Apunte de probabilidad (LM)

Tiziano Brunelli, Fran Arredondo y Fernando Ahumada

Definición: Sea Ω un conjunto. Una colección no vacía $\mathscr{A}$ de subconjuntos de Ω es una σ-Álgebra de subconjuntos de Ω si cumplen las siguientes dos propiedades:

1. si $A \in \mathscr{A} \implies A^c \in \mathscr{A}$
2. sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de elementos de $\mathscr{A}$, entonces $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$ y $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{A}$

Propiedad: Sea $\mathscr{A}$ una σ-Álgebra de subconjuntos de Ω, entonces $\Omega \in \mathscr{A}$ y $\emptyset \in \mathscr{A}$.

Definición: Sea $\mathscr{A}$ una $\sigma -$ Álgebra de subconjuntos de $Omega$. Una medida de probabilidad P es una función $P : \mathscr{A} \to \mathbb{R}$ con las siguientes propiedades:

1. $P(\Omega) = 1$
2. $P(A) \geq 0 \quad \forall A \in \mathscr{A}$
3. sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de elementos de $\mathscr{A}$,disjuntos dos a dos, entonces:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Propiedades:

1. $\forall A \in \mathscr{A} : P(A) \leq 1$.
2. $P(\emptyset) = 0$
3. $P(A^c) = 1 - P(A)$

Definición: Un Espacio de probabilidad, representado por $(\Omega, \mathscr{A}, P)$, es una tupla con un conjunto Ω, una $\sigma -$Álgebra $\mathscr{A}$ de subconjuntos de Ω y una medida de probabilidad P definida sobre $\mathscr{A}$.

Definición: En contexto de un Espacio de probabilidad, a un elemento A de una $\sigma -$ Álgebra $\mathscr{A}$ se le llama evento.

Teorema de Inclusión-Exclusión: Sean $A_1, ..., A_n$ conjuntos finitos entonces:

$$\left|\bigcup_{k=1}^{n} A_k\right| = \sum_{k=1}^{n}(-1)^{k+1}s_k$$

donde:

$$s_k = \sum_{1\le i_1<...<i_k\le n} \left|\bigcap_{j=1}^{k} A_{i_j}\right|$$

O, como me gusta escribirlo:

$$\left|\bigcup_{k=1}^{n} A_k\right| = \sum_{\substack{\sigma\subseteq\{1,...,n\}\\ \sigma\neq\emptyset}} (-1)^{|\sigma|+1} \left|\bigcap_{k\in\sigma} A_k\right|$$

Análogamente, si $A_1, \ldots, A_n$ son eventos de un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$ (nótese que ahora no se dijo nada sobre la finitud de los A_k), entonces:

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n}(-1)^{k+1}S_k$$

donde:

$$S_k = \sum_{1\le i_1<...<i_k\le n} P\left(\bigcap_{j=1}^{k} A_{i_j}\right)$$

Definición: Sean A y B dos eventos de un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$, con $P(A) > 0$, entonces se define la probabilidad de B dado A, $P(B \mid A)$, como:

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

$P(B \mid A)$ queda indefinida si $P(A) = 0$.

Teorema de probabilidad Total (TPT): En un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$, sean $A_1, ..., A_n$ eventos disjuntos dos a dos con $P(A_i) > 0$ $\forall i$, entonces $\forall B \in \mathscr{A}$:

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i)P(A_i)$$

Definición: dado un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$, se dice que una colección de eventos $\mathcal{A} = \{A_i\}_{1\le i\le n}$ es independiente si $n = 1$ o:

$$P\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} P(A_k)$$

y todo subconjunto no vacío de $\mathcal{A}$ con menos de n elementos es independiente. Además diremos que los eventos $A_1, \ldots, A_n$ son independientes.

Definición: Una variable aleatoria X en un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$es una función $X : \Omega \to \mathbb{R}$ tal que $\forall x \in \mathbb{R} : \{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{A}$.

Definición: Se dice que una v.a X es discreta si su rango es contable.

Definición: Sea $(\Omega, \mathscr{A}, P)$un Espacio de probabilidad, y X una v.a discreta en este Espacio. Entonces la función $f : \mathbb{R} \to [0, 1]$ dada por $f(x) = P(X = x) \forall x \in \mathbb{R}$ se llama función de densidad discreta de X. También se le suele denotar con $f_X(x)$.
Además diremos que x es un valor posible de X si $f(x) > 0$.

Teorema Sea $f : \mathbb{R} \to \mathbb{R}$ con:

1. $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
2. $\{x \mid f(x) \neq 0\}$ es contable, sean $x_1, x_2, \ldots$ sus elementos.
3. $\sum_i f(x_i) = 1$

Entonces f es una función de densidad discreta.

DENSIDADES DISCRETAS

Binomial:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ , si } x = 0, 1, \ldots, n.$$

$$X \sim B(n, p)$$

Hipergeométrica:

$$f(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \text{ , si } x = 0, 1, \ldots, n.$$

$$X \sim \mathscr{H}(n, M, N)$$

Geométrica:

$$f(x) = p(1-p)^x \text{ , si } x \in \mathbb{N}_0, 0 < p < 1.$$

$$X \sim G(p)$$

Binomial Negativa:

$$f(x) = p^\alpha \binom{-\alpha}{x} (-1)^x (1-p)^x \text{ , si } x \in \mathbb{N}_0, 0 < p < 1.$$

$$X \sim B^-(\alpha, p)$$

$$\binom{-\alpha}{x} := \frac{\prod_{i=0}^{x-1}(-\alpha - i)}{x!}$$

Poisson:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \text{ , si } x \in \mathbb{N}_0.$$

$$X \sim \mathcal{P}(\lambda)$$

Definición: Sea $(\Omega, \mathscr{A}, P)$ un Espacio de probabilidad, y X una v.a en este Espacio. Entonces la función $F : \mathbb{R} \to [0,1]$ dada por $F(x) = P(X \leq x) \forall x \in \mathbb{R}$ se llama función de distribución de X.

Teorema: Sea $F : \mathbb{R} \to \mathbb{R}$ tal que:

1. $\lim\limits_{x\to\infty} F(x) = 1$ y $\lim\limits_{x\to-\infty} F(x) = 0$
2. F es continua por derecha, i.e $\forall x \in \mathbb{R} : \lim\limits_{t\to x^+} F(t) = F(x)$.
3. F es no decreciente.

Entonces F es una función de distribución.

Definición: Sean $X_1, \ldots, X_r$ v.a en un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$ entonces el vector $\mathbf{X} = (X_1, \ldots, X_r)$ es una v.a $r-$dimensional. Asimismo, si $X_1, \ldots, X_r$ son v.a discretas, entonces $\mathbf{X}$ se llamará discreta.

Definición: Análogamente se define la función de densidad discreta $f : \mathbb{R}^r \to \mathbb{R}$ de una v.a discreta $\mathbf{X}$ como:

$$f(\mathbf{x}) = P(\mathbf{X} = \mathbf{x})$$

también diremos que f es la función de densidad conjunta de $X_1, \ldots, X_r$.
Llamaremos a la función de densidad discreta f_{X_i} de la $i-$ésima coordenada de $\mathbf{X}$ como la $i-$ésima función de densidad marginal de $\mathbf{X}$.

Al igual que en el caso unidimensional, se cumple el siguiente

Teorema: Sea $f : \mathbb{R}^r \to \mathbb{R}$ tal que:

1. $f(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^r$
2. $\{\mathbf{x} \mid f(\mathbf{x}) \neq 0\}$ es contable, sean $\mathbf{x}_1, \mathbf{x}_2, \ldots$ sus elementos.
3. $\sum\limits_i f(\mathbf{x_i}) = 1$

Entonces f es una función de densidad discreta.

Definición: Sean $X_1, \ldots, X_r$ v.a discretas en un Espacio de probabilidad $(\Omega, \mathscr{A}, P)$, y sea f su función de densidad conjunta, diremos que son mutualmente independientes si:

$$f(x_1, \ldots, x_r) = \prod_{i=1}^{r} f_{X_i}(x_i)$$

Funciones Generadoras

Definición: Sea X una v.a entera no negativa. La función generadora $\Phi_X : [-1,1] \to \mathbb{R}$ de X como la siguiente serie:

$$\Phi_X(t) = \sum_{x=0}^{\infty} f_X(x) t^x$$

Teorema: Sean X e Y dos v.a enteras no negativas, si $\Phi_X = \Phi_Y$ entonces $X \sim Y$.

Teorema: Sean X e Y dos v.a enteras no negativas independientes, entonces:

$$\Phi_{X+Y} = \Phi_X \Phi_Y$$

Funciones Generadoras de distribuciones:

- $X \sim B(n,p) \implies \Phi_X(t) = (pt + 1 - p)^n$
- $X \sim B^-(\alpha,p) \implies \Phi_X(t) = \left(\frac{p}{1-t(1-p)}\right)^\alpha$
- $X \sim \mathcal{P}(\lambda) \implies \Phi_X(t) = e^{\lambda(t-1)}$

VARIABLES ALEATORIAS CONTINUAS

Definición: Una variable aleatoria X se dice continua si $\forall x \in \mathbb{R} : P(X = x) = 0$

Teorema: Una variable aleatoria X es continua si y solo si su función de distribución F es continua.

Definición: Diremos que una función $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ integrable que cumpla:

$$\int_{-\infty}^{\infty} f(x) = 1$$

es una función de densidad.

Definición: Sea X una v.a continua. Si la función de distribución F de X cumple que existe una función de densidad tal que:

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

Entonces se dirá que X es *absolutamente continua*, y llamaremos a esa f como una función de densidad de X, que denotaremos como f_X.

Teorema: Sea $\varphi : I \to \mathbb{R}$ una función diferenciable estrictamente creciente o estrictamente decreciente, definida en un intervalo I,y sea φ^{-1} su inversa. Sea X una variable aleatoria continua con densidad f tal que $f(x) = 0 \forall x \notin I$. Entonces $Y = \varphi(X)$ tiene densidad g dada por $g(y) = 0 \forall y \notin \varphi(I)$ y:

$$g(y) = f(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right| \quad \forall y \in \varphi(I)$$

Definición: Una variable aleatoria X se llama *simétrica* si X y $-X$ tienen la misma función de distribución.

Teorema Una variable aleatoria absolutamente continua X es simétrica si y solo si su función de densidad f_X es simétrica.

DENSIDADES CONTINUAS

Uniforme:

$$f(x) = \begin{cases} (b-a)^{-1}, & \text{si } a < x < b \\ 0 & \text{Caso Contrario} \end{cases}$$

$$X \sim \mathcal{U}(a,b), \quad a < b$$

Exponencial:

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

$$X \sim Exp(\lambda), \quad \lambda > 0$$

Normal:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

$$X \sim N(\mu, \sigma^2), \quad \sigma > 0$$

Gamma:

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$\alpha = m \in \mathbb{N} \implies F(x) = \int_0^x f(y)dy = \int_0^x \frac{\lambda^m y^{m-1} e^{-\lambda y}}{(m-1)!} dy \overset{\text{I.P.P}}{=} 1 - \sum_{k=0}^{m-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}, \quad x > 0$$

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

$$\Gamma(\alpha) := \int_0^\infty x^{\alpha-1} e^{-x} dx, \alpha > 0$$

Cauchy:

$$f(x) = \frac{1}{\pi(1+x^2)}$$

Definición: Sean X e Y dos variables aleatorias, se define a la función de distribución conjunta $F : \mathbb{R}^2 \to \mathbb{R}$ de X e Y como:

$$F(x,y) = P(X \leq x, Y \leq y)$$

Las funciones de distribución unidimensionales F_X y F_Y se llamarán funciones de distribución marginales de X e Y.

Si además existe una función $f : \mathbb{R}^2 \to \mathbb{R}$ integrable tal que:

$$F(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(u,v) dv du$$

entonces diremos que f es una función de densidad conjunta del vector (X,Y)

Propiedades:

1.

$$F_X(x) = \lim_{y\to\infty} F(x,y)$$

$$F_Y(x) = \lim_{x\to\infty} F(x,y)$$

2.

$$P\left((X,Y)\in A\right) = \iint\limits_A f(x,y)dxdy$$

3.

$$f_X(x) = \int\limits_{-\infty}^{\infty} f(x,y)dy$$

$$f_Y(y) = \int\limits_{-\infty}^{\infty} f(x,y)dx$$

4.

$$\frac{\partial^2}{\partial x \partial y} F(x,y) = f(x,y)$$

Definición: Dos v.a X e Y se dicen independientes si para todo cuarteto $a \leq b, c \leq d$:

$$P(a < X \leq b, c < Y \leq d) = P(a < X \leq b)P(c < Y \leq d)$$

+Propiedades:

1.

$$F_{X+Y}(z) = \int\limits_{-\infty}^{\infty} f(x, z-x)dx$$

2.

$$f_{\frac{Y}{X}}(z) = \int\limits_{-\infty}^{\infty} |x| f(x, xz)dx$$

$$F_{\frac{Y}{X}}(z) = \int\limits_{-\infty}^{z} f_{\frac{Y}{X}}(v)dv$$

Definición: Sean X e Y dos variables aleatorias con densidad conjunta f. Se define la densidad condicional $f_{Y|X} : \mathbb{R} \to \mathbb{R}$ como:

$$f_{Y|X}(y \mid x) = \begin{cases} \frac{f(x,y)}{f_X(x)}, & \text{si } f_X(x) \neq 0 \\ 0, & \text{C.C} \end{cases}$$

ESPERANZA

Definición: Sea X una v.a discreta con imagen $\{x_1, x_2, \dots\}$, tal que $\sum\limits_i |x_i| f_X(x_i) < \infty$, se define el valor esperado de X, $E(X)$, como:

$$E(X) = \sum_i x_i f(x_i)$$

En caso contrario se dice que su esperanza no es finita.

Teorema: Sea X una v.a, y sea $\varepsilon > 0$, declaramos la variable aleatoria X_ε definida por:

$$X_\varepsilon = k, \quad \varepsilon k \leq X < \varepsilon(k+1)$$

entonces si X_ε tiene esperanza finita para algún $\varepsilon > 0$ entonces X_ε tiene esperanza finita para todo $\varepsilon > 0$ y el límite:

$$\lim_{\varepsilon \to 0} E(X_\varepsilon)$$

existe.

Definición: Sea X una v.a, entonces si existe un $\varepsilon > 0$ tal que X_ε tiene esperanza finita, entonces decimos que X tiene esperanza $E(X)$ finita:

$$E(X) = \lim_{\varepsilon \to 0} E(X_\varepsilon)$$

caso contrario decimos que X tiene esperanza no finita.

También definimos la media de X como $\mu = E(X)$.

Teorema: Sea X una v.a continua con densidad f, entonces X tiene esperanza finita si y solo si:

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

y en tal caso se cumple la igualdad:

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx$$

Definición: Sea X una v.a, entonces X tiene momento de orden $r \in \mathbb{N}$ si $E(X^r)$ es finito, el momento central de orden r se define como $E\left((X-\mu)^r\right)$.

Definición: Sea X una v.a con momento central de orden 2, definimos la varianza de X, $\operatorname{Var}(X)$ como:

$$\operatorname{Var}(X) = E\left((X-\mu)^2\right)$$

Asimismo se define la desviación standard $\sigma = \sqrt{\operatorname{Var}(X)}$

Definición: Sean X e Y v.a con momento de orden 2 finito. Se define la covarianza de X e Y como:

$$\text{Cov}(X,Y) = E\left(\ (X - E(X)) \ \ (Y - E(Y)) \ \right) = E(XY) - E(X)E(Y)$$

Definición: Sean X e Y dos v.a con varianza no nula, se define el coeficiente de correlación:

$$\rho(X,Y) = \frac{\text{Cov(X,Y)}}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Se dirá que X e Y no están correlacionadas si $\rho = 0$.

Teorema: Sean X e Y dos v.a con momento de segundo orden finito, entonces:

$$[E(XY)]^2 \leq E(X^2)E(Y^2)$$

Más aún, la igualdad se da si y solo si o bien $P(Y = 0) = 1$ o bien $P(X = aY) = 1$ para alguna constante a.

Corolario: $|\rho| \leq 1$, y $|\rho| = 1$ si y solo si $P(X = aY) = 1$ para alguna constante a.

Teorema débil de los números grandes:

Sean $X_1, X_2, \ldots$ v.a independientes con la misma distribución Y media finita μ.
Sea $S_n = X_1 + \cdots + X_n$, entonces para todo $\delta > 0$:

$$\lim_{n\to\infty} P\left(\left| \frac{S_n}{n} - \mu \right| \geq \delta \right) = 0$$

TEOREMA CENTRAL DEL LÍMITE (T.C.L):

Sean $X_1, X_2, \ldots$ v.a independientes con la misma distribución, media finita μ y desviación σ^2. Sea $S_n = X_1 + \cdots + X_n$, si además definimos:

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Entonces:

$$\lim_{n\to\infty} P(S_n^* \leq x) = \Phi(x)$$

Propiedades:

1. Sean X e Y dos variables aleatorias continuas con densidad conjunta f, medias μ_X, μ_Y y momentos de orden 2 finitos, entonces:

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f(x,y)\; dydx$$

2. (* De mano propia pero la demo es como la de $f_{\frac{Y}{X}}$)

$$f_{XY}(u) = \int_{-\infty}^{\infty} \frac{f\left(x, \frac{u}{x}\right)}{|x|} dx$$

$$F_{XY}(z) = \int_{-\infty}^{z} f_{XY}(u) du$$