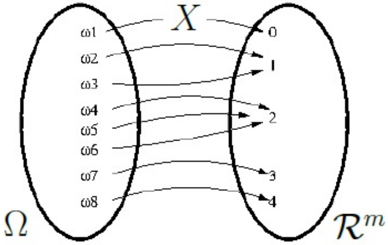
Variables aleatorias

Alan Reyes-Figueroa

Introducción a la Ciencia de Datos (Aula 04) 21.enero.2021



Definición

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

Aquí mesurable significa que si $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, entonces la preimagen de cualquier elemento en $\mathcal{B}(\mathbb{R})$ es un elemento de $\mathcal{F}$. Esto es, X^{-1} lleva conjuntos mesurables de $\mathbb{R}$ (bajo la medida de Lebesgue μ), a conjuntos mesurables en $\mathcal{F}$ (bajo la probabilidad $\mathbb{P}$).

A los elementos de $\mathcal{B}(\mathbb{R})$ se les llama los *borelianos* de $\mathbb{R}$.

Ejemplo
Elegimos al azar una persona de un grupo. De cada persona tenemos un registro de su edad, altura, peso, $\ldots$

Mapeamos cada persona ω a $X(\omega) = (X_1(\omega), \ldots, X_d(\omega))$, 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 $X_1, \ldots, X_d$ son variables aleatorias.
En este ejemplo llamaremos a X como una variable aleatoria (en realidad X es un vector aleatorio).

Variables aleatorias

Observaciones:

- una variable aleatoria determina una relación determinística.
- una variable aleatoria induce una función de probabilidad.

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

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

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

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

<u>Caso continuo</u>:

Definición

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

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

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

Propiedades

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 $X = (X_1, \ldots, X_d)$ como

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

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

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	$\mathbb{P}(X = x) = F_X(x) - F_X(x^-)$	$f_X(x) = \dfrac{d}{dx} F_X(x)$
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	$F_X(t) = 0, \forall t < \min_{x \in I(\Omega)}$	$\lim_{x \to -\infty} F_X(x) = 0$
límite superior	$F_X(t) = 1, \forall t \geq \max_{x \in I(\Omega)}$	$\lim_{x \to +\infty} F_X(x) = 1$

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*

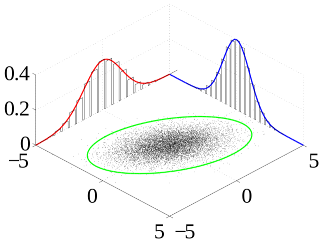
$$F_{X,Y}(X = x, Y = y) = \mathbb{P}(X \leq x, Y \leq y), \;\; \forall\, x, y \in \mathbb{R}.$$

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

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

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

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



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*

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

Distribuciones marginales

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

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

Distribuciones condicionales

Definición

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

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

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

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

Podemos escribir

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

Independencia

Definimos la independencia de variables aleatorias de la siguiente manera:

Definición

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

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

o equivalentemente, $\mathbb{P}_{X,Y} = \mathbb{P}_X \cdot \mathbb{P}_Y$.
En general, las v.a. discretas $X_1, \ldots, X_n$ son ***mutuamente independientes*** *si*

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

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

Independencia

Definición

Dos variables aleatorias continuas X y Y definidas sobre el mismo espacio Ω 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 $X_1, \ldots, X_n$ son ***mutuamente independientes*** *si*

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

Se puede mostrar que esto es equivalente a

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

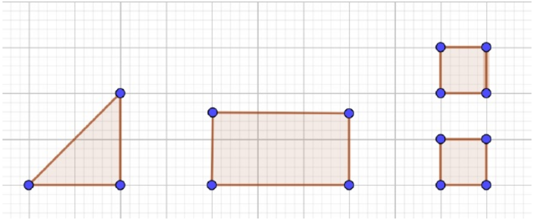
$$f_{X_1,X_2,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} f_{X_i}(x_i), \quad \forall x_1, x_2, \ldots, x_n \in \mathbb{R}.$$

Independencia

Casi siempre, una condición necesaria para que las variables X y Y sean independientes es que el *soporte* de (X, Y) (la región de $\mathbb{R}^2$ donde $\mathbb{P}_{X,Y} > 0$) sea un dominio rectangular, o producto de uniones de intervalos:

Ejemplo:

¿Cuáles variables X y Y son independientes?



UVG UNIVERSIDAD DEL VALLE DE GUATEMALA