Probabilidad y Estadística

Clase 1

Cronograma

Clase 1	Repaso Distribuciones útiles.
Clase 2	Transf de v.a.
Clase 3	V.a. condicionadas Esperanza condicional ECM y estimadores de cuadrados mínimos
Clase 4	Estimación Bayesiana Estimador de máxima verosimilitud
Clase 5	Estimación no paramétrica Intervalos de confianza
Clase 6	Intervalos de confianza Test de hipótesis
Clase 7	Repaso
Clase 8	Examen

Repaso

Eventos

Espacios equiprobables

Si estamos en presencia de un espacio **equiprobable**, es decir donde todos los elementos tienen las mismas chances de ocurrir, las probabilidades pueden calcularse como la proporción entre la cantidad de casos donde ocurre el experimento y la cantidad de elementos que existen en el espacio muestral. A esto se lo conoce como regla de **Laplace**

$$\mathbb{P}(A) = \frac{\#A}{\#\Omega} = \frac{\#\text{"casos favorables"}}{\#\text{casos totales}}$$

Probabilidades condicionales y proba. total

Def: Se llama **probabilidad condicional** de A dado B ($\mathbb{P}(A|B)$) a la probabilidad de que ocurra el evento A sabiendo que B ha ocurrido, y está definida por

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

Def: Diremos que los eventos $B_1, \ldots B_n$ forman una **partición** si $B_i \cap B_j = \emptyset \, \forall \, i, j$ y $\bigcup_{i=1}^{n} B_j = \Omega$.

Luego podemos describir al evento A como $A = (A \cap B_1) \cup \ldots \cup (A \cap B_n)$ de forma que

$$\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \ldots + \mathbb{P}(A \cap B_n) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \ldots + \mathbb{P}(A|B_n)\mathbb{P}(B_n)$$

Fórmula de probabilidad total

Teorema de Bayes e independencia

Teorema de Bayes: Sean $B_1, \dots B_n$ una partición de Ω , y A un evento con probabilidad positiva:

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

Def: Diremos que dos eventos A y B son **independientes** si y sólo sí vale que

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

Variables aleatorias

Variables aleatorias

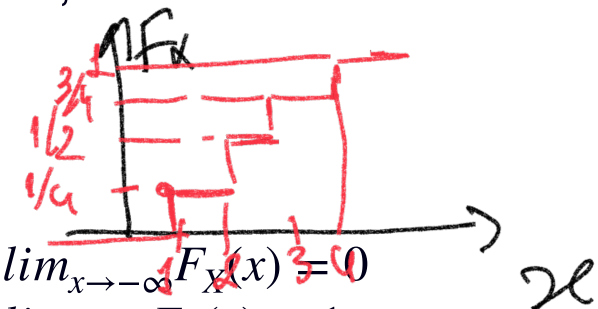
Una va. X es una función que mapea eventos a números.

$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$

Ejemplo: basto →1, Oro →2, espada →3, copa→4

X tiene asociada una función de distribución, definida como

$$F_X(x) = \mathbb{P}(X \leq x)$$



- $F_X(x) \geq 0$
- $F_X(x)$ Es no decreciente
- $lim_{x \to -\infty} F_X(x) = 0$
- $lim_{x \to \infty} F_X(x) = 1$

Tipos de v.a.

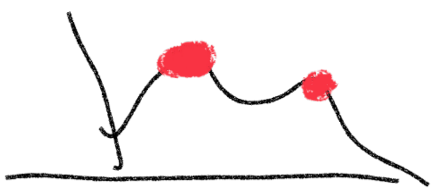
- Discretas (v.a.d): toman valores en un conjunto discreto o enumerable de puntos. Si X es **v.a.d**, tendrá además función de probabilidad dada por

Atomo (A) Conj. de valores que puede tomar X

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

$p_X(x) \geq 0$ $\qquad \sum_{x \in A} p_X(x) = 1$

- Continuas (v.a.c): toman valores en un intervalo continuo. Si X es una **v.a.c**. tendrá asociada una función de densidad $\mathbb{P}(X = x) = 0$



$$f_X(x) = \frac{dF_X(x)}{dx}$$

$\int_{\mathbb{R}} f_X(x)\, dx = 1$

$f_X(x) \geq 0$ $\qquad \mathbb{P}(a < X < b) = \int_a^b f_X(x)\,dx$

Tipos de v.a.

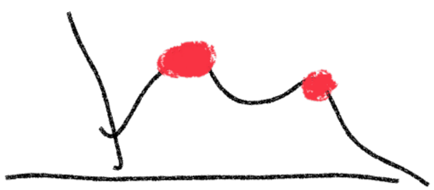
- Discretas (v.a.d): toman valores en un conjunto discreto o enumerable de puntos. Si X es **v.a.d**, tendrá además función de probabilidad dada por

Atomo (A) Conj. de valores que puede tomar X

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

$p_X(x) \geq 0$ $\quad \sum_{x \in A} p_X(x) = 1$

- Continuas (v.a.c): toman valores en un intervalo continuo. Si X es una **v.a.c**. tendrá asociada una función de densidad $\mathbb{P}(X = x) = 0$



$$f_X(x) = \frac{dF_X(x)}{dx}$$

$\int_{\mathbb{R}} f_X(x)\, dx = 1$

$f_X(x) \geq 0$ $\quad \mathbb{P}(a < X < b) = \int_a^b f_X(x)\, dx$

¿Qué propiedades debían cumplir cada una de estas funciones?

Vectores aleatorios

Distribución conjunta y marginales

Si tenemos dos variables X e Y se define su **función de distribución conjunta** como

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

En este caso, vale la regla del rectángulo

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c) + F_{X,Y}(a,c)$$

Caso continuo: $f_{X,Y}(x,y)$ es la función de densidad conjunta y se definen las funciones de densidad marginales como $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)dy$ y $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)dx$

Caso discreto: $p_{X,Y}(x,y)$ es función de probabilidad conjunta y se definen las funciones de probabilidad marginal como $p_X(x) = \sum_y p_{X,Y}(x,y)$ y $p_Y(y) = \sum_x p_{X,Y}(x,y)$

Independencia de v.a.

Diremos que dos v.a. X e Y son independientes si vale que

$$F_{X.Y}(x,y) = F_X(x)F_Y(y) \quad \forall\, x, y \in \mathbb{R}$$

Caso discreto:

$$p_{X.Y}(x,y) = p_X(x)p_Y(y) \quad \forall\, x, y \in \mathbb{R}$$

Caso continuo:

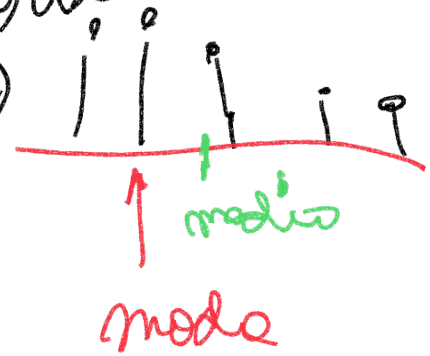
$$f_{X.Y}(x,y) = f_X(x)f_Y(y) \quad \forall\, x, y \in \mathbb{R}$$

Momentos

Momentos

Esperanza (o media):

$$\mathbb{E}[X] = \sum_{x \in A} x\, p_X(x) = \int x f_X(x)\,dx$$



orden 1

$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$
$$= \sum g(x) p_X(x)$$

Varianza:

orden 2

$$var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}(X)^2$$

Covarianza:

$$cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}(X)\mathbb{E}(Y)$$

Si X,Y son indep

$\Downarrow$ ~~$\Uparrow$~~

$cov(X,Y) = 0$

Momentos

En general, definimos el ***n*-ésimo momento** de la v.a. X como $\mathbb{E}[X^n]$

Se define la **función generadora de momentos** como

$$M_X(t) = \mathbb{E}[e^{tX}]$$

Si que $\mathbb{E}[e^{tX}]$ existe (es finita) para algún intervalo que contiene al cero, se puede calcular el n-ésimo momento de X como la derivada n-ésima de $M_X(t)$ evaluada en 0:

$$\mathbb{E}[X^n] = \frac{d}{dt} M_X(t)\Big|_{t=0}$$

Ejercicio 0

Calcular la función generadora de momentos para una variable

$$X \sim \mathcal{B}in(n, p)$$

Algunas distribuciones útiles

Variables discretas

- **Bernoulli(p)**: X = {0,1}. Asociada a la ocurrencia o no de un éxito

$$P_X(x) = p^x (1-p)^{1-x} = \begin{cases} p & x=1 \\ 1-p & x=0 \end{cases}$$

- **Binomial(n, p)**: me cuenta la cantidad de éxitos en n ensayos.

$$\binom{n}{x} p^x (1-p)^{n-x}$$

ordena la secuencia de 1 y 0

$$X \sim Bin(n,p) \qquad Y_i \sim Bn(p) \Rightarrow X = \sum_{i=1}^{n} Y_i$$

- **Geométrica(p)**: cantidad de ensayos que debo realizar hasta observar el primer éxito

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

$$\underline{0} \; \underline{0} \cdots \underline{0} \; \underline{1}$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad \frac{1}{p} \; \frac{0}{(1-p)} \; \frac{0}{(1-p)} \; \frac{1}{p} \; \cdots \; \frac{1}{p} = p^x (1-p)^{n-x}$$

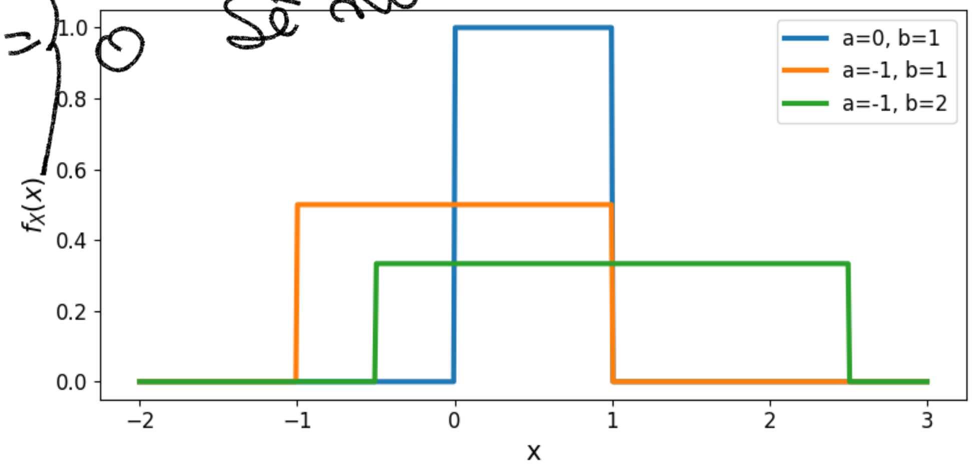
Variables continuas

$$\mathbb{I}\{x \in A\} = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si no} \end{cases}$$

- **Uniforme:** todos los puntos son equiprobables.

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

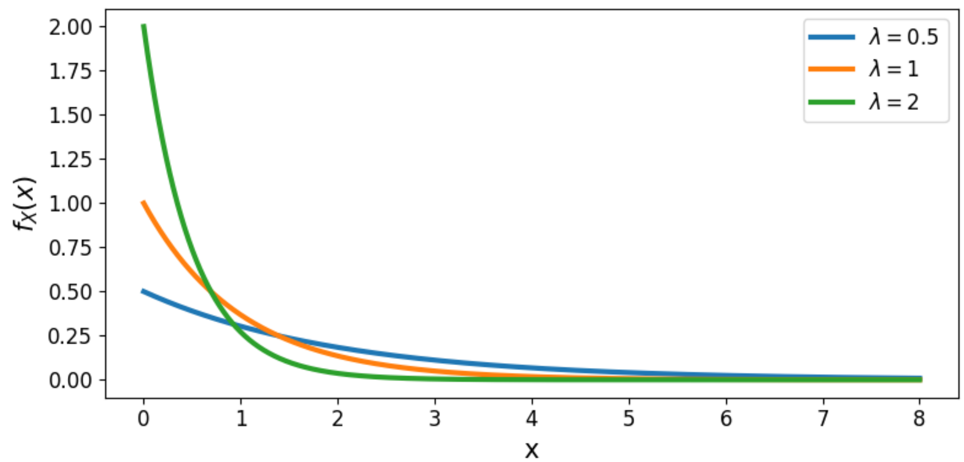
$$f_X(x) = \frac{1}{b-a}\mathbf{I}\{a < x < b\}$$



- **Exponencial:** sirve para modelar tiempos hasta eventos que no tienen memoria.

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

$$f_X(x) = \lambda e^{-\lambda x}\mathbf{I}\{x > 0\}$$

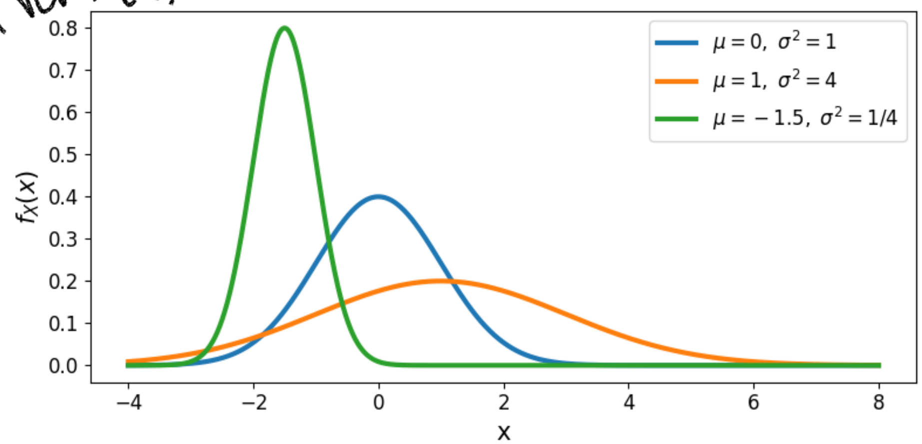


Variables continuas

$\mathbb{E}[X]$ $Var(X)$

- **Normal (gaussiana).** $X \sim \mathcal{N}(\mu, \sigma^2)$

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



Propiedades: Sean X, Y dos v.a. Independientes

$$X \sim \mathcal{N}(\mu, \sigma^2) \rightarrow \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1) \qquad \text{(estandarización)}$$

$\rightarrow \sigma = \sqrt{Var(X)}$ (desvío)

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2),\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \rightarrow aX + bY \sim \mathcal{N}(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$$

(combinación lineal de normales es normal)

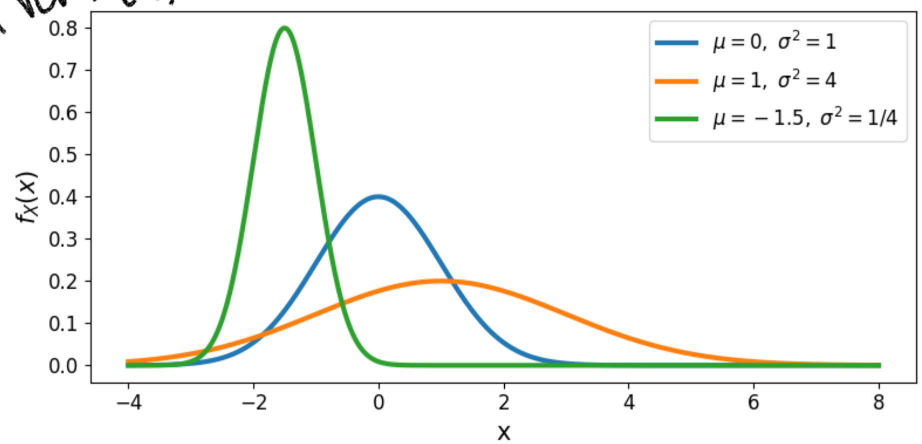
Variables continuas

$\mathbb{E}[X]$ $\mathrm{Var}(X)$

- **Normal (gaussiana).** $X \sim \mathcal{N}(\mu, \sigma^2)$

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

μ es la media
σ^2 es la varianza



Propiedades: Sean X, Y dos v.a. Independientes

$$X \sim \mathcal{N}(\mu, \sigma^2) \rightarrow \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1) \qquad \text{(estandarización)}$$

$\sigma = \sqrt{\mathrm{var}(X)}$ (des vío)

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2),\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \rightarrow aX + bY \sim \mathcal{N}(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$$

(combinación lineal de normales es normal)

Ejercicio 1

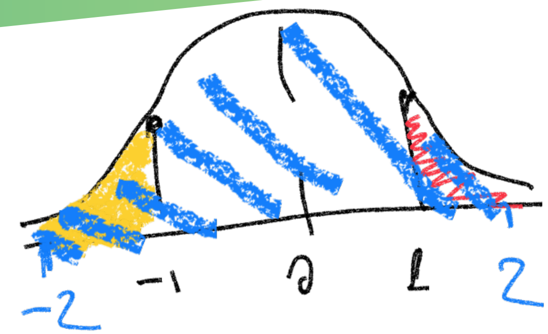
Sea X una v.a con distribución normal estándar. Hallar la probabilidad de que

1. $X>1$
2. $X<-1$
3. $|X| <2$
4. Hallar los cuantiles 0.1 y 0.9

Sea además $Y \sim N(2,9)$

1. Hallar $P(2X+Y <5)$

$X \sim N(0, 1)$ $\quad\quad \mathbb{P}(X > 1)$



(por simetria)

$\mathbb{P}(X < -1) = 0{,}1587$ $\quad = 0{,}1587$

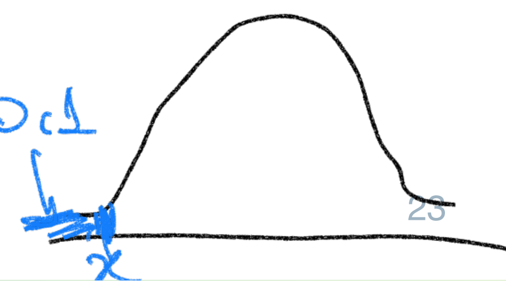
Python

$\mathbb{P}(|X| < 2) = \mathbb{P}(-2 < X < 2)$

$= \mathbb{P}(X < 2) - \mathbb{P}(X < -2) = 0{,}9544$

$x_{0,1} : \mathbb{P}(X \leq x_{0,1}) = 0{,}1$ $\quad F_X(x_{0,1}) = 0{,}1$

$\Rightarrow x_{0,1} = -1{,}2816$



$$z_{0,9} = -z_{0,1} = 1{,}2816$$

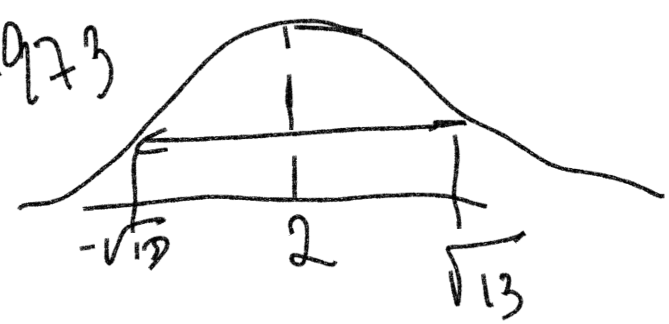
$$Y \sim N(2, 9) \rightarrow W = 2X + 5 \qquad P(W < 5)$$

$$W \sim N(2 \cdot 0 + 1 \cdot 2,\ 2^2 \cdot 1 + 1^2 \cdot 9) \equiv N(2, 13)$$

estandarizar

$$P(W < 5) = P\left(Z < \frac{5-2}{\sqrt{13}}\right) = 0{,}7973$$

$$Z \sim N(0, 1)$$



Ejercicio 2

El tiempo (en minutos) entre llamadas a un call center tiene una distribución exponencial de parámetro 1/5.

1. Calcular la probabilidad de que la primer llamada llegue después de 2 minutos
2. Calcular la probabilidad de que la primer llamada llegue después de los 5 minutos, si se sabe que en los primeros 3 minutos no se recibieron llamados

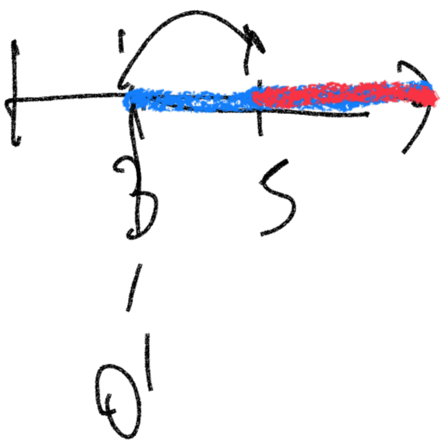
$$X \sim \mathcal{E}(1/5) \qquad \mathbb{P}(X>2) = \int_2^{\infty} 1/5\, e^{-1/5 x}\, dx$$

$$f_X(x) = \frac{1}{5} e^{-\frac{1}{5}x} \mathbb{1}\{x>0\} \qquad = e^{-1/5 \cdot 2} = e^{-2/5} = 0{,}67$$

$$F_X(x) = \left(1 - e^{-1/5 x}\right) \mathbb{1}\{x>0\}$$

$$S_X(x) = \mathbb{P}(X>x) = 1 - F_X(x) = e^{-1/5 x}$$

$$\mathbb{P}(X>5 \mid X>3) = \frac{\mathbb{P}(X>5,\, X>3)}{\mathbb{P}(X>3)} = \frac{\mathbb{P}(X>5)}{\mathbb{P}(X>3)}$$



$$= \frac{e^{-1/5 \cdot 5}}{e^{-1/5 \cdot 3}} = e^{-1/5(5-3)} = e^{-1/5 \cdot 2}$$

Pérdida de memoria

Distribución normal multivariada

Función de densidad conjunta

Sea $\underline{X} = [X_1, \ldots, X_n]^T$ un vector aleatorio continuo, diremos que X tiene distribución normal multivariada, si su función de densidad conjunta es de la forma

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1}(\underline{x}-\underline{\mu})}$$

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

donde $\underline{\mu}$ se corresponde con la media de la v.a. y Σ Es la matriz de covarianza .

$$\Sigma = \begin{bmatrix} \sigma_1^2 & cov(X_1,X_2) & \cdots & cov(X_1,X_n) \\ cov(X_2,X_1) & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}$$

Notación: $\underline{X} \sim \mathcal{N}(\underline{\mu}, \Sigma)$

Distribuciones marginales

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma) \Rightarrow X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$$

$$\boldsymbol{\mu} = [\mu_1, \ldots, \mu_n]^T$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & cov(X_1, X_2) & \ldots & cov(X_1, X_n) \\ cov(X_2, X_1) & \sigma_2^2 & \ldots & cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, X_1) & cov(X_n, X_2) & \ldots & \sigma_n^2 \end{bmatrix}$$

Ejercicio 3

Sean X,Y dos v.a. Con función de densidad conjunta

$$f_{X,Y}(x,y) = \frac{1}{2\pi 0.6} e^{-\frac{1}{2}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}^{-1}\begin{bmatrix} x \\ y \end{bmatrix}}$$

1. Calcular E[X], E[Y], var(X), var(Y), y cov(X,Y)
2. Hallar las densidades marginales de X e Y
3. Calcular P(X<2, Y<-1)

Bibliografía

Bibliografía

"Mathematical Statistics with Applications", Dennis D. Wackerly, William Mendenhall III, Richard L. Scheaffer.