



**UFM**  
UNIVERSIDAD  
FRANCISCO  
MARROQUÍN

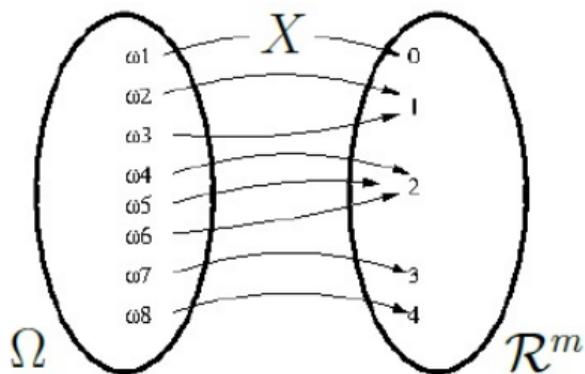
VERITAS • LIBERTAS • JUSTITIA

FACULTAD de  
**CIENCIAS ECONÓMICAS**

## **VARIABLES ALEATORIAS**

ALAN REYES-FIGUEROA  
ELEMENTS OF MACHINE LEARNING

(AULA 04) 24.ENERO.2023



## Definición

Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Una **variable aleatoria** (v.a.) es una función medible  $X : \Omega \rightarrow \mathbb{R}$ .

**Ejemplo:** Elegimos al azar una persona de un grupo. De cada persona tenemos un registro de su edad, altura, peso, ...

Mapeamos cada persona  $\omega$  a  $X(\omega) = (X_1(\omega), \dots, X_d(\omega))$ , donde por ejemplo  $X_1(\omega)$  representa su edad,  $X_2(\omega)$  su altura, etc.

Si el grupo de personas corresponde a una base de datos, entonces  $X$  regresa los campos de interés de cada registro. Las variables  $X_1, \dots, X_d$  son variables aleatorias. En este ejemplo llamaremos a  $X$  como una variable aleatoria (en realidad  $X$  es un **vector aleatorio**).

# Variables aleatorias

**Observación:** una variable aleatoria induce una función de probabilidad.

Definimos  $\mathbb{P}_X(\cdot)$  como

$$\mathbb{P}_X(\mathbf{A}) = P(\{\omega \in \Omega : X(\omega) \in \mathbf{A}\}).$$

Escribimos  $\mathbb{P}_X(\cdot)$  como  $\mathbb{P}(\cdot)$ .

Por ejemplo

- $\mathbb{P}(X = x)$  denota  $\mathbb{P}_X(X = x) = \mathbb{P}(\{\omega : X(\omega) = x\})$ .
- $\mathbb{P}(X < a)$  denota  $\mathbb{P}_X(X < a) = \mathbb{P}(\{\omega : X(\omega) < a\})$ .

Casi siempre vamos a medir qué ocurre con la probabilidad de nuestros eventos, mediante la probabilidad que describe una variable aleatoria de ellos (es más fácil medir en  $\mathbb{R}$  que en abstracto).

# Variables aleatorias

Caso discreto:

## Definición

Diremos que  $X$  es una variable aleatoria **discreta** si su contradominio  $I = X(\Omega)$  es enumerable y  $\mathbb{P}_X(i) = \mathbb{P}(X = i)$  existe para cada  $i \in I$ . (Comunmente se identifica el contradominio  $I$  con los naturales ).

## Definición

Al conjunto de probabilidades  $\{\mathbb{P}_X(i)\}_{i \in I}$  le llamamos la **distribución** de  $X$ . (En general, a  $\mathbb{P}_X$  se le llama la **función de masa de probabilidad**).

## Definición

Si  $X \in \mathbb{R}$ , llamamos a  $F_X(x) = \mathbb{P}(X \leq x)$  la **función de distribución** (acumulativa) de  $X$ .

# Variables aleatorias

Caso continuo:

## Definición

Considere la función  $F : \mathbb{R} \rightarrow \mathbb{R}$ , dada por

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

Diremos que  $X$  es una variable aleatoria **continua** si existe una función no-negativa  $f_X : \mathbb{R} \rightarrow \mathbb{R}$ , tal que

$$F_X(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t f_X(x) dx.$$

## Definición

En ese caso, a la función  $f_X$  le llamamos la **densidad de probabilidad** de  $X$ .

**Obs!** La función de densidad  $f_X$  no tiene por qué ser continua.

Ya sea en el caso discreto o continuo,

## Definición

Si  $x \in \mathbb{R}$ , llamamos a  $F_X(x) = \mathbb{P}(X \leq x)$  la **función de distribución** (acumulativa) de  $X$ .

En general, definimos la función de distribución para un vector aleatorio  $X = (X_1, \dots, X_d)$  como

$$F_X(x_1, \dots, x_d) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d), \quad \forall (x_1, \dots, x_d) \in \mathbb{R}^d.$$

En este caso, llamamos a  $F_X$  la **función de distribución conjunta** de  $X_1, \dots, X_d$ .

Otros nombres:

- distribución **univariada**, es la distribución en  $\mathbb{R}$  de una variable aleatoria.
- distribución **multivariada**, es la distribución en  $\mathbb{R}^d$  de un vector aleatorio.

# Propiedades

## Propiedades de $\mathbb{P}_X$ y $f_X$ :

Propiedad	X discreta	X continua
no-negativa	$\mathbb{P}_X(A) \geq 0$	$f_X(x) \geq 0$
suma total	$\sum_x \mathbb{P}_X(x) = 1$	$\int_{\mathbb{R}} f_X(x) dx = 1$
relación entre $f_X$ y $F_X$	$\mathbb{P}(X = x) = F_X(x) - F_X(x^-)$	$f_X(x) = \frac{d}{dx} F_X(x)$
relación entre $f_X$ y $F_X$	$F_X(x) = \sum_{t \leq x} \mathbb{P}(X = t)$	$F_X(x) = \int_{-\infty}^x f_X(t) dt$

# Propiedades

## Propiedades de $F_X$ :

Propiedad	$X$ discreta	$X$ continua
limitada	$0 \leq F_X(x) \leq 1$	$0 \leq F_X(x) \leq 1$
monotonía	$F_X$ no-decreciente	$F_X$ no-decreciente
límite inferior	$F_X(t) = 0, \forall t < \min_{x \in I(\Omega)}$	$\lim_{x \rightarrow -\infty} F_X(x) = 0$
límite superior	$F_X(t) = 1, \forall t \geq \max_{x \in I(\Omega)}$	$\lim_{x \rightarrow +\infty} F_X(x) = 1$

Además,  $F_X$  tiene la propiedad de semi-continuidad inferior:  $F_X$  es continua por la derecha, con límites por la izquierda.

# Distribuciones conjuntas

Cuando tenemos varias variables aleatorias (definida sobre el mismo espacio  $(\Omega, \mathcal{F}, \mathbb{P})$ ), podemos estudiar la distribución conjunta de dichas variables, esto es, la distribución de  $(X, Y)$ .

## Definición

La **distribución conjunta** de las v.a.  $X$  y  $Y$  se define por

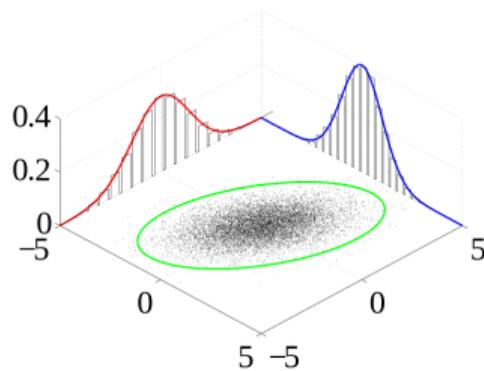
$$F_{X,Y}(X = x, Y = y) = \mathbb{P}(X \leq x, Y \leq y), \quad \forall x, y \in \mathbb{R}.$$

En el caso que  $X$  y  $Y$  son v.a. discretas, su **probabilidad conjunta** es

$$\mathbb{P}_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y), \quad \forall x, y \in \mathbb{R}.$$

En el caso en que  $X$  y  $Y$  son continuas, su **densidad conjunta** es

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial y \partial x}, \quad \forall x, y \in \mathbb{R}.$$



La normal bivariada es la distribución conjunta entre dos normales.

# Distribuciones marginales

Cuando tenemos varias variables aleatorias y su distribución conjunta, podemos “regresar” a las distribuciones originales.

## Definición

Dadas  $X$  y  $Y$  v.a. discretas y su distribución conjunta  $\mathbb{P}_{X,Y}$ , la **distribución marginal** para  $X$  y para  $Y$  son

$$\mathbb{P}_X(x) = \sum_y \mathbb{P}(x, y), \quad \mathbb{P}_Y(y) = \sum_x \mathbb{P}(x, y).$$

En el caso que  $X$  y  $Y$  son v.a. continuas, y  $f_{x,y}$  es su densidad conjunta, la **densidad marginal** de  $X$  y de  $Y$  son

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy, \quad f_Y(y) = \int_{\mathbb{R}} f(x, y) dx.$$

Ahora, si  $X, Y$  toman valores en  $[a, b] \times [c, d]$ , la **distribución marginal** se calcula como

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \int_a^x \int_b^y f_{X,Y}(s, t) ds dt.$$

luego  $F_X(x) = F_{X,Y}(x, d)$ ,  $F_Y(y) = F_{X,Y}(b, y)$ , y en el caso  $b = \infty$  ó  $d = \infty$

$$F_X(x) = \lim_{d \rightarrow \infty} F_{X,Y}(x, d), \quad F_Y(y) = \lim_{b \rightarrow \infty} F_{X,Y}(b, y).$$

## Definición

Sean  $X, Y$  v.a. discretas tales que  $\mathbb{P}(X = x) > 0$ . La **probabilidad condicional** de  $Y$  dado  $X = x$  es

$$\mathbb{P}_{Y|X}(y | x) = \mathbb{P}(Y = y | X = x) = \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)} = \frac{\mathbb{P}_{X,Y}(x, y)}{\mathbb{P}_X(x)}.$$

En el caso continuo, la **densidad condicional** de  $Y$  dado  $X$  es

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

Podemos escribir

- $\mathbb{P}_X(x) = \sum_y \mathbb{P}_{X|Y}(x | y) \mathbb{P}_Y(y).$
- $f_X(x) = \int_{\mathbb{R}} f_{X|Y}(x | y) f_Y(y) dy.$

# Independencia

Definimos la independencia de variables aleatorias de la siguiente manera:

## Definición

Dos variables aleatorias discretas  $X$  y  $Y$  definidas sobre el mismo espacio  $\Omega$  son **independientes** si

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

o equivalentemente,  $\mathbb{P}_{X,Y} = \mathbb{P}_X \cdot \mathbb{P}_Y$ .

En general, las v.a. discretas  $X_1, \dots, X_n$  son **mutuamente independientes** si

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n \mathbb{P}(X_i = x_i), \quad \forall x_1, x_2, \dots, x_n \in \mathbb{R}.$$

o equivalentemente,  $\mathbb{P}_{X_1, \dots, X_n} = \mathbb{P}_{X_1} \cdot \mathbb{P}_{X_2} \cdot \dots \cdot \mathbb{P}_{X_n}$ .

## Definición

Dos variables aleatorias continuas  $X$  y  $Y$  definidas sobre el mismo espacio  $\Omega$  son **independientes** si

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

En general, las v.a. continuas  $X_1, \dots, X_n$  son **mutuamente independientes** si

$$F_{X_1, X_2, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i), \quad \forall x_1, x_2, \dots, x_n \in \mathbb{R}.$$

Se puede mostrar que esto es equivalente a

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

$$f_{X_1, X_2, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i), \quad \forall x_1, x_2, \dots, x_n \in \mathbb{R}.$$