}\,(\sigma/\sqrt{n})$
	σ desconocida	$\lvert\overline{x} - \mu_0\rvert > t_{\{n-1;\,\alpha/2\}}\,(s/\sqrt{n})\}$
$\mu \geq \mu_0$	σ conocida	$\overline{x} < \mu_0 - z_\alpha\,(\sigma/\sqrt{n})$
	σ desconocida	$\overline{x} < \mu_0 - t_{\{n-1;\,\alpha\}}\,(s/\sqrt{n})$
$\mu \leq \mu_0$	σ.conocida	$\overline{x} > \mu_0 + z_\alpha\,(\sigma/\sqrt{n})$
	σ desconocida	$\overline{x} > \mu_0 + t_{\{n-1;\,\alpha\}}\,(s/\sqrt{n})$
$\sigma = \sigma_0$		$\frac{(n-1)s^2}{\sigma_0^2} \notin \left(\chi^2_{\{n-1;\,1-\alpha/2\}},\ \chi^2_{\{n-1;\,\alpha/2\}}\right)$
$\sigma \geq \sigma_0$		$\frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\{n-1;\,1-\alpha\}}$
$\sigma \leq \sigma_0$		$\frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{\{n-1;\,\alpha\}}$

BER(p)

Hip. nula H_0	Región de rechazo R
$p = p_0$	$\lvert\overline{x} - p_0\rvert > z_{\alpha/2}\,\sqrt{p_0(1-p_0)}/\sqrt{n}$
$p \geq p_0$	$\overline{x} < p_0 - z_\alpha\,\sqrt{p_0(1-p_0)}/\sqrt{n}$
$p \leq p_0$	$\overline{x} > p_0 + z_\alpha\,\sqrt{p_0(1-p_0)}/\sqrt{n}$

Poisson(λ)

Hip. nula H_0	Región de rechazo R
$\lambda = \lambda_0$	$\lvert\overline{x} - \lambda_0\rvert > z_{\alpha/2}\,\sqrt{\lambda_0/n}$
$\lambda \geq \lambda_0$	$\overline{x} < \lambda_0 - z_\alpha\,\sqrt{\lambda_0/n}$
$\lambda \leq \lambda_0$	$\overline{x} > \lambda_0 + z_\alpha\,\sqrt{\lambda_0/n}$

Contrastes para DOS distribuciones

Normales $N(\mu_1;\sigma_1^2), N(\mu_2;\sigma_2^2)$

Hipótesis nula H_0		Región de rechazo R
$\mu_1=\mu_2$	σ_1,σ_2 conocidas	$\lvert\overline{x}_1-\overline{x}_2\rvert > z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$
	$\sigma_1=\sigma_2$ desconocidas	$\lvert\overline{x}_1-\overline{x}_2\rvert > t_{\{n_1+n_2-2;\,\alpha/2\}}\ s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$
	$\sigma_1\neq\sigma_2$ desconocidas	$\lvert\overline{x}_1-\overline{x}_2\rvert > t_{\{f;\,\alpha/2\}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$
$\mu_1\leq\mu_2$	σ_1,σ_2 conocidas	$\overline{x}_1-\overline{x}_2 > z_{\alpha}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$
	$\sigma_1=\sigma_2$ desconocidas	$\overline{x}_1-\overline{x}_2 > t_{\{n_1+n_2-2;\,\alpha\}}\ s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$
	$\sigma_1\neq\sigma_2$ desconocidas	$\overline{x}_1-\overline{x}_2 > t_{\{f;\,\alpha\}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$
$\sigma_1=\sigma_2$		$\frac{s_1^2}{s_2^2}\notin\left(\{F_{n_1-1;\,n_2-1;\,1-\alpha/2\}},\ F_{\{n_1-1;\,n_2-1;\,\alpha/2\}}\right)$
$\sigma_1\leq\sigma_2$		$\frac{s_1^2}{s_2^2} > F_{\{n_1-1;\,n_2-1;\,\alpha\}}$

- n_1, n_2 tamaños muestrales, $\overline{x}_1, \overline{x}_2$ medias muestrales, s_1, s_2 cuasi-desviaciones típicas muestrales
- $s_p^2 = \dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$
- f es el entero más próximo a $\dfrac{(s_1^2/n_1+s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}}$

Proporciones p_1, p_2

Hipótesis nula H_0	Región de rechazo R
$p_1=p_2$	$\lvert\overline{x}_1-\overline{x}_2\rvert > z_{\alpha/2}\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$
$p_1\leq p_2$	$\overline{x}_1-\overline{x}_2 > z_{\alpha}\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$

- n_1; n_2 tamaños muestrales, $\overline{x}_1, \overline{x}_2$ proporciones muestrales
- $\overline{p} = \dfrac{n_1\overline{x}_1+n_2\overline{x}_2}{n_1+n_2}$

$\bar{x} = \frac{1}{n}\sum x_i$
$\bar{x} = val \times prob$
$V(x) = \frac{1}{n}\sum (x_i - \bar{x})^2$
$V(x) = \mathbb{E}(x^2) - \mathbb{E}(x)^2$
$(n-1)S_x^2 = n V_x$
$|\mathbb{E}(xy)|^2 \leq \mathbb{E}(x^2)\,\mathbb{E}(y^2)$

mom. no centr. $\rightarrow \mathbb{E}(x^k) = \sum x_i^k \, \mathbb{P}(x = x_i)$
mom. centr. $\rightarrow \mathbb{E}((x - \mathbb{E}(x))^k) = \sum (x_i - \mathbb{E}(x))^k \cdot \mathbb{P}(x = x_i)$
$asim_x = \frac{\mathbb{E}((x-\mu)^3)}{\sigma^3}$; curtosis $= \frac{\mathbb{E}((x-\mu)^4)}{\sigma^4}$
Des. Cheb: $\lambda > 0 \quad \mathbb{E}(|x - \mathbb{E}(x)| > \lambda) \leq \frac{V(x)}{\lambda^2}$
Des. Markov: $a > 0 \quad \mathbb{E}(x > a) \leq \frac{\mathbb{E}(x)}{a}$

$\rho_{xy} = \frac{cov_{xy}}{\sqrt{V_x}\sqrt{V_y}}$; X, Y indep $\Rightarrow cov = 0 \Rightarrow \rho = 0$
$Cov_{xy} = \frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y})$, $cov = \overline{xy} - \bar{x}\bar{y}$

$\hat{b} = \frac{cov_{xy}}{V_x}$; $\hat{a} = \bar{y} - \hat{b}\bar{x}$
$y = \hat{a} + \hat{b}x$; $y - \bar{y} = \hat{b}(x - \bar{x})$

RECORDAR:
$\mathbb{E}(ax + by) = a\mathbb{E}(x) + b\mathbb{E}(y)$
$V(X \pm Y) = V(x) + V(Y) \pm 2cov(X, Y)$
$cov(x, y)$ es bilineal

$X \sim N(\mu, \sigma^2)$
$\mathbb{E}(x^2) = \mu^2 + \sigma^2$
$\mathbb{E}(x^3) = \mu^3 + 3\mu\sigma^2$
$\mathbb{E}(x^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$
$\mathbb{E}((x-\mu)^4) = 3\sigma^4$
$\mathbb{E}((x-\mu)^3) = 0 \qquad V(x) = \sigma^2$

UNIFORME $\mathbb{E} = \frac{a+b}{2}$
$f_X(x) = \begin{cases} 1/b-a & x \in [a,b] \\ 0 & \text{resto} \end{cases}$
$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}$
$V = \frac{(b-a)^2}{12} \qquad \frac{x}{b-a} \sim Unif(0,1)$

BERNOULLI $X \sim Ber(p) \quad x_k \in \{0,1\}$
$p = \mathbb{P}(X=1)$; $\mathbb{E} = p$; $V = p(1-p)$
$\mathbb{P}(X = k) = p^k (1-p)^{1-k}$

BINOMIAL: $X \sim Bin(n, p)$
$f_X(k) = \mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
$\mathbb{E} = np$; $V = np(1-p)$

POISSON: $X \sim Poisson(\lambda = np)$
Bin con $n \to \infty \quad p \to 0$
$\mathbb{E} = \lambda$; $V = \lambda$
$\mathbb{P}(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

EXPONENCIAL: $X \sim Exp(\lambda)$
$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \qquad \mathbb{E} = \frac{1}{\lambda}$
$F_X(x) = 1 - e^{-\lambda x} \quad x \geq 0$; $V = \frac{1}{\lambda^2}$
$X \sim Exp(\lambda) \rightarrow \lambda X \sim Exp(1)$
$X \sim Exp(\lambda)$, $Y \sim Exp(\lambda) \Rightarrow X + Y \sim Gamma(\lambda, 2)$

GEOMÉTRICA: $X \sim Geo(p)$
$\mathbb{P}(X = k) = (1-p)^{k-1} p$
$\mathbb{E} = 1/p$; $V = \frac{1-p}{p^2}$

NORMAL UNI: $X \sim N(\mu, \sigma^2)$
$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}(x-\mu)^2/\sigma^2\right)$
$\mathbb{E}(x^2) = 1$ si $X \sim \Phi$
$\lambda(a\mu + b\nu, a^2\sigma^2 + b^2\tau^2)$

GAMMA: $X \sim Gamma(\lambda, t) \quad t > 0$
$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx = \int_0^\infty x^t e^{-x} \frac{dx}{x}$
$\Gamma(t) = (t-1)\Gamma(t-1)$; $\Gamma(n) = (n-1)!$
$\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$; $\Gamma(-1/2) = -2\sqrt{\pi}$
$E(x^k) = \frac{(t+k-1)!}{\lambda^k (t-1)!}$; $V(x) = t/\lambda^2$
$t > 0$, $b > 0 \quad \int_0^\infty x^t e^{-bx} \frac{dx}{x} = \frac{\Gamma(t)}{b^t}$
$f_X(x) = \frac{1}{\Gamma(t)} \lambda^t x^{t-1} e^{-\lambda x}$
$X \sim G(\lambda, s) \quad Y \sim G(\lambda, t) \Rightarrow X + Y \sim G(\lambda, t+s)$
$Gamm(\lambda, 1) = Exp(\lambda)$; $G(1/2, n/2) \sim \chi_n^2$

$\vec{m} \in \mathbb{R}^n$, $V \in M_{n\times n}$ sim. def. pos $X \sim N(\vec{m}, V)$
$f_X(x) = \frac{1}{(\sqrt{2\pi})^n} \cdot \frac{1}{\sqrt{\det(V)}} \exp\left(\frac{-(\vec{x} - \vec{m})^t}{2} \cdot V^{-1} (\vec{x} - \vec{m})\right)$
$V = UU^t$ V no singular. $\det(V) = \det(U)^2$

Sea $X \sim N(\vec{m}, V)$ $\vec{h} \in \mathbb{R}^n$, $B \in M_{n \times n}$ no sing.
$Z \sim \vec{h} + BX \rightarrow Z \sim N(\vec{h} + B\vec{m}, BVB^t)$
$Y \sim N(0, I_n)$ y $O \in$ ortonormal $\rightarrow Z = OY \sim \Phi$

$X \sim N(\vec{m}, V)$, $\vec{a} \in \mathbb{R}^n$ y $\vec{a} \neq \vec{0}$
$\sum a_i x_i = \vec{a}^t X \sim N(\vec{a}^t \vec{m}, \vec{a}^t V \vec{a})$

CHI-CUAD: $Z = X_1^2 + \dots + X_n^2 \quad X_i \sim \Phi$
$\mathbb{E}(Z) = n$; $V(Z) = 2n$; $\mathbb{E}(Z^2) = n(n+2)$

F-FISHER: $Z \sim F_{n,m} \qquad \mathbb{E}(Z) = \frac{m}{m-2} \quad m \geq 3$
$Z = \frac{(X_1^2 + \dots + X_n^2)/n}{(Y_1^2 + \dots + Y_m^2)/m} \quad X_i \sim \Phi, \; Y_i \sim \Phi \qquad V(Z) = \frac{2(m+n-2)m^2}{n(m-4)(m-2)^2} \quad m \geq 5$

T-STUDENT $n \geq 2 \quad Y \sim \Phi \quad U_n \sim \chi_n^2$
$Z := Y/\sqrt{U_n/n}$; $\mathbb{E}(Z) = 0$; $V(Z) = \mathbb{E}(Z^2) = \frac{n}{n-2}$
$n \to \infty \qquad t_n \longrightarrow N(0,1)$

ESTADÍSTICO $\bar{x}$: $\bar{x} = \frac{1}{n}\sum x_i$; $\mathbb{E}(\bar{x}) = \mathbb{E}(x)$; $V(\bar{x}) = V(x)/n$
Cheb: $\mathbb{P}(|\bar{x} - \mathbb{E}(x)| \geq \varepsilon) \leq \frac{V(x)}{n\varepsilon^2}$
- $X \sim Ber(p) \longrightarrow n\bar{x} \sim Bin(n, p)$
- $X \sim Poiss(\lambda) \quad \lambda > 0 \rightarrow n\bar{x} \sim Poiss(n\lambda)$
- $X \sim Exp(\lambda) \longrightarrow n\bar{x} \sim Gamma(\lambda, n)$
- $X \sim N(\mu, \sigma^2) \longrightarrow n\bar{x} \sim N(n\mu, n\sigma^2)$
  $\longrightarrow \bar{x} \sim N(\mu, \sigma^2/n)$

$\boxed{TCL} \Rightarrow \sqrt{n}(\bar{x} - \mathbb{E}(x)) \xrightarrow{d} N(0, V(x))$

T_n consistente $\Leftrightarrow \lim_{n\to\infty} \mathbb{P}(|T_n - \theta| \geq \varepsilon) = 0 \; \forall \varepsilon$
Con la des. Cheb, si T_n son insesgados y $\lim_{n\to\infty} V_\theta(T_n) = 0 \Rightarrow T_n$ consistente

ESTADÍSTICO $S^2 \quad \mathbb{E}(S^2) = V(X)$
$V(S^2) = \frac{1}{n}\mathbb{E}((x - \mathbb{E}(x))^4) - \frac{n-3}{n(n-1)} V(x)^2$

TEOREMA FISHER-COCHRAN: $X \sim N(\mu, \sigma^2)$
$X_1, \dots, X_n$ clones indep.
1) $\bar{X} \sim N(\mu, \sigma^2/n) \qquad$ 2) $\frac{(n-1)}{\sigma^2} S^2 \sim \chi_{n-1}^2$
3) $\bar{x}$ e S^2 son independientes
Si $X \sim N(0,1) \Rightarrow \frac{\bar{x}}{(\sqrt{S^2}/\sqrt{n})} \sim t_{n-1}$

ESTADÍSTICOS MAX/MIN:
$F_{M_n}(t) = (F_X(t))^n$; $F_{m_n}(t) = 1 - (1 - F_X(t))^n$

CRAMER-RAO: $Y = \partial_\theta \log f(x; \theta)$
$I_X(\theta) = V_\theta(Y)$
CRAMER-RAO: $V_\theta(T) \geq \frac{1}{n I_X(\theta)} \; \forall T$
T eficiente $\Leftrightarrow V(T) = \frac{1}{n I_X(\theta)}$

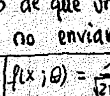
$\hat{\theta}$ = estimación de θ dados unos datos ; $sop_\theta = \{x \in \mathbb{R} : f(x; \theta) > 0\}$
El SESGO de T: $sesgo_\theta(T) = \mathbb{E}_\theta(T) - \theta$; T insesg. $\Leftrightarrow \mathbb{E}_\theta(T) = \theta \; \forall \theta$
T, T' estimadores $\longrightarrow T$ más eficiente $\Leftrightarrow V(T) < V(T')$
Estimador por momentos de orden n: $\overline{x^n} = \mathbb{E}_\theta(x^n)$
En las distr. simétricas se utilizan momentos pares. Si el momento no depende de θ entonces no hay estimador.
Método de máxima verosimilitud: $VERO(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \theta)$
La estimación de MAXVERO $\hat{\theta}$, si existe, es el máximo de VERO en Θ.
Para maximizar usar logVERO $\rightarrow \log(ab) = \log(a) + \log(b)$
Comportamiento asintótico (Método Delta) Z_n := sucesión v.a.
$\sqrt{n}(Z_n - \alpha) \xrightarrow[n\to\infty]{d} N(0, \beta^2)$ donde $\mathbb{E}(Z_n) = \alpha \quad V(Z_n) = \beta^2$
Sea $g \in C^2(a,b)$, $g'(\alpha) \neq 0 \Rightarrow \sqrt{n}(g(Z_n) - g(\alpha)) \xrightarrow{d} N(0, |g'(\alpha)|^2 \beta^2)$
Si $g'(\alpha) = 0 \wedge g''(\alpha) \neq 0 \wedge g \in C^3(a,b) \Rightarrow n(g(Z_n) - g(\alpha)) \rightarrow \frac{1}{2} g''(\alpha) \beta^2 \chi_1^2$
Intervalo asintótico de confianza: $\sqrt{n}(T_n - a) \longrightarrow N(0, b^2)$
$\mathbb{P}\left(-Z_{\alpha/2} \leq \frac{T_n - a}{b/\sqrt{n}} \leq Z_{\alpha/2}\right) \approx 1 - \alpha \Rightarrow$ Intervalo confianza $1-\alpha$: $a = \hat{t}_n \pm \frac{b}{\sqrt{n}} Z_{\alpha/2}$
Contraste de hipótesis: hipótesis nula $H_0 \equiv \theta \in \Theta_0$
Calculamos región de rechazo $\rightarrow$ rechazamos con seguridad
Error tipo 1: rechazar H_0 si es cierta $\quad \mathbb{P}(\text{error } t1) = \alpha$ (pequeño)
Error tipo 2: aceptar H_0 si es falsa
P-valor nos muestra la probabilidad de haber obtenido el resultado obtenido si suponemos H_0 cierta.
P-valor $\rightarrow$ valor de α para que $\bar{x}$ esté en el borde de la región de rechazo. Se calcula con $\alpha = p$ y despejando percentil C_p:
$F(C_p) = 1 - p \Rightarrow p = 1 - F(C_p)$. En una normal:
$\Phi(Z_{\alpha/2}) = 1 - P/2 \Rightarrow p = 2(1 - \Phi(Z_{p/2}))$ $\rightarrow$ no siempre es posible despejar $\rightarrow$ tablas a ojo
$|\bar{x} - \mu_0| = \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma \Rightarrow |\bar{x} - \mu_0| \frac{\sqrt{n}}{\sigma} = Z_{\alpha/2} \Rightarrow \alpha = 2\left(1 - \Phi\left(|\bar{x} - \mu_0| \frac{\sqrt{n}}{\sigma}\right)\right)$
Función de potencia: $\beta(\theta) = \mathbb{P}_\theta(\text{rechazar}) = 1 - \mathbb{P}_\theta(\text{aceptar})$
Significación (posibilidad de rechazar al correcto) $= \sup_{\theta \in \Theta_0} \beta(\theta)$
TEST RAZÓN VEROSIMILITUDES $H_0 \equiv \theta \in \Theta_0 \subset \Theta$
1. Construir $VERO(\theta; x_1, \dots, x_n)$
2. Hallar $\sup_{\theta \in \Theta_0} VERO(\theta)$ [1] y $\sup_{\theta \in \Theta} VERO(\theta)$ [2] $\quad$ 3. $RV = \frac{[1]}{[2]}$
4. Definir el test: Rechazo $H_0 \Leftrightarrow RV < C \in (0,1)$ calibre
5. Calcular $\beta(\theta) = \mathbb{P}_\theta(\text{rechazar}) = \mathbb{P}_\theta(RV < C)$
6. Obtener significación: $\sup_{\theta \in \Theta_0} \beta(\theta)$

$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$; $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$ (independientes)
$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$H_0: p \geq 5\% \quad n = 30 \quad \mathbb{P}_p(\text{rechazar}) = \mathbb{P}_p$(de 30 0 defectuosas) ; significación $= \beta(0'05)$
$\Rightarrow \beta(p) = \mathbb{P}(Bin(30, p) = 0) = (1-p)^{30}$
¿Máx prob de que una caja MALA salga al mercado? $\rightarrow \beta(0'05)$
¿Prob. de no enviar al mercado una caja con 3% defect.? $\rightarrow 1 - \beta(0'03)$

DATOS: $X \sim Unif(0, a) \quad M_n = \max(x_i)$ estim. de a
$\mathbb{P}(\frac{M_n}{a} \leq t) = \mathbb{P}(X \leq at)^n = t^n \quad t \in (0,1)$
Hallar C_1 y C_2 para que
$\mathbb{P}(\frac{M_n}{a} \leq C_1) = \mathbb{P}_a(\frac{M_n}{a} \geq C_2) = \alpha/2$
$C_1^n = \mathbb{P}(\frac{M_n}{a} \leq C_1) = \alpha/2 \Rightarrow$
$\Rightarrow C_1 = \sqrt[n]{\alpha/2}$
$\alpha/2 = \mathbb{P}(\frac{M_n}{a} \geq C_2) = 1 - \mathbb{P}(\frac{M_n}{a} \leq C_2)$
$= 1 - C_2^n \Rightarrow C_2 = \sqrt[n]{1 - \alpha/2}$
Entonces: $1 - \alpha = \mathbb{P}(C_1 \leq \frac{M_n}{a} \leq C_2)$

a) $n = 200 \quad \alpha = 5\% \quad H_0 \equiv$ moneda equili ¿rechazar?
$\bar{x}$ = prop. unos $\quad \rightarrow H_0 \equiv p = 1/2$
Rechazamos si $|\bar{x} - 1/2| \geq \frac{1}{2} \cdot 1'96 \cdot 1/\sqrt{200}$
b) n general $\alpha = 5\% \quad \bar{x} = 57\%$
¿valores de n para aceptar?
Si $n \geq (\frac{1}{2} \cdot \frac{1'96}{0'07})^2$ rechazamos, por el contr. aceptamos
c) ¿p-valor?
$|\bar{x} - \frac{1}{2}| = \frac{1}{2} \cdot Z_{\alpha/2} \cdot \frac{1}{\sqrt{n}} \Rightarrow \Phi(|\bar{x} - \frac{1}{2}| \cdot 2\sqrt{n}) = 1 - \alpha/2 \Rightarrow \alpha = 2(1 - \Phi(|\bar{x} - \frac{1}{2}| \cdot 2\sqrt{n}))$

$f(x; \theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/2\theta} \quad x \in \mathbb{R} \qquad H_0: \theta < \theta_0 \quad \mathbb{R}^+$
$VERO(\theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sqrt{\theta}}\right)^n \cdot \exp\left(-\frac{\sum x_i^2}{2\theta}\right) \rightarrow n\overline{x^2}$
log $\rightarrow$ der $\rightarrow = 0 \rightarrow \theta = \overline{x^2}$
$RV = \frac{\sup \theta_0}{\sup \theta} \begin{cases} \frac{VERO(\overline{x^2})}{VERO(\overline{x^2})} & \text{si } \theta_0 > \overline{x^2} \\ \frac{VERO(\theta_0)}{VERO(\overline{x^2})} & \text{si } \theta_0 < \overline{x^2} \end{cases}$

Rechazo si $\left(\frac{\overline{x^2}}{\theta_0}\right)^{n/2} \cdot \exp\left(-\frac{\overline{x^2} n}{2\theta_0} + \frac{1}{2}n\right) \leq C$ and $\overline{x^2} > \theta_0$
$\mathbb{P}_\theta(\text{rechazar}) = \beta(\theta)$
Significación $= \sup_{\Theta_0} \beta(\theta)$

DATOS: $\Theta = \{0, 1\}$
$f(x; 0) = 1 \quad f(x; 1) = \frac{1}{2\sqrt{x}}$
MAXVERO para θ
$VERO(0, x_1, \dots, x_n) = \prod f(x_i, 0) = 1$
$VERO(1, x_1 - x_n) = \prod f(x_i, 1) =$
$= \frac{1}{2^n} \cdot \frac{1}{\sqrt{\prod x_i}} \quad (*)$
$\hat{\theta} = \begin{cases} 0 & \text{si } (*) < 1 \\ 1 & \text{si } (*) > 1 \\ \text{sin estimación} & \text{si } (*) = 1 \end{cases}$

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$V(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

$V(x) = \mathbb{E}(x^2) - \mathbb{E}(x)^2$

$\rho_{x,y} = \frac{cov_{x,y}}{\sqrt{V_x}\sqrt{V_y}}$; x, y dep $\Rightarrow cov = 0$

$cov_{x,y} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

$cov_{x,y} = \overline{xy} - \bar{x}\cdot\bar{y}$

mom. no centrados $\rightarrow \mathbb{E}(x^k) = \sum x_i^k \, \mathbb{P}(x = x_i)$

mom. centrados $\rightarrow \mathbb{E}((x - \mathbb{E}(x))^k) = \sum (x_i - \mathbb{E}(x))^k \, \mathbb{P}(x = x_i)$

Des. Chebyshev: $\lambda > 0$ $\quad \mathbb{E}(|x - \mathbb{E}(x)| > \lambda) \leq \frac{V(x)}{\lambda^2}$

Des. Markov: $a > 0$ $\quad \mathbb{E}(x > a) \leq \frac{\mathbb{E}(x)}{a}$

$\mathbb{E}(\bar{x}) = \mathbb{E}(x)$ | $V(\bar{x}) = V(x)/n$

$|\mathbb{E}(xy)|^2 \leq \mathbb{E}(x^2)\,\mathbb{E}(y^2)$

$X \sim N(\mu, \sigma^2)$

$\mathbb{E}(x^2) = \mu^2 + \sigma^2$ $\quad \mathbb{E}((x-\mu)^4) = 3\sigma^4$

$\mathbb{E}(x^3) = \mu^3 + 3\mu\sigma^2$ $\quad \mathbb{E}((x-\mu)^3) = 0$

$\mathbb{E}(x^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$ $\quad V(x) = \sigma^2$

UNIFORME $f_X(x) = \begin{cases} 1/(b-a) & x \in [a,b] \\ 0 & \text{resto} \end{cases}$

$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & \text{resto} \end{cases}$; $\mathbb{E}(x) = \frac{a+b}{2}$

$V(x) = \frac{(b-a)^2}{12}$ $\quad \frac{x}{b-a} \sim Unif(0,1)$

BERNOULLI: $X \sim Ber(p)$ $\quad x_i \in \{0,1\}$

$p = \mathbb{P}(X=1)$; $\mathbb{E}(x) = p$; $V(x) = p(1-p)$

$\mathbb{P}(X=k) = p^k(1-p)^{1-k}$

BINOMIAL $X \sim Bin(n,p)$

$f_X(k) = \mathbb{P}(X=k) = \binom{n}{k}p^k(1-p)^{n-k}$

$\mathbb{E}(x) = np$; $V(x) = np(1-p)$

POISSON: $X \sim Poisson(\lambda = np)$

Bin con $n \to \infty$ $\quad p \to 0$

$\mathbb{E}(x) = \lambda$ $\quad V(x) = \lambda$

$\mathbb{P}(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

EXPONENCIAL $X \sim Exp(\lambda)$

$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$; $\mathbb{E}(x) = \frac{1}{\lambda}$

$F_X(x) = 1 - e^{-\lambda x}$ $\quad x \geq 0$; $V(x) = \frac{1}{\lambda^2}$

$X \sim Exp(\lambda) \rightarrow \lambda X \sim Exp(1)$

GEOMÉTRICA: $X \sim Geo(p)$

$\mathbb{P}(X=k) = (1-p)^{k-1}p$; $\mathbb{E}(X) = \frac{1}{p}$; $V(x) = \frac{1-p}{p^2}$

NORMAL UNI $X \sim N(\mu, \sigma^2)$

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

$\mathbb{E}(x^2) = 1$ si $X \sim \Phi$

GAMMA $X \sim Gamma(\lambda, t)$ $\quad t > 0$

$\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx = \int_0^\infty x^t e^{-x}\frac{dx}{x}$

$\Gamma(t) = (t-1)\Gamma(t-1)$; $\Gamma(n) = (n-1)!$

$\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$; $\Gamma(-1/2) = -2\sqrt{\pi}$

$\mathbb{E}(x^k) = \frac{(t+k-1)!}{\lambda^k(t-1)!}$; $V(x) = t/\lambda^2$

$\rightarrow \int_0^\infty x^t e^{-bx}\frac{dx}{x} = \frac{\Gamma(t)}{b^t}$ $\quad f(x) = 0$ si $x < 0$

$f(x) = \frac{1}{\Gamma(t)}\lambda^t x^{t-1}e^{-\lambda x}$ $\quad x > 0$

CHI CUADRADO: $Z = X_1^2 + \dots + X_n^2$ $\quad X_i \sim \Phi$ indp.

$\mathbb{E}(Z) = n$; $V(Z) = 2n$; $\mathbb{E}(Z^2) = n(n+2)$

FISHER: $Z \sim F_{n,m}$

$Z = \frac{(X_1^2 + \dots + X_n^2)/n}{(Y_1^2 + \dots + Y_m^2)/m}$ $\quad X_i \sim \Phi$ indp. $\quad Y_i \sim \Phi$ indp.

$\mathbb{E}(Z) = \frac{m}{m-2}$ $\quad m \geq 3$ $\quad V(Z) = \frac{2(m+n-2)m^2}{n(m-4)(m-2)^2}$

STUDENT $n \geq 2$ $\quad Y \sim \Phi$ $\quad U_n \sim \chi^2_n$

$Z = \frac{Y}{\sqrt{U_n/n}}$; $\mathbb{E}(Z) = 0$; $V(Z) = \mathbb{E}(Z^2) = \frac{n}{n-2}$ $\quad$ ⊛ $n \to \infty$: $t_n \to \Phi$

Un estimador es un estadístico $T = h(X_1 \dots X_n) = \hat{\theta}$

$\hat{\theta}$ = estimación de θ dados unos datos. ; $sop_\theta = \{x \in \mathbb{R} ; f(x;\theta) > 0\}$

El sesgo de T: $sesgo_\theta(T) = \mathbb{E}_\theta(T) - \theta$; T insesg $\Leftrightarrow \mathbb{E}_\theta(T) = \theta \; \forall\theta$

T, T' estimadores $\rightarrow$ T más eficiente $\Leftrightarrow V(T) < V(T')$

Estimador por momentos de orden n: $\overline{X^n} = \mathbb{E}_\theta(X^n)$

En las distribuciones simétricas, se utilizan momentos pares

Si el momento no depende de θ entonces no hay estimador

Método de máxima verosimilitud: $VERO(\theta; x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i;\theta)$

La estimación de MAX.VERO. $\hat{\theta}$, si existe, es el máximo de VERO en θ. Para maximizar usar logVERO. ($\log(ab) = \log a + \log b$)

Comportamiento asintótico (Método Delta) Z_n := sucesión v.a.

$\sqrt{n}(Z_n - \alpha) \xrightarrow[n\to\infty]{d} N(0, \beta^2)$ donde $\mathbb{E}(Z_n) = \alpha$ $\quad V(Z_n) = \beta^2$.

Sea $g \in C^2(a,b)$, $g'(\alpha) \neq 0 \Rightarrow \sqrt{n}(g(Z_n) - g(\alpha)) \rightarrow N(0, |g'(\alpha)|^2\beta^2)$

Si $g'(\alpha) = 0 \wedge g''(\alpha) \neq 0 \wedge g \in C^3 \Rightarrow n(g(Z_n) - g(\alpha)) \longrightarrow \frac{1}{2}g''(\alpha)\beta^2\chi^2_1$

CONTRASTE DE HIPÓTESIS: Hipótesis nula $H_0 \equiv \theta \in \Theta_0$

Calculamos región de rechazo $\rightarrow$ rechazamos con seguridad

Error tipo 1: rechazar H_0 si es cierta $\quad \mathbb{P}(\text{error } t1) = \alpha$ (pequeño)

Error tipo 2: aceptar H_0 si es falsa

p-valor nos muestra la probabilidad de haber obtenido el resultado obtenido si suponemos H_0 cierta.

p-valor $\rightarrow$ valor de α para que $\bar{x}$ esté en el borde de la región de rechazo. Se calcula con $\alpha = p$ y despejando el percentil c_p.

$F(c_p) = 1 - p \Rightarrow p = 1 - F(c_p)$. En una normal:

$\Phi(Z_{\alpha/2}) = 1 - p/2 \Rightarrow p = 2(1 - \Phi(Z_{p/2}))$ $\longrightarrow$ $|\bar{x} - \mu_0| = \frac{Z_{\alpha/2}}{\sqrt{n}}\sigma \Rightarrow$ ⊛

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

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

⊛ No siempre es posible despejar α por lo que habría que mirarlo numéricamente en las tablas.

Función de potencia: $\beta(\theta) = \mathbb{P}_\theta(\text{rechazar})$

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

TEST RAZÓN VEROSIMILITUDES $H_0 \equiv \theta \in \Theta_0 \subset \Theta$

1. Construir $VERO(\theta; x_1, \dots, x_n)$
2. Hallar $\sup_{\theta \in \Theta_0} VERO(\theta)$ [1] y $\sup_{\theta \in \Theta} VERO(\theta)$ [2]
3. $RV = \frac{[1]}{[2]}$
4. Definir el test: Rechazo $H_0 \Leftrightarrow RV < c \in (0,1)$ calibre
5. Calcular $\beta(\theta) = \mathbb{P}_\theta(\text{rechazar}) = \mathbb{P}_\theta(RV < c)$
6. Obtener significación: $\sup_{\theta \in \Theta_0} \beta(\theta)$

TCL: $\sqrt{n}(\bar{X}_n - \mathbb{E}(X)) \longrightarrow N(0, V(X))$

$Y = \partial_\theta \ln f(x;\theta)$; $I_X(\theta) = V_\theta(Y)$

CRAMER-RAO: $V_\theta(T) \geq \frac{1}{n I_X(\theta)}$

T eficiente $\Leftrightarrow V(T) = 1/n \, I_X(\theta)$

Obs: S^2 := cuasivarianza muestral

S := cuasidesv. típica muestral

Si la letra es minúscula son muestrales, si es mayúscula son del modelo teórico.

RECUERDO: $\mathbb{E}(S^2) = V(X)$

TEOREMA FISHER-COCHRAN

$X \sim N(\mu, \sigma^2)$ $\quad X_1, \dots, X_n$ clones indep.

1) $\bar{X} \sim N(\mu, \sigma^2/n)$

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

3) $\bar{X}$ e S^2 son independientes

⊛ $X \sim \Phi \Rightarrow \frac{\bar{X}}{(\sqrt{S^2}/\sqrt{n})}$ es t_{n-1}

ESTADÍSTICOS MAX/MIN

$F_{M_n}(t) = (F_X(t))^n$

$F_{m_n}(t) = 1 - (1 - F_X(t))^n$

$\mathbb{P}(\min(\dots) \leq t) = 1 - \mathbb{P}(\min(\dots) > t)$

INTERVALO PARA LA MEDIA DE UNA POBLACIÓN GENERAL

X v.a. $\quad \mu_\theta = \mathbb{E}_\theta(X)$; $\quad V(\theta) = V_\theta(X)$

$\bar{X} \overset{d}{\approx} N(\mu_\theta, V_\theta/n)$, esto es, $\frac{\bar{X} - \mu_\theta}{\sqrt{V_\theta/n}} \overset{d}{\approx} N(0,1)$

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

ejemplo 1: Intervalo para λ de $X \sim Poiss(\lambda)$

$\mu_\lambda = V_\lambda = \lambda$. La estimación $\hat{\lambda} = \bar{x}$ nos lleva:

Intervalo: $\left(\bar{x} - Z_{\alpha/2}\frac{\sqrt{\bar{x}}}{\sqrt{n}} , \bar{x} + Z_{\alpha/2}\frac{\sqrt{\bar{x}}}{\sqrt{n}}\right)$ confianza $1-\alpha$

ejemplo 2: Intervalo esperanza $X \sim Exp(\lambda)$

$\mu_\lambda = 1/\lambda$; $V_\lambda = 1/\lambda^2$. Planteamos $\bar{x} = \mu_\lambda$

$\Rightarrow \hat{\lambda} = 1/\bar{x}$, y de ahí, $V_{\hat{\lambda}} = \bar{x}^2$

Intervalo: $\left(\bar{x} - Z_{\alpha/2}\frac{\bar{x}}{\sqrt{n}} , \bar{x} + Z_{\alpha/2}\frac{\bar{x}}{\sqrt{n}}\right)$

para aprox. $\mathbb{E}_\lambda(x)$ con confianza $1-\alpha$

ejemplo 3: Intervalo esperanza $X \sim Ray(\theta)$

$\mu_\theta = \sqrt{\pi\theta/2}$; $V_\theta = (2 - \pi/2)\theta$.

Planteamos $\bar{x} = \mu_\theta \Rightarrow \hat{\theta} = \frac{2}{\pi}\bar{x}^2$

$\Rightarrow V_{\hat{\theta}} = (2 - \frac{\pi}{2})\hat{\theta} = (\frac{4}{\pi} - 1)\bar{x}^2$

Intervalo: $\bar{x} \pm Z_{\alpha/2}\frac{\bar{x}}{\sqrt{n}}\sqrt{4/\pi - 1}$

para estimar $\mathbb{E}_\theta(x)$ con confianza $1-\alpha$

DATOS $\theta \in [0,1]$

valor	0	1	2	3
prob	$\theta/3$	$1/3$	$1/3$	$\frac{1-\theta}{3}$

¿M? ¿EMV?

Muestra:

valor	0	1	2	3
#ap	10	27	23	40

Momentos: intentamos $\mathbb{E}_\theta(X) = \bar{X}$

$\mathbb{E}_\theta(X) = 0 \cdot \frac{\theta}{3} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1-\theta}{3} = 2 - \theta$

$\bar{x} = \frac{0.10 + 1.27 + 2.23 + 3.40}{100} = 1'93$

$\boxed{M_\theta = 2 - \bar{X} \Rightarrow \theta = 2 - \bar{x}}$ estimador

$2 - \theta = 1'93 \Rightarrow \boxed{\theta = 0'07}$ estimación

Máxima vero $\text{VERO} = \prod_{i=1}^{n} f(x_i;\theta) = \left(\frac{\theta}{3}\right)^{10} \cdot \left(\frac{1}{3}\right)^{27} \cdot \left(\frac{1}{3}\right)^{23} \cdot \left(\frac{1-\theta}{3}\right)^{40}$

$\log \text{VERO} \to$ der + max $\to \boxed{\theta = \frac{1}{5}}$

estimador $\to$ despejar θ teóricame.

DATOS $X \sim N(\mu, 1)$ $\mu \in [-1,1]$

$\text{VERO}(\theta; x_1, \dots, x_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \exp\left(-\frac{1}{2}\sum(x_i - \mu)^2\right)$

log + der $\Rightarrow \boxed{\mu = \bar{x}}$

Pero $\mu \in [-1,1]$, entonces:

$$\boxed{\hat{\theta}} = \begin{cases} \bar{x} & \text{si } \bar{x} \in (-1,1) \\ 1 & \text{si } \bar{x} > 1 \\ -1 & \text{si } \bar{x} < -1 \end{cases}$$

DATOS $\Theta = \{0,1\}$

$f(x;0) = 1$ $\quad f(x;1) = \frac{1}{2\sqrt{x}}$

Máx VERO para θ

$\text{VERO}(0; x_1, \dots, x_n) = \prod f(x_i, 0) = 1$

$\text{VERO}(1; x_1, \dots, x_n) = \prod f(x_i, 1) = \frac{1}{2^n} \cdot \frac{1}{\sqrt{\prod x_i}}$ (*)

$$\boxed{\hat{\theta}} = \begin{cases} 0 & \text{si } (*) < 1 \\ 1 & \text{si } (*) > 1 \\ \text{Sin estimación} & \text{si } (*) = 1 \end{cases}$$

DATOS: $\theta \in (0,1)$

valor	-1	0	1
prob	$\theta/4$	$1-\theta/2$	$\theta/4$

Cota Cramer-Rao

$Y = \partial_\theta \log f(x;\theta)$

$\frac{d}{d\theta} \ln f(-1;\theta) = \frac{1}{\theta}$

$\frac{d}{d\theta} \ln f(0;\theta) = \frac{1}{\theta - 2}$

$\frac{d}{d\theta} \ln f(1;\theta) = \frac{1}{\theta}$

valor	$1/\theta$	$\frac{1}{\theta-2}$	$1/\theta$
prob	$\theta/4$	$1-\theta/2$	$\theta/4$

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

$V_\theta(T) \geq \frac{1}{n V_\theta(Y)} = \frac{1}{n I_X(\theta)}$

DATOS b

$T = 2\left(\frac{1}{n}\sum x_i^2\right)$ ¿insesg? ¿min. var?

$\mathbb{E}_\theta(T) = 2\mathbb{E}_\theta(x^2) = 2\left((-1)^2 \cdot \frac{\theta}{4} + 0^2 \cdot \left(1 - \frac{\theta}{2}\right) + 1^2 \cdot (\theta/4)\right) = \theta$ ✓

$V_\theta(T) = \frac{4n}{n^2} V_\theta(x^2) = \frac{4}{n}\left(\mathbb{E}_\theta(x^4) - \mathbb{E}(x^2)^2\right)$

DATOS: $X \sim Geo(p)$ $p \in (0,1)$ $T_n = 1/\bar{X}_n$ estima p

TCL $\Rightarrow \sqrt{n}\left(\bar{X}_n - 1/p\right) \longrightarrow N\left(0, \frac{1-p}{p^2}\right)$

$g(x) = \frac{1}{x}$; $g'(x) = \frac{-1}{x^2}$; $|g'(x)|^2 = \frac{1}{x^4}$

$\Rightarrow \sqrt{n}\left(g(\bar{X}_n) - g(1/p)\right) \longrightarrow N\left(0, |g'(1/p)|^2 \cdot \frac{1-p}{p^2}\right) \Rightarrow$

$\Rightarrow \sqrt{n}\left(1/\bar{X}_n - p\right) \longrightarrow N\left(0, (1-p)p^2\right)$

DATOS: $T = \left(\frac{1}{2n}\sum x_i^2\right)^{1/2}$ normalidad asintótica

TCL $\Rightarrow \sqrt{n}\left(\overline{X^2} - \underbrace{2\theta^2}_{E(X^2)}\right) \longrightarrow N\left(0, \underbrace{20\theta^4}_{V(X^2)}\right)$

$g(x) = \sqrt{\frac{x}{2}}$; $g'(x) = \frac{1}{2\sqrt{2}\sqrt{x}}$; $|g'(x)|^2 = \frac{1}{8x}$

$\Rightarrow \sqrt{n}(T - \theta) \longrightarrow N\left(0, \frac{1}{16\theta^2} \cdot 20\theta^4\right)$

DATOS: $X \sim Ber(p)$ $\text{odds} = \frac{p}{1-p}$ $T_n = \frac{\bar{X}_n}{1 - \bar{X}_n}$

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

$g(x) = \frac{x}{1-x}$; $g'(x) = \frac{1}{(1-x)^2}$; $|g'(x)|^2 = \frac{1}{(1-x)^4}$

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

DATOS: $X \sim Unif(0,a)$ $M_n = \max(X_i)$ estimador de a

$\mathbb{P}\left(\frac{M_n}{a} \leq t\right) = \mathbb{P}(X \leq at)^n = t^n$ para $t \in (0,1)$

Hallar C_1 y C_2 para que $\mathbb{P}_a\left(\frac{M_n}{a} \leq C_1\right) = \mathbb{P}_a\left(\frac{M_n}{a} \geq C_2\right) = \alpha/2$

$C_1^n = \mathbb{P}\left(\frac{M_n}{a} \leq C_1\right) = \alpha/2 \Rightarrow$

$\Rightarrow C_1 = \sqrt[n]{\alpha/2}$

$\alpha/2 = \mathbb{P}\left(\frac{M_n}{a} \geq C_2\right) = 1 - \mathbb{P}\left(\frac{M_n}{a} \leq C_2\right) = 1 - C_2^n \Rightarrow C_2 = \sqrt[n]{1 - \alpha/2}$

Intervalo de confianza:

Probabilidad (antes experimento): $1 - \alpha = \mathbb{P}\left(C_1 \leq \frac{M_n}{a} \leq C_2\right)$

Muestra $x_1, \dots, x_n \longrightarrow m_n = \max$

confianza $1-\alpha$: $\frac{m_n}{C_2} \leq a \leq \frac{m_n}{C_1}$

a) $n = 200$ $\alpha = 5\%$ $H_0 \equiv$ moneda equilibrada ¿rechazar?

$\to H_0 \equiv p = 1/2 = p_0$

$\bar{x}$ = proporción de unos

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

$\Rightarrow$ Rechazamos si: $\left|\bar{x} - \frac{1}{2}\right| \geq \frac{1}{2} \cdot 1'96 \cdot 1/\sqrt{200}$

b) n general. $\alpha = 5\%$ proporción de caras $= 57\%$

¿Valores de n para la que rechazaríamos, y valores para los que aceptaríamos? Tenemos $\bar{x}$ y α.

Si $n \geq \left(\frac{1}{2} \cdot \frac{1'96}{0'07}\right)^2$ rechazamos, por el contrario, aceptamos

c) ¿p valor?

$\left|\bar{x} - \frac{1}{2}\right| = \frac{1}{2} Z_{\alpha/2} \cdot \frac{1}{\sqrt{n}} \Rightarrow \phi\left(|\bar{x} - 1/2| \cdot 2\sqrt{n}\right) = 1 - \alpha/2 \Rightarrow$

$\Rightarrow \alpha = 2\left(1 - \phi\left(|\bar{x} - 1/2| \cdot 2\sqrt{n}\right)\right)$

Test razón verosimilitudes: $f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/2\theta}$ $x \in \mathbb{R}$

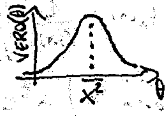
test para $H_0 = \theta < \theta_0$ con nivel α. $\theta \in (0,\infty) = \Theta$

① Calcular VERO: $\text{VERO}(\theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sqrt{\theta}}\right)^n \exp\left(-\frac{\sum x_i^2}{2\theta}\right)$ $\to n\overline{x^2}$

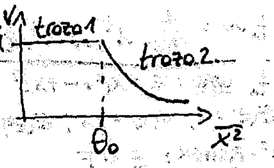
dibujar VERO, monotonía, $\text{VERO}(0^+) = 0$, $\text{VERO}(\infty) = 0$ ...

② Calcular MAX VERO: logVERO $\to$ derivar $\to$ obtenemos estimación

③ $RV = \frac{\sup_{\theta \leq \theta_0} \text{VERO}(\theta)}{\sup_{\theta \in \Theta} \text{VERO}(\theta)} = \begin{cases} \frac{\text{VERO}(\overline{x^2})}{\text{VERO}(\overline{x^2})} = 1 & \text{si } \theta_0 \geq \overline{x^2} \\ \frac{\text{VERO}(\theta_0)}{\text{VERO}(\overline{x^2})} & \text{si } \theta_0 < \overline{x^2} \end{cases}$ $\quad \theta = \overline{x^2}$



④ Rechazo si $RV \leq C$ (poco verosímil)



trozo 1: siempre $RV = 1 \to$ nunca rechazo

trozo 2: $RV = \left(\frac{\overline{x^2}}{\theta_0}\right)^{n/2} \cdot \exp\left(-\frac{1}{2}\frac{\overline{x^2}}{\theta_0} n + \frac{1}{2} n\right)$

Región rechazo: $\left(\frac{\overline{x^2}}{\theta_0}\right)^{n/2} \cdot \exp\left(-\frac{1}{2}\frac{\overline{x^2}}{\theta_0} n + \frac{1}{2} n\right) \leq C$ and $\overline{x^2} > \theta_0$

⑤ Función de potencia: $\mathbb{P}_\theta(\text{rechazar}) = \beta(\theta)$

⑥ Significación: $\sup_{\theta \in \Theta_0} \beta(\theta)$

$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$

$\bar{x}=\sum_{i=1}^{K} f_i x_i$

si $z=a+bx$, $\bar{z}=a+b\bar{x}$

$V_x=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\mathbb{E}(x^2)-\mathbb{E}(x)^2$

$V_x=\sum_{i=1}^{K} f_i(x_i-\bar{x})^2$; S_x^2 = cuasivari

$(n-1)S_x^2=nV_x$; $V_z=b^2V_x$

$\text{asim}_x=\dfrac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{V_x^{3/2}}=\frac{1}{n}\sum_{i=1}^{n}\left(\dfrac{x_i-\bar{x}}{\sqrt{V_x}}\right)^3$

$z=a+bx \rightarrow \text{asim}_z=\text{sgn}(b)\,\text{asim}_x$

$\rho_{x,y}=\dfrac{\text{cov}_{x,y}}{\sqrt{V_x}\sqrt{V_y}}$ adimensional ; $\rho_{a+bx,c+dy}=\rho_{x,y}$, $\text{sgn}(\rho)=\text{sgn}(\text{cov}_{x,y})$ tipificado

$\text{cov}_{x,y}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$

si $y=a+bx \rightarrow \text{cov}_{x,y}=bV_x$

$\text{cov}_{x,y}=\overline{xy}-\bar{x}\cdot\bar{y}$; $|\text{cov}_{x,y}|\leq\sqrt{V_x}\sqrt{V_y}$

X,Y dep $\Rightarrow$ cov $=0 \Rightarrow \rho=0$

$E(a,b)=\frac{1}{n}\sum_{i=1}^{n}(y_i-(a+bx_i))^2$

$\hat{b}=\dfrac{\text{cov}_{x,y}}{V_x}$; $\hat{a}=\bar{y}-\hat{b}\bar{x}=\bar{y}-\left(\dfrac{\text{cov}_{x,y}}{V_x}\right)\bar{x}$

$y=\hat{a}+\hat{b}x$; $y-\bar{y}=\hat{b}(x-\bar{x})$; $\dfrac{y-\bar{y}}{\sqrt{V_y}}=\rho_{x,y}\dfrac{x-\bar{x}}{\sqrt{V_x}}$

coef. determ $=\rho_{x,y}^2$

$\sqrt{E(\hat{a},\hat{b})}=\sqrt{\frac{1}{n}\sum(y_i-(\hat{a}-\hat{b}x_i))^2}=\sqrt{V_y}\sqrt{1-\rho_{x,y}^2}$

$\mathbb{E}(X)=\sum_{i=1}^{n} x_i p_i=\sum_{i=1}^{n} x_i\,\mathbb{P}(X=x_i)=\int_{\mathbb{R}} x f_X(x)\,dx$

$\rightarrow \mathbb{E}(aX+b)=a\mathbb{E}(X)+b$; $\mathbb{E}(X+Y)=\mathbb{E}(X)+\mathbb{E}(Y)$

$\rightarrow X,Y$ indep. $\Rightarrow \mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$

$\rightarrow$ Cauchy-Schwarz: $|\mathbb{E}(XY)|^2\leq\mathbb{E}(X^2)\mathbb{E}(Y^2)$

$\rightarrow Y=h(X) \rightarrow \mathbb{E}(Y)=\mathbb{E}(h(X))=\sum_{i=1}^{n} h(x_i)p_i$

$V(X)=\mathbb{E}((X-\mathbb{E}(X))^2)=\mathbb{E}(X^2)-\mathbb{E}(X)^2$; $V(\lambda X)=\lambda^2V(X)$

mom. no centrados $\rightarrow \mathbb{E}((X-\mathbb{E}(X))^K)=\sum(x_i-\mathbb{E}(X))^K\cdot P(X=x_i)$

mom. centrados $\rightarrow \mathbb{E}(X^K)=\sum x_i^K\,\mathbb{P}(X=x_i)$

$\text{asim}_x=\dfrac{\mathbb{E}((x-\mu)^3)}{\sigma(x)^3}$; $\text{curtosis}_x=\dfrac{\mathbb{E}((x-\mu)^4)}{\sigma(x)^4}$

$V(X+Y)=V(X)+V(Y)+2\text{cov}_{X,Y}$

LOG: $y=B\ln(x)+A$; $Z=\ln(x)$

EXP: $y=Ce^{Dx}$; $W=\ln(Y)$; $C=e^{\hat{a}}$; $D=\hat{b}$

ejemplo: $Y=ae^{bx} \rightarrow \ln(Y)=\ln(a)+bx$
$\rightarrow W=a'+bx \rightarrow$ sacamos $\hat{a}'$ y $\hat{b}$.

Des. Chebyshev: $\lambda>0$ $\mathbb{E}(|X-\mathbb{E}(X)|>\lambda)\leq\dfrac{V(X)}{\lambda^2}$

Des. Markov: $a>0$ $\mathbb{E}(X>a)\leq\dfrac{\mathbb{E}(X)}{a}$
↳ v.a. no-negativa

Si $Y=H(X)$: $f_Y(H(x))\cdot|H'(x)|=f_X(x)$

$f_Y(y)=f_X(H^{-1}(y))\cdot|(H^{-1})'(y)|$

Si $Z=X+Y$ (indep.) $\rightarrow f_Z(z)=\int_{-\infty}^{\infty} f_{(X,Y)}(x,z-x)\,dx$

Sean $\vec{m}\in\mathbb{R}^n$, $V\in M_{n\times n}$ simétrica def. posit. $X\sim N(\vec{m},V)$:

$f_X(\vec{x})=\dfrac{1}{(\sqrt{2\pi})^n}\cdot\dfrac{1}{\sqrt{\det(V)}}\cdot\exp\left(\dfrac{-(\vec{x}-\vec{m})^t}{2}\cdot V^{-1}\cdot(\vec{x}-\vec{m})\right)$

Como V es simétrica y def. positiva: $V=UU^t$, U matriz no singular. $\det(V)=\det(U)^2$

$V^{-1}=(U^{-1})^tU^{-1}$

$f_X(\vec{x})=\dfrac{1}{(\sqrt{2\pi})^n}\cdot\dfrac{1}{|\det(U)|}\cdot\exp\left(-\frac{1}{2}\|U^{-1}\cdot(\vec{x}-\vec{m})\|^2\right)$

H transf. lineal $Y=\vec{h}+MX$:

$f_X(x)=f_Y(h+Mx)\cdot|\det(M)|$

$f_Y(y)=\dfrac{1}{|\det(M)|}\cdot f_X(M^{-1}(y-h))$

$X\sim Geo(p)$

$F=p(1-p)^{K-1}p$

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

$V(X)=\dfrac{1-p}{p^2}$

UNIFORME: $f_X(x)=\begin{cases}1/(b-a), & x\in[a,b]\\ 0 & \text{resto}\end{cases}$

$F_X(x)=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & x\in[a,b]\\ 1 & \text{resto}\end{cases}$; $\mathbb{E}(X)=\dfrac{b-a}{2}$

$V(X)=\mathbb{E}(X^2)-\mathbb{E}(X)^2=\left(\dfrac{1}{b-a}\int_a^b x^2dx\right)-\left(\dfrac{b-a}{2}\right)^2$

$X\sim \text{unif}(a,b) \longrightarrow \dfrac{X}{b-a}\sim\text{unif}(0,1)$

BERNOULLI: $X\sim Ber(p)$ $x_k\in\{0,1\}$

$p=\mathbb{P}(X=1)$ $\mathbb{E}(X)=p$ $V(X)=p(1-p)$

$\mathbb{P}(X=K)=p^K(1-p)^{1-K}$

BINOMIAL: $X\sim Bin(n,p)$

$f_X(k)=\mathbb{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$

$E(X)=np$ $V(X)=np(1-p)$

POISSON: $X\sim$ Poisson $(\lambda=np)$.

Bin con $n\to\infty$, $p\to 0$

$E(X)=\lambda$ $V(X)=\lambda$ $\mathbb{P}(X=k)=\dfrac{e^{-\lambda}\lambda^k}{k!}$

(EX)PONENCIAL $X\sim Exp(\lambda)$

$f_X(x)=\begin{cases}\lambda e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$; $\mathbb{E}(X)=\dfrac{1}{\lambda}$

$F_X(x)=1-e^{-\lambda x}$ $x\geq 0$; $V(X)=\dfrac{1}{\lambda^2}$

$X\sim Exp(\lambda) \rightarrow \lambda X\sim Exp(1)$

$X\sim Exp(\lambda)$ $Y\sim Exp(\lambda) \Rightarrow X+Y\sim Gamma(\lambda,2)$

(NO)RMAL UNI $X\sim N(\mu,\sigma^2)$

$f_X(x)=\dfrac{1}{\sqrt{2\pi\sigma^2}}\cdot\exp\left(-\frac{1}{2}(x-\mu)^2/\sigma^2\right)$

$\mathbb{E}(X)=\mu$; $V(X)=\sigma^2$; $\mathbb{E}(X^2)=1$ si $X\sim\Phi$

$X\sim N(\mu,\sigma^2) \rightarrow Y=a+bX \rightarrow Y\sim N(b\mu,b^2\sigma^2)$

$a,b\in\mathbb{R}$ $X,Y\sim N(-,-) \rightarrow aX+bY\sim N(a\mu+b\nu, a^2\sigma^2+b^2\rho^2)$

(GA)MMA: $X\sim Gamma(\lambda,t)$ $t>0$

$\Gamma(t)=\int_0^\infty x^{t-1}e^{-x}dx=\int_0^\infty x^te^{-x}\dfrac{dx}{x}$; $\Gamma(1)=1$; $\Gamma(1/2)=\sqrt{\pi}$

$\Gamma(t)=(t-1)\Gamma(t-1)$; $\Gamma(n)=(n-1)!$; $\Gamma(-1/2)=-2\sqrt{\pi}$

$\Gamma(k+\frac{1}{2})=\dfrac{(2k)!}{4^kk!}\sqrt{\pi}$ k entero ≥ 0 ; $\mathbb{E}(X^K)=\dfrac{(t+k-1)!}{\lambda^k(t-1)!}$

$\lambda,b>0 \rightarrow \int_0^\infty x^te^{-bx}\dfrac{dx}{x}=\dfrac{\Gamma(t)}{b^t}$; $V(X)=t/\lambda^2$

$f_X(x)=\dfrac{1}{\Gamma(t)}\lambda^tx^{t-1}e^{-\lambda x}=\dfrac{1}{\Gamma(t)}(\lambda x)^te^{-\lambda x}\dfrac{1}{x}$ $x>0$

$f_X(x)=0$ para $x<0$; $X\sim Gamma(\lambda,t)$ $Y\sim Gamma(\lambda,s)$ INDEPENDIENTES

$\rightarrow X+Y\sim Gamma(\lambda,t+s)$

$Gamma(\lambda,1)\sim exp(\lambda)$; $Gamma(1/2,n/2)\sim\chi^2_n$

$X\sim N(\mu,\sigma^2) \rightarrow \mathbb{E}(X^2)=\mu^2+\sigma^2$ $V(X)=\sigma^2$

$\mathbb{E}(X^3)=\mu^3+3\mu\sigma^2$ $\mathbb{E}((X-\mu)^3)=0$

$\mathbb{E}(X^4)=\mu^4+6\mu^2\sigma^2+3\sigma^4$ $\mathbb{E}((X-\mu)^4)=3\sigma^4$

Sea $X\sim N(\vec{m},V)$. $\vec{h}\in\mathbb{R}^n$ y $B\in M_n$ no singular

$Z=\vec{h}+BX \Rightarrow Z\sim N(\vec{h}+B\vec{m}, BVB^t)$

Y es $N(\vec{0},I_n)$ y $O\in M_n$ ortogonal $\Rightarrow Z=OY$ es $N(\vec{0},I_n)$

$X\sim N(\vec{m},V)$, $\vec{a}\in\mathbb{R}^n$ y $\vec{a}\neq\vec{0}$

$\sum a_ix_i$ no es un vector aleatorio, es un vector aleatorio

$\sum a_ix_i=\vec{a}^tX\sim N(\vec{a}^t\vec{m},\vec{a}^tV\vec{a})$

CHI-CUADRADO $Z=X_1^2+\cdots+X_n^2$ donde X_i son normales estándar INDEP.

$\mathbb{E}(Z)=n$; $V(Z)=2n$; $\mathbb{E}(Z^2)=n(n+2)$

$f_Z(x)=\dfrac{1}{\Gamma(n/2)}\cdot\dfrac{1}{2^{n/2}}\cdot x^{\frac{n}{2}-1}e^{-x/2}$

$\mathbb{E}(Z^\alpha)=2^\alpha\dfrac{\Gamma(\frac{n}{2}+\alpha)}{\Gamma(n/2)}$ cuando $\frac{n}{2}+\alpha>0$

$\mathbb{E}(1/Z)=\dfrac{1}{n-2}$ cuando $n\geq 3$

$\mathbb{E}(\sqrt{Z})=\sqrt{2}\dfrac{\Gamma(\frac{n+1}{2})}{\Gamma(n/2)}$ cuando $n\geq 1$

$X_1^2\sim Gamma(1/2,1/2)$; $X_n^2\sim Gam(\frac{1}{2},\frac{n}{2})$

F DE FISHER: $Z\sim F_{n,m}$

$Z=\dfrac{(X_1^2+\cdots+X_n^2)/n}{(Y_1^2+\cdots+Y_m^2)/m}$

$Z=\dfrac{U_n/n}{V_m/m}$, $U_n\sim\chi^2_n$, $V_m\sim\chi^2_m$

$\mathbb{E}(Z)=\dfrac{m}{m-2}$ $m\geq 3$

$m=1,2 \rightarrow \mathbb{E}(Z)=\infty$

$V(Z)=\dfrac{2(m+n-2)m^2}{n(m-4)(m-2)^2}$ $m\geq 5$

$m\leq 4 \rightarrow V(Z)=\infty$

t_n DE STUDENT $n\geq 2$

$Y\sim N(0,1)$; $U_n\sim\chi^2_n$

$Z:=\dfrac{Y}{\sqrt{U_n/n}}$

$\mathbb{E}(Z)=0$; $V(Z)=\mathbb{E}(Z^2)-\overset{0}{\mathbb{E}(Z)^2}$

$\mathbb{E}(Z^2)=\underbrace{\mathbb{E}(Y^2)}_{1}\cdot\mathbb{E}\left(\dfrac{n}{U_n}\right)=\dfrac{n}{n-2}$

NOTA: $n\to\infty$: $t_n \xrightarrow{n/n-2} N(0,1)$

ESTADÍSTICOS: $X_1,\ldots,X_n$ clones $\rightarrow$ indep. si $i\neq j$; $\rightarrow X_i\overset{d}{=}X$

estadístico $T=H(X_1,\ldots,X_n)$ donde $H:\mathbb{R}^n\rightarrow\mathbb{R}$.

ESTADÍSTICO MEDIA MUESTRAL $\bar{X}$ $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

$\mathbb{E}(\bar{X})=\mathbb{E}(X)$; $V(\bar{X})=\dfrac{V(X)}{n}$; $\mathbb{P}(|\bar{X}-\mathbb{E}(X)|\geq\varepsilon)\leq\dfrac{V(X)}{n\varepsilon^2}$ CHEBYSHEV

- Si $X\sim Ber(p) \rightarrow n\bar{X}\sim Bin(n,p)$: $\mathbb{P}(n\bar{X}=k)=\binom{n}{k}p^k(1-p)^{n-k}$
- Si $X\sim Poiss(\lambda)$ $\lambda>0 \rightarrow n\bar{X}\sim Poiss(n\lambda)$: $\mathbb{P}(n\bar{X}=k)=\mathbb{P}(\bar{X}=\frac{k}{n})=e^{-\lambda n}\cdot\dfrac{(\lambda n)^k}{k!}$
- Si $X\sim Exp(\lambda) \rightarrow n\bar{X}\sim Gamma(\lambda,n)$
- Si $X\sim N(\mu,\sigma^2) \rightarrow n\bar{X}\sim N(n\mu,n\sigma^2)$ $\rightarrow \bar{X}\sim N(\mu,\frac{\sigma^2}{n})$

ESTO ESTA MAL: C

$\mathbb{P}(\bar{X}\leq\lambda/2, S\leq\lambda/2)=\mathbb{P}(\bar{X}<\lambda/2)\cdot\mathbb{P}(S\leq\lambda/2)$

$\mathbb{P}(\{|\bar{X}|\geq\sigma\}\cup\{S^2>2\sigma^2\})=(\mathbb{P}(|\bar{X}|>\sigma)+\mathbb{P}(S^2>2\sigma^2))$ del prod

$\boxed{\text{TCL}}$ $X_1,\ldots,X_n$ clones indep.

$X_1+\cdots+X_n\overset{d}{=}N(n\mathbb{E}(X),nV(X))$ $n\to\infty$

$\bar{X}-\mathbb{E}(X)\overset{d}{=}N(0,V(X)/n)$ $n\to\infty$

$\sqrt{n}(\bar{X}-\mathbb{E}(X))\overset{d}{=}N(0,V(X))$ $n\to\infty$

ESTADÍSTICO CUASIVARIANZA MUESTRAL S^2

$S^2=\frac{1}{n-1}\sum(X_i-\bar{X})^2$; $\mathbb{E}(S^2)=V(X)$

$n(n-1)S^2=(n-1)\sum_{i=1}^{n}X_i^2-\sum_{i\neq j}X_iX_j$

$V(S^2)=\frac{1}{n}\mathbb{E}((X-\mathbb{E}(X))^4)-\dfrac{n-3}{n(n-1)}V(X)^2$

$\mathbb{P}(|S^2-V(X)|>\lambda)\leq\dfrac{V(S^2)}{\lambda^2}\leq\dfrac{1}{\lambda^2}\cdot\dfrac{1}{n}\cdot\mathbb{E}((X-\mathbb{E}(X))^4)$

TEOREMA FISHER-COCHRAN: Si $X\sim N(\mu,\sigma^2)$, $X_1,\ldots,X_n$ clones indep. de X:

1) $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$ 2) $\dfrac{(n-1)}{\sigma^2}S^2\sim\chi^2_{n-1}$

3) $\bar{X}$ e S^2 son independientes

Además: si $X\sim N(0,1) \rightarrow \dfrac{\bar{X}}{(\sqrt{S^2}/\sqrt{n})}$ es t_{n-1}

ESTADÍSTICOS MAX/MIN

$F_{M_n}(t)=(F_X(t))^n$; $F_{m_n}(t)=1-(1-F_X(t))^n$

α mínimo esencial $(F_X(X<\alpha)=0 \wedge F(X>\alpha)>0)$

β máximo esencial $(F_X(X<\beta)<1 \wedge F(X>\beta)=1)$ de X

$\mathbb{P}(\alpha\leq m_n\leq r)\xrightarrow{n\to\infty}1$ $r>\alpha$

$\mathbb{P}(r\leq M_n\leq\beta)\xrightarrow{n\to\infty}1$ $r<\beta$

$\mathbb{P}(\min(X_1,\ldots,X_n)\leq t)=1-\mathbb{P}(\min(X_1,\ldots,X_n)>t)$