Variáveis aleatórias

Em muitas situações é complexo lidar com todos os resultados possíveis de Ω e com as probabilidades associadas.

Exemplo - lançamento de dois dados e registo dos pontos das faces de cada um, para se definir a soma.

É mais cómodo associar a cada acontecimento um número real, definido de acordo com o objectivo do estudo. Esta correspondência define o que se costuma designar por **variável aleatória**.

Tipos de variáveis aleatórias

- Variáveis aleatórias **discretas** se assumem um conjunto finito ou infinito numerável de valores.
  Exemplos:
  – número de pintas que sai no lançamento de um dado;
  – registo, a intervalos regulares, do número de pessoas em fila espera na caixa de um supermercado;
- Variáveis aleatórias **contínuas** são as susceptíveis de tomar qualquer valor real num dado intervalo, que pode ser a recta real(definição mais rigorosa será dada à frente)
  Exemplos:
  – o peso de um indivíduo;
  – o comprimento de um folha.

Variáveis aleatórias discretas

Chama-se **distribuição de probabilidade** da v.a. X ao conjunto de pares $(x_i, p_i)_{i=1,\cdots n}$. Habitualmente a **lei (distribuição) de probabilidade** da v.a. X dispõe-se na forma:

$$X = \begin{cases} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{cases}$$

A **função de distribuição cumulativa** de uma variável aleatória discreta é assim calculada

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} P[X = x_i].$$

Vectores Aleatórios

Muitas vezes pretendemos associar a cada resultado de uma experiência aleatória $k \geq 2$ atributos numéricos. Obtemos então um vector $(x_1, \cdots, x_k)$, realização do **vector aleatório** $(X_1, \cdots, X_k)$.

Iremos referir-nos apenas ao caso $\boldsymbol{k = 2}$.

Exemplos Pretendemos registar:

- a quantidade de precipitado P e o volume V de gás numa experiência química
- para uma árvore seleccionada ao acaso, a altura e o diâmetro do tronco à altura do peito . . .

Pares Aleatórios

Definição Chama-se par aleatório (X, Y) à aplicação

$$\begin{aligned} (X, Y) : \Omega &\longrightarrow \mathbb{R}^2 \\ \omega &\longrightarrow (x, y) \end{aligned}$$

Tipos de pares aleatórios que vamos estudar:

- Par aleatório **discreto** $\Rightarrow$ componentes são ambas variáveis aleatórias discretas;
- Par aleatório **contínuo** $\Rightarrow$ componentes são ambas variáveis aleatórias contínuas.

Pares Aleatórios discretos

(X, Y) diz-se um par aleatório **discreto** se toma os valores (x_i, y_j) com probabilidades $p_{ij} = P[X = x_i, Y = y_j]$.

Definição Chama-se **distribuição de probabilidades conjunta** do par (X, Y) aos valores (x_i, y_j) e respectivas probabilidades p_{ij}

p_{ij} é chamada **função massa de probabilidade conjunta** e deve verificar as seguintes condições:

$$p_{ij} \geq 0 \qquad \forall i, j \qquad \text{e} \qquad \sum_i \sum_j p_{ij} = 1.$$

Um modo cómodo de representar a distribuição de probabilidades conjunta de um par aleatório discreto é na forma de uma tabela.

Pares Aleatórios discretos

Dado o par aleatório discreto (X, Y) chama-se **tabela de contingência** ou **tabela de dupla entrada** ao quadro na forma

$X \backslash Y$	y_1	y_2	$\cdots$	y_n	
x_1	p_{11}	p_{12}	$\cdots$	p_{1n}	$p_{1.}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2n}	$p_{2.}$
.	.	.	$\cdots$	.	.
.	.	.	$\cdots$	.	.
.	.	.	$\cdots$	.	.
x_m	p_{m1}	p_{m2}	$\cdots$	p_{mn}	$p_{m.}$
	$p_{.1}$	$p_{.2}$	$\cdots$	$p_{.n}$	1

$p_{i.} = \sum_{j=1}^{n} p_{ij}$ e $p_{.j} = \sum_{i=1}^{m} p_{ij}$ chamam-se **probabilidades marginais** de X e Y respectivamente.

Pares Aleatórios discretos

Define-se **probabilidade condicional** de X dado $Y = y_j$ (fixo) com $P[Y = y_j] > 0$ como

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{.j}},$$

Do mesmo modo **probabilidade condicional** de Y dado $X = x_i$ (fixo) com $P[X = x_i] > 0$ e x_i como

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i.}}.$$

Pares Aleatórios Contínuos

Define-se **densidade condicional** de X dado $Y = y$, fixo, como

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}, \qquad f_Y(y) > 0$$

Analogamente **densidade condicional** de Y dado $X = x$, fixo, como

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}, \qquad f_X(x) > 0$$

Independência de variáveis aleatórias

Definição Dado o par aleatório (X, Y) diz-se que as variáveis X e Y são **independentes** se e só se

- $p_{ij} = p_{i.}\, p_{.j} \quad \forall i, j$, se (X, Y) é um par aleatório discreto
- $f(x, y) = f_X(x)\, f_Y(y) \quad \forall (x, y) \in \mathbb{R}^2$ se (X, Y) é um par aleatório contínuo.

Coeficiente de Correlação

Definição: Chama-se **coeficiente de correlação** de X e Y e representa-se por ρ ou $\rho_{X,Y}$ a

$$\rho = \rho_{X,Y} = \frac{Cov[X,Y]}{\sigma_X \, \sigma_Y}$$

($\sigma_X > 0$ e $\sigma_Y > 0$).

Propriedades do coeficiente de correlação

1. $-1 \leq \rho_{X,Y} \leq 1$
2. Se X e Y são v. a. independentes $\Longrightarrow \quad \rho_{X,Y} = 0.$
3. $\rho_{aX+b,cY+d} = \rho_{X,Y}$ se $ac > 0$

   $= -\rho_{X,Y}$ se $ac < 0$