

Proba y Estadística

Sebastián Sepúlveda A.

1. Probabilidades

1.1. Combinatoria

Reglas de suma y producto

- **Regla de la suma:** Si una tarea T puede ser llevada a cabo en paralelo por dos procesos P_1 y P_2 , tal que hay m maneras de llevar a cabo P_1 , hay n maneras de llevar a cabo P_2 , y hay p maneras de llevar a cabo tanto P_1 como P_2 , entonces T puede ser llevado a cabo de $n + m - p$ maneras.
- **Regla del producto:** Si T puede ser llevada a cabo en serie por dos procesos P_1 y P_2 , tal que hay m maneras de llevar a cabo P_1 y n maneras de llevar a cabo P_2 , entonces hay mn maneras de llevar a cabo T .

Def Permutaciones (Variaciones) Simples. Son las diferentes ordenaciones que se pueden hacer en un arreglo de tamaño r con n elementos distinguibles, donde los objetos se pueden usar sólo una vez.

$$P_r^n = \frac{n!}{(n-r)!}$$

Obs: Usualmente se le llama permutación al caso especial $r = n$ donde: $P_n = n!$

Def Permutaciones con elementos repetidos. Si se tienen n elementos divididos en k grupos, con n_i : cant. de objetos tipo i , tq $\sum_{i=1}^k n_i = n$. El total de permutaciones posibles es:

$$\frac{n!}{\prod_{i=1}^k n_i}$$

Def Combinaciones: Dada una agrupación de n elementos, la cantidad de subconjuntos de k elementos de dicha agrupación está dada por:

$$C_k^n = \binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Obs: El número total de subconjuntos no vacíos que se pueden formar es: $2^n - 1$

1.2. Axiomatica de probabilidades

Def Una medida de probabilidad sobre un espacio Γ cumple.

$$(A1) \quad 0 \leq P(A) \leq 1$$

$$(A2) \quad P(\Omega) = 1, P(\emptyset) = 0$$

(A3) $A_{nn} \in \mathbb{N}$ una familia disjunta de eventos, entonces

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

(1) **Monotonía:** Sean A, B eventos. Si $A \subseteq B$, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$

(2) **Subaditividad:** Sea $\{A_i\}_{i=1}^n$ una familia finita de eventos, entonces $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

(3) **Propiedad del complemento:** $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$

(4) **Principio de Inclusión y Exclusión:**

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

1.3. Probabilidades en eventos

Def Probabilidad condicional Sean A, B eventos tales que $\mathbb{P}(B) > 0$. La probabilidad de A condicionado por B se define por:

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

Teo Fórmula de Bayes: Dados A, B eventos, se tiene que:

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

Teo Probabilidades Totales: Sea Ω un espacio muestral y $\{A_n\}_{n \in \mathbb{N}}$ partición de Ω , entonces:

$$\mathbb{P}(A) = \sum_{n \in \mathbb{N}} \mathbb{P}(A|A_n)\mathbb{P}(A_n)$$

Def Independencia: Diremos que dos eventos A y B son independientes si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

1.4. Variable aleatoria:

Def Variable aleatoria: Una variable aleatoria se dice *continua* si existe una función $f_x : \mathbb{R} \rightarrow [0, \infty)$, llamada densidad de probabilidad de X , tal que $\forall A \subseteq \mathbb{R}$:

$$\mathbb{P}(x \in A) = \int_A f_X(x) dx$$

Una variable aleatoria se dice *discreta* si $\forall A \subseteq \mathbb{R}$:

$$\mathbb{P}(x \in A) = \sum_{k \in A \cap \mathbb{R}_X} p_X(k)$$

Donde:

$$p_X(k) = \mathbb{P}(X = k)$$

Def Función de distribución acumulada: Dada X variable aleatoria continua, definimos su función de distribución acumulada ($F_X(x) = \mathbb{P}(X \leq x)$) como:

$$F_X(x) = \int_{-\infty}^x f_X(z)dz \Rightarrow f_X(x) = \frac{dF_X(x)}{dx}$$

Caso discreto:

$$F(b) = \sum_{x_i \leq b} p(x_i)$$

Def Esperanza La esperanza es como el centro de masa de la probabilidad y se define:

$$\mathbb{E}(x) = \begin{cases} \sum_{k \in R_X} k \cdot p_X(k) & \text{con } k_{va} \text{ discreta} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{con } x_{va} \text{ continua} \end{cases}$$

Para todos los valores que puede tomar x tq $p(x) \geq 0$

$$(1) \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

$$(2) \mathbb{E}(1_A) = \mathbb{P}(A)$$

$$(3) \mathbb{E}(\alpha X + \beta) = \alpha \mathbb{E}(X) + \beta$$

$$(4) \mathbb{E}(\alpha) = \alpha$$

$$(5) \text{ Si } X \perp\!\!\!\perp Y \Rightarrow \mathbb{E}(XY) = E(X)E(Y)$$

Def Covarianza Sean X e Y va. Supongamos que la distribución conjunta existe y que $Var(X), Var(Y)$ existen.

$$\begin{aligned} Cov(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(X, Y) - E(X)E(Y) \end{aligned}$$

Obs: Si X, Y son independientes, entonces $Cov(X, Y) = 0$, el recíproco, en general, es falso.

Def Varianza Representa como el momento de inercia y se define como:

$$\begin{aligned} Var(x) &= E((X - E(X))^2) \\ &= E(X^2) - E(X)^2 \end{aligned}$$

Además:

$$Var(X, Y) = Var(X) + Var(Y) + 2Cov(X, Y)$$

$$(1) \text{ Desviación estándar: } \sigma(X) = \sqrt{Var(X)}$$

$$(2) Var(\alpha X + \beta) = \alpha^2 \cdot Var(X)$$

$$(3) X \perp\!\!\!\perp Y \Rightarrow Var(X, Y) = Var(X) + Var(Y)$$

1.5. Tipos de V.A Discretas

Def V.A Bernoulli Esta variable aleatoria se caracteriza por tener dos posibilidades, cuando $X = 1$ (éxito) o $X = 0$ (fracaso). Sea p la probabilidad de éxito y $(1 - p)$ la de fracaso, así:

$$p(0) = \mathbb{P}(X = 0) = 1 - p$$

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

$$\blacksquare \mathbb{E}(X) = p$$

$$\blacksquare Var(X) = p(1 - p)$$

$$\blacksquare \begin{aligned} F(0) &= \mathbb{P}(X \leq 0) = 1 - p \\ F(1) &= \mathbb{P}(X \leq 1) = 1 \end{aligned}$$

Def V.A Binomial Esta variable aleatoria busca i éxitos en n . Escribimos $X \sim Bin(n, p)$:

$$p(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad i = 0, \dots, n$$

$$\blacksquare \mathbb{E}(X) = np$$

$$\blacksquare Var(X) = np(1 - p)$$

$$\blacksquare F(a) = \mathbb{P}(X \leq a) = \sum_a^k \binom{n}{k} p^k (1 - p)^{n-k}$$

Def V.A poisson Es una variable aleatoria que toma los valores $X = 0, 1, 2, \dots$ con parámetro λ se escribe $X \sim poisson(\lambda)$

$$p_X(k) = \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad i = 0, \dots, n$$

$$\blacksquare \mathbb{E}(X) = \lambda$$

$$\blacksquare Var(X) = \lambda$$

$$\blacksquare \mathbb{P}(X \leq k) = \frac{\lambda}{k} \mathbb{P}(X = i - 1) \text{ Importante para calcular la función distribución.}$$

$$\blacksquare \text{ Recordar que } e^\lambda = \sum_{j=1}^{\infty} \frac{\lambda^j}{j!}$$

Def V.A Geométrica: modela la probabilidad de obtener un primer éxito en el intento k , con probabilidad de éxito $p \in (0, 1)$. $X \sim Geom(p)$:

$$p_X(k) = (1 - p)^{k-1} p$$

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

$$\blacksquare Var(X) = \frac{1 - p}{p^2}$$

$$\blacksquare F(a) = \mathbb{P}(X \leq a) = \sum_{i \leq a} (1 - p)^{k-1} p = 1 - (1 - p)^a$$

1.6. Tipos de V.A Continuas

Def V.A Uniforme Una variable aleatoria uniforme en un intervalo (a,b) tiene como función densidad a:

$$f(x) = \begin{cases} \frac{1}{a-b} & \text{si } a \leq x \leq b \\ 0 & \text{en otro caso.} \end{cases}$$

- $\mathbb{E}(X) = \frac{a+b}{2}$
- $Var(X) = \frac{(b-a)^2}{12}$
- $F(j) = \int_{-\infty}^j f(x)dx = \frac{j-a}{b-a}$

Def V.A Normal X sigue una variable aleatoria normal de media μ y varianza σ^2 ($X \sim \mathcal{N}(\mu, \sigma^2)$). Su función de densidad esta dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$$

- $\mathbb{E}(X) = \mu$
- $Var(X) = \sigma^2$
- $\phi(-t) = 1 - \phi(t)$
- $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Def V.A Exponencial X sigue una variable aleatoria exponencial de parámetro $\lambda > 0$ ($X \sim Exp(\lambda)$). Su función densidad esta definida como:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{en otro caso.} \end{cases}$$

- $\mathbb{E}(X) = \frac{1}{\lambda}$
- $Var(X) = \frac{1}{\lambda^2}$
- $\mathbb{P}(X > s+t \mid X > t) = \mathbb{P}(X > s)$. La exponencial cumple la propiedad de perdida de memoria.

1.7. F.G.M. y Proceso de Poisson

Def Función generadora de momentos (f.g.m): Sea X v.a., se define su f.g.m como:

$$M_x : \mathbb{R} \rightarrow [0, \infty] \\ t \rightarrow M_X(t) = \mathbb{E}(e^{tx})$$

Prop Se define el k -ésimo momento de una va como:

$$\frac{d^k M_X(0)}{dt^k} = \mathbb{E}(X^k) \quad \forall k = \{1, 2, \dots\}$$

Prop Sea $X \perp\!\!\!\perp Y$ v.a.'s $\Rightarrow M_{X+Y}(t) = M_X(t)M_Y(t)$

Teo X, Y v.a.'s, con M_X, M_Y finitas en un intervalo en torno a 0, entonces:

X, Y tienen la misma distribución $\Leftrightarrow M_X = M_Y \quad \forall t \in \mathbb{R}$

Def Función Gamma: Dado $\theta > 0$:

$$\Gamma(\theta) = \int_0^\infty e^{-z} z^{\theta-1} dz$$

Obs: Si integramos por partes, resulta:

$$\Gamma(\theta) = (\theta - 1)\Gamma(\theta - 1)$$

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

Def Distribución Gamma: $\theta > 0, \lambda > 0, X \sim Gamma(\theta, \lambda)$

$$f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{\theta-1}}{\Gamma(\theta)} \mathbb{1}_{[0, \infty]}(x)$$

Prop X_1, \dots, X_n v.a independientes con $X_i \sim exp(\lambda)$, entonces:

$$X = X_1, \dots, X_n \Rightarrow X \sim Gamma(n, \lambda)$$

Def Proceso de Poisson: Modela la ocurrencia de sucesos que ocurren uno después del otro "sin condición". Ejemplo: llegada de clientes a un banco.

Denotamos $\forall i = 1, 2, \dots$:

- $T_i :=$ Tiempo que transcurre entre la llegada del cliente $i = 1$ y el i -ésimo
- $S_k :=$ Instante en que llega el cliente k -ésimo $= \sum_{i=1}^k T_i$ ($S_0 = 0$)

Supuesto del modelo:

- Los T_i son **independientes**, $T_i \sim exp(\lambda)$, donde es la tasa de llegada de clientes
- Lo anterior implica que $S_k \sim Gamma(k, \lambda)$

Para $t \geq 0$, definimos la va:

- $N_t :=$ La cantidad de clientes que han llegado hasta el instante $t = |\{k \mid S_k \leq t\}|$

Visto como una función (aleatoria) de t , la colección $(N_t)_{t \geq 0}$ se llama proceso de Poisson.

Prop $\forall t \geq 0 \quad N_t \sim Poisson$:

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

Prop $(N_t)_{t \geq 0}, (M_t)_{t \geq 0}$ procesos de Poisson independientes con tasa λ y μ respectivamente, entonces: $(N_t + M_t)_{t \geq 0}$ es un proceso de poisson con tasa $\lambda + \mu$

1.8. Vectores Aleatorios

Un vector aleatorio es una función $\vec{X} : \Omega \rightarrow \mathbb{R}^n$, donde (Ω, \mathbb{P}) es un espacio de probabilidad, implícito. Se escribe como $\vec{X} = (X_1, \dots, X_n)$ cuyas componentes son v.a. independientes entre ellas.

Def Distribución Conjunta Sea (X, Y) v.a. Entonces la distribución del \vec{v}_a es:

(I) **Discreta**

$$\begin{aligned} p_{X,Y}(n, v) &= \mathbb{P}(X = n, Y = v) \\ &= \mathbb{P}((X = n) \cap (Y = v)) \end{aligned}$$

(II) **Continua**

$$\mathbb{P}((X, Y) \in A) = \int_A \int f_{X,Y}(u, v) du dv$$

Def Distribución Acumulada (Conjunta) Dado un vector aleatorio (X, Y) conjuntamente continuo, es posible ver que: $F = F_{X_1, \dots, X_n}$ de un vector aleatorio (X_1, \dots, X_n) es definido por:

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) \\ &= \mathbb{P}((X, Y) \in \{(-\infty, x] \times \mathbb{R}\}) \\ &= \int_{-\infty}^x \int_{-\infty}^{\infty} f_{X,Y}(u, v) dy du \end{aligned}$$

Def Densidades Marginales: Al aplicar el TFC en la formula anterior $\left(\frac{dF_X(x)}{dx}\right) = f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$

$$\begin{aligned} \text{(I)} \quad p_X(u) &= \sum_{v \in \text{Rec}(Y)} p_{X,Y}(u, v) \\ p_Y(v) &= \sum_{u \in \text{Rec}(X)} p_{X,Y}(u, v) \end{aligned}$$

$$\begin{aligned} \text{(II)} \quad f_X(u) &= \int_{-\infty}^{\infty} f_{X,Y}(u, v) dv \\ f_Y(v) &= \int_{-\infty}^{\infty} f_{X,Y}(u, v) du \end{aligned}$$

Prop Sea (X, Y) v.a conjuntamente continuo, y sea $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, entonces:

$$\mathbb{E}(g(x, y)) = \int \int_{\mathbb{R}^2} g(x, y) f_{X,Y}(x, y) dx dy$$

Def Independencia Las variables aleatorias X_1, \dots, X_n se dicen independientes si su distribución acumulada conjunta se factoriza como producto de sus distribuciones marginales.

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot \dots \cdot F_{X_n}(x_n)$$

Def Convolución Sean $g, h : \mathbb{R} \rightarrow \mathbb{R}$, su convolución es la función:

$$\begin{aligned} g * h : \mathbb{R} &\rightarrow \mathbb{R} \\ z \rightarrow (g * h)(z) &= \int_{-\infty}^{\infty} g(x) h(z - x) dx \end{aligned}$$

Prop : Dados $X \perp\!\!\!\perp Y$ v.a continuas, entonces:

$$\begin{aligned} f_{X+Y} &= f_X * f_Y \\ \mathcal{L}[g * f](s) &= \mathcal{L}[g](s) \cdot \mathcal{L}[f](s) \end{aligned}$$

Teo Método del Cambio de variable Sean X, Y v.a con densidad conjunta $f_{X,Y}(x, y)$ conocida. Sea $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ g suave, invertible y g^{-1} suave, entonces $\forall (u, v) \in \mathbb{R}^2$:

$$\begin{aligned} f_{U,V}(u, v) &= \frac{f_{X,Y}(x, y)}{|\det(J_g(x, y))|} \\ &= f_{X,Y}(g^{-1}(u, v)) \cdot |\det(J_{g^{-1}}(u, v))| (*) \end{aligned}$$

Recordar: Jacobiano, sea:

$$h(x, y) = \begin{pmatrix} h_1(x, y) \\ h_2(x, y) \end{pmatrix}$$

Entonces:

$$J_h(x, y) = \begin{pmatrix} \frac{\partial h_1(x, y)}{\partial x} & \frac{\partial h_1(x, y)}{\partial y} \\ \frac{\partial h_2(x, y)}{\partial x} & \frac{\partial h_2(x, y)}{\partial y} \end{pmatrix}$$

Pasos para resolver problema estándar con Cambio de Variable:

- (I) Dados (X, Y) v.a independientes con distribuciones iguales o distintas cada una, nos solicitan encontrar la densidad conjunta de dos v.a (U, V) que dependen de X e Y de distinta manera
- (II) Definimos $g(X, Y) = (U, V) = (f(X), g(Y))$
- (III) Luego despejamos X e Y , y nos definimos $g^{-1}(U, V) = (X, Y) = (g_1^{-1}, g_2^{-1})$.
- (IV) Calculamos su Jacobiano:

$$J_{g^{-1}}(u, v) = \begin{pmatrix} \frac{\partial g_1^{-1}(u, v)}{\partial u} & \frac{\partial g_1^{-1}(u, v)}{\partial v} \\ \frac{\partial g_2^{-1}(u, v)}{\partial u} & \frac{\partial g_2^{-1}(u, v)}{\partial v} \end{pmatrix}$$

- (v) Calculamos el determinante del jacobiano, y su valor absoluto
- (vi) Finalmente obtenemos la densidad conjunta ocupando (*) como:

$$f_{U,V}(u, v) = f_{X,Y}(g^{-1}(u, v)) \cdot |\det(J_{g^{-1}}(u, v))|$$

(vii) *Obs:* Notar que estos pasos son análogos cuando se analiza una sola v.a, considerando un vector de 1×1

Def Correlación Si X e Y son v.a de modo que la distribución conjunta existe y que $\text{Var}(X), \text{Var}(Y)$ existen y no nulas, entonces:

$$\rho_{X,Y}(x, y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(x)\text{Var}(y)}}$$

Def Distribución condicional

$$f_{X|Y=y} = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

análogamente se obtiene la fórmula de Bayes:

$$p_{Y|X=x(y)} = \frac{p_{X|Y=y}(x) \cdot p_Y(y)}{p_X(x)}$$

Def Esperanza Condicional

- $\mathbb{E}(X|Y=y) = \sum_{x \in R_X} x \cdot p_{X|Y=y}(x)$
- $\mathbb{E}(X|Y=y) = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y}(x) dx$
- $\mathbb{E}(g(x)|Y=y) = \int g(x) \cdot f_{X|Y=y}(x) dx$
- RPT continua para esperanzas:

$$\mathbb{E}(x) = \int_{-\infty}^{\infty} \mathbb{E}(X|Y=y) f_Y(y) dy$$

- Con $\mathbb{P}(A|Y=y) = \mathbb{E}(1|Y=y)$

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A|Y=y) f_Y(y) dy$$

Teo $\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$ **Def** Varianza Condicional

- $V(Y|X=x) = \sum_{y \in R_Y} (y - \mathbb{E}(Y|X=x))^2 p_{Y|X=x}(y)$
- $V(Y|X=x) = \int_{-\infty}^{\infty} (y - \mathbb{E}(Y|X=x))^2 f_{Y|X=x}(y) dy$

Prop Sea (X, Y) v.a.:

- (a) $\mathbb{E}(y) = \mathbb{E}(\mathbb{E}(Y|X))$
- (b) $Var(y) = \mathbb{E}(Var(Y|X)) + Var(\mathbb{E}(Y|X))$

Prop Desigualdad de Markov Sea X v.a. que toma valores (no negativos). Entonces $\forall a > 0$:

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

Prop Desigualdad de Chebyshev Sea X v.a. con $\mu = \mathbb{E}(X)$ y $\sigma^2 = Var(X)$ finitos, entonces $\forall b > 0$

$$\mathbb{P}(|X - \mu| \geq b) \leq \frac{\sigma^2}{b^2}$$

1.9. Convergencia

Def Sea $\{Y_n\}$ una sucesión de v.a. y Y una v.a. Entonces:

- (I) Decimos que Y_n converge a Y **casi seguramente**, anotado $Y_n \xrightarrow[n]{C.S.} Y$, si:

$$\mathbb{P}(Y_n \rightarrow Y) = 1$$

- (II) Decimos que Y_n converge a Y en **probabilidad**, anotado $Y_n \xrightarrow[n]{\mathbb{P}} Y$, si:

$$\mathbb{P}(|Y_n - Y| \geq \epsilon) \xrightarrow[n]{} 0 \quad \forall \epsilon > 0$$

- (III) Decimos que Y_n converge a Y en **distribución, o en ley** anotado así $Y_n \xrightarrow[n]{\mathcal{L}} Y$ si:

$$F_{Y_n}(y) \xrightarrow[n]{\mathcal{L}} F_Y(y) \quad \forall y \in \mathbb{R}$$

en que $F_Y(\cdot)$ es continua

Prop

$$Y_n \xrightarrow[n]{C.S.} Y \Rightarrow Y_n \xrightarrow[n]{\mathbb{P}} Y \Rightarrow Y_n \xrightarrow[n]{\mathcal{L}} Y$$

Def Ley Fuerte de los Grandes Números Sea $\{X_i\}_{i \in [1, \dots, n]}$ una sucesión de v.a. i.i.d., de modo que $\mathbb{E}(X_i) = \mu \neq \infty, \forall n \in \mathbb{N}$, con $\mu \in \mathbb{R}$:

$$\bar{X}_n \xrightarrow[n]{C.S.} \mu$$

Teo Teorema del limite central: Sean X_1, X_2, \dots secuencia de v.a. i.i.d. con $\mu = \mathbb{E}(X_i)$, $\sigma^2 = Var(X_i)$ y sea: $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Suponiendo que μ y σ son finitos, entonces:

$$Z_n = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n]{\mathcal{L}} \mathcal{N}(0, 1) \quad (*)$$

- Este limite no depende del tipo de distribución de los X_i , es siempre igual para todas.
- En la práctica (*) significa que las probabilidades de Z_n se aproximan a la de $Z \sim \mathcal{N}(0, 1)$, ie.
 $\mathbb{P}(Z_n \in A) \xrightarrow[n]{} \mathbb{P}(Z \in A) \quad \forall A \subseteq \mathbb{R}$ “razonable”

2. Estadística

Def Muestra aleatoria simple: Una m.a.s es una colección X_1, \dots, X_n de variables aleatorias independientes o idénticamente distribuidas.

Obs: Los X_1, \dots, X_n corresponden a los datos *antes de saber su valor*. Si nos referimos a los valores que toman anotaremos x_1, \dots, x_n . Además supondremos que la distribución de la cual provienen los datos poseen un parámetro $\theta \in \mathbb{R}$ el cual se desea aproximar.

2.1. Estimadores

Def Estadístico: Es una función de la muestra:

$$e = e(X_1, \dots, X_n)$$

Def Estimador Un estimador $\hat{\theta}$ para el parámetro θ es un estadístico:

$$\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$$

que se usa para aproximar θ

Def Estimador insesgado: Un estimador se dice *insesgado* si:

$$\mathbb{E}(\hat{\theta}) = \theta$$

Prop La cantidad $[\mathbb{E}(\hat{\theta}) - \theta]$ se llama **sesgo**

Prop Sean $\mu = \mathbb{E}(X_i)$, $\sigma^2 = \text{Var}(X_i)$, son estimadores insesgados:

- $\bar{X} = \sum_{i=1}^n$ para μ
- $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ para σ^2

Obs:

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

Def Error cuadrático medio ($\hat{\theta}$):

$$ECM(\hat{\theta}) = \mathbb{E}((\hat{\theta} - \theta)^2)$$

Teo $ECM(\hat{\theta}) = \text{Var}(\hat{\theta}) + \text{sesgo}(\hat{\theta})^2$

Def Consistencia: Se dice que $\hat{\theta}$ es *consistente* si $\hat{\theta} \xrightarrow[n]{\mathbb{P}} \theta$, es decir, si:

$$\mathbb{P}(|\hat{\theta} - \theta| > \epsilon) \xrightarrow[n \rightarrow \infty]{} 0$$

Prop Si $\hat{\theta}$ es *insesgado* y $\text{Var}(\hat{\theta}) \xrightarrow[n \rightarrow \infty]{} 0$ entonces $\hat{\theta}$ es *consistente*

Def Función de verosimilitud: Suponiendo que la distribución de los X_1, \dots, X_n posee densidad $f = f(x, \theta) \forall x \in \mathbb{R}$, luego definimos la FV como:

$$L(X_1, \dots, X_n; \theta) := \prod_{i=1}^k f(x_i; \theta)$$

Donde la segunda expresión es la densidad conjunta del vector $\vec{X} = (X_1, \dots, X_n)$ (con va. iid)

Def Método de máxima verosimilitud: método que permite encontrar estimadores. Propone estimar θ mediante lo siguiente: si x_1, \dots, x_n son los números específicos obtenidos en la muestra, tenemos:

$$\hat{\theta}_{MV} = \underset{\theta \in \mathbb{R}}{\operatorname{argmax}} L((x_1, \dots, x_n); \theta)$$

Obs: $\hat{\theta}_{MV} = \underset{\theta \in \mathbb{R}}{\operatorname{argmax}} \log L((x_1, \dots, x_n); \theta)$ Obs: Si hay p

parámetros desconocidos ($\vec{\theta} = (\theta_1, \dots, \theta_p)$) el método es el mismo: $\hat{\theta}_{MV} = \underset{\theta \in \mathbb{R}^p}{\operatorname{argmax}} \log L((x_1, \dots, x_n); \theta)$

Obs: Si **L es suave**, para maximizar L c/r a $\vec{\theta}$ se buscan puntos críticos:

$$\nabla_{\vec{\theta}} L(x_1, \dots, x_n; \theta) = 0$$

Def Método de los momentos: Se definen $\forall k$

$$m_k = \mathbb{E}(X_1^k) \quad \text{k-ésimo momento real}$$

$$\hat{m}_k = \frac{1}{n} \sum_{i=1}^n X_i^k \quad \text{k-ésimo momento muestral}$$

donde X_1, \dots, X_n en la m.a.s. Suponemos que hay p parámetros desconocidos $\theta_1, \dots, \theta_p$. El método de los momentos propone imponer: $m_i = \hat{m}_i$. Esto genera p ecuaciones para las p incógnitas. La solución corresponde a los estimadores del método de los momentos.

Def Intervalo de confianza: Deseamos proveer de un intervalo $[\hat{\theta}_L, \hat{\theta}_U]$ (dependientes de la muestra), llamado *intervalo de confianza*, tal que el parámetro desconocido θ esté en él con alta probabilidad, es decir:

$$\mathbb{P}(\theta \in [\hat{\theta}_L, \hat{\theta}_U]) = 1 - \alpha$$

Donde $\alpha = 5\%, \alpha = 1\%$ son usuales.

Método gral para encontrar i.d.c para θ :

- Trabajar con un estadístico U cumpliendo que el estadístico involucre el parámetro desconocido θ pero su distribución no dependa de θ y sea desconocida.
- Imponer el nivel deseado a un intervalo para U (posiblemente con una condición de simetría).
- De una tabla se obtienen los extremos del intervalo. Finalmente se despeja θ .

Teo Sea X_1, \dots, X_n m.a.s. con distribución común $\mathcal{N}(\mu, \sigma)$ con μ y σ^2 desconocidos. Para encontrar el i.d.c para μ al nivel $(1 - \alpha)$, entonces:

$$\mu \in \left[\bar{x} - c \frac{\sigma}{\sqrt{n}}, \bar{x} + c \frac{\sigma}{\sqrt{n}} \right]$$

Donde c cumple $\mathbb{P}(\mathcal{N}(0, 1) > c) = \frac{\alpha}{2}$

Teo Si la distribución común es genérica (no necesariamente normal) y se busca un i.d.c para $\mu = \mathbb{E}(X_1)$ con $\sigma^2 = \text{Var}(X_1)$ conocida, el estadístico:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

será aproximadamente $\mathcal{N}(0, 1)$ para n grande (por TCL). Este método entrega un i.d.c aproximado para μ

Def Variable aleatoria T-student Una va. T se dice *T-student* con n grados de libertad si su densidad es:

$$f_T(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \quad \forall x$$

La anotamos como $T \sim t_n$

Obs: Notar qe para n grande, $f_t(x) = ce^{-\left(\frac{x^2}{2}\right)}$

Prop Sea X_1, \dots, X_n m.a.s de una $\mathcal{N}(\mu, \sigma^2)$, entonces:

$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

Donde $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Para encontrar el *Intervalo de confianza* para μ con μ y σ desconocidos, a un nivel $(1 - \alpha)$ y n genéricos tenemos la siguiente fórmula:

$$\mathbb{P}(t_{n-1} > c) = \frac{\alpha}{2}$$

Donde c se obtiene de la tabla *t-student* en la posición $[t_{n-1}, \frac{\alpha}{2}]$

Estimación de proporción

Prop Si $Y_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} Y$ y $A_n \xrightarrow[n \rightarrow \infty]{\text{C.S.}} 1$, entonces:

$$A_n Y_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} Y$$

Estimación de idc para una Bernoulli

Dada una m.a.s X_1, \dots, X_n de una *Bernoulli*(p) con p desconocido queremos encontrar un i.d.c para p .

Def Estimador:

$$\hat{p} = \frac{\sum 1's}{n} = \bar{X}$$

Por TCL y Prop Anterior, definimos el estadístico:

$$Z = \frac{\hat{p} - p}{\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}} = \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Donde para n grande, ocupamos Z para encontrar el intervalo de confianza para p , que es:

$$\left[\hat{p} - c \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}, \hat{p} + c \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right] = [b_1, b_2]$$

donde c cumple $\mathbb{P}(\mathcal{N}(0, 1) > c) = \frac{\alpha}{2}$

Si los extremos se salen del intervalo $[0, 1]$ pueden truncarse, obteniendo:

$$\mathbb{P}(p \in \max(b_1, 0), \min(b_2, 1)) \approx (1 - \alpha)$$

Estimación de idc para una varianza

Dado una m.a.s X_1, \dots, X_n con distribución común $\mathcal{N}(\mu, \sigma^2)$ con μ, σ^2 desconocidos, queremos encontrar un i.d.c para σ al nivel $(1 - \alpha)$

Def Chi-cuadrado: Sean Z_1, \dots, Z_n i.i.d $\mathcal{N}(0, 1)$. La va. $U = \sum_{i=1}^n Z_i^2$ se llama chi-cuadrado con n -grados de libertad, anotado $U \sim \chi_n^2$. Se puede mostrar que:

$$f_n(x) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{1}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \mathbb{1}_{[0, \infty]}(x)$$

Prop Sea X_1, \dots, X_n de $\mathcal{N}(\mu, \sigma^2)$, con $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, entonces:

$$(n-1) \frac{s^2}{\sigma^2} \sim \chi_{n-1}^2$$

Def Estadístico:

$$U = (n-1) \frac{s^2}{\sigma^2} \sim \chi_{n-1}^2$$

Tabla Chi-cuadrado: Conocido $(n-1)$ y α podemos obtener (a, b) de:

$$\begin{aligned} \mathbb{P}(\chi_n^2 > a) &= 1 - \frac{\alpha}{2} \\ \mathbb{P}(\chi_n^2 > b) &= \frac{\alpha}{2} \end{aligned}$$

Con esto sabemos que:

$$\begin{aligned} 1 - \alpha &= \mathbb{P}\left(a \leq (n-1) \frac{s^2}{\sigma^2} \leq b\right) \\ 1 - \alpha &= \mathbb{P}\left((n-1) \frac{s^2}{b} \leq \sigma^2 \leq (n-1) \frac{s^2}{a}\right) \end{aligned}$$

Por tanto el i.d.c para la varianza es:

$$\left[(n-1) \frac{s^2}{b}, (n-1) \frac{s^2}{a} \right]$$

Test de hipótesis

Dado un parámetro desconocido θ , nos interesa comparar dos afirmaciones sobre θ y ver en base a los datos si es razonable descartar una en favor de la otra.

- H_0 : *Hipotesis nula*. Es la afirmación de base
- H_1 : *Hipotesis alternativa*. sobre la cual se constrasta H_0

	Se rechaza H_0	Se rechaza H_1
Si H_0 es cierta	Error tipo 1	No hay error
Si H_1 es cierta	No hay error	Error tipo 2

Tipicamente se controla:

- $\alpha = \mathbb{P}(\text{error tipo I})$
- $\alpha = \mathbb{P}(\text{Rechaza } H_0 \mid H_0 \text{ es cierta})$
- $\alpha = \mathbb{P}(\text{condenar a un inocente})$

Def Hipótesis simple: una hipótesis se dice simple si ella determina completamente la distribución de los datos. En otro caso se dice **compuesta**.

Tipicamente una hipótesis se dice **simple** cuando es de la forma " $\theta = \theta_0$ " para $\theta_0 \in \mathbb{R}$ conocido, donde θ es el único parámetro desconocido. Por ejemplo, si la muestra tiene distribución común $\exp(\lambda)$:

- $H: \lambda = 2$ es simple
- $\tilde{H}: \lambda = 2$ es simple

Siempre trabajaremos con hipótesis nula H_0 de tipo simple.

Def Potencia: La potencia de un *test* con H_1 simple es:

$$\begin{aligned} \text{Potencia} &= \mathbb{P}(\text{rechaza } H_0 \mid H_1) \\ &= 1 - \mathbb{P}(\text{No rechaza } H_0 \mid H_1) \\ &= 1 - \mathbb{P}(\text{Error tipo II}) \\ &= 1 - \beta \end{aligned}$$

Tipicamente, ante dos test con el mismo α se prefiere aquel con menor β , es decir, el test con mayor potencia.

Def Región de rechazo: Dado a que un *test* es un criterio para decidir cuando rechaza, decimos que un *test* corresponde a un $R \subseteq \mathbb{R}^n$ llamada **Región de rechazo**, tal que el test se rechaza sí y sólo sí el vector de datos $(x_1, \dots, x_n) = \vec{x}$ pertenece a R . Es de la forma:

$$R = \{\vec{x} \in \mathbb{R}^n : \mathcal{U}(\vec{x} > c)\}$$

donde $U = \mathcal{U}(X_1, \dots, X_n)$ es un estadístico adecuado y c es una constante que se adecua para lograr el nivel deseado.

Teo Lema de Neyman - Pearson: Sea una m.a.s X_1, \dots, X_n cuya distribución común tiene un único parámetro desconocido θ y sean:

- $H_0 : \theta = \theta_0$
- $H_1 : \theta = \theta_1$

Entonces dado α fijo, el test con la potencia máxima tiene región de rechazo de la forma:

$$R = \left\{ \vec{x} \in \mathbb{R}^n : \frac{\mathcal{L}(\vec{x}; \theta_0)}{\mathcal{L}(\vec{x}; \theta_1)} \leq cte \right\}$$

donde \mathcal{L} es de la función de verosimilitud:

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

donde $f(x_i; \theta)$ es la densidad común de los x_i

Ejemplo útil

Sea una m.a.s X_1, \dots, X_n de una $N(\mu, \sigma^2)$ con σ^2 conocido. Sean:

- $H_0 : \mu = \mu_0$
- $H_1 : \mu = \mu_1$

con $\mu_0 < \mu_1$ Encontrar el test más potente a nivel α

Sol:

1. Definir función de verosimilitud:

$$\begin{aligned} \Leftrightarrow L(\vec{x}; \mu) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \\ \Leftrightarrow L(\vec{x}; \mu) &= (\sqrt{2\pi}\sigma)^{-n} e^{-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2} \end{aligned}$$

2. Desarrollamos la condición sobre la región para que el test sea potencia máxima:

$$\begin{aligned} \Leftrightarrow \frac{\mathcal{L}(\vec{x}; \theta_0)}{\mathcal{L}(\vec{x}; \theta_1)} &\leq cte \\ \Leftrightarrow &\dots \\ \Leftrightarrow 2(\mu_0 - \mu_1)n\bar{x} &\leq \widetilde{cte} \quad (*) \\ \Leftrightarrow \bar{x} &\geq \frac{\widetilde{cte}}{2(\mu_0 - \mu_1)n} \\ \Leftrightarrow \bar{x} &\geq \widetilde{\widetilde{cte}} \end{aligned}$$

3. La forma de la región de rechazo es:

$$\begin{aligned} R &= \{\vec{x} \in \mathbb{R}^n \mid \bar{x} \geq a\} \\ &= \{\vec{x} \in \mathbb{R} \mid \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \geq c\} \quad (**) \end{aligned}$$

4. Lo cual permite despejar c imponiendo que:

$$\mathbb{P}(\vec{x} \in \mathbb{R} \mid H_0) = \alpha$$

y sabiendo que $\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$

Obs 1: La forma de $(**)$ es la misma para todo $\mu_0 < \mu_1$ En tal caso decimos que el test es **uniformemente más potente** para el caso $\mu_0 < \mu_1$

Obs 2: Si $\mu_0 > \mu_1$ la región de rechazo cambia a:

$$R = \left\{ \vec{x} \in \mathbb{R}^n \mid \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq c \right\}$$

Cuadro 1: Resumen Distribuciones

nombre	parámetros	notacion	tipo	soporte	distribucion	esperanza	varianza	f.g.m
bernoulli	$p \in (0, 1)$	$X \sim \text{Bernoulli}(p)$	discreta	$\{0, 1\}$	$p_X(0) = 1 - p$ $p_X(1) = p$	p	$p(1 - p)$	$1 - p + pe^t$
binomial	$n \in \mathbb{N}^*, p \in (0, 1)$	$X \sim \text{bin}(n, p)$	discreta	$\{0, 1, \dots, n\}$	$p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$	np	$np(1 - p)$	$(1 - p + pe^t)^n$
geométrica	$p \in (0, 1)$	$X \sim \text{geom}(p)$	discreta	$\{1, 2, 3, \dots\}$	$p_X(k) = (1 - p)^{k-1} p$	$\frac{1}{p}$	$\frac{1 - p}{p^2}$	$\frac{pe^t}{1 - (1 - p)e^t}$
binomial negativa	$r \in \mathbb{N}^*, p \in (0, 1)$	$X \sim \text{BN}(r, p)$	discreta	$\{r, r + 1, \dots\}$	$p_X(k) = \binom{k-1}{r-1} (1 - p)^{k-r} p^r$	$\frac{r}{p}$	$\frac{r(1 - p)}{p^2}$	$\left(\frac{pe^t}{1 - (1 - p)e^t} \right)^r$
Poisson	$\lambda > 0$	$X \sim \text{Poisson}(\lambda)$	discreta	$\{0, 1, 2, \dots\}$	$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ	$e^{\lambda(e^t - 1)}$
uniforme	$a, b \in \mathbb{R}, a < b$	$X \sim \text{unif}(a, b)$	continua	$[a, b]$	$f_X(x) = \frac{1}{b - a} \mathbb{1}_{[a, b]}(x)$	$\frac{a + b}{2}$	$\frac{(b - a)^2}{12}$	$\frac{e^{tb} - e^{ta}}{t(b - a)}$
exponencial	$\lambda > 0$	$X \sim \text{exp}(\lambda)$	continua	$[0, \infty)$	$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0, \infty)}(x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t} \quad \forall t < \lambda$
normal	$\mu \in \mathbb{R}, \sigma > 0$	$X \sim \mathcal{N}(\mu, \sigma^2)$	continua	\mathbb{R}	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu t + \frac{1}{2}\sigma^2 t^2}$
gamma	$\theta > 0, \lambda > 0$	$X \sim \text{gamma}(\theta, \lambda)$	continua	$[0, \infty)$	$f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{\theta-1}}{\Gamma(\theta)} \mathbb{1}_{[0, \infty)}(x)$	$\frac{\theta}{\lambda}$	$\frac{\theta}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - t} \right)^\theta \quad \forall t < \lambda$