81.09 PROBABILIDAD

Sean Ω un conjunto no vacío cuyos elementos ω serán llamados EVENTOS ELEMENTALES y $\mathcal{A}$ una familia de subconjuntos de Ω que serán llamados EVENTOS.

→ $\mathcal{A}$ es un ÁLGEBRA DE EVENTOS si contiene a Ω y es cerrada por complementos y uniones finitas

I) $\Omega \in \mathcal{A}$

II) $A \in \mathcal{A}$ implica $\overline{A} \in \mathcal{A}$

III) $A, B \in \mathcal{A}$ implica $A \cup B \in \mathcal{A}$

→ Una MEDIDA DE PROBABILIDAD P sobre $(\Omega, \mathcal{A})$ es una función $P: \mathcal{A} \to \mathbb{R}$ que satisface los axiomas siguientes:

I) Para cada $A \in \mathcal{A}$, $P(A) \geq 0$

II) $P(\Omega) = 1$

III) Si $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$

IV) Para cada sucesión decreciente de eventos

$$A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$$

tal que $\bigcap_{n=1}^{\infty} A_n = \emptyset$ vale que $\lim_{n\to\infty} P(A_n) = 0$

De este último axioma se deducen algunos teoremas:

I) Si $\bigcap_{n=1}^{\infty} A_n = A$, entonces $P(A) = \lim_{n\to\infty} P(A_n)$

(siendo $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$)

NOTA

II) sean $A_1, A_2, \dots, A_n$ una sucesión de eventos disjuntos a pares, si $\bigcup_{n=1}^{\infty} A_n$ es un evento en $\mathcal{A}$, entonces

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

RECORDAR: $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ si y solo si $|x|<1$

→ Un ESPACIO DE PROBABILIDAD es una terna $(\Omega, \mathcal{A}, P)$ formada por un conjunto no vacío Ω, llamado el ESPACIO MUESTRAL; un álgebra $\mathcal{A}$ de subconjuntos de Ω, llamados los EVENTOS ALEATORIOS; y una medida de probabilidad P definida sobre los eventos aleatorios.

- $\rightarrow A \cup B = B \cup A$
- $\rightarrow A \cap B = B \cap A$
- $\rightarrow A \cup (B \cup C) = (A \cup B) \cup C$
- $\rightarrow A \cap (B \cap C) = (A \cap B) \cap C$
- $\rightarrow A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\rightarrow A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $\rightarrow A \cap \emptyset = \emptyset$
- $\rightarrow A \cup \emptyset = A$
- $\rightarrow \overline{(A \cup B)} = \bar{A} \cap \bar{B}$
- $\rightarrow \overline{(A \cap B)} = \bar{A} \cup \bar{B}$

→ si $\emptyset$ es el conjunto vacío, entonces $P(\emptyset)=0$

→ $P(A\cup B) = P(A) + P(B) - P(A\cap B)$

→ $P(A\cup B\cup C) = P(A)+P(B)+P(C) - P(A\cap B) - P(A\cap C) -$
$- P(B\cap C) + P(A\cap B\cap C)$

→ se considera un conjunto finito Ω y una función $p: \Omega \to [0,1]$ tal que $\sum_{\omega\in\Omega} p(\omega) = 1$. La función p se llama FUNCIÓN DE PROBABILIDAD y los números $p(\omega)$ se llaman las PROBABILIDADES DE LOS EVENTOS ELEMENTALES $\omega\in\Omega$

→ El álgebra de eventos $\mathcal{A}$ se toma como el conjunto de todos los subconjuntos de Ω y para cada $A\in\mathcal{A}$ se define $P(A) = \sum_{\omega\in A} p(\omega)$

ANÁLISIS COMBINATORIO

Cuando se estudian juegos de azar, procedimientos muestrales, problemas de orden y ocupación, se trata con espacios muestrales finitos Ω en los que a todos los eventos elementales se les atribuye igual probabilidad. Para calcular la probabilidad de un evento A tenemos que dividir la cantidad de eventos elementales contenidos en A (CASOS FAVORABLES) entre la cantidad total de eventos elementales contenidos en Ω (CASOS POSIBLES)

→ $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} \Rightarrow \#(\Omega) = n$ es el CARDINAL

→ Como álgebra de eventos $\mathcal{A}$, tomamos el conjunto de todos los subconjuntos de $\Omega \Rightarrow$ si $\#(\Omega)=K$, entonces hay 2^K subconjuntos

REGLA DEL PRODUCTO

Sean A y B dos conjuntos cualesquiera. El producto cartesiano de A y B se define por $A \times B = \{(a,b) : a \in A \text{ y } b \in B\}$. Si A y B son finitos, entonces $\#(A \times B) = \#(A) . \#(B)$.
Y se extiende a n conjuntos.

MUESTRAS ORDENADAS

Para una población de n elementos y un tamaño de muestra prefijado k, existen n^k diferentes muestras con reposición y nPk muestras sin reposición.

→ $nPk = \frac{n!}{(n-k)!}$ es la cantidad de subconjuntos ordenados de k elementos elegidos entre n

SUBPOBLACIONES

Una población de n elementos tiene nCk diferentes subpoblaciones de tamaño $k \leq n$

→ $nCk = \binom{n}{k} = \frac{n!}{(n-k)!\,k!}$ es la cantidad de subconjuntos sin ordenar de r elementos elegidos entre n

PROBABILIDAD CONDICIONAL

Sea $A \subset \Omega$ un evento de probabilidad positiva, para cada evento B definimos

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

Propiedades

I) para cada $B \in \mathcal{A}$, $P(B|A) \geq 0$

II) $P(\Omega|A) = 1$

III) si $B \cap C = \emptyset$, entonces $P(B \cup C|A) = P(B|A) + P(C|A)$

IV) para cada sucesión decreciente de eventos $B_1 \supset B_2 \supset \ldots$ tal que $\bigcap_{n=1}^{\infty} B_n = \emptyset$ vale que $\lim_{n \to \infty} P(B_n|A) = 0$

$$\rightarrow P(A \cap B) = P(B|A) . P(A)$$

$$\rightarrow P(A \cap B \cap C) = P(C|A \cap B) . P(B|A) . P(A)$$

PROBABILIDAD TOTAL

Sea $A_1, A_2, \ldots$ una sucesión de eventos disjuntos a pares tal que $\bigcup_{i=1}^{n} A_i = \Omega$, si $P(A_i) > 0$, para cada $B \in \mathcal{A}$ vale:

$$\rightarrow P(B) = \sum_{i=1}^{n} P(B|A_i) . P(A_i)$$

REGLA DE BAYES

Sean A y B dos eventos de probabilidad positiva:

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

$$P(B_1 \cap B_2) = P(B_2 | B_1 \cap A_i) . P(B_1 | A_i) . P(A_i)$$

$$P(B_1 \cap B_2 \cap B_3) = P(B_3 | B_2 \cap B_1 \cap A_i) . P(B_2 | B_1 \cap A_i) . P(B_1 | A_i) . P(A_i)$$

$$P(B) = P(B|A) . P(A) + P(B|\bar{A}) . P(\bar{A})$$

INDEPENDENCIA

Dos eventos A_1 y A_2 son independientes si satisfacen la ecuación

→ $P(A_1 \cap A_2) = P(A_1) . P(A_2)$

→ $P(A_2 | A_1) = P(A_2)$ y $P(A_1 | A_2) = P(A_1)$

VARIABLES ALEATORIAS

Una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$ es una función de Ω en $\mathbb{R}$ tal que las ocurrencias de los valores de la función corresponden a eventos en $\mathcal{A}$ y por lo tanto se les puede asignar probabilidad

FUNCIÓN DISTRIBUCIÓN

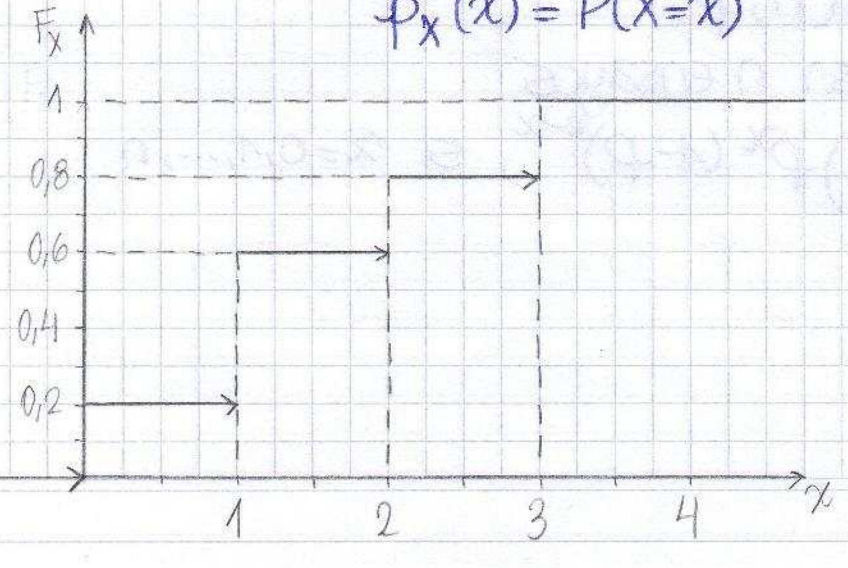
* $F_X : \mathbb{R} \to [0,1] \,/\, F_X(x) = P(X \leq x)$
* $P(a < X \leq b) = F_X(b) - F_X(a)$
* $P(X = a) = F_X(a) - \lim\limits_{x \to a^-} F_X(x)$

* Si la función es continua en $x_0 \Rightarrow P(X = x_0) = 0$
* Los únicos puntos con probabilidad no nula son aquellos donde el gráfico de F_X presenta saltos (puntos de discontinuidad)
* A estos puntos de discontinuidad se los llama ÁTOMOS

Una V.A. es DISCRETA si y solo si $\sum\limits_{a \in A} P(X = a) = 1$

La función que permite conocer la probabilidad puntual se denomina FUNCIÓN DE PROBABILIDAD p_X

$$p_X(x) = P(X = x)$$



$P(X=0) = 0,2$

$P(X=1) = 0,4$

$P(X=2) = 0,2$

$P(X=3) = 0,2$

NOTA

Una V.A. es <u>continua</u> si y solo si F_X es continua

Una V.A. es <u>MIXTA</u> si y solo si no es continua ni discreta

Una V.A. X continua es absolutamente continua si y solo si existe una función simbolizada f_X y llamada <u>DENSIDAD DE X</u> tal que $f_X : \mathbb{R} \to \mathbb{R}^+$ y $P(a < X \leq b) = \int_a^b f_X(x)\,dx$

$$* \; F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)\,dx$$

$$* \; \frac{dF_X(x)}{dx} = f_X(x) \qquad * \int_{-\infty}^{+\infty} f_X(x)\,dx = 1$$

DISTRIBUCIONES DISCRETAS

BERNOULLI $\sim BER(p)$

X es una V.A. con distribución de Bernoulli de parámetro p con $0 < p < 1$ si X es discreta con función de probabilidad dada por:

$P(X=0) = 1-p$ ("FRACASO")

$P(X=1) = p$ ("ÉXITO")

$\Rightarrow p_X(x) = p^x (1-p)^{1-x}$, si $x = 0, 1$

BINOMIAL $\sim BIN(n, p)$

número de éxitos en n ensayos

$\Rightarrow p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$, si $x = 0, 1, \dots, n$

NOTA

VARIABLES ALEATORIAS

Una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$ es una función de Ω en $\mathbb{R}$ tal pue las ocurrencias de los valores de la función corresponden a eventos en $\mathcal{A}$ y por lo tanto se les puede asignar probabilidad

FUNCIÓN DISTRIBUCIÓN

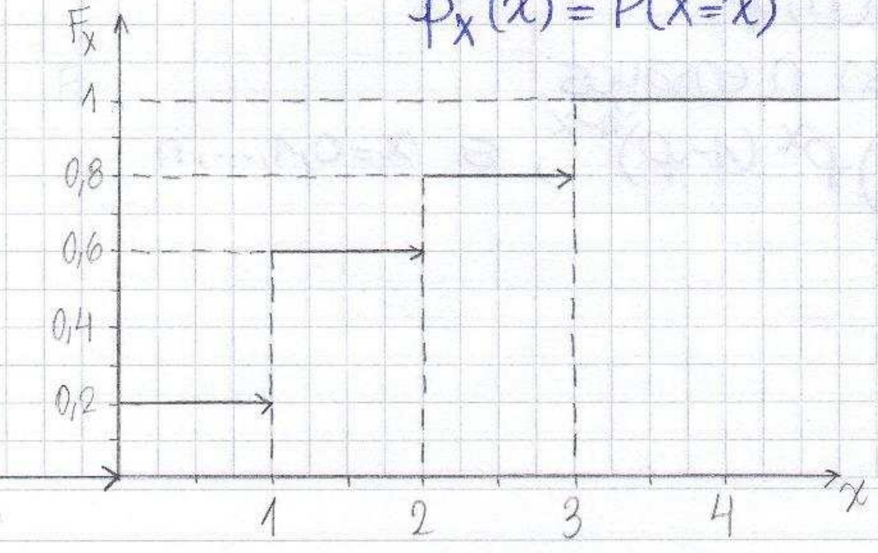
* $F_X : \mathbb{R} \to [0,1] \,/\, F_X(x) = P(X \leq x)$
* $P(a < X \leq b) = F_X(b) - F_X(a)$
* $P(X=a) = F_X(a) - \lim\limits_{x \to a^-} F_X(x)$

* Si la función es continua en $x_0 \Rightarrow P(X=x_0)=0$
* Los únicos puntos con probabilidad no nula son aquellos donde el gráfico de F_X presenta saltos (puntos de discontinuidad)
* A estos puntos de discontinuidad se los llama ÁTOMOS

Una V.A. es DISCRETA si y solo si $\sum\limits_{a \in A} P(X=a) = 1$

La función que permite conocer la probabilidad puntual se denomina FUNCIÓN DE PROBABILIDAD p_X

$$p_X(x) = P(X=x)$$



$P(X=0)=0{,}2$

$P(X=1)=0{,}4$

$P(X=2)=0{,}2$

$P(X=3)=0{,}2$

Una V.A. es <u>continua</u> si y solo si F_X es continua

Una V.A. es <u>MIXTA</u> si y solo si no es continua ni discreta

Una V.A. X continua es absolutamente continua si y solo si existe una función simbolizada f_X y llamada <u>DENSIDAD DE X</u> tal que $f_X : \mathbb{R} \to \mathbb{R}^+$ y $P(a < X \leq b) = \int_a^b f_X(x)\,dx$

$$* \quad F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)\,dx$$

$$* \quad \frac{dF_X(x)}{dx} = f_X(x) \qquad * \quad \int_{-\infty}^{+\infty} f_X(x)\,dx = 1$$

DISTRIBUCIONES DISCRETAS

BERNOULLI $\sim BER(p)$

X es una V.A. con distribución de Bernoulli de parámetro p con $0 < p < 1$ si X es discreta con función de probabilidad dada por:

$P(X=0) = 1-p$ ("FRACASO")

$P(X=1) = p$ ("ÉXITO")

$\Rightarrow p_X(x) = p^x (1-p)^{1-x}$, si $x = 0,1$

BINOMIAL $\sim BIN(n, p)$

número de éxitos en n ensayos

$\Rightarrow p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$, si $x = 0, 1, \dots, n$

NOTA

GEOMÉTRICA ~ GEOM(p)

número de ensayos hasta el primer éxito

$\Rightarrow p_N(n) = p(1-p)^{n-1}, \quad n \in \mathbb{N}$

* $P(N>n) = (1-p)^n$
* $P(N \leq n) = 1-(1-p)^n$

PASCAL ~ PASCAL(r, p)

número de ensayos hasta r éxitos

$\Rightarrow p_X(x) = \binom{x-1}{r-1} p^r (1-p)^{n-r}, \quad n = r, r+1, r+2, \ldots$

POISSON ~ POISSON(μ)

$$\Rightarrow p_N(n) = \frac{\mu^n e^{-\mu}}{n!}, \quad n = 0, 1, 2, \ldots$$

HIPERGEOMÉTRICA ~ $\mathcal{H}(N, r, n)$

N = total poblacional

r = cantidad de "buenos" en la población

n = cantidad de elementos extraídos

$$\Rightarrow p_X(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

DISTRIBUCIONES CONTINUAS

UNIFORME ~ $U(a, b)$

$$\Rightarrow F_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otros } x \end{cases}$$

$$\Rightarrow F_X(x) = \frac{1}{b-a} \mathbb{1}\{a < x < b\}$$

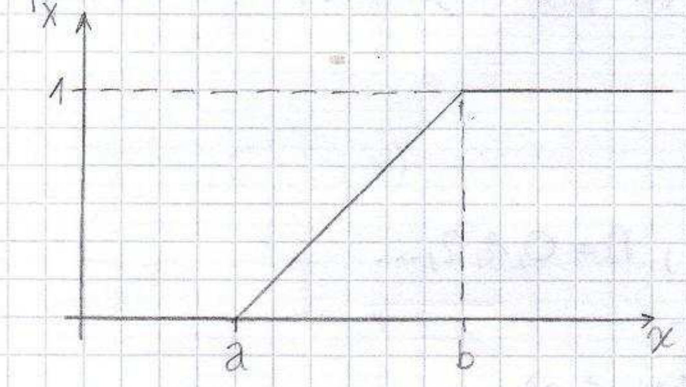
NOTA

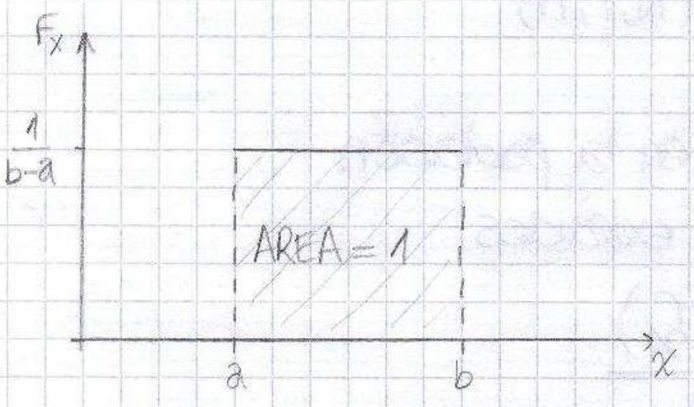
$* \ P(a<X<c) = \int_a^c \frac{1}{b-a}\,dx = \frac{c-a}{b-a}$

$F_X(x) = P(X \leq x)$ para $X \sim U(a,b)$, $\int_{-\infty}^{x} F_X(x)\,dx = F_X(x)$

$x \leq a$	$a < X \leq b$	$x > b$
$P(X \leq x) = 0$	$P(X \leq x) = \frac{x-a}{b-a}$	$P(X \leq x) = 1$

$* \ F_X(x) = \frac{x-a}{b-a} \mathbb{1}\{a < x \leq b\} + \mathbb{1}\{x > b\}$





EXPONENCIAL $\sim EXP(\lambda)$

$\Rightarrow F_X(x) = \lambda e^{-\lambda x} \mathbb{1}\{x > 0\}$

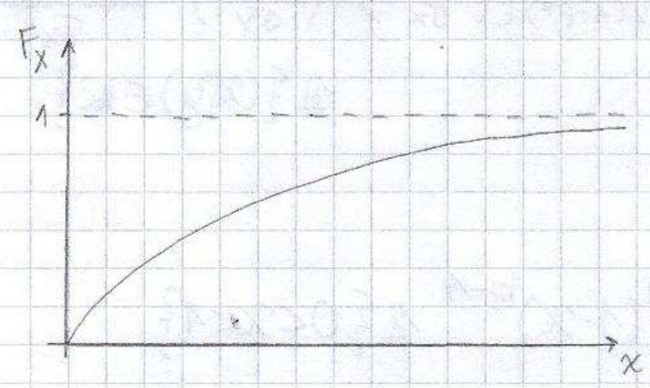
$x \leq 0 \rightarrow P(X \leq x) = \int_{-\infty}^{x} F_X(x)\,dx = 0$

$x > 0 \rightarrow P(X \leq x) = \int_{-\infty}^{x} F_X(x)\,dx = \int_{-\infty}^{0} F_X(x)\,dx + \int_{0}^{x} F_X(x)\,dx$

$= 0 + \int_0^x \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda x}$

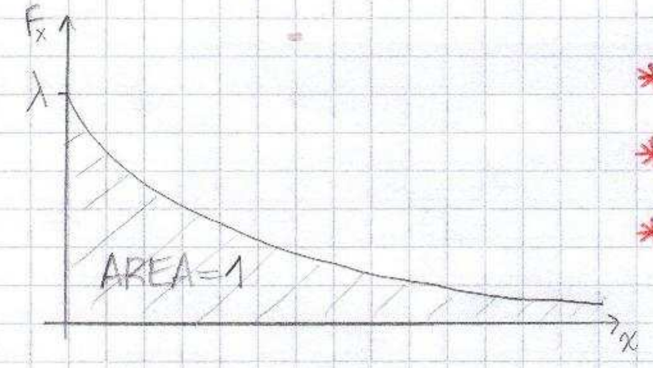
NOTA

$\Rightarrow F_X(x) = (1 - e^{-\lambda x})\, \mathbb{1}\{x>0\}$



$\lambda(t)$ FUNCIÓN INTENSIDAD DE FALLA

$$F_T(t) = 1 - e^{-\int_0^t \lambda(u)\,du}\, \mathbb{1}\{t>0\}$$



* $P(X \leq x_0) = 1 - e^{-\lambda x_0}$
* $P(X > x_0) = e^{-\lambda x_0}$
* $P(x_1 < X < x_2) = e^{-\lambda x_1} - e^{-\lambda x_2}$

GAMMA $\sim$ GAMMA(n, λ)

$$\Rightarrow F_X(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}\quad \mathbb{1}\{x>0\}$$

Propiedad: si $X_1, X_2, \ldots, X_n$ son n variables independientes identicamente distribuidas (i.i.d.), todas con la misma distribución exponencial de parámetro λ, entonces:

$$Y = X_1 + X_2 + \ldots + X_n \sim \text{GAMMA}(n, \lambda)$$

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

$$F_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\quad \mathbb{1}\{x \in \mathbb{R}\}$$

NORMAL ESTANDAR $\sim N(0,1)$

NOTA

NORMAL BIVARIADA

coeficiente de correlación (pag 17)

$(X,Y) \sim N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$

$$F_{xy}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2+\left(\frac{y-\mu_y}{\sigma_y}\right)^2-\frac{2(x-\mu_x)(y-\mu_y)\rho}{\sigma_x\ \sigma_y}\right]} \mathbb{1}\{(x,y)\in\mathbb{R}\}$$

BETA $\sim \beta(a,b)$

$$F_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \mathbb{1}\{0<x<1\}$$

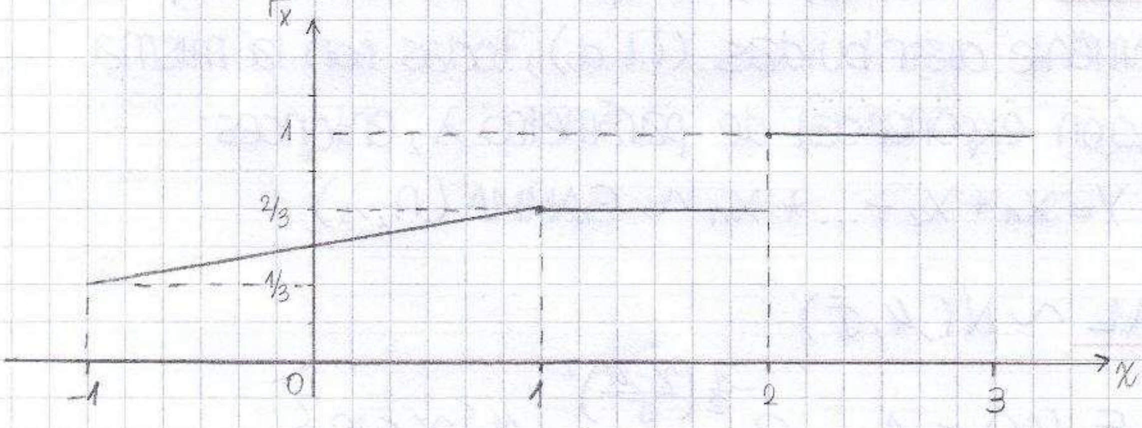
CAUCHY $\sim C(0,\lambda)$

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

CUANTILES

Sea $\alpha\in[0,1]$, un cuantil-α de X es cualquier número $x_\alpha\in\mathbb{R}$ tal que

$$\Rightarrow P(X<x_\alpha)\leq\alpha\leq P(X\leq x_\alpha)$$



$x_{0,5}=0$ $\quad x_{0,1}=-1$ $\quad x_{0,8}=2$

PÉRDIDA DE MEMORIA

Se dice que una variable aleatoria T no tiene memoria, o tiene pérdida de memoria, si:

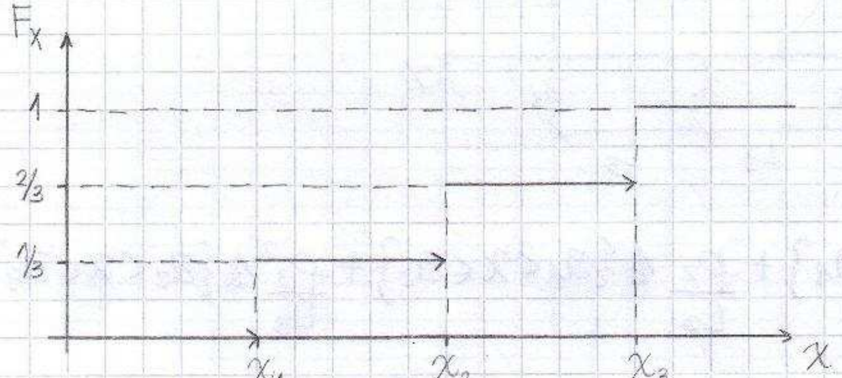
$$P(T > s+t \mid T > t) = P(T > s)$$

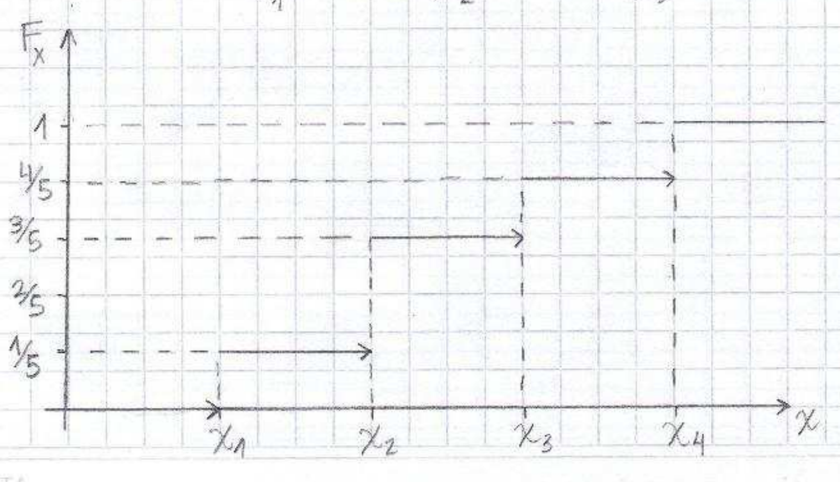
Hay solo dos variables con esta propiedad, la geométrica y la exponencial

DISTRIBUCIÓN EMPÍRICA

Sea X una V.A. generada por observaciones en algún experimento, si se repite n veces el experimento y se obtienen n valores $x_1, x_2, \ldots, x_n$, la FUNCIÓN DE DISTRIBUCIÓN EMPÍRICA basada en estos n valores es una función escalera con saltos de valor $\frac{1}{n}$ en cada x_i

$\{x_1, x_2, x_3\}$





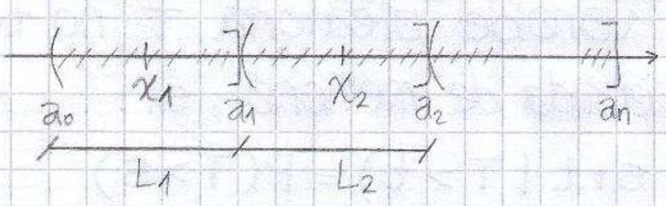
$\{x_1, x_2, x_2, x_3, x_4\}$

NOTA

FUNCIÓN HISTOGRAMA

$\{x_1, x_2, \dots, x_n\}$ n puntos

m intervalos con extremos $a_0, a_1, a_2, \dots, a_m$

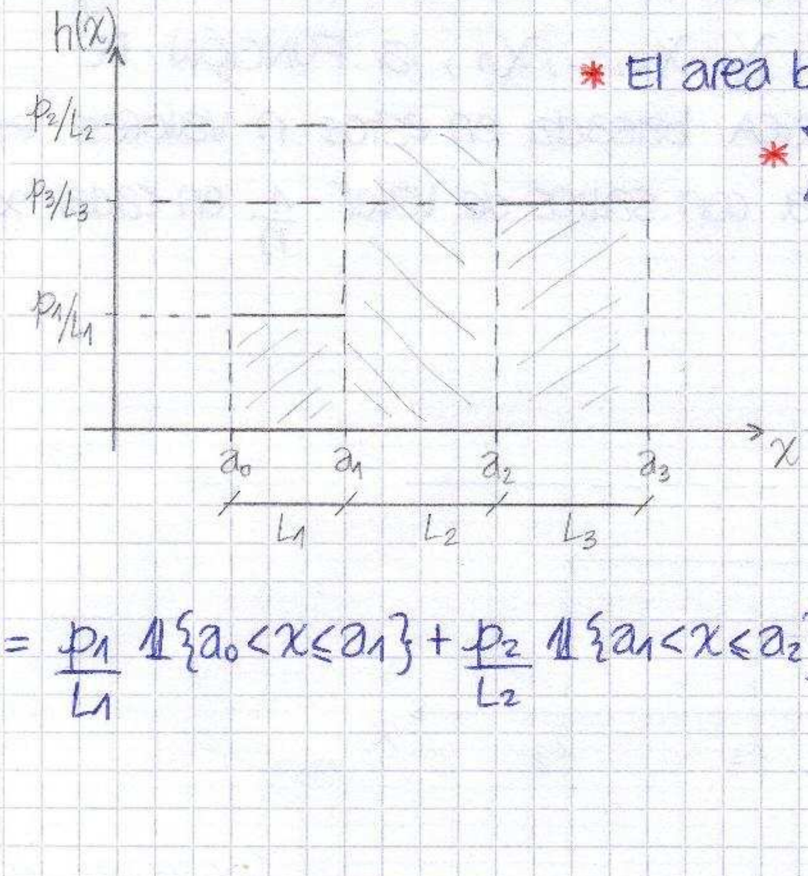


L_j = amplitud del intervalo $(a_{j-1}, a_j]$

p_j = Frecuencia relativa del intervalo $(a_{j-1}, a_j]$

$$p_j = \frac{\text{n° de puntos } x_i \text{ en el intervalo}}{n}$$

La FUNCIÓN HISTOGRAMA vale p_j/L_j sobre cada intervalo



* El area bajo la curva es p_i

* $\sum p_j = 1$

$$h(x) = \frac{p_1}{L_1} \mathbb{1}\{a_0 < x \leq a_1\} + \frac{p_2}{L_2} \mathbb{1}\{a_1 < x \leq a_2\} + \frac{p_3}{L_3} \mathbb{1}\{a_2 < x \leq a_3\}$$

SIMULACIÓN

Defino la INVERSA GENERALIZADA de F mediante:

$$F^{-1}(u) = \min\{x \in \mathbb{R} / u \leq F(x)\} \quad u \in (0,1)$$

Defino X mediante:

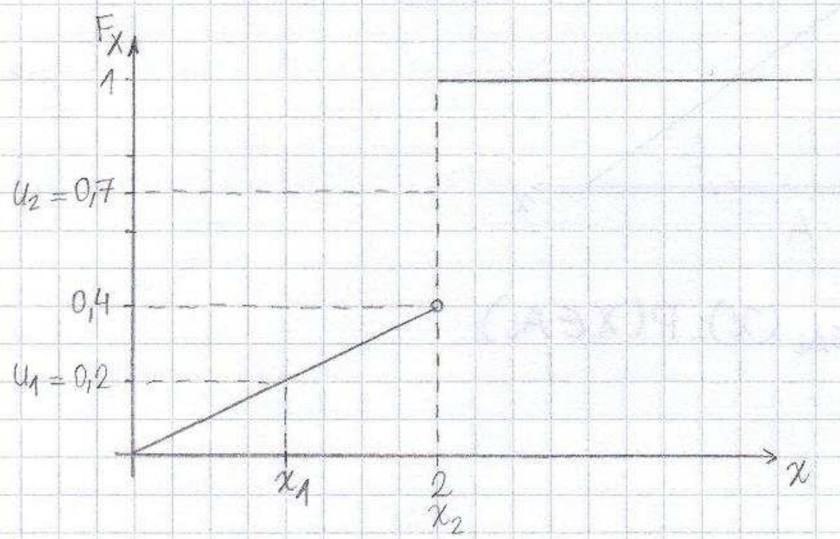
$$X = F^{-1}(U) \quad \text{donde } U \sim U(0,1)$$

Observo que vale la equivalencia $F^{-1}(u) \leq x \Leftrightarrow u \leq F(x)$

y deduzco que $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

<u>Observación</u>: si la función F es continua, la inversa generalizada es simplemente la inversa

<u>Ejemplo</u>: Usando los números aleatorios $u_1 = 0,2$ y $u_2 = 0,7$, simular dos valores de la V.A. X cuya función de distribución tiene el siguiente gráfico:



x_1: $\quad u_1 = 0,2 = 0,2x_1 \Rightarrow x_1 = 1$

x_2: $\quad F_X^{-1}(u_2) = \min\{x \in \mathbb{R} / F_X(x) \geq u_2\}$

$$\Rightarrow x_2 = F_X^{-1}(0,7) = \min\{x \in \mathbb{R} / F_X(x) \geq 0,7\}$$

$$= \min(2, +\infty) = 2$$

NOTA

VARIABLES TRUNCADAS

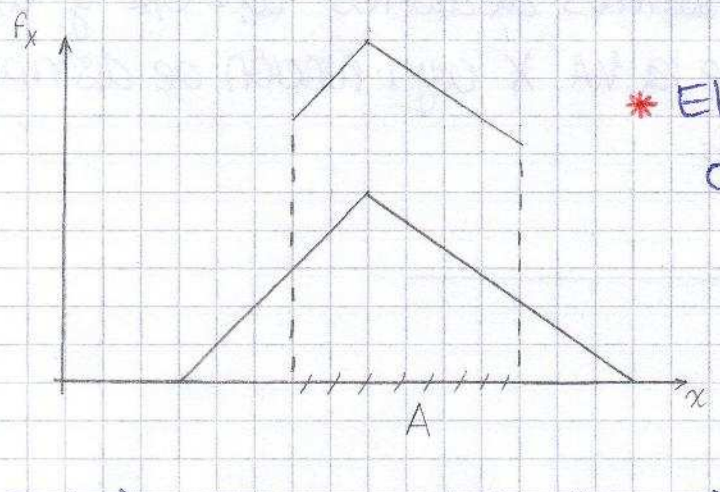
Llamamos truncar la variable X al conjunto A a restringir los valores de X al conjunto A:

$$\Rightarrow F_{X|X\in A}(x) = P(X \leq x \mid X\in A) = \frac{P(X\leq x \wedge X\in A)}{P(X\in A)}$$

Si X es continua con densidad f_X, entonces:

$$f_{X|X\in A}(x) = \frac{\int_{-\infty}^{x} f_X(x)\, \mathbb{1}\{X\in A\}\, dx}{P(X\in A)}$$

$$\Rightarrow f_{X|X\in A}(x) = \frac{f_X(x)}{P(X\in A)}\, \mathbb{1}\{X\in A\}$$



* El área bajo ambas curvas es 1

$$\Rightarrow F_X(x) = \sum F_{X|X\in A_i}(x) \cdot P(X\in A_i)$$

VECTORES ALEATORIOS

Si X e Y son V.A. $\Rightarrow$ (X,Y) es un VECTOR ALEATORIO si puede asignarse probabilidad a sucesos conjuntos de ambas variables

FUNCIÓN DE DISTRIBUCIÓN CONJUNTA DE X e Y

$$\Rightarrow F_{XY}(x,y) = P(X \leq x, Y \leq y)$$

I) (X,Y) es DISCRETO si y solo si existe un conjunto de átomos tal que $\sum_{(x,y)\in A} P(X=x, Y=y) = 1$

La FUNCIÓN DE PROBABILIDAD CONJUNTA de X e Y es:

$$\Rightarrow p_{XY}(x,y) = P(X=x, Y=y)$$

II) (X,Y) es CONTINUO si y solo si existe una función de densidad conjunta de X e Y tal que

$$\Rightarrow F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(x,y)\,dx\,dy$$

$$\Rightarrow \iint_{\mathbb{R}^2} f_{XY}(x,y)\,dx\,dy = 1$$

III) (X,Y) es MIXTO si y solo si no es discreto ni continuo

NOTA

DISTRIBUCIÓN UNIFORME

Sea $R \subset \mathbb{R}^2$ un conjunto acotado tal que $area(R) = |R|$, se dice que (X,Y) tiene distribución uniforme sobre la región R si y solo si (X,Y) tiene densidad conjunta dada por:

$$\Rightarrow F_{XY}(x,y) = \frac{1}{|R|} \mathbb{1}\{(x,y) \in R\}$$

$$\Rightarrow P((X,Y) \in A) = \iint_A \frac{1}{|R|} dx dy = \frac{1}{|R|} \iint_A dx dy = \frac{|A|}{|R|}$$

DISTRIBUCIONES MARGINALES

Si (X,Y) es un vector aleatorio, X e Y son V.A. con funciones de distribución $F_X(x) = P(X \leq x)$ y $F_Y(y) = P(Y \leq y)$ llamadas distribuciones marginales de X e Y respectivamente

1) Si (X,Y) es discreto con función de probabilidad conjunta $p_{XY}(x,y) = P(X=x, Y=y)$, X e Y son discretas tal que:

FUNCIÓN DE PROBABILIDAD MARGINAL DE X $\Rightarrow$ $p_X(x) = \sum_{\forall y} p_{XY}(x,y)$

FUNCIÓN DE PROBABILIDAD MARGINAL DE Y $\Rightarrow$ $p_Y(y) = \sum_{\forall x} p_{XY}(x,y)$

NOTA

II) Si (X,Y) es continuo con densidad conjunta f_{XY}, X e Y son continuas tal que:

$$\Rightarrow F_X(x) = P(X \leq x) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{XY}(s,y)\, dy\, ds$$

DENSIDAD MARGINAL DE X $\Rightarrow$ $f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy$

$$\Rightarrow F_Y(y) = P(Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{+\infty} f_{XY}(x,t)\, dx\, dt$$

DENSIDAD MARGINAL DE Y $\Rightarrow$ $f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dx$

X e Y son independientes si:

$$\Rightarrow F_{XY}(x,y) = F_X(x) \cdot F_Y(y)$$

I) caso discreto: $p_{XY}(x,y) = p_X(x) \cdot p_Y(y)$

II) caso continua: $f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$

MOMENTOS

Si X es una V.A., su valor esperado, media o esperanza, simbolizada $E[X]$, es:

I) $E[X] = \sum_{x \in A} x \, P(X=x)$ si X es discreta

II) $E[X] = \int_{-\infty}^{+\infty} x \, F_X(x) \, dx$ si X es continua

III) $E[X] = \sum_{x \in A} x \, P(X=x) + \int_{\mathbb{R}-A} x \, F_X(x) \, dx$ si X es mixta

* $X \sim Ber(p) \Rightarrow E[X] = p$
* $X \sim BIN(n, p) \Rightarrow E[X] = n \cdot p$
* $X \sim GEOM(p) \Rightarrow E[X] = \frac{1}{p}$
* $X \sim PAS(r, p) \Rightarrow E[X] = \frac{r}{p}$
* $X \sim POISSON(\mu) \Rightarrow E[X] = \mu$
* $X \sim \mathcal{H}(N, r, n) \Rightarrow E[X] = \frac{nr}{N}$
* $X \sim U(a, b) \Rightarrow E[X] = \frac{a+b}{2}$
* $X \sim EXP(\lambda) \Rightarrow E[X] = \frac{1}{\lambda}$
* $X \sim GAMMA(n, \lambda) \Rightarrow E[X] = \frac{n}{\lambda}$
* $X \sim N(\mu, \sigma) \Rightarrow E[X] = \mu$

NOTA

PROPIEDADES DE LA ESPERANZA

→ Si X es no negativa ⇒ $E[X]$ es no negativa

→ Si X V.A. ⇒ $Y = aX$ es otra V.A. con $E[Y] = E[aX] = aE[X]$

→ el valor esperado de una constante es $E[K] = K$

→ $E[aX+b] = aE[X]+b$

Si X es una V.A. y $Y = g(X)$ es otra V.A., entonces:

→ $E[Y] = E[g(X)] = \sum_{x \in A} g(x) . P(X=x)$ si X es discreta

→ $E[Y] = E[g(X)] = \int_{-\infty}^{+\infty} g(x) F_X(x) dx$ si X es continua

→ $E[Y] = E[g(X)] = \sum_{x \in A} g(x) P(X=x) + \int_{\mathbb{R}-A} g(x) F_X(x) dx$

si (X,Y) es un vector aleatorio, $Z = h(X,Y)$ es una V.A. tal que:

→ $E[Z] = E[h(X,Y)] = \sum_{(x,y) \in A} h(x,y) P(X=x, Y=y)$ si (X,Y) discreto

→ $E[Z] = E[h(X,Y)] = \iint_{\mathbb{R}^2} h(x,y) F_{XY}(x,y) dy dx$ si (X,Y) continua

en particular, si $h(X,Y) = X$ ó $h(X,Y) = Y$, entonces

$E[X] = \sum_{(x,y) \in A} x P(X=x, Y=y)$	$E[X] = \iint_{\mathbb{R}^2} x F_{XY}(x,y) dx dy$
$E[Y] = \sum_{(x,y) \in A} y P(X=x, Y=y)$	$E[Y] = \iint_{\mathbb{R}^2} y F_{XY}(x,y) dx dy$

NOTA

Teniendo las densidades marginales:

$$E[X]=\int_{\mathbb{R}} x\, f_X(x)\,dx \qquad E[Y]=\int_{\mathbb{R}} y\, f_Y(y)\,dy$$

→ $E[aX+bY]=aE[X]+bE[Y]$

→ si X e Y son independientes $\Rightarrow E[XY]=E[X].E[Y]$

VALOR ESPERADO DE UNA VARIABLE TRUNCADA

→ $$E[X|X\in A]=\frac{E[X\,\mathbb{1}\{X\in A\}]}{P(X\in A)}$$

Sea X una V.A. y sean $A_1, A_2, \ldots, A_n$ una partición de $\mathbb{R}$:

→ $$E[X]=\sum_{i=1}^{n} E[X|X\in A_i].\,P(X\in A_i)$$

VARIANZA

Es la dispersión de los valores de una variable alrededor de su media

$$V(X)=E[(X-E[X])^2]$$

→ $V(X)=E[X^2]-E^2[X]$

CASO ESPECIAL

Sea T una V.A. con distribución exponencial, entonces:

→ $E[T|T>a]=a+E[T]$

NOTA

* $X \sim BER(p) \Rightarrow V(X) = p(1-p)$
* $X \sim BIN(n,p) \Rightarrow V(X) = np(1-p)$
* $X \sim GEOM(p) \Rightarrow V(X) = \dfrac{1-p}{p^2}$
* $X \sim PAS(r,p) \Rightarrow r\dfrac{(1-p)}{p^2} = V(X)$
* $X \sim POISSON(\mu) \Rightarrow V(X) = \mu$
* $X \sim \mathcal{H}(N,r,n) \Rightarrow V(X) = \dfrac{nr}{N}\dfrac{(N-r)}{N}\dfrac{(N-n)}{(N-1)}$
* $X \sim U(a,b) \Rightarrow V(X) = \dfrac{(b-a)^2}{12}$
* $X \sim EXP(\lambda) \Rightarrow V(X) = \dfrac{1}{\lambda^2}$
* $X \sim GAMMA(n,\lambda) \Rightarrow V(X) = \dfrac{n}{\lambda^2}$
* $X \sim N(\mu,\sigma) \Rightarrow V(X) = \sigma^2$

PROPIEDADES DE LA VARIANZA

→ $V(K) = 0$
→ $V(aX+b) = a^2 V(X)$
→ $V(aX \pm bY) = a^2V(X) + b^2V(Y) \pm 2ab\,COV(X,Y)$

COVARIANZA

$$COV(X,Y) = E[(X-E[X])(Y-E[Y])]$$

→ $COV(X,Y) = E[XY] - E[X]E[Y]$

Si X e Y son independientes $\Rightarrow COV(X,Y) = 0$
El recíproco es falso

NOTA

PROPIEDADES DE LA COVARIANZA

→ $COV(X,X) = V(X)$

→ $COV(aX,Y) = a\,COV(X,Y)$

→ $COV(X,Y+Z) = COV(X,Y) + COV(X,Z)$

→ $COV(X,Y) = COV(Y,X)$

→ $COV(X,k) = 0$

DESVIO ESTÁNDAR

* $\sigma_x = \sqrt{V(X)}$

NOTACIÓN USUAL

* $E[X] = \mu_x$ * $V(X) = \sigma_x^2$

DESIGUALDAD DE MARKOV

$$P(X>t) \leq \frac{E[X]}{t}$$

DESIGUALDAD DE TCHEBBYCHEFF

$$P(|X-E[X]| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

RECTA DE REGRESIÓN

$$\tilde{y}(x) = \frac{COV(X,Y)}{V(X)}(x-E[X]) + E[Y]$$

NOTA

FUNCIONES DE VARIABLES ALEATORIAS

Sea X una VA y $Y = g(X)$ otra VA.

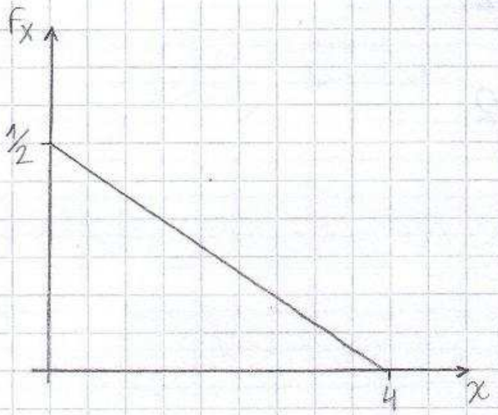
MÉTODO BÁSICO: EVENTOS EQUIVALENTES

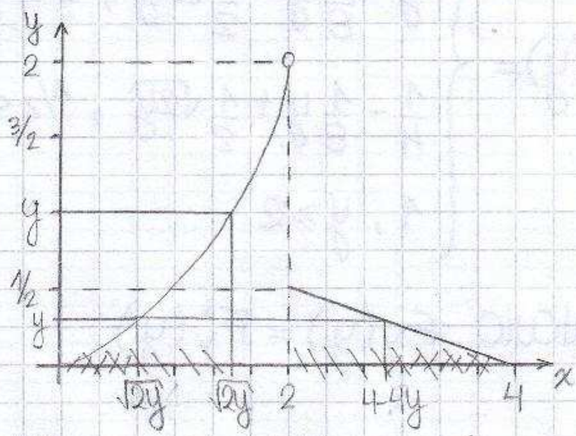
Si queremos hallar la función de distribución de $Y = g(X)$:

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

Ejemplo: Sea X con densidad $F_X(x) = \left(-\frac{1}{8}x + \frac{1}{2}\right) \mathbb{1}\{0 \leq x \leq 4\}$

$$\text{Sea } Y = \begin{cases} \frac{1}{2}X^2, & 0 < X < 2 \\ -\frac{1}{4}X + 1, & 2 \leq X < 4 \\ 0, & \text{otras } X \end{cases}$$





$P(0 \leq X \leq 4) = P(0 \leq Y \leq 2) = 1$

Calculo $F_X(x) = \int_{-\infty}^{x} F_X(x)\,dx = \int_0^x -\frac{1}{8}x + \frac{1}{2} = -\frac{1}{16}x^2 + \frac{1}{2}x \ \mathbb{1}\{0 \leq x \leq 4\}$

$$\Rightarrow F_X(x) = \left(-\frac{1}{16}x^2 + \frac{1}{2}x\right)\mathbb{1}\{0 \leq x \leq 4\} + \mathbb{1}\{x > 4\}$$

$F_Y(y) = P(Y \leq y) =$

I) $y < 0 \Rightarrow P(Y \leq y) = 0$ II) $y \geq 2 \Rightarrow P(Y \leq y) = 1$

NOTA

III) $0 \leq y \leq 1/2 \Rightarrow P(Y \leq y) = P(Y<0) + P(0 \leq y \leq y) =$

$= P(0 \leq x \leq \sqrt{2y}) + P(4-4y \leq x \leq 4) =$

$= F_x(\sqrt{2y}) - F_x(0) + F_x(4) - F_x(4-4y) =$

$= y^2 - \frac{1}{8}y + \frac{1}{2}\sqrt{2y}$

IV) $1/2 < y < 2 \Rightarrow P(Y \leq y) = P(Y \leq 1/2) + P(1/2 < Y \leq y) =$

$= \left(\frac{1}{2}\right)^2 - \frac{1}{8}\cdot\frac{1}{2} + \frac{1}{2}\sqrt{2\cdot\frac{1}{2}} + P(1 < X \leq \sqrt{2y}) =$

$= \frac{11}{16} + F_x(\sqrt{2y}) - F_x(1) =$

$= \frac{1}{4} - \frac{1}{8}y + \frac{1}{2}\sqrt{2y}$

$$F_Y(y) = \begin{cases} 0, & y<0 \\ y^2 - \frac{1}{8}y + \frac{1}{2}\sqrt{2y}, & 0 \leq y \leq 1/2 \\ \frac{1}{4} - \frac{1}{8}y + \frac{1}{2}\sqrt{2y}, & 1/2 < y < 2 \\ 1, & y \geq 2 \end{cases}$$

y calculo $f_Y(y) = F'_Y(y)$

TEOREMA: sea X una V.A. continua con función de distribución creciente, entonces $Y = F_X(X) \sim U(0,1)$

dem: $F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y))$

$= y$

NOTA

FUNCIONES INYECTIVAS SUAVES

Sea X una VA continua y $Y = g(X)$ una función monótona con derivada no nula. Entonces Y es continua y admite una densidad de probabilidades de la forma

$$\Rightarrow f_Y(y) = \left.\frac{f_X(x)}{|g'(x)|}\right|_{x = g^{-1}(y)}$$

dem: $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(g^{-1}(y))}{dy} = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}$$

FUNCIONES DE VECTORES ALEATORIOS

Sea (X,Y) un vector aleatorio y $Z = g(X,Y)$ una VA.

MÉTODO BÁSICO: EVENTOS EQUIVALENTES

Si queremos hallar la función de distribución de $Z = g(X,Y)$

$$F_Z(z) = P(Z \leq z) = P((X,Y) \leq g^{-1}(z)) = F_{XY}(g^{-1}(z))$$

MÉTODO DEL JACOBIANO

Sea (X,Y) continuo con densidad conjunta f_{XY} con soporte R acotado, sea $(U,V) = T(X,Y)$ una función de $\mathbb{R}^2$ en $\mathbb{R}^2$ uno a uno:

$$\Rightarrow \iint_R f_{XY}(x,y)\,dx\,dy = 1 = \iint_D f_{XY}(x(u,v), y(u,v)) \times \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv$$

NOTA

$$\Rightarrow f_{UV}(u,v) = \frac{f_{XY}(x(u,v), y(u,v))}{\left|\frac{\partial(U,V)}{\partial(X,Y)}\right|}$$

Si $(U,V) = T(X,Y)$ no es uno a uno, entonces:

$$f_{UV}(u,v) = \frac{f_{XY}(T_1^{-1}(u,v))}{J_1} + \frac{f_{XY}(T_2^{-1}(u,v))}{J_2}$$

MÍNIMO Y MÁXIMO DE DOS EXPONENCIALES INDEPENDIENTES

Sean X_1 y X_2 dos V.A. independientes

$X_1 \sim EXP(\lambda_1) \qquad X_2 \sim EXP(\lambda_2)$

$U = \min(X_1, X_2) \qquad V = \max(X_1, X_2) \qquad W = V - U$

$J = \mathbb{1}\{U = X_1\} + 2\,\mathbb{1}\{U = X_2\}$

Entonces:

1) $U \sim EXP(\lambda_1 + \lambda_2)$

2) $P(J = i) = \dfrac{\lambda_i}{\lambda_1 + \lambda_2}$

3) U y J son independientes

4) $f_W(w) = f_{X_2}(w) \cdot P(J=1) + f_{X_1}(w) \cdot P(J=2)$

$$= \lambda_2 e^{-\lambda_2 w} \cdot \frac{\lambda_1}{\lambda_1+\lambda_2} + \lambda_1 e^{-\lambda_1 w} \cdot \frac{\lambda_2}{\lambda_1+\lambda_2} \; \mathbb{1}\{w > 0\}$$

5) U y W son independientes

PROPIEDADES DE ALGUNAS VARIABLES

* Sean $X_1, X_2, \dots, X_n$ V.A. independientes, sea $U = \min(X_1, X_2 \dots, X_n)$ entonces $P(U > u) = P(X_1 > u, X_2 > u, \dots, X_n > u) =$

$$= P(X_1 > u) \cdot P(X_2 > u) \dots P(X_n > u)$$

$\Rightarrow F_U(u) = P(U \leq u) = 1 - P(U > u)$

* Sean $X_1, X_2, \dots, X_n$ V.A. independientes, sea $V = \max(X_1, X_2, \dots, X_n)$ entonces $P(V \leq v) = P(X_1 \leq v, X_2 \leq v, \dots, X_n \leq v)$

$$= P(X_1 \leq v) \cdot P(X_2 \leq v) \dots P(X_n \leq v)$$

$$\Rightarrow F_V(v) = F_{X_1}(v) \cdot F_{X_2}(v) \dots F_{X_n}(v)$$

* Sean X_1 y X_2 V.A. independientes con distribuciones GAMMA(n_1, λ) y GAMMA(n_2, λ) respectivamente.

Sean $Y_1 = X_1 + X_2$ y $Y_2 = \dfrac{X_1}{X_1 + X_2}$

$\Rightarrow Y_1$ e Y_2 son independientes y sus distribuciones son $Y_1 \sim \text{GAMMA}(n_1 + n_2, \lambda)$ y $Y_2 \sim \beta(n_1, n_2)$

* Sean $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$ independientes, entonces $Y = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

* Sean $X_1 \sim P_o(\mu_1)$ y $X_2 \sim P_o(\mu_2)$ independientes, entonces $Y = X_1 + X_2 \sim P_o(\mu_1 + \mu_2)$

NOTA

DISTRIBUCIONES CONDICIONALES

CASO DISCRETO

$$\Rightarrow F_{Y|X=x}(y) = \frac{P(Y \leq y, X=x)}{P(X=x)}$$

$$\Rightarrow p_{Y|X=x}(y) = P(Y=y \mid X=x) = \frac{P(Y=y, X=x)}{P(X=x)}$$

$$= \frac{p_{XY}(x,y)}{p_X(x)}$$

$$E[Y|X=x] = \sum_{\forall y} y \, p_{Y|X=x}(y)$$

CASO CONTINUO

Si se quiere definir $P(Y \leq y \mid X=x)$ NO PODEMOS UTILIZAR LA DEFINICIÓN DE PROBABILIDAD CONDICIONAL

$$\cancel{F_{Y|X=x}(y) = P(Y \leq y \mid X=x) = \frac{P(Y \leq y \wedge X=x)}{P(X=x)}} = \frac{0}{0}$$

Para superar esta dificultad, consideramos el intervalo $(x-h, x+h)$ tal que $P(x-h \leq X \leq x+h) > 0$

$$P(Y \leq y \mid x-h \leq X \leq x+h) = \frac{P(Y \leq y, x-h \leq X \leq x+h)}{P(x-h \leq X \leq x+h)} =$$

$$= \frac{\int_{x-h}^{x+h} \int_{-\infty}^{y} f_{XY}(s,t)\, dt\, ds}{\int_{x-h}^{x+h} f_X(s)\, ds}$$

NOTA

$$\lim_{h\to 0} P(Y\leq y \mid x-h \leq X \leq x+h) = P(Y\leq y \mid X=x)$$

$$= \frac{\int_{-\infty}^{y} f_{XY}(x,t)\,dt}{f_X(x)}$$

La expresión obtenida define una función de distribución de $Y|X=x$

$$\Rightarrow F_{Y|X=x}(y) = P(Y\leq y \mid X=x) = \frac{\int_{-\infty}^{y} f_{XY}(x,t)\,dt}{f_X(x)}$$

$$\Rightarrow f_{Y|X=x}(y) = \frac{f_{XY}(x,y)}{f_X(x)}$$ DENSIDAD CONDICIONAL DE $Y|X=x$

Análogamente:

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

$$\Rightarrow E[Y|X=x] = \int_{-\infty}^{+\infty} y\, f_{Y|X=x}(y)\,dy$$

FUNCIÓN DE REGRESIÓN DE $Y|X=x$

$$\varphi(x) = E[Y|X=x]$$

COEFICIENTE DE CORRELACIÓN

$$\rho(X,Y) = \frac{COV(X,Y)}{\sigma_X \times \sigma_Y} = \frac{COV(X,Y)}{\sqrt{V(X)\,V(Y)}}$$

* si vale cero no hay correlación entre las variables
* si $|\rho|=1 \Rightarrow P(Y=aX+b)=1 \Rightarrow$ relación lineal

NOTA

IMPORTANTE: si $\varphi(x)$ es de la forma $ax+b$ (lineal) entonces la RECTA DE REGRESIÓN ES LA FUNCIÓN DE REGRESIÓN

$$\Rightarrow a = \frac{COV(X,Y)}{V(X)}$$

VARIANZA CONDICIONAL DE $Y|X=x$

$$V(Y|X=x) = E[(Y-E[Y|X=x])^2|X=x]$$

$$V(Y|X=x) = E[Y^2|X=x] - E^2[Y|X=x]$$

MEZCLA DE VARIABLES ALEATORIAS

Sea M una V.A. discreta tal que $M=\{m_1, m_2, \dots, m_n\}$, sean $X_{m_1}, X_{m_2}, \dots, X_{m_n}$ n variables aleatorias independientes de M, la variable X es la mezcla de las variables X_{m_i} si se define $X|M=m \sim X_m$

(La distribución de probabilidades de M indica la proporción en que deben mezclarse las variables X_m; para cada $m \in M$, la probabilidad $p_M(m)$ representa la proporción con que la variable X_m participa de la mezcla X)

La función de distribución de la mezcla X se obtiene utilizando la fórmula de probabilidad total:

$$\Rightarrow F_X(x) = P(X \leq x) = \sum_{m \in M} P(X \leq x \mid M=m) \cdot P(M=m)$$

$$= \sum_{m \in M} P(X_m \leq x \mid M=m) \cdot P(M=m)$$

$$= \sum_{m \in M} P(X_m \leq x) \cdot P(M=m)$$

$$\Rightarrow F_X(x) = \sum_{m \in M} F_{X_m}(x) \cdot P(M=m)$$

* Si las X_m son discretas:

$$\Rightarrow p_X(x) = \sum_{m \in M} p_{X_m}(x) \cdot p_M(m)$$

* Si las X_m son continuas:

$$\Rightarrow f_X(x) = \sum_{m \in M} f_{X_m}(x) \cdot p_M(m)$$

SOBRE LA REGLA DE BAYES (caso continuo)

$$\cancel{P(M=m \mid X=x) = \frac{P(M=m \wedge X=x)}{P(X=x)} = \frac{0}{0}}$$

$\Rightarrow$ se resuelve considerando un intervalo $(x-h, x+h)$ y calculando $P(M=m \mid x-h \leq X \leq x+h)$ y tomando el límite $h \to 0$ se obtiene

$$\Rightarrow P(M=m \mid X=x) = \frac{f_{X_m}(x) \cdot p_M(m)}{f_X(x)}$$

FÓRMULA DE BAYES PARA MEZCLA DE VARIABLES CONTINUAS Y CONDICIÓN DE PROBABILIDAD CERO

NOTA

ESPERANZA CONDICIONAL Y|X

$\Rightarrow \varphi(X) = E[Y|X]$ (ES UNA V.A.)

PROPIEDADES DE LA ESPERANZA CONDICIONAL

I) $E[E[Y|X]] = E[Y]$

II) $E[g(X) \cdot Y | X] = g(X) E[Y|X]$

III) Si X e Y son independientes $E[Y|X] = E[Y]$

IV) $E[aY_1 + bY_2 | X] = aE[Y_1|X] + bE[Y_2|X]$

VARIANZA CONDICIONAL

$\Rightarrow \psi(X) = V(Y|X)$

TEOREMA DE PITÁGORAS

$\Rightarrow V(Y) = E[V(Y|X)] + V(E[Y|X])$

SUMAS ALEATORIAS DE VARIABLES ALEATORIAS

Sea $X_1, X_2, \dots, X_n$ una sucesión de variables aleatorias identicamente distribuidas de media μ y varianza σ^2.
Sea N una V.A. discreta a valores en $\mathbb{N}$ independiente de las X_i.
Sea $S = \sum_{i=1}^{n} X_i$, hallar su media y varianza.

$\Rightarrow E[S] = E[E[S|N]]$ $\Rightarrow V(S) = E[V(Y|X)] + V(E[Y|X])$

* $E[S|N=n] = E\left[\sum_{i=1}^{n} X_i \,\middle|\, N=n\right] = E\left[\sum_{i=1}^{n} X_i\right] =$

$$= \sum_{i=1}^{n} E[X_i] = n\mu$$

* $E[S|N] = N\mu \Rightarrow E[E[S|N]] = E[N\mu] = \mu E[N]$

* $V(S|N=n) = V\left(\sum_{i=1}^{n} X_i \,\middle|\, N=n\right) = V\left(\sum_{i=1}^{n} X_i\right) =$

$$= \sum_{i=1}^{n} V(X_i) = n\sigma^2$$

* $V(S|N) = N\sigma^2 \Rightarrow E[V(S|N)] = E[N\sigma^2] = \sigma^2 E[N]$

$$\Rightarrow V(E[S|N]) = V(N\mu) = \mu^2 V(N)$$

$\Rightarrow E[S] = \mu E[N]$

$\Rightarrow V(S) = \sigma^2 E[N] + \mu^2 V(N)$

ESPERANZA Y VARIANZA DE UNA MEZCLA

$\Rightarrow E[X] = E[E[X|M]]$ $\quad \Rightarrow V(X) = E[V(X|M)] + V(E[X|M])$

* $E[X] = E[E[X|M]] = \sum_{m \in M} E[X|M=m]\, P(M=m) = \sum_{m\in M} E[X_m]\, p_M(m)$

* $E[V(X|M)] = \sum_{m\in M} V(X|M=m)\, P(M=m) = \sum_{m \in M} V(X_m)\, p_M(m)$

* $V(E[X|M]) = E[(E[X|M] - E[X])^2] =$

$$= \sum_{m\in M} (E[X|M=m] - E[X])^2\, P(M=m) =$$

$$= \sum_{m \in M} (E[X_m] - E[X])^2\, p_M(m)$$

NOTA

$$\Rightarrow E[X] = \sum_{m \in M} E[X_m]\, p_M(m)$$

$$\Rightarrow V(X) = \sum_{m \in M} V(X_m)\, p_M(m) + \sum_{m \in M} (E[X_m] - E[X])^2\, p_M(m)$$

Para el caso especial de que X sea mezcla de dos variables X_1 y X_2 con una variable M con dos átomos, por ejemplo $\{1,2\}$, tal que $X|M=1 \sim X_1$ y $X|M=2 \sim X_2$, la fórmula anterior resulta:

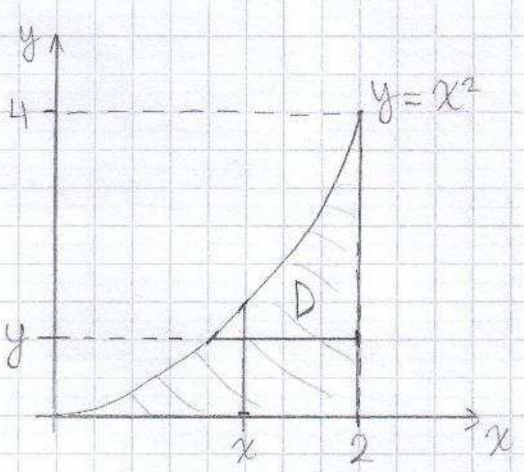
$$\Rightarrow V(X) = V(X_1)P(M=1) + V(X_2)P(M=2) + P(M=1)P(M=2)(E[X_1] - E[X_2])^2$$

De todas formas, siempre puede calcularse usando:

$$\Rightarrow V(X) = E[X^2] - E^2[X]$$

EJERCICIO IMPORTANTE

Sea $(X,Y) \sim U(D)$ con D acotado



$\Rightarrow Y|X=x \sim U(0, x^2)$

$\Rightarrow X|Y=y \sim U(\sqrt{y}, 2)$

NOTA

PROCESOS DE BERNOULLI

Un proceso de Bernoulli es el modelo que se asocia a una secuencia de ensayos de Bernoulli independientes con probabilidad de éxito constante p. A este proceso se asocia una sucesión de variables $X_i \sim Ber(p)$ independientes.

En un proceso de Bernoulli podemos definir 3 variables aleatorias:

1) V.A. BINOMIAL: nro de éxitos en n ensayos

$X \sim BIN(n, p)$

$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad , x = 0, 1, 2, \ldots, n$

$E[X] = np \qquad V(X) = np(1-p)$

NEWTON $\rightarrow (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \; , n \in \mathbb{N}$

$X = \sum_{k=1}^{n} X_k$ ← Ber(p)

2) V.A. GEOMÉTRICA: nro de ensayos hasta el primer éxito

$N \sim GEOM(p)$

$p_N(n) = p(1-p)^{n-1} \; , n \in \mathbb{N}$

$E[N] = \frac{1}{p} \qquad V(N) = \frac{1-p}{p^2}$

$P(N > n) = (1-p)^n$

NOTA

3) V.A. PASCAL : nro de ensayos hasta r éxitos

$N \sim PAS(r, p)$

$$\underbrace{\_\ \_\ E\ \_\ \_\ \_\ E\ \_\ \_\ |\ E}_{n \text{ tiros}}$$

$$p_N(n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \quad n = r, r+1, r+2, \ldots$$

Propiedad:

$$N \sim PAS(r, p) \Rightarrow N = \sum_{k=1}^{r} \underset{\uparrow\, GEOM(p)}{N_k}$$

$$\underbrace{\_\ \_\ E}_{N_1}\ |\ \underbrace{\_\ \_\ \_\ \_\ E}_{N_2}\ |\ \underbrace{\_\ \_\ E}_{N_3} \quad \rightarrow r \text{ veces}$$

Entonces: $E[N] = \frac{r}{p}$ $\qquad V(N) = r\left(\frac{1-p}{p^2}\right)$

RELACIÓN ENTRE LAS DISTRIBUCIONES BINOMIAL Y PASCAL

$X \sim BIN(n, p)$

$N \sim PAS(r, p)$

$$\Rightarrow P(N > n) = P(X < r) = P(X \leq r-1) = \sum_{x=0}^{r-1} \binom{n}{x} p^x (1-p)^{n-x}$$

$P(N > n)$: nro de ensayos hasta r exitos sea mayor a n

$P(X < r)$: nro de éxitos en n ensayos sea menor a r

Es decir, decir que en n ensayos Bernoulli ocurren por lo menos r éxitos es lo mismo que decir que el tiempo de espera hasta observar el r-ésimo éxito no supera a n $(P(N \leq n) = P(X \geq r))$

NOTA

VECTOR ALEATORIO MULTINOMIAL

$(X_1, X_2, \ldots, X_r) \sim \text{MULTINOMIAL}(n, p_1, p_2, \ldots, p_r)$

X_k es la cantidad de veces que ocurre el resultado k en n ensayos. La probabilidad de que en n ensayos el resultado 1 ocurra n_1 veces, el resultado 2 ocurra n_2 veces, etc. es:

$$\rightarrow P(X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r) = \frac{n!}{n_1! n_2! \ldots n_r!} p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r}$$

con $n = n_1 + n_2 + \ldots + n_r$

Propiedades:

$\rightarrow X_i \sim \text{BIN}(n, p_i)$

$\rightarrow X_i \mid X_j = x \sim \text{BIN}\left(n - x, \frac{p_i}{1 - p_j}\right)$

TÉRMINO CENTRAL

El valor de x que hace máxima la probabilidad de una binomial es:

$$\rightarrow (n+1)p - 1 \leq x_{MAX} \leq (n+1)p$$

siendo $X \sim \text{BIN}(n, p)$

NOTA

DISTRIBUCIÓN DE POISSON

Una VA N a valores enteros no negativos sigue una distribución de Poisson de parámetro $\mu > 0$ si y solo si N tiene una función de probabilidad dada por:

$$\rightarrow p_N(n) = P(N=n) = \frac{\mu^n e^{-\mu}}{n!} \quad , \; n = 0, 1, 2, \ldots$$

$\rightarrow N \sim POISSON(\mu)$

$\rightarrow E[N] = \mu$

$\rightarrow V(N) = \mu$

$$\rightarrow \frac{P(N=n)}{P(N=n-1)} = \frac{\frac{\mu^n e^{-\mu}}{n!}}{\frac{\mu^{n-1} e^{-\mu}}{(n-1)!}} = \frac{\mu}{n} \Rightarrow p_n = \frac{\mu}{n} p_{n-1}$$

Propiedades

1) La suma de VA de Poisson independientes es otra VA de Poisson

$N_1 \sim POISSON(\mu_1) \wedge N_2 \sim POISSON(\mu_2)$ INDEPENDIENTES

$\Rightarrow N_1 + N_2 \sim POISSON(\mu_1 + \mu_2)$

$$\rightarrow P(N_1 + N_2 = n) = \sum_{m=0}^{n} P(N_1 = m, N_2 = n-m)$$

NOTA

2) Competencia de variables de Poisson independientes (distribución conjunta de variables de Poisson independientes condicional a la suma)

→ Sean $N_1, N_2, \dots, N_m$ VA independientes tal que $N_i \sim \text{POISSON}(\mu_i)$

→ Sea $S = N_1 + N_2 + \dots + N_m$, entonces

$$(N_1, N_2, \dots, N_m) \mid S = n \sim \text{MULTINOMIAL}\left(n, \frac{\mu_1}{\mu}, \frac{\mu_2}{\mu}, \dots, \frac{\mu_m}{\mu}\right)$$

$$\Rightarrow P(N_1 = n_1, N_2 = n_2, \dots, N_r = n_r) = \frac{n!}{n_1! n_2! \dots n_r!} p_1^{n_1} p_2^{n_2} \dots p_r^{n_r}$$

donde $n = n_1 + n_2 + \dots + n_r$ y $p_i = \frac{\mu_i}{\mu}$

$$\Rightarrow (N_1, N_2, \dots, N_r) \sim \text{MULTINOMIAL}\left(n, \frac{\mu_1}{\mu}, \frac{\mu_2}{\mu}, \dots, \frac{\mu_r}{\mu}\right)$$

$$\Rightarrow N_i \sim \text{BIN}\left(n, \frac{\mu_i}{\mu}\right)$$

$$\Rightarrow N_i \mid S = n \sim \text{BIN}\left(n, \frac{\mu_i}{\mu}\right)$$

3) Adelgazamiento

Sea $N \sim \text{POISSON}(\mu)$ y sea M una VA discreta tal que $M \mid N = n \sim \text{BIN}(n, p)$, entonces:

→ $M \sim \text{POISSON}(p\mu)$

→ $N - M \sim \text{POISSON}((1-p)\mu)$

(N es una VA Poisson típicamente utilizada para "contar" eventos)

NOTA

PROCESOS DE POISSON

Un proceso de Poisson se encuadra dentro de lo que es un PROCESO PUNTUAL, el cual se define como una sucesión de variables S_n, $n \geq 0$, tal que S_n es el tiempo hasta el n-ésimo arribo. Las S_n satisfacen $S_0 = 0$

$$0 < S_1 < S_2 < \dots < S_n$$

$$\lim_{n \to \infty} S_n = +\infty$$

→ Tiempo entre arribos: $T_k = S_k - S_{k-1}$

→ $S_n = \sum_{k=1}^{n} T_k$

→ $N(t)$: nro de eventos en $(0, t]$

→ $N(t_1, t_2]$: nro de eventos en $(t_1; t_2]$

Relación fundamental entre las variables S_n y $N(t)$:

→ $P(S_n > t) = P(N(t) < n)$

Un proceso de Poisson está caracterizado por una constante positiva λ llamada "intensidad del proceso"

Un proceso de Poisson de intensidad λ es proceso puntual tal que:

→ para todo conjunto de variables $N(t_1, t_2]$ asociados a intervalos no superpuestos, se cumple que dichas variables son independientes

NOTA

→ las variables $N(t_1, t_2]$ siguen una distribución de Poisson de media $\lambda.\Delta t$. Entonces la distribución de estas variables no depende de la posición del intervalo sino solo de su amplitud.

De estos últimos dos puntos se desprende que si el proceso se reinicia a partir de cualquier punto, el proceso resultante no depende de todo lo anterior, es decir, el proceso no tiene memoria

Propiedades

1) Los tiempos T_K entre eventos consecutivos son variables independientes exponenciales de parámetro λ

$$T_K \sim EXP(\lambda)$$

Sea $T_1, T_2, \dots, T_n$ una sucesión de VA exponenciales independientes todas de parámetro λ; si se define un proceso de arribos tal que $S_n = \sum_{K=1}^{n} T_K$ entonces este proceso es un proceso de Poisson de intensidad λ

2) $S_n = \sum_{K=1}^{n} T_K \sim GAMMA(n, \lambda)$

NOTA

3) Sea Π un proceso de Poisson de intensidad λ. Dado el evento $N(t)=n$ (en el intervalo $(0,t]$ ocurrieron n arribos), los n puntos se distribuyen al azar en el intervalo $(0,t]$, lo cual significa que la posición de cada arribo es una variable uniforme en el intervalo.

Esta propiedad se conecta directamente con la propiedad vista de competencia

$N(t) \sim \text{POISSON}(\lambda t)$ $\qquad$ $N(\Delta) \sim \text{POISSON}(\lambda \Delta)$

$$\Rightarrow N(\Delta) \mid N(t)=n \sim \text{BIN}(n, \Delta/t)$$

Cada uno de los n arribos tiene una probabilidad $\frac{\Delta}{t}$ de ocurrir en $(0,\Delta]$

4) Sea Π un proceso de Poisson de intensidad λ

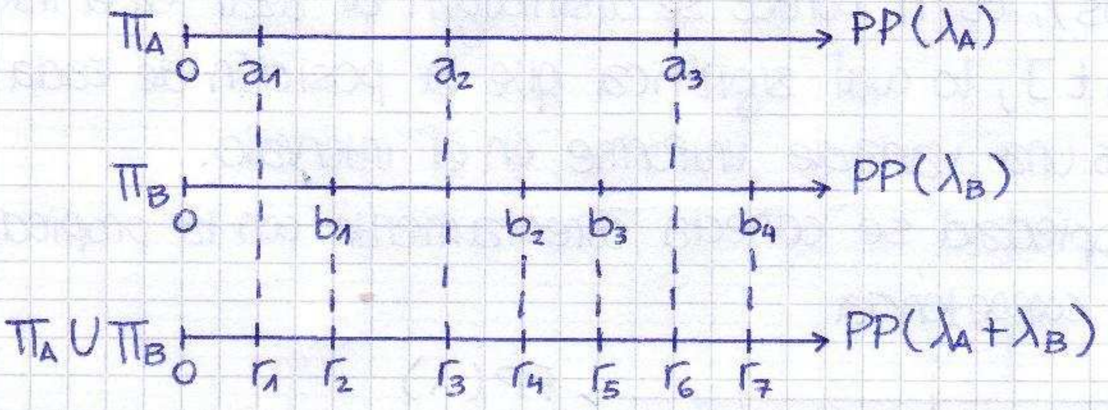
ADELGAZAMIENTO: si cada evento o punto del proceso tiene una probabilidad p de ser "éxito" y $1-p$ de ser "fracaso" con independencia de todos los puntos. Entonces el proceso Π_E de arribo de éxitos y el proceso Π_F de arribo de fracasos son ambos procesos de Poisson independientes:

$\rightarrow \Pi_E \sim PP(p\lambda)$ $\qquad$ $\rightarrow \Pi_F \sim PP((1-p)\lambda)$

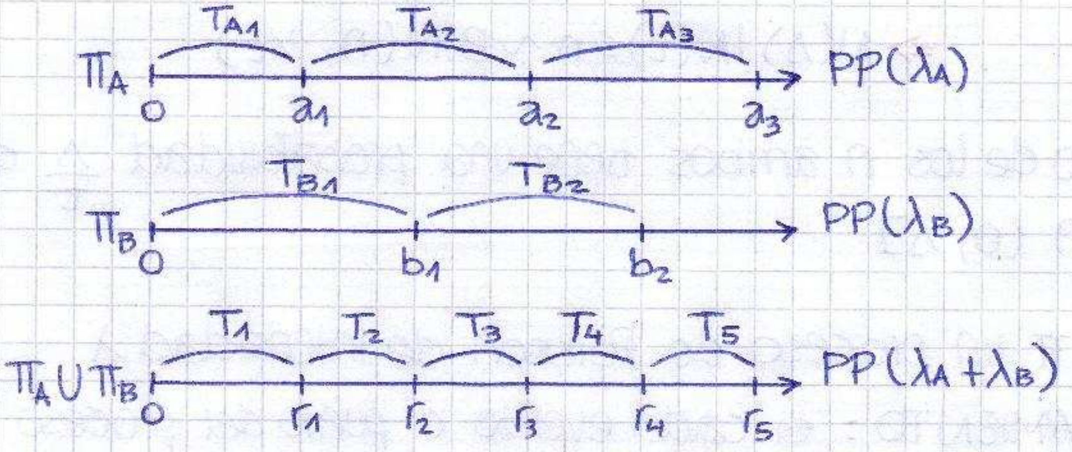
NOTA

5) Superposición de procesos de Poisson independientes:

Sean $\Pi_A \sim PP(\lambda_A)$ y $\Pi_B \sim PP(\lambda_B)$ dos procesos de Poisson independientes. Entonces la unión de ambos procesos $\Pi_A \cup \Pi_B$ es un proceso de Poisson de intensidad $\lambda_A + \lambda_B$



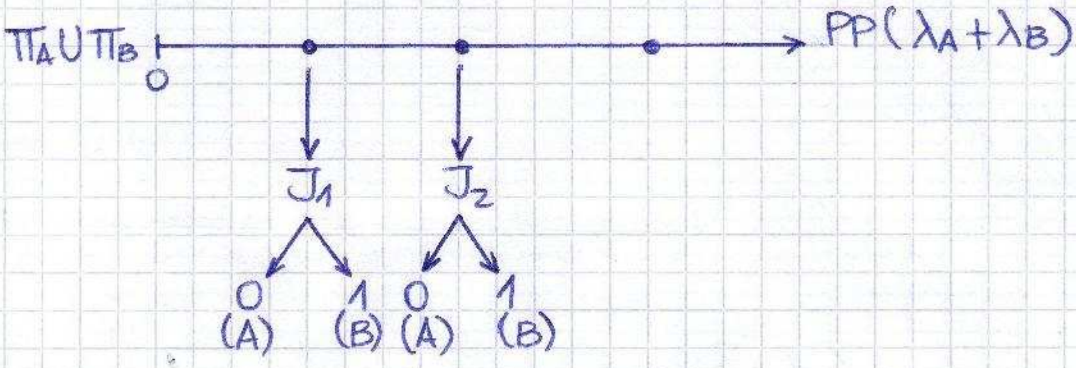
P("ocurran simultáneamente un arribo de Π_A y un arribo de Π_B") = 0



$\rightarrow T_{Ak} \sim EXP(\lambda_A)$ $\quad\rightarrow T_{Bk} \sim EXP(\lambda_B)$

$\Rightarrow T_k \sim EXP(\lambda_A + \lambda_B)$

notar que T_k sigue la distribución del mínimo entre T_{Ak} y T_{Bk}



$\rightarrow P(J_K = 0) = P(T_{A_i} < T_{B_j}) = \dfrac{\lambda_A}{\lambda_A+\lambda_B}$

Para cada evento definimos una VA J_K que indica si el evento proviene de Π_A o de Π_B. Las J_K resultan independientes entre sí $\Rightarrow$ al proceso de Poisson se le asocia una sucesión de variables BER(p)

ENTONCES EXISTE UN PROCESO DE BERNOULLI ASOCIADO INDEPENDIENTE DEL PROCESO DE POISSON.

<u>RECORDAR</u>: J_K y T_K son independientes

$\rightarrow P(J_K = 0 \wedge T_K > t) =$

$= P(J_K = 0) . P(T_K > t) =$

$= \dfrac{\lambda_A}{\lambda_A+\lambda_B} . e^{-(\lambda_A+\lambda_B)t}$

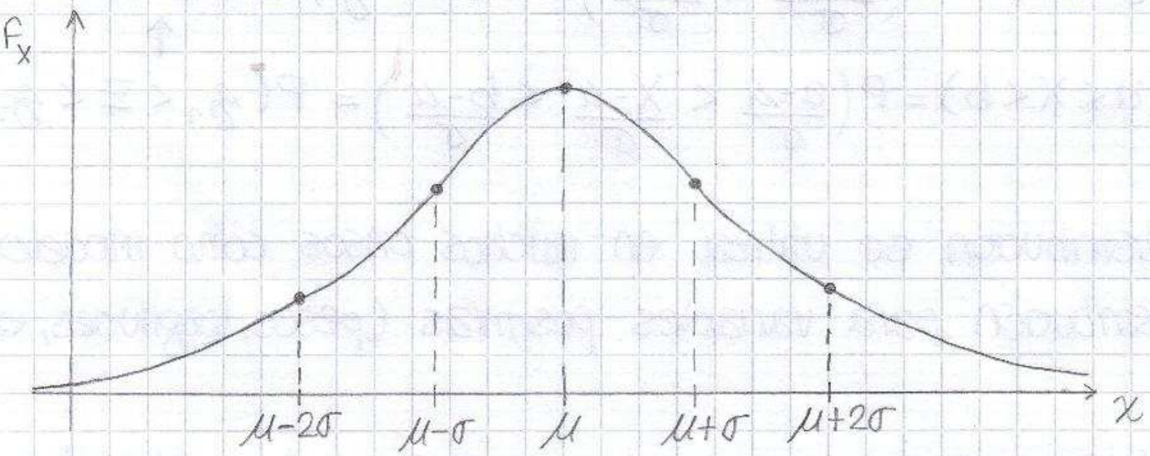
NOTA

DISTRIBUCIÓN NORMAL

$$X \sim N(\mu, \sigma^2) \Rightarrow F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R}$$

→ $E[X] = \mu$

→ $V(X) = \sigma^2$



→ $Z \sim N(0,1)$ VA NORMAL ESTÁNDAR

Propiedad

1) $X \sim N(\mu, \sigma^2) \Rightarrow Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$

2) $X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Suma de VA normales independientes:

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, X_1 y X_2 indep.

$$\Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

y se extiende a n variables

NOTA

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, X_1 y X_2 indep.

$\Rightarrow aX_1 + bX_2 \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$

NOTACIÓN: $\rightarrow Z \sim N(0,1)$

$\rightarrow \varphi(z) = f_Z(z)$

$\rightarrow \Phi(z) = F_Z(z) = P(Z \leq z)$

$\rightarrow P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P(Z \leq z) = \text{TABLA}$

$\rightarrow P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) = P(z_1 < Z < z_2)$

La distribución se utiliza en muchos casos como modelo de distribución para variables positivas (pesos, longitudes, etc.)

El modelo se considera aceptable siempre que la probabilidad de la parte negativa sea prácticamente cero.

$\rightarrow P(-x_1 < X < x_2) = P(-z < Z < z) = 2\Phi(z) - 1$

$\rightarrow P(\mu - \sigma < X < \mu + \sigma) = P(-1 < Z < 1) \cong 0{,}68$

$\rightarrow P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < Z < 2) \cong 0{,}95$

$\rightarrow P(\mu - 3\sigma < X < \mu + 3\sigma) = P(-3 < Z < 3) \cong 0{,}99$

$\rightarrow \Phi(-x) = 1 - \Phi(x)$

NOTA

TEOREMA CENTRAL DEL LÍMITE (TCL)

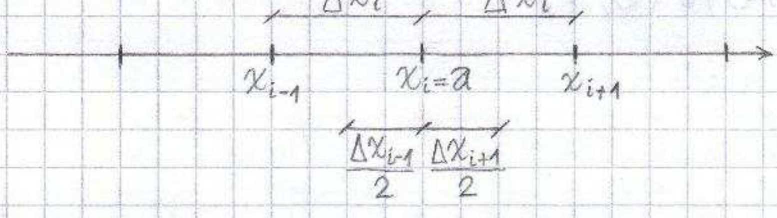
Sea $X_1, X_2, \ldots, X_n$ una sucesión de variables aleatorias independientes idénticamente distribuidas, cada una con media μ y varianza σ^2. Sea $S_n = \sum_{k=1}^{n} X_k$, entonces la distribución de S_n tiende a la distribución normal de media $n\mu$ y varianza $n\sigma^2$ cuando n tiende a infinito.

$$\rightarrow S_n \approxeq N(n\mu, n\sigma^2)$$

Es decir que $\dfrac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$ tiende a la normal estándar cuando $n \rightarrow \infty$

CORRECCIÓN POR CONTINUIDAD

Cuando se aproxima una VA discreta con una variable normal, se puede mejorar la aproximación utilizando la llamada CORRECCIÓN POR CONTINUIDAD, la cual consiste en asignar a cada punto de la variable discreta una probabilidad de la siguiente manera:



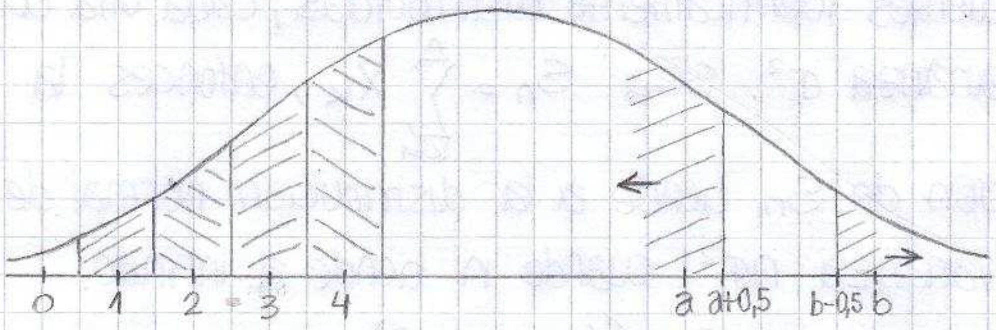
$$\rightarrow \underbrace{P(X = a)}_{\text{VA. DISCRETA}} = P\left(a - \frac{\Delta x_{i-1}}{2} \leq \underbrace{X'}_{\text{VA. NORMAL}} \leq a + \frac{\Delta x_{i+1}}{2}\right)$$

Por ejemplo, si X es binomial, todos los Δx_i son iguales a 1

Entonces si $X \sim BIN(n,p) \approx N(np, np(1-p))$

→ $P(X=a) = P(a-0{,}5 \leq X' \leq a+0{,}5)$

Luego debe estandarizarse X' y sacar los valores de tabla



→ $P(X \leq a) = P(X' \leq a+0{,}5)$

→ $P(X \geq b) = P(X' \geq b-0{,}5)$

APROXIMACIÓN POR LA DENSIDAD NORMAL

Sea $S_n \sim BIN(n,p)$, entonces

→ $P(S_n = k) \cong \dfrac{1}{\sqrt{np(1-p)}}\, \varphi\left(\dfrac{k-np}{\sqrt{np(1-p)}}\right)$

con $\varphi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

* Si $N \sim POISSON(\lambda)$ con λ "grande", entonces

→ $\dfrac{N-\lambda}{\sqrt{\lambda}} \sim N(0,1)$