

Capítulo III: Vectores Aleatórios Reais

Probabilidades e Aplicações

Licenciatura em Matemática
Licenciatura em Ciências da Computação
Universidade do Minho
Ano Letivo 2025/2026

1. Definição e função distribuição de um vector aleatório real

Em muitas situações práticas, é frequente haver interesse em associar aos resultados decorrentes de uma experiência aleatória **mais do que uma** característica numérica. Em tais situações, interessa estudar, em simultâneo, várias v.a.r.'s e a existência de relações entre elas.

Nestes casos, a cada resultado elementar da experiência aleatória será associado um **vector de números reais**.

Definição [Vector aleatório real ($\vec{v.e.a.r.}$) e respetivas margens]

Sejam $n \in \mathbb{N}$ e (Ω, \mathcal{A}, P) o espaço de probabilidade associado a uma experiência aleatória. Chama-se vector aleatório real (abrevia-se por $\vec{v.e.a.r.}$) de dimensão n a todo o n -uplo (X_1, X_2, \dots, X_n) formado por n v.a.r.'s definidas sobre (Ω, \mathcal{A}, P) .

As v.a.r.'s

$$X_1, X_2, \dots, X_n$$

são designadas de margens (ou marginais) do $\vec{v.e.a.r.}$ (X_1, X_2, \dots, X_n) .

1. Definição e função distribuição de um vector aleatório real

Definições [σ -álgebra de Borel sobre \mathbb{R}^n e lei de probabilidade de um $\vec{v.e.a.r.}$]

- i) Chamamos σ -álgebra de Borel sobre \mathbb{R}^n , denota-se por $\mathcal{B}_n(\mathbb{R})$, à σ -álgebra gerada pela seguinte família, \mathcal{S} , de subconjuntos de \mathbb{R}^n

$$\mathcal{S} = \{B_1 \times B_2 \times \dots \times B_n : B_i \in \mathcal{B}(\mathbb{R}), i = 1, \dots, n\},$$

com $\mathcal{B}(\mathbb{R})$ a σ -álgebra de Borel sobre \mathbb{R} .

- ii) Seja (X_1, X_2, \dots, X_n) um $\vec{v.e.a.r.}$ definido sobre o espaço de probabilidade (Ω, \mathcal{A}, P) . Chama-se lei de probabilidade do (X_1, X_2, \dots, X_n) , e denota-se por $P_{(X_1, X_2, \dots, X_n)}$, à medida de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ definida por:

$$\begin{aligned} B \in \mathcal{B}_n(\mathbb{R}), P_{(X_1, X_2, \dots, X_n)}(B) &= P((X_1, X_2, \dots, X_n)^{-1}(B)) \\ &\equiv P(\{\omega \in \Omega : (X_1, X_2, \dots, X_n)(\omega) \in B\}) \\ &\equiv P((X_1, X_2, \dots, X_n) \in B) \end{aligned}$$

Obs.: $\mathcal{B}_n(\mathbb{R})$ não coincide com $\mathcal{P}(\mathbb{R}^n)$, i.e., $\mathcal{B}_n(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R}^n)$. No entanto, os subconjuntos de \mathbb{R}^n com que lidamos habitualmente são elementos de $\mathcal{B}_n(\mathbb{R})$.

1. Definição e função distribuição de um vector aleatório real

Nota: Sejam $n \in \mathbb{N}$ e (Ω, \mathcal{A}, P) o espaço de probabilidade associado a uma experiência aleatória. É possível mostrar que:

$$(X_1, X_2, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$$

é um *v.e.a.r.* se e só se

$$\forall B \in \mathcal{B}_n(\mathbb{R}), \quad (X_1, X_2, \dots, X_n)^{-1}(B) \in \mathcal{A},$$

com (Ω, \mathcal{A}, P) o espaço de probabilidade.

1. Definição e função distribuição de um vector aleatório real

Tal como acontece com as v.a.r.'s, a lei de probabilidade de um *v.e.a.r.* é caracterizada pela respectiva função de distribuição, que vamos definir de seguida.

Definição [Função de distribuição de um *v.e.a.r.*]

Seja (X_1, X_2, \dots, X_n) um *v.e.a.r.* definido sobre o espaço de probabilidade (Ω, \mathcal{A}, P) . Chama-se função de distribuição de (X_1, X_2, \dots, X_n) à função de distribuição da respectiva lei de probabilidade $P_{(X_1, X_2, \dots, X_n)}$, i.e., à função $F : \mathbb{R}^n \rightarrow [0, 1]$ definida por

$$\begin{aligned} F((c_1, c_2, \dots, c_n)) &= P_{(X_1, X_2, \dots, X_n)} (]-\infty, c_1] \times]-\infty, c_2] \times \dots \times]-\infty, c_n]) \\ &\equiv P(X_1 \leq c_1, X_2 \leq c_2, \dots, X_n \leq c_n) \end{aligned}$$

Nota: Entre v.a.r.'s, é usual a “,” para substituir a “ \cap ”. Assim,

$$P(X_1 \leq c_1, X_2 \leq c_2, \dots, X_n \leq c_n) \equiv P((X_1 \leq c_1) \cap (X_2 \leq c_2) \cap \dots \cap (X_n \leq c_n)).$$

2. Vector aleatório real discreto

Definição [$\overrightarrow{ve}.a.r.$ discreto e respetivo contradomínio]

Sejam (Ω, \mathcal{A}, P) um espaço de probabilidade e (X_1, X_2, \dots, X_n) um $\overrightarrow{ve}.a.r.$ com lei de probabilidade $P_{(X_1, X_2, \dots, X_n)}$. Este vector diz-se discreto se a respetiva lei, $P_{(X_1, X_2, \dots, X_n)}$, for uma lei discreta sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$, i.e., se existir um elemento M de $\mathcal{B}_n(\mathbb{R})$, finito ou infinito numerável, tal que

$$P_{(X_1, X_2, \dots, X_n)}(M) = 1.$$

Ao menor conjunto M que satisfaz esta condição chamamos contradomínio de (X_1, X_2, \dots, X_n) e denotamos por $C_{(X_1, X_2, \dots, X_n)}$.

2. Vector aleatório real discreto

Se (X_1, \dots, X_n) é um *v.e.a.r.* discreto, de contradomínio C , então a sua lei de probabilidade, $P_{(X_1, \dots, X_n)}$, é caracterizada por uma função

$$f : \mathbb{R}^n \rightarrow [0, 1],$$

designada de *função de probabilidade conjunta de* (X_1, \dots, X_n) , definida por

$$f((a_1, \dots, a_n)) = \begin{cases} P_{(X_1, \dots, X_n)}(\{(a_1, \dots, a_n)\}) & \text{se } (a_1, \dots, a_n) \in C \\ 0 & \text{se } \dots \text{ c.c.} \end{cases}$$

Nota: Observe que $P_{(X_1, \dots, X_n)}(\{(a_1, \dots, a_n)\})$ representa

$$P((X_1, \dots, X_n) = (a_1, \dots, a_n))$$

que se reduz a

$$P(X_1 = a_1, \dots, X_n = a_n).$$

2. Vector aleatório real discreto

Se (X_1, \dots, X_n) é um *v.e.a.r.* discreto, de contradomínio C e função de probabilidade conjunta $f : \mathbb{R}^n \rightarrow [0, 1]$, então, para todo $B \in \mathcal{B}_n(\mathbb{R})$, tem-se

$$P((X_1, X_2, \dots, X_n) \in B) = \sum_{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n : (a_1, a_2, \dots, a_n) \in B \cap C} f((a_1, a_2, \dots, a_n)).$$

Nota: No caso dos pares aleatórios ($n = 2$), é usual apresentar a função de probabilidade conjunta através de uma tabela de dupla entrada. Esta tabela permite aceder muito facilmente ao cálculo de probabilidades do tipo acima referido. Permite também fácil acesso às funções de probabilidade das duas v.a.r.'s marginais.

2. Vector aleatório real discreto

~ Como obter as leis de probabilidade das margens de um $\overrightarrow{v.e.a.r.}$ discreto?

Seja (X_1, X_2, \dots, X_n) um $\overrightarrow{v.e.a.r.}$ discreto, de contradomínio C e função de probabilidade conjunta f . Então as v.a.r.'s X_1, X_2, \dots, X_n são discretas e, para a i -ésima v.a.r. marginal, X_i , tem-se que:

- o contradomínio é dado por

$$C_{X_i} = \{a \in \mathbb{R} : (x_1, x_2, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \in C\}$$

- a função de probabilidade, $f_{X_i} : \mathbb{R} \rightarrow [0, 1]$, é dada por

$$f_{X_i}(a) = \sum_{x_1 \in C_{X_1}, \dots, x_{i-1} \in C_{X_{i-1}}, x_{i+1} \in C_{X_{i+1}}, \dots, x_n \in C_{X_n}} f((x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)),$$

para $a \in C_{X_i}$.

3. Vector aleatório real absolutamente contínuo

Definição [v.e.a.r. difuso]

Um v.e.a.r. (X_1, X_2, \dots, X_n) diz-se difuso se a sua lei de probabilidade for uma lei de probabilidade difusa sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$, i.e., se

$$P_{(X_1, X_2, \dots, X_n)}(\{(a_1, a_2, \dots, a_n)\}) = 0, \quad \forall (a_1, a_2, \dots, a_n) \in \mathbb{R}^n. \quad (1)$$

Obs.: Note que a igualdade (1) é equivalente a

$$\forall (a_1, a_2, \dots, a_n) \in \mathbb{R}^n, P((X_1, X_2, \dots, X_n) = (a_1, a_2, \dots, a_n)) = 0,$$

e ainda a

$$\forall (a_1, a_2, \dots, a_n) \in \mathbb{R}^n, P(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n) = 0.$$

Tal como no caso univariado (apresentado no Cap. II), nesta UC vamos estudar apenas as leis de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ difusas que são absolutamente contínuas. Estas leis são agora caracterizadas através de uma função densidade de probabilidade sobre \mathbb{R}^n .

3. Vector aleatório real absolutamente contínuo

Definição [Função densidade de probabilidade sobre \mathbb{R}^n]

Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diz-se uma função densidade de probabilidade sobre \mathbb{R}^n se satisfaz as seguintes condições:

- i) $f((x_1, x_2, \dots, x_n)) \geq 0$, para todo o $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$,
- ii) f é integrável em \mathbb{R}^n e

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f((x_1, x_2, \dots, x_n)) dx_1 dx_2 \dots dx_n = 1.$$

3. Vector aleatório real absolutamente contínuo

Definição [$\overrightarrow{ve.a.r.}$ absolutamente contínuo]

Um $\overrightarrow{ve.a.r.} (X_1, X_2, \dots, X_n)$ diz-se absolutamente contínuo se a sua lei de probabilidade for uma lei de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ absolutamente contínua, i.e., se existir $f : \mathbb{R}^n \rightarrow \mathbb{R}$ função densidade probabilida-de sobre \mathbb{R}^n , tal que, para todos os

$$a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n, \text{ com } a_i < b_i, i = 1, \dots, n,$$

se tem

$$\begin{aligned} P(a_1 < X_1 < b_1, a_2 < X_2 < b_2, \dots, a_n < X_n < b_n) &= \\ &= \int_{a_1}^{b_1} \left[\int_{a_2}^{b_2} \cdots \left[\int_{a_n}^{b_n} f((x_1, x_2, \dots, x_n)) dx_n \right] \cdots dx_2 \right] dx_1. \end{aligned}$$

A função f é designada de função densidade de probabilidade conjunta do $\overrightarrow{ve.a.r.} (X_1, X_2, \dots, X_n)$.

3. Vector aleatório real absolutamente contínuo

De um modo geral, se (X_1, X_2, \dots, X_n) é um *v.e.a.r.* absolutamente contínuo, com função densidade de probabilidade conjunta $f : \mathbb{R}^n \rightarrow \mathbb{R}$, tem-se, para todo $B \in \mathcal{B}_n(\mathbb{R})$,

$$P((X_1, X_2, \dots, X_n) \in B) = \int_B \int \dots \int f((x_1, x_2, \dots, x_n)) dx_1 dx_2 \dots dx_n.$$

3. Vector aleatório real absolutamente contínuo

~ Como obter as leis de probabilidade das margens de um $\overrightarrow{v.e.a.r.}$ absolutamente contínuo?

Se (X_1, X_2, \dots, X_n) é um $\overrightarrow{v.e.a.r.}$ absolutamente contínuo, com função densidade de probabilidade conjunta $f : \mathbb{R}^n \rightarrow \mathbb{R}$, então as margens X_1, X_2, \dots, X_n são v.a.r.'s absolutamente contínuas e, para a i -ésima v.a.r. marginal, X_i , a função densidade de probabilidade $f_{X_i} : \mathbb{R} \rightarrow \mathbb{R}$ é dada por

$$f_{X_i}(x) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f((x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n,$$

para $i = 1, \dots, n$.

4. Variáveis aleatórias reais independentes

Definição [Independência de v.a.r's]

Sejam X_1, X_2, \dots, X_n v.a.r's, todas definidas sobre o mesmo espaço de probabilidade (Ω, \mathcal{A}, P) , com leis de probabilidade $P_{X_1}, P_{X_2}, \dots, P_{X_n}$, respectivamente.

X_1, X_2, \dots, X_n dizem-se v.a.r's independentes se, para todos os $B_1, B_2, \dots, B_n \in \mathcal{B}(\mathbb{R})$,

$$P_{(X_1, X_2, \dots, X_n)}(B_1 \times B_2 \times \dots \times B_n) = P_{X_1}(B_1)P_{X_2}(B_2)\dots P_{X_n}(B_n), \quad (2)$$

com $P_{(X_1, X_2, \dots, X_n)}$ a lei de probabilidade do *v.e.a.r.* (X_1, X_2, \dots, X_n) .

Nota: Observe que a condição (2) pode ser escrita do seguinte modo:

$$P(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n) = P(X_1 \in B_1)P(X_2 \in B_2)\dots P(X_n \in B_n).$$

4. Variáveis aleatórias reais independentes

Teorema

Sejam X_1, X_2, \dots, X_n v.a.r.'s independentes e g_1, g_2, \dots, g_n funções tais que, para $i = 1, \dots, n$,

- $g_i : D_i \subseteq \mathbb{R} \rightarrow \mathbb{R}$,
- $g_i(X_i)$ é uma v.a.r..

Então as v.a.r.'s $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$ também são independentes.

[Sem demonstração] (Ver Lopes & Gonçalves)

Exemplos: Da utilização deste teorema resulta, por exemplo, que: se X e Y são v.a.r.'s independentes então

- para todo $k \in \mathbb{N}$, as v.a.r.'s X^k e Y^k também são independentes;
- desde que $Y > 0$, as v.a.r.'s e^X e $\log(Y)$ também são independentes;
- quaisquer que sejam as constantes reais a, b, c e d , as v.a.r.'s

$$aX + b \text{ e } cY + d$$

também são independentes.

4. Variáveis aleatórias reais independentes

~ Condições necessárias e suficientes para independência de v.a.r's:

Provar, por definição, que X_1, X_2, \dots, X_n são v.a.r's independentes é extremamente difícil. Precisamos de ter "formas alternativas" para o fazer, i.e., de condições necessárias e suficientes para estabelecer a independência de v.a.r's.

1) X_1, X_2, \dots, X_n são v.a.r's independentes se e só se a função de distribuição conjunta do *v.e.a.r.* (X_1, X_2, \dots, X_n) , F verifica a seguinte condição:

$$\forall (c_1, c_2, \dots, c_n) \in \mathbb{R}^n, \quad F((c_1, c_2, \dots, c_n)) = \prod_{i=1}^n F_{X_i}(c_i),$$

onde F_{X_i} representa a função de distribuição de X_i , $i = 1, \dots, n$.

[Demonstração] (\Rightarrow) é consequência imediata de definição de independência de v.a.r's; ver Lopes & Gonçalves para (\Leftarrow).

4. Variáveis aleatórias reais independentes

Condições necessárias e suficientes para independência de v.a.r's (cont.)

2) Sejam X_1, X_2, \dots, X_n v.a.r's discretas, cujos contradomínios são $C_{X_1}, C_{X_2}, \dots, C_{X_n}$, respectivamente. Neste caso, X_1, X_2, \dots, X_n são v.a.r's independentes se e só

$$\forall a_1 \in C_{X_1}, a_2 \in C_{X_2}, \dots, a_n \in C_{X_n}, \quad P(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n) = \prod_{i=1}^n P(X_i = a_i). \quad (3)$$

[Demonstração] (\Rightarrow) é consequência imediata de definição de independência de v.a.r's; ver Lopes & Gonçalves para (\Leftarrow).

Nota: A condição (3) equivale a dizer que a função de probabilidade conjunta do *v.e.a.r.* (X_1, X_2, \dots, X_n) é igual ao produto das funções de probabilidade das v.a.r's marginais.

4. Variáveis aleatórias reais independentes

Condições necessárias e suficientes para independência de v.a.r's (cont.)

3) Uma condição necessária e suficiente para que um $\overrightarrow{v.e.a.r}$. absolutamente contínuo, (X_1, X_2, \dots, X_n) , tenha as margens independentes é que uma função densidade de probabilidade conjunta, $f_{(X_1, X_2, \dots, X_n)}$, seja da forma

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

onde $f_{X_1}, f_{X_2}, \dots, f_{X_n}$ são funções densidade de probabilidade das margens X_1, X_2, \dots, X_n , respectivamente.

[Sem demonstração] (Ver Lopes & Gonçalves)

4. Variáveis aleatórias reais independentes

Antes de prosseguir com mais resultados relativos à independência de v.a.r.'s, vejamos a definição de esperança matemática (ou valor médio) de uma função de um $\overrightarrow{v.e.a.r.}$.

Definição [Esperança matemática de uma função de um $\overrightarrow{v.e.a.r.}$ discreto]

Sejam (X_1, \dots, X_n) um $\overrightarrow{v.e.a.r.}$ e $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função tal que $\phi((X_1, \dots, X_n))$ é uma v.a.r..

- i) Se (X_1, \dots, X_n) é um $\overrightarrow{v.e.a.r.}$ discreto, de contradomínio C e função de probabilidade conjunta f , e tal que

$$\sum_{(x_1, \dots, x_n) \in C} |\phi((x_1, \dots, x_n))| f((x_1, \dots, x_n)) < +\infty$$

então $E[\phi((X_1, \dots, X_n))]$ existe e é dada por

$$E[\phi((X_1, \dots, X_n))] = \sum_{(x_1, \dots, x_n) \in C} \phi((x_1, \dots, x_n)) f((x_1, \dots, x_n)).$$

Obs.: Recorde que, neste caso discreto,

$$f((x_1, \dots, x_n)) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n).$$

4. Variáveis aleatórias reais independentes

Definição [Esperança matemática de uma função de um $\overrightarrow{v.e.a.r.}$ contínuo]

- ii) Se (X_1, \dots, X_n) é um $\overrightarrow{v.e.a.r.}$ absolutamente contínuo, com função densidade de probabilidade conjunta f , e tal que

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} |\phi((x_1, \dots, x_n))| f((x_1, \dots, x_n)) dx_1 \dots dx_n < +\infty$$

então $E[\phi((X_1, \dots, X_n))]$ existe e é dada por

$$E[\phi((X_1, \dots, X_n))] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \phi((x_1, \dots, x_n)) f((x_1, \dots, x_n)) dx_1 \dots dx_n$$

Nota: Se, nestas duas últimas definições, usarmos a função $\phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$\phi_i((x_1, x_2, \dots, x_i, \dots, x_{n-1}, x_n)) = x_i,$$

somos obviamente conduzidos a $E[X_i]$, $i \in \{1, 2, \dots, n\}$.

4. Variáveis aleatórias reais independentes

Observação:

A partir destas duas últimas definições, facilmente se conclui (TPC) que, se (X_1, \dots, X_n) é um *v.e.a.r.* possuindo esperança matemática (i.e., $E[X_i]$ existe para todo $i = 1, \dots, n$), então

$$E[a_1X_1 + a_2X_2 + \dots + a_nX_n] = \sum_{i=1}^n a_iE[X_i],$$

para quaisquer constantes reais a_1, a_2, \dots, a_n .

4. Variáveis aleatórias reais independentes

Teorema

Seja (X_1, \dots, X_n) um $\overrightarrow{v.e.a.r.}$ possuindo esperança matemática (i.e., $E[X_i]$ existe para todo $i = 1, \dots, n$) e com margens independentes. Então

$$E \left[\prod_{i=1}^n X_i \right] = \prod_{i=1}^n E [X_i].$$

Demonstração: Vai ser feita apenas para o caso em que (X_1, X_2, \dots, X_n) é um $\overrightarrow{v.e.a.r.}$ absolutamente contínuo (TPC: caso discreto).

Comecemos por provar que

$$E \left[\prod_{i=1}^n X_i \right]$$

existe.

4. Variáveis aleatórias reais independentes

Sejam f a função densidade de probabilidade conjunta do v.e.a.r. e f_{X_i} a função densidade de probabilidade da v.a.r. X_i , $i = 1, \dots, n$.

$$\begin{aligned} E[|X_1 \dots X_n|] &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |x_1 \dots x_n| f((x_1, \dots, x_n)) dx_1 \dots dx_n \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |x_1| \dots |x_n| f_{X_1}(x_1) \dots f_{X_n}(x_n) dx_1 \dots dx_n \\ &= \underbrace{\int_{-\infty}^{+\infty} |x_1| f_{X_1}(x_1) dx_1}_{< +\infty} \dots \underbrace{\int_{-\infty}^{+\infty} |x_n| f_{X_n}(x_n) dx_n}_{< +\infty} \\ &< +\infty. \end{aligned}$$

E de modo análogo tem-se

$$\begin{aligned} E[X_1 X_2 \dots X_n] &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} x_1 \dots x_n f((x_1, \dots, x_n)) dx_1 \dots dx_n \\ &= \underbrace{\int_{-\infty}^{+\infty} x_1 f_{X_1}(x_1) dx_1}_{E[X_1]} \dots \underbrace{\int_{-\infty}^{+\infty} x_n f_{X_n}(x_n) dx_n}_{E[X_n]} \end{aligned}$$

c.q.d

5. Covariância e correlação

Recordar que, para caracterizar a dispersão de uma v.a.r., em torno do seu valor médio, fazemos uso de variância e/ou o desvio-padrão.

Para um $\vec{v.e.a.r.}$ a questão é mais complexa e é preciso introduzir um parâmetro que permita avaliar o tipo e o grau de dependência entre os pares de variáveis marginais.

Para esse efeito, serão aqui apresentados duas tais medidas, e respetivas propriedades: a *covariância* e o *coeficiente de correlação*.

5. Covariância e correlação

Definição [Covariância]

Sejam X e Y v.a.r.'s definidas sobre o mesmo espaço de probabilidade (Ω, \mathcal{A}, P) e ambas admitindo momento de segunda ordem. Chama-se covariância entre X e Y , denota-se por $Cov(X, Y)$, a

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))].$$

De certo modo, a covariância mede a dispersão do par aleatório (X, Y) relativamente ao par $(E[X], E[Y])$.

Observação: $Cov(X, X) = Var[X]$.

5. Covariância e correlação

~ Propriedades da covariância:

Sejam $X, Y, X_1, X_2, \dots, X_n$ v.a.r's, todas admitindo momento de segunda ordem (e, sempre que necessário, definidas sobre o mesmo espaço de probabilidade).

1) $\forall a, b, c, d \in \mathbb{R}$, $Cov(aX + c, bY + d) = abCov(X, Y)$.

Demonstração:

$$\begin{aligned} Cov(aX + c, bY + d) &= E[(aX + c - E(aX + c))(bY + d - E(bY + d))] \\ &= E[a(X - E(X))b(Y - E(Y))] \\ &= abE[(X - E(X))(Y - E(Y))] \\ &= abCov(X, Y) \end{aligned} \quad \text{c.q.d.}$$

5. Covariância e correlação

2) $Cov(X, Y) = E[XY] - E[X]E[Y]$

Demonstração:

$$\begin{aligned} Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E[XY] - E[X]E(Y) - E[Y]E(X) + E(X)E(Y) \\ &= E[XY] - E[X]E[Y] \quad \text{c.q.d.} \end{aligned}$$

Nota: Na dedução acima usamos o facto de $E[XY]$ existir. É possível mostrar (ver Lopes & Gonçalves) que, se X e Y admitem momento de segunda ordem, então $E[|XY|] < +\infty$, pelo que $E[XY]$ existe.

5. Covariância e correlação

3) $\text{Var} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \text{Var}[X_i] + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_i, X_j)$

Demonstração:

$$\begin{aligned}\text{Var} \left[\sum_{i=1}^n X_i \right] &= E \left[\left(\sum_{i=1}^n X_i - E \left[\sum_{i=1}^n X_i \right] \right)^2 \right] = E \left[\left(\sum_{i=1}^n (X_i - E[X_i]) \right)^2 \right] \\ &= E \left[\sum_{i=1}^n (X_i - E[X_i])^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n (X_i - E[X_i])(X_j - E[X_j]) \right] \\ &= \sum_{i=1}^n E[(X_i - E[X_i])^2] + 2 \sum_{i=1}^n \sum_{j=i+1}^n E[(X_i - E[X_i])(X_j - E[X_j])] \\ &= \sum_{i=1}^n \text{Var}[X_i] + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_i, X_j)\end{aligned}$$

c.q.d

Observação Importante: Se X_1, X_2, \dots, X_n são v.a.r.'s **independentes**, tem-se que $\text{Cov}(X_i, X_j) = 0$, para $i \neq j$, e, consequentemente, tem-se

$$\text{Var} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \text{Var}[X_i].$$

5. Covariância e correlação

Usando a covariância entre duas v.a.r.'s, define-se um outro parâmetro, o *coeficiente de correlação*, que mede o grau e o tipo de dependência existente entre as v.a.r.'s.

Definição [Coeficiente de correlação]

Sejam X e Y v.a.r.'s, definidas sobre o mesmo espaço de probabilidade (Ω, \mathcal{A}, P) , com variâncias σ_X^2 e σ_Y^2 estritamente positivas. Chama-se coeficiente de correlação entre X e Y , denota-se por $\rho(X, Y)$, ao número real

$$\rho(X, Y) = Cov \left(\frac{X - E(X)}{\sigma_X}, \frac{Y - E(Y)}{\sigma_Y} \right),$$

em que σ_X e σ_Y denotam, respetivamente, o desvio-padrão de X e o desvio-padrão de Y .

5. Covariância e correlação

~Algumas observações sobre $\rho(X, Y)$:

1) Usando propriedades da covariância, é simples deduzir que

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

2) É possível mostrar (ver Lopes & Gonçalves) que

$$|\rho(X, Y)| \leq 1$$

e que

$$|\rho(X, Y)| = 1 \Leftrightarrow \exists_{a,b,c \in \mathbb{R}: ab \neq 0} : P(aX + bY + c = 0) = 1.$$

Este resultado permite-nos concluir que o coeficiente de correlação quantifica a existência de uma **relação linear** entre X e Y .

3) Quando $\rho(X, Y) = 0$ dizemos que X e Y não estão correlacionadas (significa apenas que não existe uma relação linear entre X e Y).

4) Se X e Y são independentes tem-se, obviamente, $\rho(X, Y) = 0$.

Mas o recíproco é falso: podemos ter v.a.r.'s não correlacionadas mas que não são independentes (ver exercícios da Folha Prática 8).

6. Algumas leis de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ mais conhecidas

Entre as leis discretas multivariadas, a mais conhecida é a distribuição *Multinomial*, que generaliza a distribuição Binomial (univariada).

Entre as leis absolutamente contínuas multivariadas, a mais conhecida é a distribuição *Normal multivariada*, que generaliza a distribuição Normal (univariada).

6. Algumas leis de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ mais conhecidas

~Lei Multinomial: (no que se segue, r e n são inteiros ≥ 2)

Considere a experiência aleatória, ξ , na qual ocorrem, **em alternativa**, “sucesso de tipo 1”, “sucesso de tipo 2”, ..., “sucesso de tipo ($r-1$)” ou “insucesso” (i.e., não ocorre qualquer tipo de sucesso), com probabilidades p_1, p_2, \dots, p_{r-1} e $p_r = 1 - (p_1 + p_2 + \dots + p_{r-1})$, respectivamente. Seja $(X_1, X_2, \dots, X_{r-1})$ o $\vec{\text{v.e.a.r.}}$ em que X_i é a v.a.r. que representa o número de vezes que ocorreu o sucesso de tipo i em n repetições independentes de ξ , $i = 1, \dots, r-1$.

A função de probabilidade conjunta deste $\vec{\text{v.e.a.r.}}$ é dada por:

$$P(X_1 = n_1, X_2 = n_2, \dots, X_{r-1} = n_{r-1}) = \frac{n!}{n_1! n_2! \dots n_{r-1}! n_r!} p_1^{n_1} p_2^{n_2} \dots p_{r-1}^{n_{r-1}} p_r^{n_r},$$

em que $n_i \in \mathbb{N}_0$, $i = 1, \dots, n_r$, e, adicionalmente, satisfazem as condições

$$n_1 + n_2 + \dots + n_{r-1} \leq n \quad \text{e} \quad n_r = n - (n_1 + n_2 + \dots + n_{r-1}).$$

Nestas condições, diz-se que o $\vec{\text{v.e.a.r.}}$ $(X_1, X_2, \dots, X_{r-1})$ segue a lei Multinomial, com parâmetros n e p_1, p_2, \dots, p_{r-1} , e abrevia-se por

$$(X_1, X_2, \dots, X_{r-1}) \sim M(n; p_1, p_2, \dots, p_{r-1}).$$

6. Algumas leis de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ mais conhecidas

Exemplo/Exercício da Folha Prática 8:

Numa lotaria com 10.000 bilhetes, numerados de 0000 a 9999, qual é a probabilidade de o primeiro prémio ir para um bilhete com exatamente dois dígitos ímpar e exatamente um zero?

Sugestão: Recorra a um $\vec{v.e.a.r.}$ (de dimensão 2) e que siga uma lei Multinomial apropriada.

6. Algumas leis de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ mais conhecidas

↪ Lei Normal Multivariada:

Diz-se que um *v.e.a.r.*, (X_1, X_2, \dots, X_p) , absolutamente contínuo segue a lei Normal Multivariada, com parâmetros $\mathbf{u} = [\mu_1 \ \mu_2 \ \dots \ \mu_p]^\top \in \mathbb{R}^p$ e $\Sigma = [\sigma_{ij}]_{i,j=1}^p$, se a função densidade de probabilidade conjunta do vetor é dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2} \sqrt{\det(\Sigma)}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{u})^\top \Sigma^{-1} (\mathbf{x} - \mathbf{u}) \right\},$$

para todo o $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]^\top \in \mathbb{R}^p$, com Σ é uma matriz real, quadrada de ordem p , invertível, simétrica e positiva definida. Abrevia-se por

$$(X_1, X_2, \dots, X_p) \sim N_p(\mathbf{u}, \Sigma).$$

É possível mostrar que \mathbf{u} é o vetor valor médio do *v.e.a.r.* (X_1, X_2, \dots, X_p) e que Σ é a respectiva matriz das covariâncias, i.e,

$$\sigma_{ij} = \text{Cov}(X_i, X_j), \quad i, j \in \{1, \dots, p\}.$$

6. Algumas leis de probabilidade sobre $(\mathbb{R}^n, \mathcal{B}_n(\mathbb{R}))$ mais conhecidas

Exemplo/Exercício da Folha Prática 8:

Sejam X_1, X_2, \dots, X_p v.a.r.'s independentes e tais que

$$X_i \sim N(\mu_i, \sigma_i^2), \quad i = 1, \dots, p.$$

Mostre que o *v.e.a.r.* (X_1, X_2, \dots, X_p) é tal que

$$(X_1 \ X_2 \ \dots \ X_p) \sim N_p(\mathbf{u}, \Sigma),$$

com $\mathbf{u} = [\mu_1 \ \mu_2 \ \dots \ \mu_p]^\top$ e $\Sigma = Diag(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$.