

MPEI

Variáveis aleatórias multidimensionais

Motivação

- Trabalhamos frequentemente com grupos de variáveis relacionadas

120 million photoreceptors



256 EEG sensors



- Exemplos:
 - Peso e altura das pessoas
 - Número de temporais em vários meses

$X1$ = número de temporais em Junho (0, 1, ou 2)

$X2$ = número de temporais em Julho (0, 1, ou 2)

Variáveis aleatórias multidimensionais

- Frequentemente temos situações em que os resultados possíveis são conjuntos de várias variáveis aleatórias, X_1, X_2, \dots
- Dois tipos de casos:
 - Experiência aleatória produz várias saídas
 - Repetições da experiência aleatória (com uma única saída)
- A um vector n-dimensional em que as componentes são as variáveis aleatórias X_1, X_2, \dots, X_n chama-se **vector aleatório** ou v.a. Vectorial

$$\mathbf{X} = (X_1 \quad X_2 \quad \dots \quad X_n)$$

Vector aleatório

- **Um vector aleatório** X é uma função que atribui um vector de números reais a todos os resultados ζ em S , o espaço de amostragem da experiência aleatória.
- Exemplo: $\mathbf{X} = (H \ W \ A)$ com
$$H(\zeta) = \text{altura do estudante } \zeta \text{ em metros,}$$
$$W(\zeta) = \text{peso do estudante } \zeta \text{ em Kg, e}$$
$$A(\zeta) = \text{idade do estudante } \zeta \text{ em anos.}$$

Como caracterizar estas variáveis aleatórias com n -dimensões ?

Funções de distribuição conjuntas

- Para lidar com estas situações envolvendo 2 ou mais variáveis, definem-se, estendendo as definições para uma variável:
 - Função massa de probabilidade conjunta
 - Função de distribuição cumulativa conjunta
 - Função de densidade de probabilidade conjunta

Função probabilidade de massa conjunta

- Para duas variáveis discretas, X e Y :

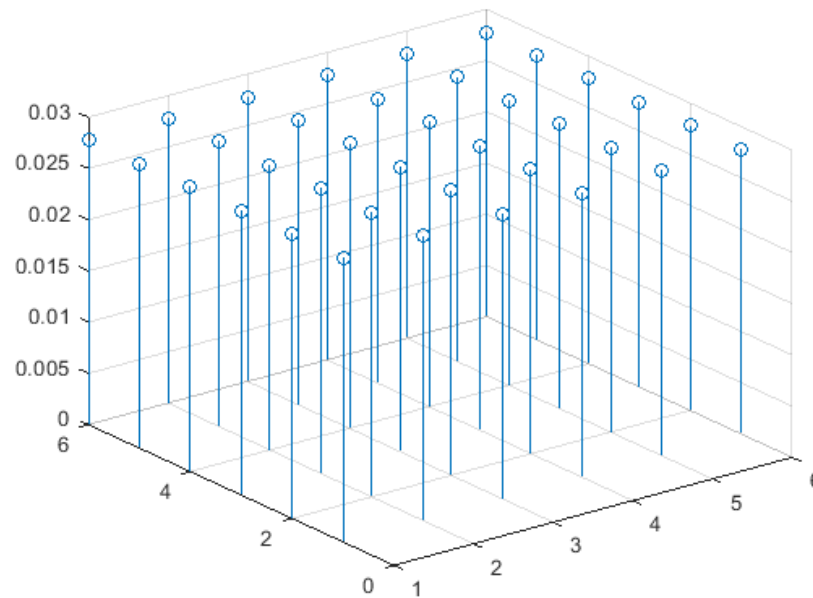
- $p_{X,Y}(i, j) = P(X = i \wedge Y = j)$

- Exemplo: X = dado 1; Y = dado 2

$$p_{X,Y}(1,1) = p_{X,Y}(1,2) = \dots = p_{X,Y}(6,6) = 1/36$$

Exemplo (continuação)

- Representação 3D



Função massa de probabilidade conjunta

- A expressão **generaliza para mais de 2 variáveis**:
- $p_{X_1, X_1, \dots, X_n}(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$
- Uma função em \mathbb{R}^n , não-negativa
- $\sum_{x_1, x_2, \dots, x_n} p_{X_1, X_1, \dots, X_n}(x_1, x_2, \dots, x_n) = 1$

função de distribuição acumulada conjunta

- Tal como no caso escalar, pode definir-se uma função de distribuição acumulada conjunta
 - Simples extensão

- Para duas variáveis, X e Y :

$$F_{X,Y}(x, y) = P(X \leq x \wedge Y \leq y)$$

- Para n variáveis:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n).$$

No caso discreto é uma função em terraços ...

Exemplo 1

- Caso discreto

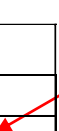
$Y_1 =$ número de temporais em Junho (0, 1, ou 2)

$Y_2 =$ número de temporais in Julho (0, 1, ou 2)

- Tabela com probabilidades

		Julho (y_2)		
Junho (y_1)		0	1	2
	0	0.05	0.1	0.15
	1	0.1	0.15	0.20
	2	0.15	0.05	0.05

$p_{y_1 y_2} (0, 2)$



Distribuição de cada uma das variáveis

- A distribuição de cada uma das variáveis **pode ser obtida da distribuição conjunta**
- Por exemplo, no caso com duas variáveis X e Y :
 - $F_X(a) = P(X \leq a)$
 - $= P(X \leq a, Y < \infty)$
 - $= F_{X,Y}(a, \infty)$
- De forma similar:
 - $F_Y(b) = P(Y \leq b) = F_{X,Y}(\infty, b)$

Funções de probabilidade marginais

- Também se pode obter facilmente a função de massa de probabilidade de cada uma das variáveis
- As fórmulas para o caso discreto são:
- $p_X(x) = \sum_y p_{X,Y}(x, y)$
- $p_Y(y) = \sum_x p_{X,Y}(x, y)$

Funções de probabilidade marginais

- No caso de duas variáveis (X e Y):
- Para obter a função massa de probabilidade de X somamos as linhas apropriadas da tabela representando a função de probabilidade conjunta
- De forma similar obtém-se Y somando as colunas

Exemplo 1

- Para o exemplo introduzido antes..

		Julho (y_2)			
Junho (y_1)		0	1	2	$p(y_1)$
	0	0.05	0.1	0.15	0.30
	1	0.1	0.15	0.20	0.45
	2	0.15	0.05	0.05	0.25
	$p(y_2)$	0.30	0.30	0.40	1.00

$$p_{Y1}(y1) =$$

y_1	$p_{Y1}(y_1)$
0	0.30
1	0.45
2	0.25
TOTAL	1.00

y_2	$p_{Y2}(y_2)$
0	0.30
1	0.30
2	0.40
TOTAL	1.00

Generalização

- O caso de n variáveis discretas é uma generalização simples
- Se X_1, X_1, \dots, X_n são variáveis aleatórias discretas no mesmo espaço de amostragem com função massa de probabilidade conjunta:

$$p_{X_1, \dots, X_n}(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n)$$

- A função de probabilidade marginal para X_1 é:

$$p_{X_1}(x_1) = \sum_{x_2, \dots, x_n} p_{X_1, \dots, X_n}(x_1, \dots, x_n)$$

- A função (bidimensional) para a função de probabilidade marginal de X_1 e X_2 :

$$p_{X_1 X_2}(x_1, x_2) = \sum_{x_3, \dots, x_n} p_{X_1, \dots, X_n}(x_1, x_2, x_3, \dots, x_n)$$

Independência

- Duas variáveis aleatórias X e Y são independentes se, para qualquer a, b se verificar
- $P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$
- Ou seja, são independentes se os eventos $E_a = \{X \leq a\}$ e $E_b = \{Y \leq b\}$ são independentes

Independência

- Em termos de função de distribuição acumulada conjunta:

X e Y são independentes **se e só se**

$$F_{X,Y}(a, b) = F_X(a)F_Y(b)$$

qualquer que sejam a e b

- Também, no caso discreto, X e Y são independentes **se e só se**

$$p(x, y) = p_X(x) p_Y(y)$$

- E no caso contínuo $f_{XY}(x, y) = f_X(x) f_Y(y)$

Generalização – independência de n variáveis aleatórias

- n variáveis X_1, X_2, \dots, X_n são independentes se

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

Exemplo 1

- Y_1 e Y_2 são independentes ?

		Julho (y_2)			
Junho (y_1)		0	1	2	$P(y_1)$
	0	0.05	0.1	0.15	0.30
	1	0.1	0.15	0.20	0.45
	2	0.15	0.05	0.05	0.25
	$f(y_2)$	0.30	0.30	0.40	1.00

		Julho (y_2)			
Junho (y_1)		0	1	2	$p(y_1)$
	0	0.09			0.30
	1				0.45
	2				0.25
	$p(y_2)$	0.30	0.30	0.40	1.00

Esperança matemática

Extensão das definições

- Os momentos de ordem j, k das variáveis X, Y definem-se como sendo,

- Caso discreto:

$$E[X^j Y^k] = \sum_m \sum_n x_m^j y_n^k p_{XY}(x_m, y_n)$$

- Caso contínuo:

$$E[X^j Y^k] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^j y^k f_{XY}(x, y) dx dy$$

- Se $j=1$ e $k=0$ ou $j=0$ e $k=1$ temos os valores médios de X e Y
- Se $j=2$ e $k=0$ ou $j=0$ e $k=2$ temos os valores quadráticos médios

...

- Os momentos centrais conjuntos de ordem j, k das variáveis X, Y definem-se como:

$$E[(X - E[X])^j (Y - E[Y])^k]$$

- Para $j=2$ e $k=0$ ou $j=0$ e $k=2$ obtemos as variâncias de X e Y

Correlação

- O momento de ordem $j=k=1$, $E[XY]$, é designado de **correlação** das variáveis X e Y
- Quando $E[XY] = 0$ as variáveis são **ortogonais**

$E[XY]$ e Independência

- Sendo X e Y independentes

$$E[XY] = E[X]E[Y]$$

- Demonstração (caso discreto):

$$E[XY] = \sum_{x,y} xy p_{X,Y}(x, y)$$

$$= \sum_{x,y} xy p_{X,Y}(x, y)$$

$$= \sum_{x,y} xy p(x)p_Y(y)$$

$$= [\sum_x x p_X(x)] [\sum_y y p_Y(y)]$$

$$= E[X]E[Y]$$

Covariância

- A **covariância** de duas variáveis X e Y é o seu momento central de ordem $j = k = 1$
 - Ou seja $E[(X - E[X])(Y - E[Y])]$
 - Designa-se por $\text{Cov}(X, Y)$
- $$\begin{aligned}\text{Cov}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY - XE[Y] - YE[X] + E[X]E[Y]] \\ &= E[XY] - 2E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y]\end{aligned}$$
- $E[X] = 0$ ou $E[Y] = 0 \quad \Rightarrow \quad \text{Cov}(X, Y) = E[XY]$

Covariância

- É uma generalização da Variância

$$\begin{aligned} Cov(X, X) &= E[(X - E[X])(X - E[X])] \\ &= Var(X) \end{aligned}$$

- A covariância é uma medida de relação linear entre as variáveis aleatórias
- Se a relação for não linear, a covariância pode não ser sensível à relação.

Covariância e independência

- Se X e Y são independentes então $\text{Cov}(X, Y) = 0$
- “Demonstração”:
- Como vimos $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$
- X e Y são independentes implica
$$E[XY] = E[X]E[Y]$$

Nota: o contrário não é verdadeiro

pode ter-se $\text{Cov}(X, Y) = 0$ e as variáveis não serem independentes

Propriedades da Covariância

- $Cov(X, X) = Var(X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(cX, Y) = c Cov(X, Y)$
- $Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$

Demonstração:

$$\begin{aligned} &= E[X(Y + Z)] - E[X]E[Y + Z] = \\ &= E[XY] + E[XZ] - E[X]E[Y] - E[X]E[Z] \\ &= E[XY] - E[X]E[Y] + E[XZ] - E[X]E[Z] \\ &= Cov(X, Y) + Cov(X, Z) \end{aligned}$$

- Generalização:
$$Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j)$$

Covariância de n variáveis

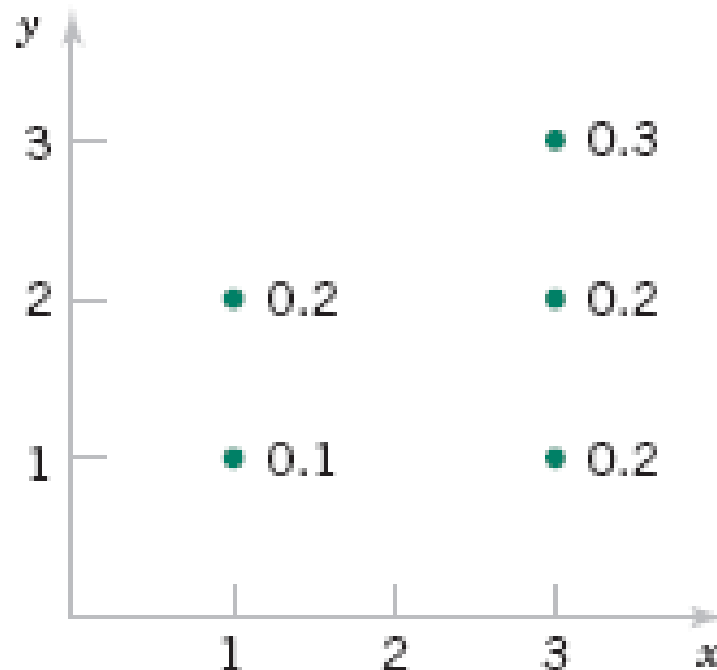
- Se tivermos um vector de n variáveis aleatórias $Y = (Y_1, Y_2, \dots, Y_n)$

- $$\text{Cov}(Y) = \begin{bmatrix} \text{Cov}(Y_1, Y_1) & \cdots & \text{Cov}(Y_1, Y_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(Y_n, Y_1) & \cdots & \text{Cov}(Y_n, Y_n) \end{bmatrix}$$

- $$= \begin{bmatrix} \text{Var}(Y_1) & \cdots & \text{Cov}(Y_1, Y_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(Y_1, Y_n) & \cdots & \text{Var}(Y_n) \end{bmatrix}$$

Exemplo

- Considere a seguinte distribuição conjunta de X e Y e calcule $Cov(X, Y)$



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

- $E(X) = ?$
 $= 1 \times 0,3 + 3 \times 0,7 = 2,4$
- $E(Y) = ?$
 $= 1 \times 0,3 + 2 \times 0,4 + 3 \times 0,3 = 2,0$
- $\text{Cov}(X,Y) = E[(X-E[X]) (Y-E[Y])]$
- $= (1-2,4)(1-2,0) \times 0,1 + (1-2,4)(2-2,0) \times 0,2$
- $+ (3-2,4)(1-2,0) \times 0,2 + (3-2,4)(2-2,0) \times 0,2$
- $+ (3-2,4)(3-2,0) \times 0,3 = 0,2$

Coeficiente de correlação

- A **coeficiente de correlação** de duas variáveis X e Y é:

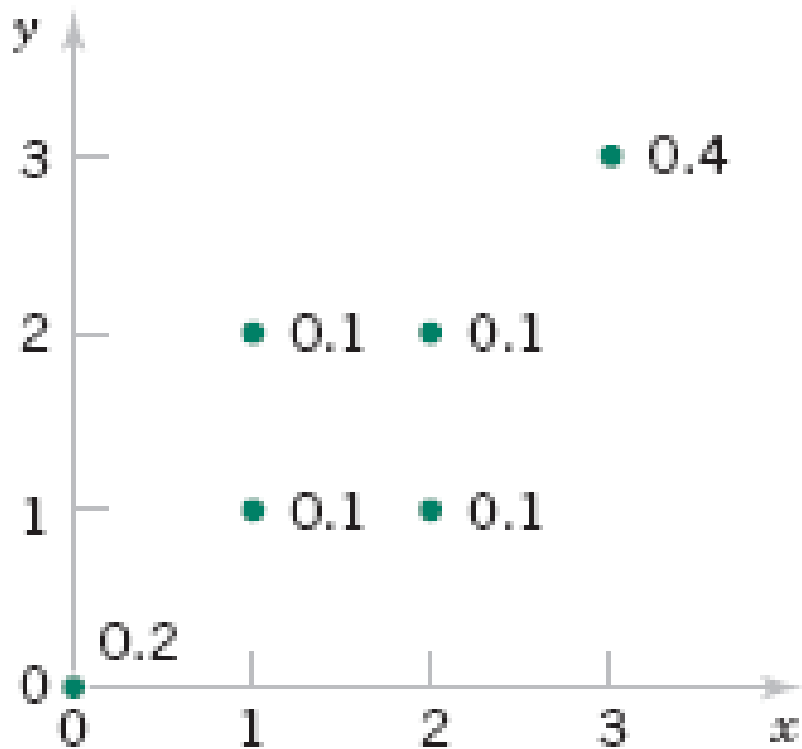
$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

- Demonstra-se que $-1 \leq \rho_{XY} \leq 1$
- E que os valores extremos (1 e -1) se obtém para a relação linear $Y = a X + b$ com $a > 0$ ou $a < 0$, respectivamente

Coeficiente de correlação

- Se $\rho_{XY} = 0$ as variáveis dizem-se **descorrelacionadas**
- Como se viu, se X e Y são independentes, a sua covariância é nula e portanto são descorrelacionadas
 - Mas o contrário não é verdadeiro

Exemplo de cálculo de ρ_{XY}



x	y	P(x,y)
0	0	0,2
1	1	0,1
1	2	0,1
2	1	0,1
2	2	0,1
3	3	0,4
	SOMA	1,0

Cálculo de $E[XY]$, $E[X]$ e $E[Y]$

x	y	P(x,y)	xy P(x,y)	x P(x)	y P(y)	$x^2 P(x)$
0	0	0,2	0x0x0,2=0	0	0	0
1	1	0,1	1x1x0,1=0,1	0,1	0,1	0,1
1	2	0,1	0,2	0,1	0,2	0,1
2	1	0,1	0,2	0,2	0,1	0,4
2	2	0,1	0,4	0,2	0,2	0,4
3	3	0,4	3,6	1,2	1,2	3,6
	SOMA	1,0	4,5	1,8	1,8	4,6

Exemplo de cálculo de ρ_{XY}

- $\text{Var}(X) = E[X^2] - (E[X])^2 = 4,6 - 3,24 = 1,36$
- $\text{Var}(Y)$ é igual à de X
- $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$
- $= 4,5 - (1,8)(1,8) = 1,26$
- Finalmente:
- $\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{1,26}{(\sqrt{1,36})(\sqrt{1,36})} = 0,926$