1. Vectores aleatorios (Variables aleatorias multidimensionales)

Definición Sea $\tilde{X} = (X_1, X_2, ..., X_n)^T$ un vector en $\mathbb{R}^n$. Se dice que $\tilde{X}$ es un VECTOR ALEATORIO sí y sólo si $X_1, X_2, ..., X_n$ son variables aleatorias.

$$\tilde{X} : \Omega \to \mathbb{R}^n \omega \to \tilde{X}(\omega) = (X_1(\omega), X_2(\omega), ..., X_n(\omega))$$

Existen tres tipos de vectores

1. **Discretos:** $X_1, X_2, ..., X_n$ son discretos.
2. **Continuos:** $X_1, X_2, ..., X_n$ son continuos.
3. **Mixtos:** $X_1, ..., X_p$ son discretos, y $X_{p+1}, ..., X_n$ son continuos.

1.1. Tipos de vectores aleatorios

Sea $\tilde{X} = (X_1, ..., X_n)$ un vector aleatorio. Se dice que $\tilde{X}$ es DISCRETO sí y sólo si:

$$\text{Rec}\tilde{X} = RecX_1 \times RecX_2 \times ... \times RecX_n \subseteq \mathbb{R}^n$$

es un conjunto numerable.
Luego se define la FUNCIÓN DE PROBABILIDAD CONJUNTA (o función de cuantía conjunta) como:

$$f_{\tilde{X}}(\vec{x}) = \mathbb{P}\{X_1 = x_1, X_2 = x_2, ..., X_n = x_n\} \text{ si } (x_1, x_2, ..., x_n) \in \mathbb{R}^n$$

La cual cumple:

1. $f_{\tilde{X}}(\vec{x}) \geq 0; \forall \vec{x} \in \mathbb{R}^n$
2. $\sum_{x_1=-\infty}^{\infty} \sum_{x_2=-\infty}^{\infty} \cdots \sum_{x_n=-\infty}^{\infty} f_{\tilde{X}}(\vec{x}) = 1$

Ejemplo Suponga que se dispone de 10 personas, de las cuales tres son *chilenos*, tres *argentinos* y cuatro *peruanos*. Defina las variables:

- X: Número de chilenos seleccionados.
- Y: Número de peruanos escogidos.

Suponga que se desea escoger un grupo de tres personas, sin reposición. Calcule la función de probabilidad conjunta del vector $(X, Y)^T$.
En este caso el recorrido asociado al vector es: $\text{Rec}\tilde{X} = \text{Rec}X \times \text{Rec}Y = \{0, 1, 2, 3\} \times \{0, 1, 2, 3, 4\}$ sujeto a $x + y \leq 3$.
Por lo tanto $\text{Rec}(\tilde{X}) = \{(x, y) \in \{0, 1, 2, 3\} \times \{0, 1, 2, 3, 4\} / x + y \leq 3\}$ Luego:

$f_{\tilde{X}}(\vec{x}) = f_{X,Y}(x,y) \qquad\qquad (x,y) \in \mathbb{R}^2$

$f_{X,Y}(0,0) = \mathbb{P}\{X=0, Y=0\} = \dfrac{\binom{3}{0}\binom{4}{0}\binom{3}{3}}{\binom{10}{3}} = \dfrac{1}{120}$

$f_{X,Y}(0,1) = \dfrac{\binom{3}{0}\binom{4}{1}\binom{3}{2}}{\binom{10}{3}} = \dfrac{12}{120}$

$f_{X,Y}(0,2) = \dfrac{\binom{3}{0}\binom{4}{2}\binom{3}{1}}{\binom{10}{3}} = \dfrac{18}{120}$

$f_{X,Y}(0,3) = \dfrac{\binom{3}{0}\binom{4}{3}\binom{3}{0}}{\binom{10}{3}} = \dfrac{4}{120}$

Se puede ver que:

$$f_{X,Y}(x,y) = \mathbb{P}\{X=x, Y=y\} = \frac{\binom{3}{x}\binom{4}{y}\binom{3}{3-(x+y)}}{\binom{10}{3}} \qquad \mathbb{I}_A(x,y)$$

con $A = \{(x,y) \in \{0,1,2,3\} \times \{0,1,2,3,4\} / x+y \leq 3$

Alternativamente:

X/Y	0	1	2	3
0	$\frac{1}{120}$	$\frac{12}{120}$	$\frac{18}{120}$	$\frac{4}{120}$
1	$\frac{9}{120}$	$\frac{36}{120}$	$\frac{18}{120}$	0
2	$\frac{9}{120}$	$\frac{12}{120}$	0	0
3	$\frac{1}{120}$	0	0	0

Sea $\tilde{X} = (X_1, X_2, ..., X_n)^T$ un vector aleatorio. Se dice que $\tilde{X}$ es CONTINUO sí y sólo si su recorrido Rec$\tilde{X} \subseteq \mathbb{R}^n$ es un conjunto no numerable y si existe una función no negativa $f_{\tilde{X}} : \mathbb{R}^n \to \mathbb{R}$ tal que para cualquier evento $B \subseteq \mathbb{R}^n$ se tiene:

$$\mathbb{P}(B) = \iint \cdots \int_B f_{\tilde{X}}(\vec{x}) d\vec{x}$$

La función $f_{\tilde{X}}$ se llama FUNCIÓN DE DENSIDAD DE PROBABILIDAD CONJUNTA, la cual cumple:

1. $f_{\tilde{X}}(\vec{x}) \geq 0; \qquad \forall \vec{x} \in \mathbb{R}^n$
2. $\iiint \cdots \int_{\mathbb{R}^n} f_{\tilde{X}}(\vec{x}) d\vec{x} = 1$

Ejemplo Sea $(X,Y)^T$ un vector aleatorio bidimensional con función de densidad conjunta dada por:

$$f_{X,Y}(x,y) = c \cdot y \qquad \mathbb{I}_A(x,y) \qquad A = \{(x,y) \in \mathbb{R}^2 / 0 \leq x \leq y \leq 1\}$$

Buscando el valor de $c \in \mathbb{R}$ tal que $f_{X,Y}$ sea densidad:

$$\iint_A f_{X,Y}(x,y) dxdy = 1$$

$$\int_0^1 \int_x^1 cy \; dydx = 1 \quad \Rightarrow \quad \frac{c}{2} \int_0^1 (1-x^2) \; dx = 1 \quad \Rightarrow \quad c = 3$$

Defina el evento $B = \{(x,y) \in \mathbb{R}^2 / x + y \leq 1\}$

Luego $\mathbb{P}(B) = \displaystyle\int_0^{1/2} \int_0^y 3y \; dxdy + \int_{1/2}^1 \int_0^{1-y} 3y \; dxdy = ?$

Alternativamente $\mathbb{P}(B) = \displaystyle\int_0^{1/2} \int_x^{1-x} 3y \; dydx$

1.2. Distribuciones marginales

Sea $\tilde{X} = (X_1, X_2, ..., X_n)^T$ un vector aleatorio con función de cuantía o densidad $f_{\tilde{X}}$.

El proceso de MARGINALIZACIÓN corresponde a una reducción de la dimensión del vector.

Se definen las distribuciones marginales como:

$$f_{X_1,...X_m}(x_1,...,x_m) = \sum_{x_{m+1}=-\infty}^{\infty} \cdots \sum_{x_n=-\infty}^{\infty} f_{\tilde{X}}(\vec{x}) \quad \text{para } \tilde{X} \text{ discreto.}$$

$$f_{X_1,...X_m}(x_1,...,x_m) = \iint \cdots \int_{\mathbb{R}^{n-m}} f_{\tilde{X}}(\vec{x}) \; dx_{m+1}...dx_n \quad \text{para } \tilde{X} \text{ continuo.}$$

que entrega la distribución de probabilidad de un vector de dimensión m, con $m < n$.

En particular, en el caso bidimensional, $n = 2$, las distribuciones marginales asociadas a un vector (X,Y) son:

$$f_Y(y) = \sum_{x=-\infty}^{\infty} f_{X,Y}(x,y) \qquad \text{caso discreto}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \; dx \qquad \text{caso continuo}$$

Análogo para f_X

Por ejemplo, en el caso discreto presentado, se tiene:

Marginal de Y

Y	0	1	2	3
$f_Y(y)$	$\frac{20}{120}$	$\frac{60}{120}$	$\frac{36}{120}$	$\frac{4}{120}$

$$f_Y(y) = \sum_{x=0}^{3} f_{X,Y}(x,y)$$

Para el caso del vector continuo, la distribución marginal de Y es:

$$f_Y(y) = \int_0^y 3y\ dx = 3y^2 \quad \mathbb{I}_{[0,1]}(y)$$

En el caso de X:

$$f_X(x) = \int_x^1 3y\ dy = \frac{3}{2}(1-x^2) \quad \mathbb{I}_{[0,1]}(x)$$

1.3. Definiciones condicionales

Sea $\tilde{X} = (X_1, X_2, ..., X_n)^T$ un vector aleatorio. Considere una partición de $\tilde{X}$ de la siguiente forma:

$$\tilde{X} = \begin{pmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{pmatrix}$$

Donde:

$$\tilde{X}_1 = (X_1, ..., X_p) \quad ; \quad \tilde{X}_2 = (X_{p+1}, ..., X_n) \qquad \text{con } p < n$$

Se define la DISTRIBUCIÓN CONDICIONAL de $\tilde{X}_1$ dado que $\tilde{X}_2 = \vec{x}_2$ como:

$$f_{\tilde{X}_1/\tilde{X}_2=\vec{x}_2}(\vec{x}_1) = \frac{f_{\tilde{X}}(\vec{x})}{f_{\tilde{X}_2}(\vec{x_2})}$$

En particular si $\tilde{X} = (X, Y)$, entonces la distribución de Y dado $X = x$ es:

$$f_{Y/X=x_0}(y) = \frac{f_{X,Y}(x_0, y)}{f_X(x_0)}$$

Observación: Es importante identificar claramente el recorrido de $Y/X = x_0$.
Nota: Definición análoga para la condicional de X dado $Y = y_0$.

1.4. Valores esperados

Definición Sea $g : \mathbb{R}^2 \to \mathbb{R}$ una función escalar y sea $\tilde{X} = (X, Y)$ un vector aleatorio con cuantía o densidad conjunta $f_{X,Y}$. Se define el VALOR ESPERADO de $g(X, Y)$ como:

$$\mathbb{E}[g(X,Y)] = \sum_{y=-\infty}^{\infty} \sum_{x=-\infty}^{\infty} g(x,y) \cdot f_X(x,y) \qquad \text{en caso } (X,Y) \text{ discreto.}$$

$$\mathbb{E}[g(X,Y)] = \iint_{\mathbb{R}^2} g(x,y) \cdot f_X(x,y)\ dxdy \qquad \text{en caso } (X,Y) \text{ continuo.}$$

1.4.1. Casos particulares

(Corregir en caso de que la notación esté mala, el profe lo anotó super mal y cuando le pregunté me respondió otra cosa)

1. Si $g(X,Y) = x \Rightarrow \mathbb{E}[g(X,Y)] = \mathbb{E}[X] \leftarrow$ Esperanza marginal de X

2. Si $g(X,Y) = y \Rightarrow \mathbb{E}[g(X,Y)] = \mathbb{E}[Y] \leftarrow$ Esperanza marginal de Y

   Nótese que, por ejemplo en el caso continuo:

$$\mathbb{E}[X] = \iint_{\mathbb{R}^2} x \cdot f_{X,Y}(x,y)\ dx\ dy = \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy}_{f_X(x) \text{ Marginal de } X}\ dx$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\ dx \quad \leftarrow \quad \text{Esperanza de } X \text{ en var. aleatoria}$$

3. Si $g(X,Y) = (x - \mathbb{E}[X])^2$ entonces:

$$\mathbb{E}[g(X,Y)] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{V}[X] \quad \leftarrow \quad \text{Varianza marginal de } X$$

4. Si $g(X,Y) = (y - \mathbb{E}[Y])^2$ entonces:

$$\mathbb{E}[g(X,Y)] = \mathbb{E}[(Y - \mathbb{E}[Y])^2] = \mathbb{V}[Y] \quad \leftarrow \quad \text{Varianza marginal de } Y$$

1.5. Esperanza condicional

Sea $\tilde{X} = (X,Y)^T$ un vector aleatorio con función de masa conjunta $f_{X,Y}$. Se define la ESPERANZA CONDICIONAL de Y *dado que* $X = x$ como:

$$\mathbb{E}[Y/X = x] = \sum_{y=-\infty}^{\infty} y \cdot f_{Y/X=x}(y) \qquad \text{en caso } Y \text{ discreto.}$$

$$\mathbb{E}[Y/X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y/X=x}(y)\ dy \qquad \text{en caso } Y \text{ continuo.}$$

De modo análogo se puede calcular la VARIANZA CONDICIONAL de Y *dado* $X = x$

$$\mathbb{V}[Y/X = x] = \mathbb{E}[(Y - \mathbb{E}[Y])^2/X = x]$$

$$\mathbb{V}[Y/X = x] = \mathbb{E}[Y^2/X = x] - \{\mathbb{E}[Y/X = x]\}^2$$

1.5.1. Propiedades

1. La esperanza condicional de $Y/X = x$ se conoce como el MEJOR PREDICTOR LINEAL de Y *dado que* $X = x$.

2. Cuando $X = x$ no es fijo, entonces $\mathbb{E}[Y/X = x]$ es una variable aleatoria en términos de X.

3. **Esperanza iterada**

$$\mathbb{E}[\mathbb{E}[Y/X=x]] = \int_{-\infty}^{\infty} \mathbb{E}[Y/X=x]\cdot f_X(x)\,dx$$

$$\cdots \text{ desarrollo pajero}$$

$$\mathbb{E}[\mathbb{E}[Y/X=x]] = \mathbb{E}[Y]$$

4. **Descomposición de la varianza de Y**

$$\mathbb{V}[Y] = \begin{cases} \text{Varianza dentro de cada condición (nivel X)} \\ \text{Varianza entre los niveles X.} \end{cases}$$

$$\mathbb{V}[Y] = \mathbb{E}[\mathbb{V}[Y/X=x]] + \mathbb{V}[\mathbb{E}[Y/X=x]]$$

5. Sea $g:\mathbb{R}\to\mathbb{R}$ una función real, luego: $\mathbb{E}[g(X)\cdot Y/X=x_0] = g(x_0)\cdot\mathbb{E}[Y/X=x_0]$

   que en general es:

$$\left.\begin{array}{l} g:\mathbb{R}\to\mathbb{R} \\ h:\mathbb{R}\to\mathbb{R} \end{array}\right\} \quad \text{funciones}$$

   Luego: $\mathbb{E}[g(X)\cdot h(Y)/X=x_0] = g(x_0)\cdot\mathbb{E}[h(Y)/X=x_0]$

1.6. Asociación y correlación

Sea $(X,Y)^T$ un vector aleatorio. Se define la COVARIANZA entre las variables X y Y como:

$$\text{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$$

la cual corresponde a una medida de asociación entre X y Y. De ahí:

- Si $\text{Cov}(X,Y) > 0 \quad \Rightarrow \quad$ Asociación directa entre X y Y $\begin{cases} X\uparrow & \Rightarrow & Y\uparrow \\ X\downarrow & \Rightarrow & Y\downarrow \end{cases}$
- Si $\text{Cov}(X,Y) < 0 \quad \Rightarrow \quad$ Asociación inversa entre X y Y $\begin{cases} X\uparrow & \Rightarrow & Y\downarrow \\ X\downarrow & \Rightarrow & Y\uparrow \end{cases}$
- Si $\text{Cov}(X,Y) = 0 \quad \Rightarrow \quad$ No hay información.

1.6.1. Propiedades

1. Si $c\in\mathbb{R}$ es una **constante**, entonces:

$$\text{Cov}(X,c) = 0$$

2. **Simetría**:

$$\text{Cov}(X,Y) = \text{Cov}(Y,X)$$

3. **Si** $Y = X$ entonces:

$$\text{Cov}(X,X) = \mathbb{V}(X)$$

4. **Bilinealidad**:

   - Linealidad por componentes:
   $$\text{Cov}(aX + b,\ Y) = a\ \text{Cov}(X, Y)$$
   - De otro modo:
   $$\text{Cov}(X,\ cY + d) = d\ \text{Cov}(X, Y)$$
   - Caso general:
   $$\text{Cov}(aX_1 + bX_2,\ Y) = a\ \text{Cov}(X_1, Y) + b\ \text{Cov}(X_2, Y)$$

5. **Desigualdad de Cauchy-Schwarz**:
   $$|\text{Cov}(X, Y)| \leq \sqrt{\mathbb{V}(X) \cdot \mathbb{V}(Y)}$$

6. **Identidad**
   $$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$$
   donde
   $$\mathbb{E}[XY] = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} xy \cdot f_X(x, y) \qquad \text{en caso discreto.}$$
   $$\mathbb{E}[XY] = \iint_{\mathbb{R}^2} xy \cdot f_X(x, y)\ dxdy \qquad \text{en caso continuo.}$$

Nota Ahora es posible conocer un resultado pendiente.

- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
- $\mathbb{V}(X + Y) \neq \mathbb{V}(X) + \mathbb{V}(Y)$

Pero...
$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \cdot \text{Cov}(X, Y)$$

Caso general:
$$\mathbb{V}(aX + bY) = a^2\mathbb{V}(X) + b^2\mathbb{V}(Y) + 2ab \cdot \text{Cov}(X, Y)$$

Observación Como $\text{Cov}(X, Y) \in \mathbb{R}$ no es informativa respecto a qué tan intensa es la asociación entre dos variables. Debido a esto suge un índice en un rango acotado que soluciona el problema.

1.7. Coeficiente de correlación lineal

Sea $(X, Y)^T$ un vector aleatorio con $\text{Cov}(X, Y); \mathbb{V}(X); \mathbb{V}(Y)$. Se define el COEFICIENTE DE CORRELACIÓN LINEAL DE PEARSON como:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}(X) \cdot \mathbb{V}(Y)}}$$

Que tiene la particularidad que:

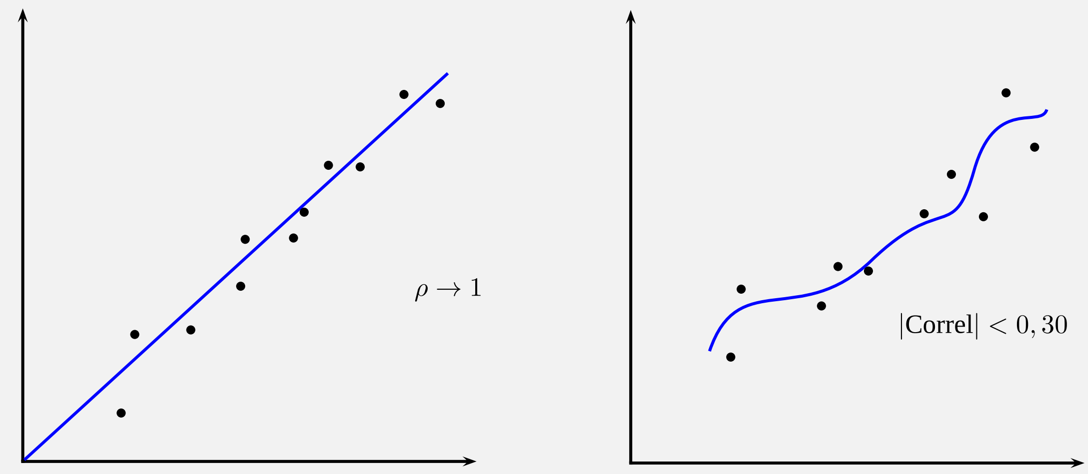
$$-1 \leq \rho \leq 1$$

Por lo cual:

- Si $\rho > 0$ y $\rho \to 1 \quad \Rightarrow \quad$ Asociación lineal directa e intensa.
- Si $\rho < 0$ y $\rho \to -1 \quad \Rightarrow \quad$ Asociación lineal invrsa e intensa.
- Si $\rho \to 0 \quad \Rightarrow \quad$ No hay una asociación lineal "decente".

Caso Considere datos:

$$(X_1, Y_1); (X_2, Y_2); ...; (X_n, Y_n)$$



1.8. Independencia

Sean $X_1, X_2, ..., X_n$ una colección de n variables aleatorias. Se dice que son (ESTOCÁSTICAMENTE) INDEPENDIENTES sí y sólo si:

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot ... \cdot f_{X_n}(x_n) \qquad \forall (x_1, ..., x_n) \in \mathbb{R}^n$$

...en particular:

$$X \text{ es independiente de } Y \Leftrightarrow f_{X,Y}(x, Y) = f_X(x) \cdot f_Y(y) \qquad \forall (x, y) \in \mathbb{R}^2$$

Notación: $X \perp Y$

Teorema Sea $(X, Y)^T$ un vector aleatorio tal que X es INDEPENDIENTE de Y, entonces:

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y) \qquad \text{en donde...}$$

$$\mathbb{E}[XY] = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} xy \cdot f_{X,Y}(x, y) \qquad \text{en caso discreto.}$$

$$\mathbb{E}[XY] = \iint_{\mathbb{R}^2} xy \cdot f_{X,Y}(x, y)\, dxdy \qquad \text{en caso continuo.}$$

Al ser independientes...

$$\mathbb{E}[XY] = \sum_{x=-\infty}^{\infty} x \cdot f_X(x) \sum_{y=-\infty}^{\infty} y \cdot f_Y(y) \qquad \text{en caso discreto.}$$

$$\mathbb{E}[XY] = \int x f_X(x)\ dx \cdot \int y f_Y(y)\ dy \qquad \text{en caso continuo.}$$

$$= \mathbb{E}(X) \cdot \mathbb{E}(Y) \Rightarrow \text{Cov}(X, Y) = 0$$

$$X \perp Y \Rightarrow \text{Cov}(X, Y) = 0$$

Ejemplo Sea $X \sim \text{Unif}(0, 2\pi)$. Defina como $U = \cos(X), V = \sin(X)$.
Calcule $\text{Cov}(U, V)$. ¿Son independientes?
$\text{Cov}(U, V) = \mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V)$ donde...

$\mathbb{E}(U) = \mathbb{E}(cosX) = \int_{-\infty}^{\infty} cosx \cdot f_X(x)\ dx$

Como $X \sim \text{Unif}(0, 2\pi) \Rightarrow f_X(x) = \frac{1}{2\pi} \quad \mathbb{I}_{0,2\pi}(x)$

Entonces $\mathbb{E}(U) = \int_0^{2\pi} \frac{1}{2\pi} cos(X) dx = \frac{1}{2\pi} \left[sin(X)\right]_0^{2\pi}$

$\mathbb{E}(UV) = \mathbb{E}(\cos X \cdot \sin X) = 0$

$$\text{Cov}(U, V) = 0$$

No son independientes, puesto que $U^2 + V^2 = 1$

1.9. Distribución normal bivariada

Sea $(X, Y)^T$ un vector aleatorio tal que:

$$\mu_x = \mathbb{E}(X) \quad ; \quad \sigma_x^2 = \mathbb{V}(X)$$

$$\mu_y = \mathbb{E}(Y) \quad ; \quad \sigma_y^2 = \mathbb{V}(Y)$$

$\rho = \text{Correl}(X, Y)$ $\sigma_{xy} = \text{Cov}(X, Y)$
Se dice que $(X, Y)^T$ sigue una DISTRIBUCIÓN NORMAL BIVARIADA sí y sólo si:

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x,y)} \qquad \mathbb{I}_{\mathbb{R}\times\mathbb{R}}(x, y)$$

donde:

$$Q(x, y) = \frac{1}{1-\rho^2} \left\{ \left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x}\right) \left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 \right\}$$

Se denota:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$$

1.9.1. Propiedades

1. Distribución cerrada **bajo marginalización**:

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

2. Distribución cerrada **bajo condicionamientos**:

$$Y/X = x \sim N(\mu_{Y/X}, \sigma_{Y/X}^2) \qquad \text{donde...}$$

$$\mu_{Y/X} = \mathbb{E}[Y/X = x] = \mu_y + \frac{\text{Cov}(X,Y)}{\sigma_x^2}(x - \mu_x)$$

$$\sigma_{Y/X}^2 = \mathbb{V}[Y/X = x] = \sigma_y^2(1 - \rho^2)$$

Análogo para $X/Y = y$

3. **Considere** $\rho = 0$

$$Q(x,y) = Q(x) + Q(y) = \left(\frac{x - \mu_x}{\sigma_x}\right)^2 + \left(\frac{y - \mu_y}{\sigma_y}\right)^2$$

y además...

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}(Q(x)+Q(y))} \quad \mathbb{I}_{\mathbb{R}\times\mathbb{R}}(x,y) \\ &= \underbrace{\frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}Q(x)}}_{f_X(x)} \underbrace{\frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2}Q(y)}}_{f_Y(y)} \quad \mathbb{I}_{\mathbb{R}}(x) \cdot \mathbb{I}_{\mathbb{R}}(y) \\ & X \sim N(\mu_x, \sigma_x^2) \qquad\qquad Y \sim N(\mu_y, \sigma_y^2) \end{aligned}$$

Por lo tanto $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \Rightarrow X$ independiente de $Y (X \perp Y)$

O sea:

$$X, Y \sim N(\mu, \sigma^2)$$

$$X \perp Y \Leftrightarrow \rho = 0 \qquad (\text{Cov} = 0)$$

4. La distribución es **cerrada bajo combinaciones lineales**. O sea:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2(\,\cdot\,,\,\cdot\,)$$

$$Y = X_1 + X_2 \sim N(\mu_y, \sigma_y^2)$$

con

$$\begin{cases} \mu_y = \mathbb{E}(Y) = \mathbb{E}(X_1 + X_2) = \mu_1 + \mu_2 \\ \sigma_y^2 = \mathbb{V}(Y) = \mathbb{V}(X_1 + X_2) = \sigma_1^2 + \sigma_2^2 + 2\mathrm{Cov}(X_1, X_2) \end{cases}$$

Observación: Es importante recordar esta propiedad, pues luego bastará ajustar los parámetros usando el hecho de que el modelo es conocido.

1.10. Distribución normal multivariada

Definición Sea $\tilde{X} = (X_1, ..., X_p)^T$ un vector aleatorio tal que $\mathbb{E}(X_i) = \mu_i$, con $i = 1, ..., p$. Se define el VECTOR DE MEDIAS como:

$$\tilde{\mu} = \mathbb{E}(\tilde{X}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} \qquad \text{con } \tilde{\mu} \in \mathbb{R}^p$$

Definición Sea $\tilde{X} = (X_1, ..., X_p)^T$ un vector aleatorio tal que:

$$\sigma_i^2 = \mathbb{V}(X_i) \qquad ; i = 1, ..., p$$

$$\sigma_{ij} = \mathrm{Cov}(X_i, X_j) \qquad ; i, j = 1, ..., p (i \neq j)$$

Se dfefine la MATRIZ DE VARIANZAS Y COVARIANZAS como:

$$\Sigma = \mathbb{V}(\tilde{X}) = \mathbb{E}[(\tilde{X} - \mathbb{E}(\tilde{X}))(\tilde{X} - \mathbb{E}(\tilde{X}))^T]$$

La cual tendrá una estructura de la siguiente forma:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{p1} & \cdots & \cdots & \sigma_p^2 \end{pmatrix}$$

Como la covarianza es simétrica $\Rightarrow \sigma_{ij} = \sigma_{ji}$
Por lo tanto la matriz Σ es **simétrica**.
Σ definida positiva:

$$\tilde{X}^T \Sigma \tilde{X} > 0 \qquad \forall \tilde{X} \in \mathbb{R}^p$$

$$\Sigma \in M(p \times p; \mathbb{R})$$

Observación En el caso de:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2(\tilde{\mu}, \Sigma)$$

se tiene:

$$\tilde{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad \Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}$$

Definición Sea $\tilde{X} = (X_1, X_2, ..., X_p)^T$ un vector aleatorio continuo tal que:

$$\tilde{\mu} = \mathbb{E}(\tilde{X}) \qquad \text{y} \qquad \Sigma = \mathbb{V}(\tilde{X})$$

Se dice que $\tilde{X}$ sigue una DISTRIBUCIÓN NORMAL P-VARIADA sí y sólo si:

$$f_{\tilde{X}}(\vec{x}) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \ e^{-\frac{1}{2}(\vec{x}-\tilde{\mu})^T \Sigma^{-1} (\vec{x}-\tilde{\mu})}$$

Observación la función $Q(\vec{x}) = (\vec{x} - \tilde{\mu})^T \ \Sigma^{-1}(\vec{x} - \tilde{\mu})$ es una forma cuadrática que posee distribución X_p^2

1.11. Propiedades

En general son análogas a las del modelo normal bivariado:

1. $X_i \sim N(\mu_i, \sigma_i^2) \qquad\qquad i = 1, ..., p$
2. Si $\Sigma = I_p$, entonces los componentes del vector son independientes.
3. Condicionales son normales, en particular, considere una partición del vector $\tilde{X}$.

   $$\tilde{X} = \begin{pmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \end{pmatrix}; \qquad \tilde{\mu} = \begin{pmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{pmatrix}; \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \qquad \text{con } \mathbb{V}(\tilde{Y}_j) = \Sigma_{jj}$$

   $\text{Dim}(\tilde{Y}_1) = n \qquad\qquad \text{Dim}(\tilde{Y}_2) = p - m$

   Luego, la distribución condicional de

   $$\tilde{Y}_1 / \tilde{Y}_2 = \vec{y}_2 \sim N_m(\tilde{\mu}_{1/2}; \Sigma_{1/2})$$

   $$\tilde{\mu}_{1/2} = \mathbb{E}(\tilde{Y}_1/\tilde{Y}_2 = \vec{y}_2) = \tilde{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\vec{y}_2 - \tilde{\mu}_2)$$
   $$\Sigma_{1/2} = \mathbb{V}(\tilde{Y}_1/\tilde{Y}_2 = \vec{y}_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

4. El modelo es **cerrado bajo combinaciones lineales**.

   Sea:

   $$\left.\begin{array}{l} A = A_{q\times p} \\ \vec{b} = b_{q\times 1} \end{array}\right\} \quad \text{constantes.}$$

   Luego se tiene: $\tilde{Y} = A\tilde{X} + \vec{b} \sim N_q(\tilde{\mu}_Y; \Sigma_Y)$ donde:

   $$\tilde{\mu}_y = A\tilde{\mu} + \vec{b}$$

$$\Sigma_Y = A\Sigma A^T$$

Obs: $\tilde{X} \sim N_p(\tilde{\mu}, \Sigma)$

Ejemplo Sea $\tilde{X} = (X_1, X_2, X_3)^T \sim N_3(\tilde{\mu}, \Sigma)$ tal que:

$$\tilde{\mu} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 4 & 6 & 3 \\ 6 & 16 & 0 \\ 3 & 0 & 9 \end{pmatrix}$$

1. Obtenga la distribución de $X_2/X_1 = 2$ y calcule $\mathbb{P}(X_2 \leq 1/X_1 = 2)$

2. Defina:

$$\begin{cases} Y_1 = 3X_1 + X_2 + 4 \\ Y_2 = 2X_2 - X_3 + 5 \end{cases}$$

   Calcule la distribución de $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$

1.

$$\left.\begin{matrix} X_2 \sim N(0, 16) \\ X_1 \sim N(1, 4) \end{matrix}\right\} \Rightarrow X_2/X_1 = 2 \sim N(\mu_{2/1}, \sigma^2_{2/1})$$

Donde:

$$\mu_{2/1} = \mathbb{E}(X_2/X_1 = 2$$

$$= \mu_2 + \frac{\text{Cov}(X_1, X_2}{\sigma^2_{X_1}} \cdot (x_1 - \mu_1)$$

$$= 0 + \frac{6}{4}(2-1) = 1,5$$

Además

$$\sigma^2_{2/1} = \sigma^2_2(1 - \rho^2_{1,2})$$

$$\rho_{1,2} = \frac{6}{\sqrt{4}\sqrt{16}} = 0,75$$

$$\sigma^2_{2/1} = 16(1 - 0,75^2) = 7$$

Por lo tanto, se puede modelar dicha probabilidad condicional como una normal univariable.

$$X_2/X_1 = 2 \sim N(\frac{3}{2}; 7)$$

$$\mathbb{P}(X_2 \leq 1/X_1 = 2) = \mathbb{P}(Z \leq \frac{1 - \frac{3}{2}}{\sqrt{7}}) = \Phi(\frac{-1}{2\sqrt{7}}) = ?$$

2. Hay dos alternativas.

- **Paso a paso:** Como X_1, X_2, X_3 son normales entonces:

$$\left.\begin{matrix} Y_1 \sim N \\ Y_2 \sim N \end{matrix}\right\} \Rightarrow \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2$$

  Para $Y_1 = 3X_1 + X_2 + 4$

  $\mathbb{E}(Y_1) = 7$

  $\mathbb{V}(Y_1) = 3^2\sigma_1^2 + \sigma_2^2 + 2 \times 3\sigma_{12} = 9 \times 4 + 16 + 6 \times 6 = 88$

$$Y_1 \sim N(7, 88)$$

  Para $Y_2 = 2X_2 - X_3 + 5$

  $\mathbb{E}(Y_2) = 2\mu_2 - \mu_3 + 5 = 3$

  $\mathbb{V}(Y_2) = 4\sigma_2^2 + \sigma_3^2 - 2 \times 2\sigma_{23} = 4 \times 16 + 9 - 0 = 73$

$$Y_2 \sim N(3, 73)$$

  Por lo tanto...

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 7 \\ 3 \end{pmatrix}, \begin{pmatrix} 88 & 59 \\ 59 & 73 \end{pmatrix} \right)$$

  Luego: $\text{Cov}(Y_1, Y_2) = \text{Cov}(3X_1 + X_2 + 4, 2X_2 - X_3 + 5)$

  $\text{Cov}(Y_1, Y_2) = \text{Cov}(3X_1 + X_2, 2X_2 - X_3)$

  $\text{Cov}(Y_1, Y_2) = \text{Cov}(3X_1, 2X_2 - X_3) + \text{Cov}(X_2, 2X_2 - X_3)$

  $\text{Cov}(Y_1, Y_2) = (\text{Cov}(3X_1, 2X_2) - \text{Cov}(3X_1, X_3)) + (\text{Cov}(X_2, 2X_2) - \text{Cov}(X_2, X_3))$

  $\text{Cov}(Y_1, Y_2) = 6\sigma_{12} - 3\sigma_{13} + 2\sigma_2^2 - \sigma_{23} = 36 - 9 + 32 - 0$

$$\text{Cov}(Y_1, Y_2) = 59$$

- Alternativa:

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix} \qquad \tilde{b} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

$$\tilde{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2(\mu_y, \Sigma_y)$$

$$\mu_y = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

Para $\Sigma_y = A\ \mathbb{V}(\tilde{X})\ A^T$:

$$\Sigma_y = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 4 & 6 & 3 \\ 6 & 16 & 0 \\ 3 & 0 & 9 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 1 & 2 \\ 0 & -1 \end{pmatrix}$$

$$\Sigma_y = \begin{pmatrix} 88 & 59 \\ 59 & 73 \end{pmatrix}$$

1.12. Sumas de variables aleatorias

Definición Sea $X_1, ..., X_n$ una colección de n variables aleatorias. Se dice que $X_1, ..., X_n$ constituyen una MUESTRA ALEATORIA sí y sólo si las variables son INDEPENDIENTES E IDÉNTICAMENTE DISTRIBUIDAS. O sea:

$$X_1, ..., X_n \text{ es muestra aleatoria - m.a. (n)} \Rightarrow X_1, ..., X_n \text{ indep. e idénticamente distribuidas (iid)}$$

Considere $X_1, ..., X_n$ una m.a (n) y suponga que el interés es conocer la distribución del estadístico:

$$S(\tilde{X}) = \sum_{i=1}^{n} X_i$$

La distribución de S es posible de conocer en algunos casos específicos. Por ejemplo:

1. Si $X_1, ..., X_n \overset{\text{iid}}{\sim} \text{Ber}(p)$ entonces:
$$S = \sum_{i=1}^{n} X_i \sim \text{Bin}(n, p)$$

2. Si $X_1, ..., X_n \overset{\text{iid}}{\sim} \text{Bin}(m, p)$ entonces:
$$S = \sum_{i=1}^{n} X_i \sim \text{Bin}(nm, p)$$

3. Si $X_1, ..., X_n \overset{\text{iid}}{\sim} \text{Geo}(p)$ entonces:
$$S = \sum_{i=1}^{n} X_i \sim BN(n, p)$$

4. Si $X_1, ..., X_n \overset{\text{iid}}{\sim} \text{Exp}(\theta)$ entonces:
$$S = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \theta)$$

5. Si $X_1, ..., X_n \overset{\text{iid}}{\sim} \text{Gamma}(r, \theta)$ entonces:
$$S = \sum_{i=1}^{n} X_i \sim \text{Gamma}(nr, \theta)$$

6. Si $X_1, ..., X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ entonces:

$$S = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

pues: $\mathbb{E}[S] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\mu$

Análogamente: $\mathbb{V}[S] = \mathbb{V}[\sum_{i=1}^{n} X_i] \overset{\text{indep.}}{=} \sum_{i=1}^{n} \mathbb{V}[X_i] = n\sigma^2$

Observación Es de interés frecuentemente conocer la distribución de la media muestral: $\overline{X}$ pues se puede probar que:

$$\widehat{\mathbb{E}(X)} = \overline{X}$$

En particular, el caso de mayor interés es cuando $X_1, ..., X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Similarmente, en este escenario, se puede probar que:

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

Y también:

$$\frac{(n-1) \cdot S^2}{\sigma^2} \sim X^2(n-1)$$

Muchas distribuciones no arrojan modelos conocidos cuando las variables iid son sumadas. Por ejemplo: $X_1, ..., X_n \sim U(0,1) \Rightarrow \sum_{i=1}^{n} X_i \sim ???$
En este caso se puede usar **muestras grandes**.

1.13. Teorema del límite central

Sean $X_1, ..., X_n$ una m.a.(n) desde una población con distribución F_x tal que X es cuadrado integrable ($\mathbb{E}(X^2) < \infty$). Luego, para $S(\tilde{X}) = \sum_{i=1}^{n} X_i$ se tiene:

$$Z = \frac{S - \mathbb{E}[S]}{\sqrt{\mathbb{V}[S]}} \underset{n\to\infty}{\sim} N(0,1)$$

Observación Considere $g : \mathbb{R} \to \mathbb{R}$ una función real tal que $\mathbb{E}[(g(X))^2] < \infty$. Luego, para la estadística:

$$T(\tilde{X}) = \sum_{i=1}^{n} g(X_i)$$

Se tiene:

$$Z = \frac{T(\tilde{X}) - \mathbb{E}[T(\tilde{X})]}{\sqrt{\mathbb{V}(T(\tilde{X}))}} \underset{n\to\infty}{\sim} N(0,1)$$

1.14. Desigualdad de Markov

Sea X una variable aleatoria y sea $g : \mathbb{R} \to \mathbb{R}$ una función no negativa. Se verifica en consecuencia, que:

$$\mathbb{P}(g(x) > \epsilon) \leq \frac{\mathbb{E}[g(x)]}{\epsilon} \qquad \forall \epsilon > 0$$

Demostración Sea $g : \mathbb{R} \to \mathbb{R}$ una función no negativa, y considere el evento:

$A = \{x \in \mathbb{R} / g(x) \geq \epsilon\} \qquad \epsilon > 0$

Se tiene: $\mathbb{E}[g(x)] = \displaystyle\int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx$

Al ser un rectángulo mayor: $\mathbb{E}[g(x)] \geq \displaystyle\int_A \epsilon \cdot f_X(x)\, dx$

$\mathbb{E}[g(x)] \geq \epsilon \cdot \displaystyle\int_A f_X(x)\, dx = \epsilon \mathbb{P}(g(X) > \epsilon)$

Por lo tanto:

$$\mathbb{P}(g(x) > \epsilon) \leq \frac{\mathbb{E}[g(X)]}{\epsilon}$$

Caso particular Desigualdad de Chebyshev

Sea una X variable aleatorial tal que $\mathbb{E}(X) = \mu$ y $\mathbb{V}(X) = \sigma^2$. Entonces, $\forall k > 0$ se tiene que:

$$\mathbb{P}(|X - \mu| > k\sigma) \leq \frac{1}{k^2}$$

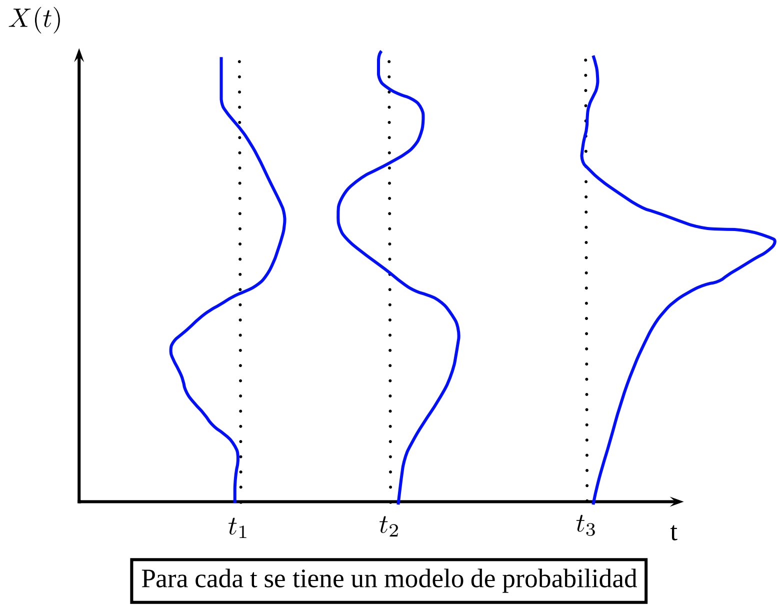
Demostración Se toma $g(x) = |x - \mu|$ y $\epsilon = k\sigma$

2. Procesos estocásticos

Hasta ahora, en los modelos, se ha mantenido una premisa por defecto: Las características aleatorias permanecen constantes en el tiempo.

En esta sección se incluirá una variable temporal que permitirá alterar las características aleatorias a través del tiempo.

Definición Un PROCESO ESTOCÁSTICO es una colección de variables aleatorias indexadas por un parámetro temporal (t). O sea: $\{X_t : t \in T\}$, donde t es no aleatorio y se suele asociar con "tiempo".



2.1. Estados

Posibles valores que puede asumir la variable aleatoria:

- Espacios de estado DISCRETOS
- Espacios de estado CONTINUOS

Como el conjunto T puede ser de características discretas o continuas, se tiene:

1. Procesos de estados discretos a tiempo discreto (*Cadenas de Markov*)
2. Procesos de estados discretos a tiempo continuo (*Procesos de saltos puros - Proceso de Poisson*)
3. Procesos de estados continuos a tiempo discreto (*Series de tiempo discreto*)
4. Procesos de estados continuos a tiempo continuo (*Procesos continuos - Series de tiempo*)

2.2. Procesos de estados discretos

Secuencia de variables que indique el valor del proceso en instantes sucesivos suele representarse como:

$$\{X_0 = x_0; X_1 = x_1, ..., X_n = x_n\}$$

2.2.1. Probabilidades de interés

- **Ocupación de estados:** $\mathbb{P}[X_n = x_n] \quad \leftarrow \quad$ depende del instante actual.
- **Probabilidad de transición:** $\mathbb{P}[X_n = x_n / \underbrace{X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_1 = x_1}_{\text{Depende de la historia previa}}]$

2.2.2. Propiedad de Markov

Una cadena se dice que cumple la PROPIEDAD DE MARKOV si:

$$\mathbb{P}[X_n = x_n / X_{n-1} = x_{n-1}, ..., X_1 = x_1, X_0 = x_0] = \mathbb{P}[X_n = x_n / \underbrace{X_{n-1} = x_{n-1}}_{\text{Corta memoria}}]$$

Observación La probabilidad anterior se suele denotar como:

$$P_{ij}(n) = \mathbb{P}[X_n = j / X_{n-1} = i] \quad \text{Probabilidad de transición}$$

Observación Cualquier cadena que cumpla la propiedad de Markov se denomina CADENA DE MARKOV.

Nota Cuando la probabilidad de transición no depennde del instante en que se está evaluando, se dice que la cadena es ESTACIONARIA. Es decir:

$$\mathbb{P}[X_n = j / X_{n-1} = i] = \mathbb{P}[X_1 = j / X_0 = i] \quad \forall n \in T$$

2.2.3. Cadenas de Markov

Considere un proceso que posee k estados: $S_1, S_2, ..., S_k$ y que las probabilidades de transición son estacionarias. Para cada $i, j = 1, ..., , k$, denote:

$$P_{ij} = \mathbb{P}[X_n = S_j / X_{n-1} = S_i$$

Entonces se define la MATRIZ DE TRANSICIÓN como:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1k} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{k1} & p_{k2} & p_{k3} & \cdots & p_{kk} \end{pmatrix}_{k \times k} \qquad \text{con } \sum_{j=1}^{k} P_{ij} = 1$$

Observación La matriz de transición presentada se dice que es a UN PASO, por lo cual:

$$P_{ij}(n) = \mathbb{P}[X_n = S_j / X_{n-1} = S_i]$$

Se le denomina PROBABILIDAD DE TRANSICIÓN A UN PASO.

Nota El trabajo con la matriz de transición P facilita el cálculo de las probabilidades de ocupación. De ahí es que $P \cdot P = P$ corresponde a la matriz de transición a dos pasos. En general $P^m, m \in \mathbb{N}$ corresponde a la matriz de transición a m pasos.

1. **Distribución inicial de una cadena de Markov**

   Vector de ocupación inicial de estados:

   $$\tilde{\Pi}_0 = (\pi_0(1), \pi_0(2), \pi_0(3), ..., \pi_0(k))$$

   donde: $\pi_0(j) = \mathbb{P}[X_0 = S_j]$

   con: $\displaystyle\sum_{j=1}^{k} \pi_0(j) = 1 \qquad \pi_0(j) \geq 0 \qquad \forall j = 1, ..., k$

   Cada una de las componentes de $\tilde{\Pi}_0$ entrega la probabilidad de estar en cada uno de los k estados existentes para el proceso.

2. **Distribución de las probabilidades de ocupación en un instante cualquiera**

   Vector de ocupación en un instante "n".

   $$\tilde{\Pi}_n = (\pi_n(1), \pi_n(2), ..., \pi_n(k))$$

   donde: $\pi_n(j) = \mathbb{P}[X_n = S_j]$

   con: $\displaystyle\sum_{j=1}^{k} \pi_n(j) = 1 \qquad \pi_n(j) \geq 0 \qquad \forall j = 1, ..., k$

En general se tiene la relación:

$$\pi_n(j) = \mathbb{P}[X_n = S_j] = \mathbb{P}[X_n = S_j/X_{n-1} = S_1] \cdot \mathbb{P}[X_{n-1} = S_1] + ... + \mathbb{P}[X_n = S_j/X_{n-1} = S_k] \cdot \mathbb{P}[X_{n-1} = S_k]$$

$$\pi_n(j) = \sum_{i=1}^{k} \mathbb{P}[X_n = S_j/X_{n-1} = S_i] \cdot \mathbb{P}[X_{n-1} = S_i]$$

$$\pi_n(j) = \sum_{i=1}^{k} \mathbb{P}[X_n = S_j/X_{n-1} = S_i] \cdot \sum_{l=1}^{k} \mathbb{P}[X_{n-1} = S_i/X_{n-2} = S_l] \cdot \mathbb{P}[X_{n-2} = S_l]$$

Ejemplo Considere las siguientes cadenas de Markov:

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \frac{1}{2} & \frac{2}{3} \\ 0 & 1 \end{pmatrix}$$

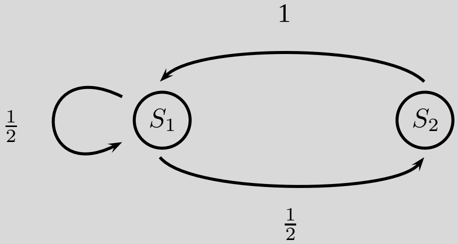
$$C = \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{5} & 0 & \frac{4}{5} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

1) $S = \{S_1, S_2\}$

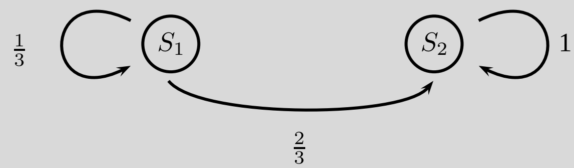
$\mathbb{P}[X_n = S_1/X_{n-1} = S_1] = 0,5]$ $\qquad$ $\mathbb{P}[X_n = S_1/X_{n-1} = S_2] = 1$

$\mathbb{P}[X_n = S/X_{n-1} = S_1] = 0,5]$ $\qquad$ $\mathbb{P}[X_n = S_2/X_{n-1} = S_2] = 0$

Diagramas de transición:



El estado S_2 se dice absorbente



Suponga las probabilidades iniciales:

$$\left.\begin{array}{l} \pi_0(1) = \mathbb{P}[X_0 = S_1] = \frac{1}{4} \\ \pi_0(2) = \mathbb{P}[X_0 = S_2] = \frac{3}{4} \end{array}\right\} \quad \tilde{\Pi} = \left(\frac{1}{4}\ ;\ \frac{3}{4}\right)$$

Consideran un paso en cada proceso:

1. $\tilde{\Pi} = \left(\dfrac{1}{4}\ ;\ \dfrac{3}{4}\right)\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{pmatrix} = \left(\dfrac{7}{8}\ ;\ \dfrac{1}{8}\right)$

2. $\tilde{\Pi} = \left(\dfrac{1}{4}\ ;\ \dfrac{3}{4}\right)\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & 0 \end{pmatrix} = \left(\dfrac{1}{12}\ ;\ \dfrac{11}{12}\right)$

Observación La distribución estacionaria se evalúa cuando $n \to \infty$. Interesa conocer:

$$P^* = \lim P^n$$

La matriz de estados estacionarios.

$$P^* = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_k \\ \pi_1 & \pi_2 & \cdots & \pi_k \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_k \end{pmatrix} \qquad \pi_i = \mathbb{P}(X_n = S_i),\ n \to \infty$$

Donde $\mathbb{P}(X_n = S_i)$ es la *probabilidad de ocupación del estado S_i a largo plazo.*

Para los valores π_i de la matriz P^* hay que considerar:

$$P^* \cdot P = P^*$$

Con lo cual:

$$\begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_k \\ \pi_1 & \pi_2 & \cdots & \pi_k \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_k \end{pmatrix} \cdot \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1k} \\ P_{21} & P_{22} & \cdots & P_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k1} & P_{k2} & \cdots & P_{kk} \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_k \\ \pi_1 & \pi_2 & \cdots & \pi_k \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_k \end{pmatrix}$$

De este modo, se genera un sistema de k ecuaciones.

$$\pi_1 \cdot P_{11} + \pi_2 \cdot P_{21} + \pi_3 \cdot P_{31} + \cdots + \pi_k \cdot P_{k1} = \pi_1 \quad \Rightarrow \quad \sum_{i=1}^{k} \pi_i \cdot P_{i1} = \pi_1$$

En general, las k ecuaciones se resumen:

$$\sum_{i=1}^{k} \pi_i \cdot P_{ij} = \pi_j \qquad\qquad j = 1, ..., k$$

Sin embargo, hay que agregar una condición adicional, puesto que:

$$\pi_1 + \pi_2 + ... + \pi_k = 1$$

Con lo cual se tienen $k+1$ ecuaciones y k incógnitas. Para evitar caer en inconsistencias, basta eliminar una de las primeras k ecuaciones.

Ejemplo Considere una Cadena de Markov con la siguiente matriz de transición:

$$P = \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix}$$

Encuentre la distribución estacionaria de la cadena de Markov.

Solución $S = \{S_1, S_2\}$. Supongamos $\pi_0 = (0,5; 0,5)$

$$\pi_0 = (0,5;0,5) \quad \Rightarrow \quad \pi_1 = (0,5;0,5) \cdot \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix} = (0,4;0,6)$$

$$\pi_1 = (0,4;0,6) \quad \Rightarrow \quad \pi_2 = (0,4;0,6) \cdot \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix} = (0,44;0,56)$$

$$\pi_2 = (0,44;0,56) \quad \Rightarrow \quad \pi_3 = (0,44;0,56) \cdot \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix} = (0,424;0,576)$$

Y así sucesivamente.
Continuando con el proceso, se obtiene la distribución estacionaria. Pero este método es poco eficiente.
Alternativa: $P^* = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix}$

$$P^* \cdot P = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix} \cdot \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix} = \begin{pmatrix} 0,2\pi_1 + 0,6\pi_2 & 0,8\pi_1 + 0,4\pi_2 \\ 0,2\pi_1 + 0,6\pi_2 & 0,8\pi_1 + 0,4\pi_2 \end{pmatrix}$$

Como $P^* \cdot P = P^*$

$$\begin{array}{l|} (1)\ 0,2\pi_1 + 0,6\pi_2 = \pi_1 \\ (2)\ 0,8\pi_1 + 0,4\pi_2 = \pi_2 \\ \hline \end{array}$$

Y la condición para que P^* sea estocástica:

$$(3)\ \pi_1 + \pi_2 = 1$$

Tenemos un sistema de tres ecuaciones y dos incógnitas, una es redundante por lo que trabajaremos sólo con (2) y (3).

$$\begin{array}{r|} 0,8\pi_1 + 0,4\pi_2 = \pi_2 \\ \pi_1 + \pi_2 = 1 \\ \hline \end{array}$$

Resolviendo se tiene:

$$\pi_1 = 0,4286 \qquad \wedge \qquad \pi_2 = 0,5714$$

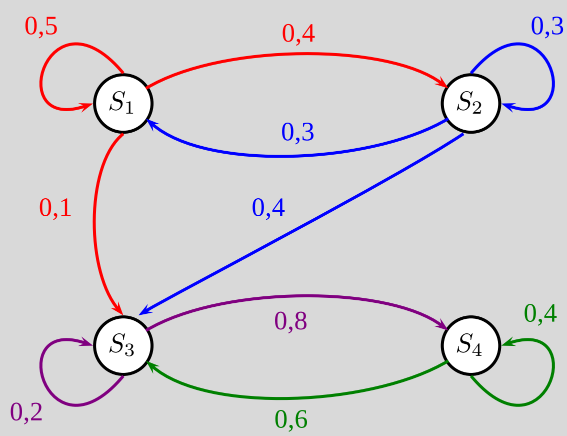
Distribución estacionaria:

$$\tilde{\Pi} = (0,4286; 0,5714)$$

Ejemplo Considere la siguiente matriz de transición:

$$P = \begin{pmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0,3 & 0,3 & 0,4 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,6 & 0,4 \end{pmatrix} \qquad S = \{S_1, S_2, S_3, S_4\}$$

Diagrama de transición:



En este caso, al conjunto $C_1 = \{S_1, S_2\}$ se le denomina CLASE DE PASO.
Al conjunto $C_2 = \{S_3, S_4\}$ se le denomina CLASE FINAL.

A raíz de lo anterior, para conocer la distribución estacionaria, se tiene:

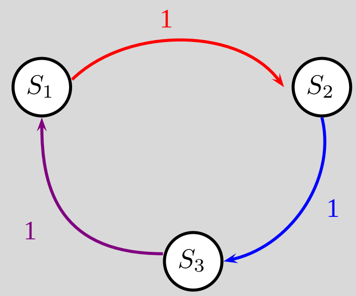
$$\left.\begin{aligned} 0,5\pi_1 + 0,3\pi_2 &= \pi_1 \\ 0,4\pi_1 + 0,3\pi_2 &= \pi_2 \\ 0,1\pi_1 + 0,4\pi_2 + 0,3\pi_3 + 0,6\pi_4 &= \pi_3 \\ 0,8\pi_3 + 0,4\pi_2 &= \pi_4 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \end{aligned}\right|$$

Como S_1 y S_2 son clases de paso, entonces $\pi_1 = 0$ y $\pi_2 = 0$
Luego $\pi_3 = \frac{3}{7}$ y $\pi_4 = \frac{4}{7}$.
Por lo tanto $\tilde{\Pi} = (0, 0, \frac{3}{7}, \frac{4}{7})$

Nota Cuando se tiene clases de paso, en la distribución estacionaria las probablidades de ocupación son 0 en cada uno de sus estados, lo cual reduce el sistema de ecuaciones a resolver.

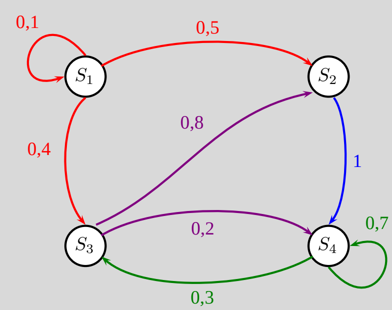
Cadena de Markov:

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$



Camino: $S_1 \to S_2 \to S_3 \to S_1$. **Corresponde a un ciclo de longitud 3.**
Cadena de Markov:

$$P = \begin{pmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0,3 & 0,3 & 0,4 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,6 & 0,4 \end{pmatrix}$$



Casos identificables:
$S_3 \to S_4 \to S_3$ (2)
$S_3 \to S_2 \to S_4 \to S_3$ (3)
$S_1 \to S_1$ (1)
$S_4 \to S_4$ (1)
$\text{MCD}\{1, 2, 3\} = 1 \Rightarrow p = 1$
Es periódica con período 1.

La PERIODICIDAD de una cadena de Markov viene dada por su período.
$P = \text{MCD}\{\text{Longitud de los ciclos distinguibles}\}$
$P > 1 \Rightarrow$ CM es periódica de período P

Ciclos:

$S_1 \to S_3 \to S_1$ (2)

$S_2 \to S_1 \to S_2$ (2)

$S_3 \to S_4 \to S_3$ (2)

$S_4 \to S_2 \to S_4$ (2)

$S_1 \to S_2 \to S_4 \to S_3 \to S_1$ (4)

$p = \mathrm{MCD}\{2, 4\} = 2$

Por lo tanto es periódica con $p = 2$.