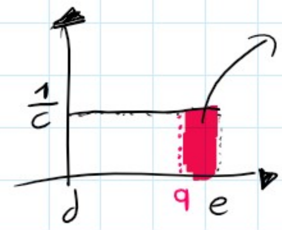
Recuerde: Sea $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $E(M_n) = \mu$, $Var(M_n) = \frac{\sigma^2}{n}$

Desigualdad de markov



$P(X \geq a) \leq \frac{E(X)}{a}$

si $a \to e \Rightarrow P(X \geq a) \to 0$

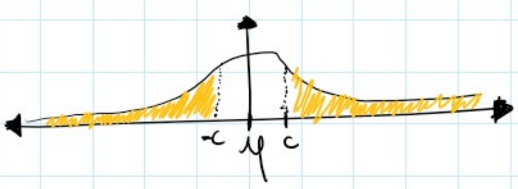
Si una v.a. $X \geq 0$, entonces

$$P(X \geq a) \leq \frac{E(X)}{a}, \quad \forall a > 0$$

Desigualdad de chebyshev

sea X v.a con media μ y var σ^2, entonces

$$P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}, \quad \forall c > 0.$$

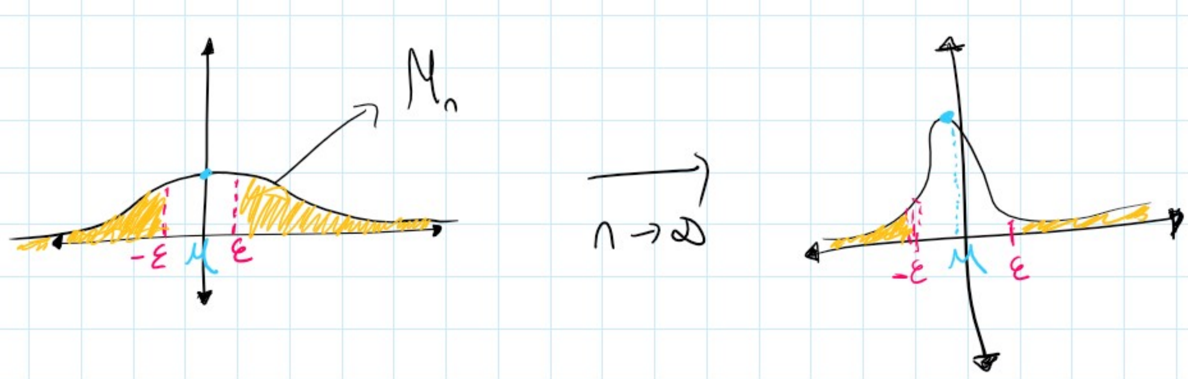


Ley débil de los grandes números

Sea $\{X_n\}$ una secuencia de v.a's iid con media μ.

$\forall \varepsilon > 0$, se tiene:

$$P(|M_n - \mu| \geq \varepsilon) = P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \varepsilon\right) \longrightarrow 0, \quad n \to \infty$$



Convergencia en probabilidad

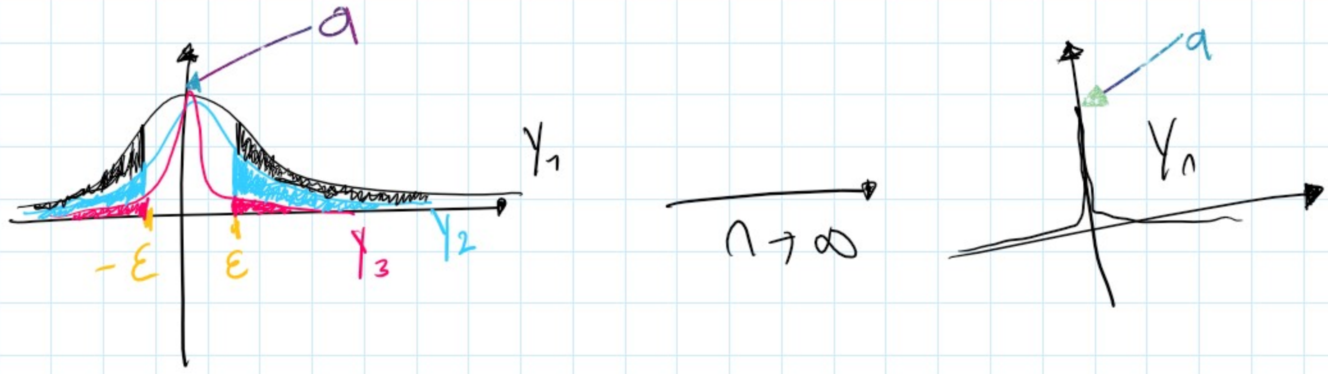
Sea $\{Y_n\}$ una secuencia de v.a.s (no necesariamente indep).

Y sea $a \in \mathbb{R}$.

Si $\forall \varepsilon > 0$, $\lim_{n\to\infty} P(|Y_n - a| \geq \varepsilon) = 0$

Si $\forall \varepsilon > 0, \lim_{n\to\infty} P(|Y_n - a| < \varepsilon) = 1$

$\Rightarrow$ Y_n converge en probabilidad a 'a'.



Teorema del límite Central

Sea $\{X_n\}$ una secuencia de iid v.a.'s con media μ y varianza σ^2

$$Z_n = \frac{\left(\sum_{i=1}^{n} X_i\right) - n \cdot \mu}{\sigma \sqrt{n}}$$

entonces la CDF de Z_n converge a la CDF N.E.

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{z} e^{-x^2/2}\, dx$$

Aproximaciones basadas en el TLC

Sea $S_n = \sum_{i=1}^{n} X_i$ de X_i iid con media μ y var σ^2, para n grande S_n se puede tratar como normal así:

$$\left.\begin{aligned} &P(S_n \leq c) \\ &P\left(Z \leq \frac{c - n\mu}{\sigma \sqrt{n}}\right) \approx \Phi(z) \end{aligned}\right\} \text{Normalización.}$$

Aproximación de Moivre - laplace

Una binomial es $S_n = \sum_{i=1}^{n} X_i$, donde cada X_i es bernoulli iid.

$E(X_i) = p$, $var(X_i) = p \cdot (1-p)$.

$$P(k \leq S_n \leq l) \approx \Phi\left(\frac{l + \frac{1}{2} - np}{\sqrt{n \cdot p \cdot (1-p)}}\right) - \Phi\left(\frac{k - \frac{1}{2} - np}{\sqrt{n \cdot p \cdot (1-p)}}\right)$$

Ley fuerte de los grandes números

grandes números

Sea $\{X_n\}$ iid. con media μ, entonces

$$M_n = \frac{\sum_{i=1}^{n} X_i}{n} \longrightarrow \mu$$, con probabilidad 1 es decir:

$$P\left(\lim_{n \to \infty} \frac{\sum_{i=1}^{n} X_i}{n} = \mu\right) = 1$$

Muestreo de una población normal

Sea $\{X_n\}$ una muestra aleatoria de una $N(\mu, \sigma^2)$ distrib.

y sean $\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$, $S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$, entonces:

a) $\bar{X}$ y S^2 son independientes v.a's.

b) $\bar{X}$ es normal distribuida con media μ y var $\frac{\sigma^2}{n}$

c) $\frac{(n-1) \cdot S^2}{\sigma^2}$ es chi^2 con $(n-1)$ g.l.

Chi^2, χ^2

$\chi^2_p \to$ g.l.

Normal estándar

a) si Z es $N(0,1)$, entonces $Z^2 \approx \chi^2$

b) Sean $\{X_n\}$ independientes y $X_i \sim \chi^2_{p_i} \Rightarrow \sum_{i=1}^{n} X_i \sim \chi^2_{\sum_{i=1}^{n} p_i}$

Una suma de chi^2's indep. es una chi^2 con Σp_i g.l.

T student

$\longrightarrow$ se usa para estimar μ si no se conoce σ^2.

Sean $\{X_n\}$ iid $X_i \sim N(\mu, \sigma^2)$

$$\Rightarrow T = \frac{(\bar{X} - \mu)}{\frac{S}{\sqrt{n}}}$$ es T-student con $(n-1 = p)$ g.l.

la PDF es:

$$f_T(t) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)} \cdot \frac{1}{(p\pi)^{1/2}} \cdot \frac{1}{\left(1 + \frac{t^2}{p}\right)^{\frac{p+1}{2}}}$$

otra forma de obtener una T student es con:

$U = N(0,1)$

$V = \chi^2_p$ indep.

$$\Rightarrow T = \frac{U}{\sqrt{V/p}}$$

$V = \chi^2_p$ indep. $\qquad \sqrt{V/p}$

- $E(T_p) = 0$, $p > 1$.
- $Var(T_p) = \frac{p}{p-2}$, $p > 2$.

F $\longrightarrow$ usado para comparar variabilidades

Sea $\{y_i\}$ una muestra aleatoria de una población con media μ_y y var σ_y^2, y sean $\{x_i\}, \mu_x, \sigma_x^2$ definidos análogamente.

$$F = \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} \quad \begin{matrix}] \; n-1 \text{ g.l.} \\ ] \; m-1 \text{ g.l.} \end{matrix}$$

la PDF es:

$$f_F(x) = \frac{\Gamma\left(\overbrace{\frac{p+q}{2}}^{\text{g.l.s}}\right)}{\Gamma\left(\frac{p}{2}\right)\cdot\Gamma\left(\frac{q}{2}\right)} \cdot \left(\frac{p}{q}\right)^{p/2} \cdot \frac{x^{(p/2 - 1)}}{\left(1+\frac{p}{q}\right)^{\left(\frac{p+q}{2}\right)}}$$

1. $E(F_{n-1,m-1}) = \frac{m-1}{m-3}$
2. si $X \sim F_{p,q}$, $\frac{1}{X} \sim F_{p,q}$
3. $X \sim t_q$, $X^2 \sim F_{1,q}$

Propiedades de estimadores

Eficiencia relativa

sean $\hat{\theta}_1$ y $\hat{\theta}_2$ estimadores insesgados de θ con $Var(\hat{\theta}_1) = \sigma^2_{\hat{\theta}_1}$ y $var(\hat{\theta}_2) = \sigma^2_{\hat{\theta}_2}$ entonces

$$eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$$

Si $eff(\hat{\theta}_1, \hat{\theta}_2) > 1$, $\hat{\theta}_2$ tiene más varianza luego es preferible $\hat{\theta}_1$.

Consistencia

Sea una muestra de tamaño n, $\hat{\theta}_n$ es consistente si:

converge en prob.

Sea una muestra de tamaño n, θ_n es consistate si:

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1, \quad \hat{\theta}_n \xrightarrow{P \to \text{tiende en prob.}} \theta$$

Otra forma de verlo es:

Sea $\hat{\theta}_n$ un estimador insesgado de θ. $\hat{\theta}_n$ es consistente si

$$\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$$

$\longrightarrow$ $\overline{Y}_n$ es un estimador consistente

i) $\overline{Y}_1 - \overline{Y}_2$ también para $\mu_1 - \mu_2$

ii) $\hat{p}_1 - \hat{p}_2$ también para $p_1 - p_2$

Suficiencia

Sea $\{Y_n\}$ una muestra aleatoria de una distrib. con parámetro θ, entonces $U = g(Y_1 \dots Y_n)$ es suficiente para θ si la distribución condicional $Y_1 \dots Y_n$ dado $U = u$ no depende de θ.

Si $P(Y_1 \dots Y_n \mid \theta, U = u)$ no depende de θ

$\Rightarrow$ $g(Y_1 \dots Y_n)$ es suficiente para θ.

Verosimilitud

$$\underbrace{L(Y_1 \dots Y_n \mid \theta)}_{L(\theta)} = \begin{cases} P_{Y_1 \dots Y_n}(Y_1 \dots Y_n \mid \theta) & \text{si } Y_i \text{ discreta} \\ f_{Y_1 \dots Y_n}(Y_1 \dots Y_n \mid \theta) & \text{si } Y_i \text{ continua} \end{cases}$$

Criterio de factorización

Sea U un estimador basado en una muestra aleatoria

$$\text{si } L(\theta) = \underbrace{g(u, \theta)}_{\text{depende de } U \text{ y de } \theta} \cdot \underbrace{h(Y_1 \dots Y_n)}_{\text{no depende de } \theta}$$

Teorema de Rao-Blackwell

Si se tiene $\hat{\theta}$ insesgado y U suficiente Aplicar Rao-Blackwell da el MVUE

Sea $\hat{\theta}$ un estimador para θ tal que $Var(\hat{\theta}) < \infty$
Si U es un estimador suficiente para θ.

Sea $\hat{\theta}^* = E(\hat{\theta} \mid U)$

$$\Rightarrow E(\hat{\theta}^*) = E(E(\hat{\theta} \mid U))$$

$$\Rightarrow E(\hat{\theta}^*) = E(E(\hat{\theta}|U))$$
$$= E(\hat{\theta})$$
$$= \theta$$

$$Var(\hat{\theta}) = Var(E(\hat{\theta}|U)) + E(Var(\hat{\theta}|U))$$
$$= Var(\hat{\theta}^*) + \underbrace{E(Var(\hat{\theta}|U))}_{\geq 0}$$

$$Var(\hat{\theta}) \geq Var(\hat{\theta}^*) \geq 0$$

MVUE ⟶ estimador insesgado de mínima varianza.

Método de Momentos ⟶ para obtener estimadores puntuales de parámetros poblacionales.

el método de momentos usa los momentos de una distribución de prob. para generar a partir de allí un sistema de eqns. de donde por solución salga el estimador para el o los parámetros.

M_K es el K-ésimo momento de una v.a

$$M_K = E(X^K) \longrightarrow \text{momentos poblacionales}$$

m_K es el estimador natural de M_K que es:

$$m_K = \frac{1}{n} \cdot \sum_{i=1}^{n} (Y_i)^K = E(Y^K) \rightsquigarrow \text{momentos muestrales}$$

La idea es entonces solucionar las eqns para los momentos poblacionales y sus correspondientes muestrales.

las eqns. son: $\hat{M}_K = m_K \quad k = 1, \ldots, t \longrightarrow$ # de parámetros a estimar

la idea aquí es que los m_K son buenos estimadores de los M_K y se pueden igualar y solucionar.

Método de máxima Verosimilitud